



## OPEN ACCESS

## EDITED BY

Eleonore Fröhlich,  
Medical University of Graz, Austria

## REVIEWED BY

Janez Mavri,  
National Institute of Chemistry, Slovenia  
Pablo Andrei Nogara,  
Federal University of Santa Maria, Brazil  
Qin Xu,  
Shanghai Jiao Tong University, China

## \*CORRESPONDENCE

Mercedes Alfonso-Prieto,  
✉ m.alfonso-prieto@fz-juelich.de

## SPECIALTY SECTION

This article was submitted to  
Predictive Toxicology,  
a section of the journal  
Frontiers in Pharmacology

RECEIVED 16 December 2022

ACCEPTED 17 February 2023

PUBLISHED 02 March 2023

## CITATION

Mueller NPF, Carloni P and  
Alfonso-Prieto M (2023), Molecular  
determinants of acrylamide neurotoxicity  
through covalent docking.  
*Front. Pharmacol.* 14:1125871.  
doi: 10.3389/fphar.2023.1125871

## COPYRIGHT

© 2023 Mueller, Carloni and Alfonso-  
Prieto. This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License](#)  
(CC BY). The use, distribution or  
reproduction in other forums is  
permitted, provided the original author(s)  
and the copyright owner(s) are credited  
and that the original publication in this  
journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Molecular determinants of acrylamide neurotoxicity through covalent docking

Nicolas Pierre Friedrich Mueller<sup>1,2</sup>, Paolo Carloni<sup>1,3</sup> and Mercedes Alfonso-Prieto<sup>1,4\*</sup>

<sup>1</sup>Institute for Advanced Simulations IAS-5, Institute of Neuroscience and Medicine INM-9, Computational Biomedicine, Forschungszentrum Jülich, Jülich, Germany, <sup>2</sup>Faculty of Mathematics and Natural Sciences, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, <sup>3</sup>Department of Physics, RWTH Aachen University, Aachen, Germany, <sup>4</sup>Cécile and Oskar Vogt Institute for Brain Research, Medical Faculty, University Hospital Düsseldorf, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

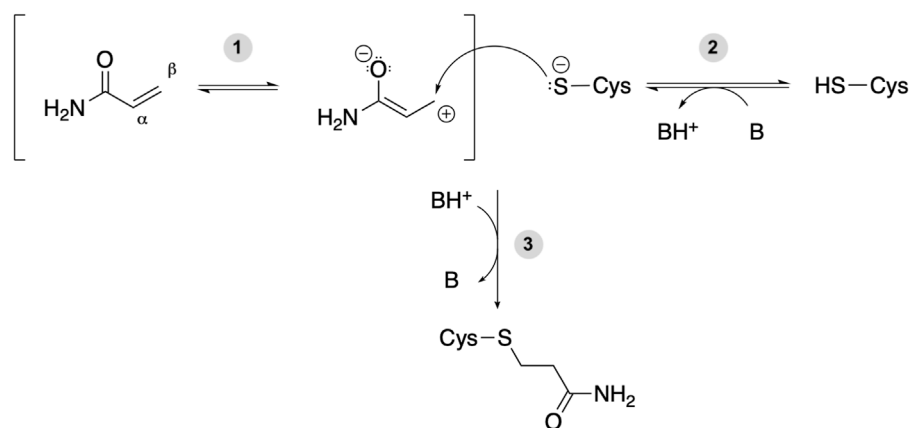
Acrylamide (ACR) is formed during food processing by Maillard reaction between sugars and proteins at high temperatures. It is also used in many industries, from water waste treatment to manufacture of paper, fabrics, dyes and cosmetics. Unfortunately, cumulative exposure to acrylamide, either from diet or at the workplace, may result in neurotoxicity. Such adverse effects arise from covalent adducts formed between acrylamide and cysteine residues of several neuronal proteins *via* a Michael addition reaction. The molecular determinants of acrylamide reactivity and its impact on protein function are not completely understood. Here we have compiled a list of acrylamide protein targets reported so far in the literature in connection with neurotoxicity and performed a systematic covalent docking study. Our results indicate that acrylamide binding to cysteine is favored in the presence of nearby positively charged amino acids, such as lysines and arginines. For proteins with more than one reactive Cys, docking scores were able to discriminate between the primary ACR modification site and secondary sites modified only at high ACR concentrations. Therefore, docking scores emerge as a potential filter to predict Cys reactivity against acrylamide. Inspection of the ACR-protein complex structures provides insights into the putative functional consequences of ACR modification, especially for non-enzyme proteins. Based on our study, covalent docking is a promising computational tool to predict other potential protein targets mediating acrylamide neurotoxicity.

## KEYWORDS

acrylamide, covalent adduct, Michael addition reaction, neurotoxicity, covalent docking

## 1 Introduction

Acrylamide (CH<sub>2</sub> = CH-C(O)NH<sub>2</sub>, PubChem CID 6579) is used in variety of industrial processes, including water waste treatment, manufacture of paper, fabrics, dyes or cosmetics (Swan et al., 2007; Pennisi et al., 2013; Bušová et al., 2020). In addition, it is a by-product of the food industry, formed by Maillard reaction of reduced sugars and amino acids (Mottram et al., 2002) and present in food items processed at high temperatures (e.g. coffee, french fries and baked and roasted potatoes) (Reynolds, 2002; Guenther et al., 2007; Schouten et al., 2020). Due to the potential toxic effects of acrylamide in the human body (Semla et al., 2017; Kumar et al., 2018), in 2018 new



**FIGURE 1**

Michael addition reaction between acrylamide and a Cys residue of a target protein. B and BH<sup>+</sup> represent a Brønsted-Lowry acid-base pair, either a protein residue or a water molecule. (1) The electron-withdrawing effect of the carbonyl group makes the β-carbon of acrylamide an electrophilic site. (2) The deprotonated cysteine side chain can act as nucleophile and react with a soft electrophile such as acrylamide. (3) Michael addition reaction between a protein thiolate and acrylamide yields a covalent adduct potentially affecting protein function.

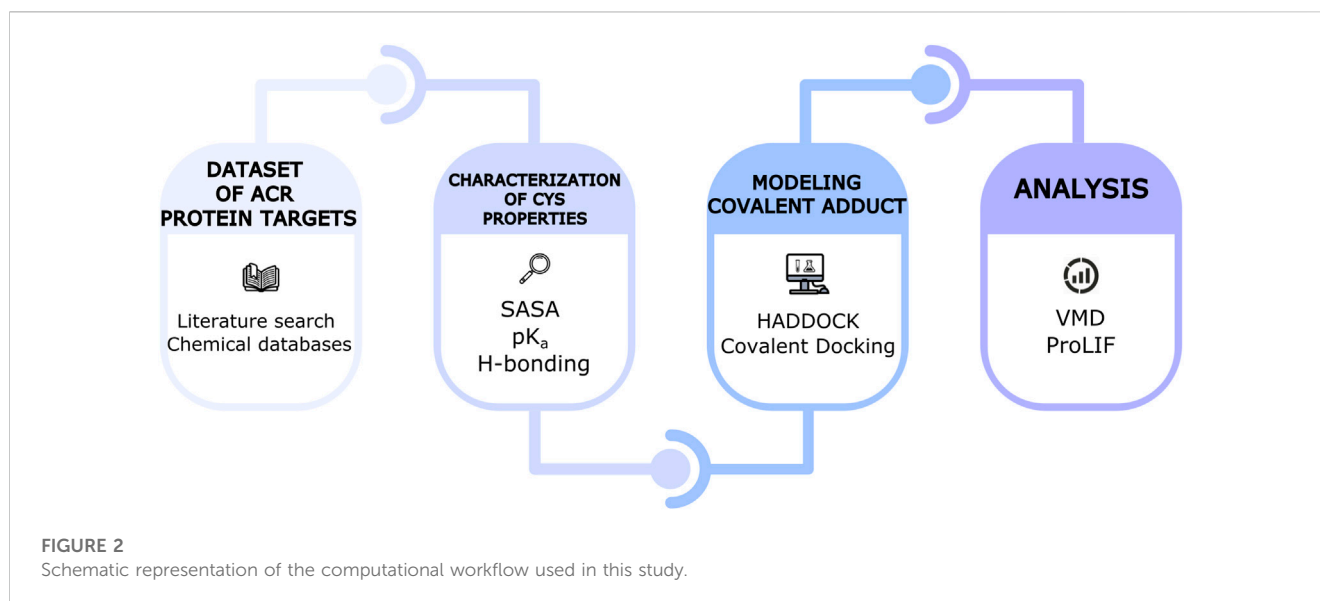
European Union wide regulations entered into force (EU Commission, 2017) to prevent and/or reduce acrylamide formation in foodstuffs, e.g. during frying, baking or roasting.

Cumulative exposure to acrylamide, either from diet or at the workplace, may result in toxicity, especially at the level of the central nervous system. Animal and clinical studies suggest that acrylamide neurotoxicity could mimic the symptoms or even contribute to the etiology of neurodegenerative disorders like Parkinson's disease (LoPachin and Gavin, 2012; Erkekoglu and Baydar, 2014; Li et al., 2015; Murray et al., 2020), as well as result in depression and anxiety-like behavioral effects (Faria et al., 2018, 2019; Raldúa et al., 2020). Three possible mechanisms have been proposed for acrylamide neurotoxicity: (i) inhibition of fast axonal transport, (ii) alteration of neurotransmitter levels, and (iii) direct inhibition of neurotransmission (LoPachin et al., 2007, 2008; Pruser and Flynn, 2011; LoPachin and Gavin, 2012; Erkekoglu and Baydar, 2014; Semla et al., 2017). In addition, acrylamide has been shown to indirectly increase oxidative stress by depleting the levels of the antioxidant glutathione (Catalgol et al., 2009; Kopanska et al., 2015; Raldúa et al., 2020).

Acrylamide (ACR) contains an α, β-unsaturated carbonyl that acts as an electrophile and thus is able to react with nucleophilic amino acids (Figure 1). The electron-withdrawing effect of the carbonyl group on the alkene makes the β-carbon the most electrophilic site (Figure 1, step 1). Based on the hard and soft acids and bases (HSAB) theory, a soft electrophile as ACR is expected to preferentially react with soft nucleophiles, such as the thiolate group of deprotonated cysteine residues (Koutsidis et al., 2009) (Figure 1, step 2). Although the intrinsic pK<sub>a</sub> of the Cys side chain is 8.6 (and thus it is expected to be present as a thiol at physiological pH), several factors, such as hydrogen bonding or the presence of positively charged amino acids in the vicinity, can decrease its pK<sub>a</sub> value. Therefore, certain protein microenvironments can favor the formation of a negatively charged thiolate (Roos et al., 2013). Michael addition reaction of such deprotonated Cys with acrylamide results in the formation of a

covalent adduct (Figure 1, step 3). The amide group of the covalent adduct is able to act both as hydrogen bond acceptor (*via* the carbonyl group) and donor (through the amino group). Such H-bonds with nearby protein residues will help stabilize the covalent adduct and may result in alterations in protein function.

The molecular determinants of acrylamide reactivity with cysteines in proteins are not completely understood. Mass spectrometry-based proteomics analysis has identified proteins that can be modified upon acrylamide incubation (Barber et al., 2007; Feng and Lu, 2011; Nagashima et al., 2019; Zhao et al., 2019). However, such studies are often performed with high ACR concentrations, resulting in e.g. ACR reaction with N-terminal residues, whose modification is unlikely to alter significantly protein function and thus have toxic effects. *In vitro* and biochemical studies focusing on individual protein targets have pinpointed Cys residues whose modification by ACR may have an impact on protein function. However, such assays are not always readily available, especially for non-enzyme proteins. Screening of large protein databases for ACR targets and prediction of the most likely reactive sites in a given protein could be sped up using computational approaches. In this regard, the webservers Cypreds (Soylu and Marino, 2015), Cpipe (Soylu and Marino, 2017) and pCysMod (Li et al., 2021) have been successfully used to predict Cys reactivity for disulfide bridge formation, metal binding, enzymatic catalysis and/or post-translational modifications. However, the underlying algorithms are not tailored to predict cysteines modified by ACR and do not provide information on the impact of the ACR covalent adduct on protein function. Hence, here we have assessed whether covalent docking could be used as a computational tool to characterize ACR reactivity and its (neuro) toxic effects at the molecular level. In particular, as a follow-up of previous computational works by some of us (Ferreira de Lima and Carloni, 2011; Papamokos et al., 2014), we compiled a list of acrylamide protein targets and performed a systematic molecular study using covalent docking. We focused on neuronal protein targets associated with neurotoxic symptoms of acrylamide, as



well as acrylamide-modified proteins detectable in plasma and liver that can be used as biomarkers to monitor ACR exposure. Based on the analysis of the modeled acrylamide-protein complexes, we conclude that acrylamide modification is favored in the presence of nearby positively charged amino acids, such as lysine and arginine. Most likely, such microenvironment facilitates the Michael addition reaction and stabilizes the resulting adduct, consistently with previous proposals (Dennehy et al., 2006). For proteins with more than one reactive Cys residue, the obtained docking scores were able to discriminate between the primary ACR binding site and secondary sites modified only at high ACR concentrations. Therefore, docking scores emerge as a potential filter to predict Cys reactivity against acrylamide. Finally, inspection of the ACR-protein complex, combined with available experimental information, provided insights into the putative functional consequences of ACR modification, especially for non-enzyme proteins, for which *in vitro* or cellular assays assessing the impact of covalent adduct formation may not be readily available. Therefore, we expect that application of covalent docking to other proteins proposed to be targeted by ACR will help discern the most likely Cys reactive site and unravel the functional consequences of ACR adduct formation.

## 2 Methods

Figure 2 shows the workflow of the present systematic study, which is further explained in the next sections.

### 2.1 Generation of the dataset of acrylamide protein targets

We first searched in the literature for human proteins targeted by acrylamide that have been associated to neurotoxic effects (Ferreira de Lima and Carloni, 2011; Papamokos et al., 2014) or used to biomonitor ACR exposure (Barber et al., 2001; Basile et al., 2008). In addition, we also included proteins listed as acrylamide

targets in several chemical and toxicology databases, namely ChEMBL (Davies et al., 2015; Mendez et al., 2018) and T3DB (Lim et al., 2009; Wishart et al., 2015), respectively. Our dataset contains 19 proteins; the full list is given in Supplementary Tables S1–S2.

### 2.2 Characterization of cysteine properties

For eight proteins in our dataset, the location of the cysteine residue modified by acrylamide or by closely related electrophilic agents, such as N-ethylmaleimide (NEM), is known (LoPachin and Gavin, 2012). For the other 11 proteins, we employed physicochemical, conservation and functional data to rank the cysteines most likely to react with acrylamide (see Supplementary Material S1). For all the Cys residues present in dataset proteins, we calculated their solvent-accessible surface area (SASA, using the corresponding tool in VMD (Humphrey et al., 1996) with a probe of radius 1.4 Å), as well as predicted their pK<sub>a</sub> values (using the H<sup>++</sup> webserver (Anandakrishnan et al., 2012) with default settings, in particular assuming pH 7). Those Cys residues that are more solvent exposed and/or have more acidic pK<sub>a</sub> values are expected to be more reactive (Marino and Gladyshev, 2011). The resulting SASA and pK<sub>a</sub> values are listed in Supplementary Table S1. Moreover, we inspected UniProt (The UniProt Consortium, 2019) entries in search for functional annotations regarding candidate Cys residues for all proteins within the dataset. On one hand, Cys residues subject to redox post-translational modifications (e.g. sulfoxidation or S-nitrosylation) indicate reactive cysteines whose modification regulates protein function (Marino et al., 2013). On the other, Cys natural variants or Cys site-directed mutants that affect protein function can suggest a link between protein inactivation and acrylamide modification. The corresponding information is included in Supplementary Material S2. Complementary, we also considered whether the structural location of the candidate Cys is near a functionally relevant site of the protein (e.g. an enzyme active site); this information is listed in

Supplementary Table S1. Further details are provided in the Supplementary Material S1.

## 2.3 Modeling of the covalent acrylamide-protein complexes

Acrylamide binding to cysteines of the proteins in Supplementary Table S1 was investigated with a covalent docking approach.

### 2.3.1 Ligand and protein structures

The product of the corresponding Michael addition reaction, i.e. propionamide (Figure 1), was used as ligand. The respective 3D structure was obtained from PubChem (Sunghwan et al., 2019) (CID 6578).

The 3D structures of the human proteins in Supplementary Table S1 were taken from the Protein Data Bank (Berman et al., 2000; Burley et al., 2021). When more than one structure was available, the one at the highest resolution was chosen (see Supplementary Table S1). Protein structures with missing residues were retrieved from the SWISS-MODEL repository (Bienert et al., 2017) or generated with SWISS-MODEL (Waterhouse et al., 2018), by selecting templates structures with the same sequence as the targets. When experimental structures of the human protein were not available, we generated homology models (see Sections 3.2.4–3.2.7). Target-template sequence alignments were obtained with either BLAST (Camacho et al., 2009) or HHblits (Steinegger et al., 2019), as implemented in the SWISS-MODEL webserver. Templates with the highest sequence identity and the highest resolution were selected and models were generated with SWISS-MODEL (Waterhouse et al., 2018) (see Supplementary Table S1). Protein structures were processed with MolProbity (Chen et al., 2010; Williams et al., 2018) to add missing hydrogen atoms, assign histidine protonation states and perform His/Gln/Asn flips, if recommended. The reactive cysteines were modeled as already deprotonated (Foloppe et al., 2001), as expected for the Michael addition reaction to take place (see Figure 1, step 2). Therefore, our computational protocol does not take into account the energetic cost of Cys deprotonation, i.e.  $\Delta G = \ln(10) \times kT \times (pK_a - pH)$ . Moreover, we have assumed a default pH of 7, even though the protein targets in our dataset exhibit different optimal pH ranges (see Supplementary Table S3) and the Michael addition reaction is favored at basic pH (Lutolf et al., 2001; LoPachin et al., 2007; Nair et al., 2014). However, even if the Cys pK<sub>a</sub> (calculated here with the H++ webserver at a default pH of 7) may predict population of the thiolate state smaller than the thiol one, reaction with acrylamide is expected to shift the acid-base equilibrium (step 2 in Figure 1) towards the deprotonated form.

### 2.3.2 Covalent docking protocol

Covalent docking to the reactive cysteine(s) of each target protein was performed using Haddock (version 2.2.) (De Vries et al., 2010; van Zundert et al., 2016). We followed the standard covalent docking protocol of Haddock (HADDOCK developer team, 2018). Such protocol was initially tested for covalent inhibitors of cathepsin K (HADDOCK developer team, 2018) and here we have validated it using experimental protein

structures containing Cys-ACR covalent adducts (see Section 1.2 in Supplementary Material S1). The covalent bond between Cys and the ligand is modeled by scaling down the van der Waals radius of the Cys sulfur atom 10-fold and introducing two distance restraints: (i) between the sulfur atom of the targeted cysteine and the reactive carbon atom of the ligand, set to  $1.8 \pm 0.1$  Å (i.e. the average length of a single C-S bond) and (ii) between the cysteine C<sub>β</sub> atom and the ligand carbon atom adjacent to the reactive carbon, set to  $2.8 \pm 0.1$  Å (i.e. the same as between the C<sub>γ</sub> and C<sub>ε</sub> atoms of methionine, to model the proper angular geometry). The docking procedure (Kurkcuoglu et al., 2018; Koukos et al., 2019) consisted in the following three different stages: (1) A rigid body docking was performed with all geometrical parameters treated as fixed and allowing 180° rotations to generate 1,000 initial poses. After minimization, the best scored 200 poses were selected for further refinement. (2) A semi-flexible simulated annealing simulation (SA) in torsion angle space was applied to introduce gradually flexibility to the system. SA can be further divided into three steps. (2a) First, a rigid body simulated annealing was performed to optimize orientations of the interacting partners. (2b) Then, the system underwent 1000 molecular dynamics (MD) steps from 500K to 50K, with a 2 fs timestep, in which ligand and protein side chain movement was allowed. (2c) Finally, flexibility was introduced to both protein side chains and backbone, besides the ligand. 1000 MD steps (with a 2 fs timestep) were performed with a stepwise temperature decrement from 300K to 50K. It should be noted that flexibility was only applied to the ligand and protein residues within a range of 5 Å. (3) The final stage of the docking protocol was a refinement in explicit water. Namely, three MD-based steps (with a 2 fs timestep) were carried out: (3a) A heating phase of 100 MD steps from 100K to 200K and to 300K, (3b) 1250 MD steps at a constant temperature of 300K, and (3c) a cooling down phase of 500 MD steps to a final temperature of 100K. In stage (3), both ligand and protein were fully flexible, with the exception of protein backbone atoms. The HADDOCK score settings recommended for small molecule docking were used across the whole protocol (Kurkcuoglu et al., 2018; Koukos et al., 2019). The obtained 200 docking poses were clustered based on their positional protein-ligand interface root-mean-square deviation (iL-RMSD) with a cutoff of 1.0 Å. The Haddock score of each cluster was calculated as the average of the top four structures, as done by the Haddock webserver (De Vries et al., 2010; van Zundert et al., 2016), using the equation:  $HADDOCK\ score = 1.0\ E_{vdw} + 0.1\ E_{elec} + 1.0\ E_{desol} + 0.1\ E_{air}$ , where  $E_{vdw}$  and  $E_{elec}$  are the van der Waals and electrostatic intermolecular energies, respectively,  $E_{desol}$  is the desolvation energy and  $E_{air}$  is the distance restraints energy; the weights of the different terms were parameterized for scoring of protein-ligand complexes (Kurkcuoglu et al., 2018; Koukos et al., 2019). Further analysis was performed for the top cluster (i.e. the one with the best average Haddock score) and, if present, also for other clusters with Haddock scores within the standard deviation of the top cluster. For proteins with more than one reactive cysteine, we performed independent dockings for each of the Cys residues. This approximation is valid provided that these cysteines are far enough apart that they (or their ACR covalent adducts) cannot interact with each other. However, in the case of one of the target proteins, creatine kinase (Supplementary Table S1), the two Cys are within 7.1 Å (C<sub>α</sub>-C<sub>α</sub> distance) and thus we also considered the possibility

that the two Cys could be targeted simultaneously by ACR (see [Section 3.2.2](#)). In this case, binding of two ligand molecules at the same time was modeled using the multibody docking approach (Karaca et al., 2010) implemented in HADDOCK. Namely, a so-called molecule interaction matrix is used to define partners that interact with each other. In particular, we defined the subsequent interacting pairs: protein-ACR molecule 1, protein-ACR molecule two and ACR molecule 1-ACR molecule 2.

## 2.4 Structure-based analysis of the modeled acrylamide-protein complexes

Hydrophobic interactions were investigated using VMD (Humphrey et al., 1996) (version 1.9.3.) and in-house scripts. Namely, such contacts were defined as interactions between either of the two carbon atoms of the ligand and “apolar” protein carbon atoms (i.e. with CHARMM-based point charges below 0.15 electrons) located within the distance cutoff of 4.0 Å. The hydrogen bond (HB) interactions with both the amide and the carbonyl group of the ligand were analyzed with ProLIF (Bouysset and Fiorucci, 2021) (version 1.0.0). The donor-acceptor distance cutoff was set to 4.1 Å and the donor-hydrogen-acceptor angle tolerance to at least 100°. Each docking was analyzed separately, regardless of whether the reactive cysteines belong to the same protein or different protein targets. The protein-ligand interaction frequency is calculated as the percentage of poses belonging to the top (best scored) cluster that exhibit such interaction. When additional clusters with HADDOCK scores within the standard deviation of the top cluster are present, their poses were also included in the analysis, but a weighted average of the interaction frequencies was calculated, based on the size of each of the clusters analyzed. 2D representations of the protein-ligand interactions for each of the docking clusters considered were generated using ProLIF (Bouysset and Fiorucci, 2021) (version 1.0.0) and are shown in the [Supplementary Material S1](#).

The covalent docking approach used here aims at predicting the most likely configuration or binding pose of the Cys-acrylamide adduct. However, the Michael addition reaction starts with the deprotonation of the reactive Cys. Hydrogen bonding to the Cys sulfur atom is crucial for thiolate formation and stabilization of the transition state of the subsequent reaction (Mazmanian et al., 2016). Moreover, the Michael addition reaction involves an enolate-type intermediate in which the ligand oxygen atom acquires negative charge (see step 2 in [Figure 1](#)) and thus hydrogen bonding or a positively charged microenvironment could stabilize this intermediate, facilitating adduct formation (Ha et al., 2011; Weber et al., 2013). Hence, we additionally analyzed protein residues either near the reactive Cys (in the initial X-ray structure of the protein target, i.e. before the Michael addition reaction occurs) or ligand (in the best structure of the top docking cluster, i.e. after covalent adduct formation). First, we checked H-bonded protein residues. These could act as potential proton acceptors to deprotonate the Cys sulfur atom or may stabilize the intermediate and/or product of the Michael addition reaction. Next, we visually inspected other nearby protein residues in the binding cavity that could have favorable, yet longer-range, electrostatic effects on thiolate or adduct formation. In particular,

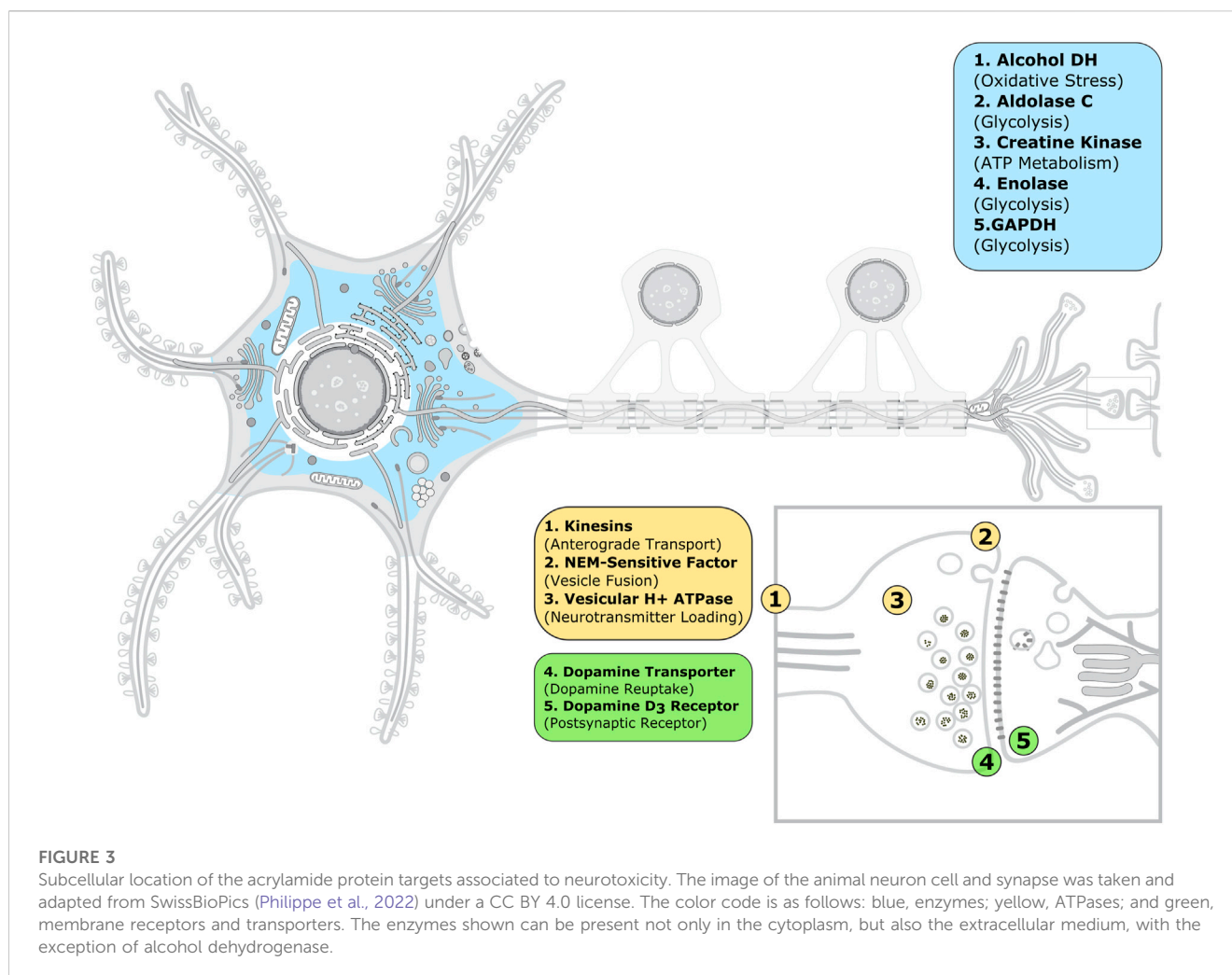
we focused on His, Asp, Glu, Arg and Lys. Histidine is one of the most interesting residues regarding acid-base properties, since its intrinsic  $pK_a$  of ~6 is the closest value to the physiological pH of around 7, as well as to the intrinsic  $pK_a$  of Cys of ~8.6. Hence, the imidazole side chain can be either singly or doubly protonated and thus serve as both proton acceptor and as positively charged residue stabilizing the thiolate formed upon Cys deprotonation. Aspartic and glutamic acids have lower intrinsic  $pK_a$  values (~4.0 and ~4.4); however, their  $pK_a$  can shift to higher values depending on their microenvironment. Hence, Asp and Glu can also be potentially responsible for Cys deprotonation in some cases. Instead, the positively charged Lys and Arg are expected to stabilize the negatively charged thiolate (or the enolate-type intermediate formed during the Michael addition reaction), either by forming a salt bridge or electrostatically. The results of this analysis of the Cys microenvironment are presented in [Supplementary Table S2](#).

## 3 Results

### 3.1 Dataset of acrylamide protein targets

Our literature and chemical database search (see [Section 2.1](#)) rendered a total of 19 proteins modified by acrylamide that have been experimentally validated to mediate ACR (neuro)toxicity or to correlate with ACR exposure as biomarkers. The full list, together with additional protein and candidate Cys information, is given in [Supplementary Tables S1-S3](#). The precise Cys modified by ACR (or closely related sulfhydryl agents, such as NEM) is known for eight of these protein targets and is indicated in bold in [Supplementary Table S1](#). Noteworthy, incubation with high ACR concentrations and/or for longer times can result in more than one Cys being modified for some protein targets (e.g., glyceraldehyde-3-phosphate dehydrogenase and hemoglobin).

Because experimental assays showing ACR-mediated inhibition are often performed *in vitro* using purified proteins, we analyzed next the functional classification and subcellular location of the proteins in our dataset in order to make a connection with the ACR toxic effects observed at the neuronal and systemic levels. Functionally, the ACR protein targets include enzymes (37%), ATPases (21%) and membrane receptors and transporters (16%), as well as plasma proteins (26%); see [Supplementary Table S3](#); [Supplementary Figure S3](#). We speculate that the predominance of enzymes and ATPases is due to the easier availability of purification protocols and functional assays to test ACR-mediated inhibition for these protein classes compared to membrane proteins. Given the diversity of protein classes targeted by ACR, we surmise that several subcellular mechanisms might contribute to acrylamide toxicity. At the neuronal level, ACR-mediated inhibition of cytosolic and extracellular enzymes (in blue in [Figure 3](#)) will cause metabolic imbalance and oxidative stress. In addition, ACR-mediated impaired function of synaptic proteins (ATPases and membrane proteins in yellow and green, respectively, in [Figure 3](#)) will result in inhibition of fast axonal transport (through kinesins), alteration of neurotransmitter levels (dopamine transporter, NEM-sensitive factor and vesicular proton ATPase) and direct inhibition of neurotransmission (dopamine D3 receptor). Therefore, our dataset contains synaptic proteins related to each of the three subcellular mechanisms proposed in reference (Erkekoglu and Baydar, 2014) to explain ACR neurotoxicity, even though the selection of



proteins to include in our dataset was only based on experimental evidence of ACR modification. Among plasma proteins, the levels of ACR-modified albumin and hemoglobin have been used to monitor ACR exposure (Barber et al., 2001; Noort and Hulst, 2003; Basile et al., 2008).

The physiological relevance of ACR modification of the remaining eight proteins in our dataset is less clear. In addition to the genotoxic effects caused by direct reactivity of ACR and its metabolite glycidamide with DNA (Galesa et al., 2008; European Food Safety Authority (EFSA) et al., 2022), ACR-mediated inhibition of DNA topoisomerase and kinesins, as well as of estrogen receptor and sex hormone binding globulin, could further promote carcinogenicity and reproductive toxicity. In the case of the immunoglobulins in our dataset (Supplementary Table S1), we speculate that their ACR-mediated impairment may contribute to the observed ACR immunotoxicity and generation of reactive oxygen species (Kumar et al., 2018).

Interestingly, ACR modification of plasma proteins involves targets with only one free cysteine (for albumin) or cysteines usually involved in disulfide bonds (for sex hormone binding globulin and immunoglobulins, see Supplementary Tables S1-S2). In the latter case, reaction with ACR would require the existence of (at least a small) population of Cys in the free state, besides the

disulfide bond-forming one. We surmise that ACR modification would then shift the equilibrium between the two redox states. In this regard, free thiols and chemical modification of disulfide bridges have been experimentally detected for immunoglobulins (Liu and May 2012).

### 3.2 Acrylamide protein targets with known reactive cysteine

The effects of ACR modification on the 19 proteins in our dataset were further investigated using covalent docking. Considering that some of the ACR protein targets have more than one potential reactive Cys (Supplementary Table S1), 34 covalent docking calculations were performed, following the protocol described in Section 2.3). Below we present the results for the eight ACR protein targets for which the reactive Cys is known (see Table 1), following the alphabetical order of the protein name. For the remaining eleven protein targets, detailed discussion can be found in the next section and in Supplementary Material S1. In all cases, we combined our computational results with previously published experimental data to surmise the possible functional consequences of ACR modification.

**TABLE 1 Covalent docking results for the subset of acrylamide protein targets with experimentally known reactive cysteine.**

For each considered Cys, the Haddock score and size of the top docking cluster are shown. The latter corresponds to the number of docking poses belonging to the top cluster upon clustering of the total 200 poses.

Protein name		Reactive Cys	Score (a.u.)	Size
<i>Albumin</i>		C34	-32.3	63
<i>Creatine Kinase</i>		C74	-21.3	12
		C141	-31.7	63
		C146	-28.2	22
		C254	-23.3	34
		C283	-32.8	61
<i>Dopamine D3 Receptor</i>		C114	-28.6	130
<i>Dopamine Transporter</i>	<i>outward</i>	C342	-2.0	21
	<i>inward</i>	C135	-11.9	79
		C342	-15.1	87
<i>Glyceraldehyde-3-phosphate dehydrogenase</i>		C152	-12.3	163
		C156	46.1	1
		C247	9.9	14
<i>Hemoglobin</i>		C93	-13.9	74
		C104	-3.3	75
<i>NEM-sensitive factor</i>		C264	-73.4	85
<i>Vesicular proton ATPase</i>		C254	-0.1	110

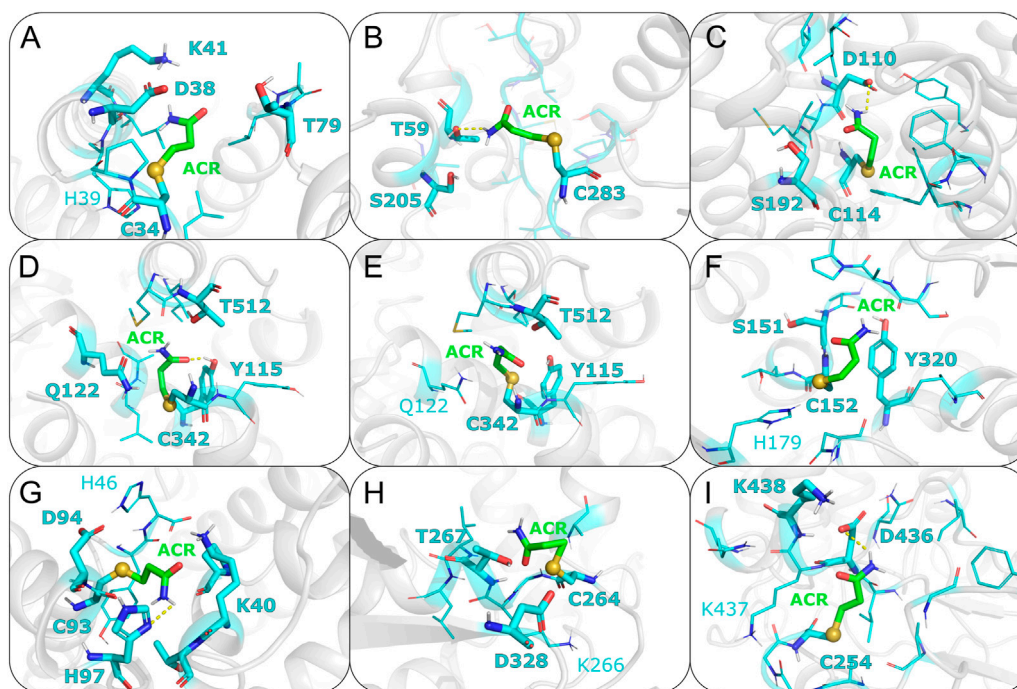
### 3.2.1 Human serum albumin (HSA)

Albumin is a plasma protein able to bind chemically diverse ligands, from hemein and fatty acids to drugs, acting as their plasma carrier/transporter (Fasano et al., 2005). Liquid chromatography–tandem mass spectrometry (LC-tandem MS) experiments have shown that C34 binds covalently acrylamide (Noort and Hulst, 2003; Tong et al., 2004). HSA contains 35 cysteine residues and all form disulfide bridges except C34 (Carter and Ho, 1994). This single free Cys is solvent exposed, with a SASA value of 8.6Å<sup>2</sup>, and has a calculated pK<sub>a</sub> value of 10.2 (see Supplementary Table S1). This is in line with spectroscopic measurements showing HSA Cys34 to be more acidic than a normal Cys, with a pK<sub>a</sub> around 7 (Pedersen and Jacobsen, 1980). The difference between the computational and experimental pK<sub>a</sub> values can be ascribed to the known limitations of computational pK<sub>a</sub> predictors when dealing with Cys residues (Awoonor-Williams and Rowley, 2016), as well as uncertainties in the experimental estimation of pK<sub>a</sub> values using spectroscopic methods. For instance, the pK<sub>a</sub> of Cys34 changes by 1.5 pH units depending on the ionic strength of the buffer used (Pedersen and Jacobsen, 1980). Covalent docking of ACR to C34 resulted in two similar clusters in terms of both score (-31.3 and -32.3 a.u., respectively) and cluster size (69 and 63, see Table 1). Mapping of C34 onto the HSA structure also revealed that this cysteine has two putative proton acceptors in the vicinity (H39 and D38) that can deprotonate the thiol group, as well as a positively charged residue (K41) that could stabilize the transition state and/or product of that reaction (see Figure 4A). Comparison with available functional information (Sampath et al., 2001; Fasano et al., 2005)

suggests that acrylamide covalent binding to Cys34 might affect the drug binding properties of albumin. In particular, infrared spectroscopy has shown that Cys34 is linked allosterically with Sudlow's site I for anesthetics such as halothane, propofol and chloroform (Sampath et al., 2001). Hence, formation of a covalent adduct at Cys34 can be transmitted to this site and modulate anesthetic binding.

### 3.2.2 Creatine kinase (CK)

Creatine kinase is an enzyme responsible for converting creatine to phosphocreatine reversibly using adenosine triphosphate (ATP). CK is inhibited by acrylamide and such inhibition exhibits a biphasic behavior with respect to acrylamide concentration (Sheng et al., 2009), suggesting that more than one Cys residue within CK might be modified. C283 has been proposed as the primary site of ACR modification in CK. Based on site-directed mutagenesis, C283 was shown to be essential for enzymatic activity (Lin et al., 1994). Furthermore, experimental, studies indicated that C283 has a pK<sub>a</sub> around 5.7 and thus this cysteine can be present as thiolate. This is probably required to constrain the position of the guanidinium group of the creatine substrate (Wang et al., 2006). The low pK<sub>a</sub> of C283 and its role in CK enzymatic activity makes it a good candidate for the main reactive Cys targeted by acrylamide. In contrast, the secondary site of ACR modification is unclear. A combined experimental and computational study (Sheng et al., 2009) suggested that acrylamide can bind to C283, as well as the nearby C74, but the results were not conclusive. Therefore, we performed docking for all five solvent exposed cysteines in CK (C74, C141, C146, C254 and C283; see Table 1). C283 is the most



**FIGURE 4**

Representative covalent binding poses of ACR for each of the protein targets discussed in the main text. Acrylamide and its surrounding residues are represented as sticks, with carbon atoms colored in green and cyan, respectively. The sulfur atom between the reactive cysteine residue and the adduct is shown as a sphere. Residues forming hydrogen bonds (HBs) with ACR are displayed with thicker sticks and with bold labels. HBs present in more than 60% of the docking poses are shown with a dashed line. Nearby residues (i.e. within 5 Å) potentially favoring the Michael addition reaction are shown with thinner lines, with positively charged residues and putative proton acceptors labeled. (A) Albumin; (B) Creatine kinase; (C) Dopamine D3 receptor; (D) Dopamine transporter (inward conformation); (E) Dopamine transporter (outward conformation); (F) Glyceraldehyde-3-phosphate dehydrogenase; (G) Hemoglobin; (H) NEM-sensitive factor; (I) Vesicular proton ATPase.

solvent exposed cysteine ( $37.2\text{\AA}^2$ ) and with the lowest predicted  $\text{pK}_a$  value (around 9, see [Supplementary Table S1](#)). Although the calculated  $\text{pK}_a$  of C283 ( $\sim 9$ ) differs from the experimentally measured value (5.7), we ascribed such difference to the known limitations of computational (implicit solvent-based) predictors when estimating the  $\text{pK}_a$  values of Cys residues, with RMSDs between 3.41 and 4.72 pKa units ([Awoonor-Williams and Rowley, 2016](#)). Given this uncertainty, we decided to use the calculated  $\text{pK}_a$  values only to rank by relative acidity the Cys residues within the same protein, i.e. C283 is the most acidic Cys in CK. In addition, the C283 docking yields the best score ( $-32.8$  a.u.), as shown in [Table 1](#). Furthermore, C283 is located in the catalytic site of CK, whereby nearby residues, such as R96 (distance of sulfur atom to  $\zeta$ -carbon atom of  $7.48\text{\AA}$ ), R132 ( $10.71\text{\AA}$ ) and R236 ( $10.70\text{\AA}$ ), can electrostatically stabilize the enolate intermediate of the Michael addition reaction (see [Supplementary Table S3](#)). Instead, docking at C74, previously proposed as acrylamide binding site ([Sheng et al., 2009](#)), gives a less favorable docking score ( $-21.3$  a.u.). Together with most of the C74 poses showing distance values between the ligand  $\text{C}_\beta$  atom and sulfur atom outside the defined covalent bond range, this suggests that modification of C283 is preferred over binding to C74. Due to the proximity of C74 to C283, we also explored the possibility of two acrylamide molecules binding simultaneously to both C283 and C74, using a multibody docking approach ([Karaca et al., 2010](#)). The resulting docking poses indicate that adduct formation

with one acrylamide molecule already occupies fully the pocket lined by C283 and C74 and thus will preclude binding of a second acrylamide molecule (see [Supplementary Figure S12](#)). Thus, C74 is unlikely to be modified by ACR, either alone or in combination with C283.

In contrast, docking to other cysteine residues revealed more suitable candidate for the secondary site of ACR modification in CK. The results for C141 and C146 yielded docking values closer to those of C283 (see [Table 1](#)), suggesting that modification of these two cysteines by acrylamide might be possible. Out of these two cysteines, C141 has a slightly more favorable docking score ( $-31.7$  a.u.) than C146 ( $-28.2$  a.u.), as well as a higher solvent exposed surface area ( $11.1$  compared to  $2.6\text{\AA}^2$ ), suggesting a slightly higher preference of ACR for C141 over C146. Additionally, C141 has two nearby residues (H145 and E150) that could facilitate thiolate and/or adduct formation (see [Supplementary Figures S6 and S7](#)), whereas C146 is hydrogen bonded to P143 (see [Supplementary Figures S6–S11](#)). Based on our covalent docking results and the biphasic time dependent inactivation of CK observed in enzymatic assays ([Sheng et al., 2009](#)), we propose a molecular model in which ACR modification of C283 ([Figure 4B](#)) occurs first and is the primary site responsible for enzyme inactivation. Adduct formation at C283, located in the enzyme active site ([Wang et al., 2006](#)), will hinder creatine binding. At longer times, C141 might also be modified by



ACR, further contributing to enzyme inactivation by thiol depletion (Meng et al., 2001; Lü et al., 2009).

### 3.2.3 Dopamine D3 receptor (D3R)

Acrylamide exposure has been shown to result in decreasing dopamine concentrations by altering postsynaptic dopamine receptors (Erkekoglu and Baydar, 2014). Site-directed mutagenesis data showed that electrophilic compounds, such as NEM, blocked ligand binding to the dopamine D3 receptor (D3R) by modifying C114 (Alberts et al., 2009). Furthermore, functional data compiled in GPCRdb (Kooistra et al., 2020) indicates that C114 is involved in both ligand binding and receptor activation. Taken together the C114 reactivity and functional data, we considered C114 as the most likely candidate for acrylamide modification. Modeling of the covalent C114-ACR adduct further revealed how ACR modification can impair D3R signaling. The ligand interacts with D110 (Figure 4C); this aspartate is essential for ligand binding in aminergic GPCRs (Michino et al., 2015), such as D3R. Together with the aforementioned functional roles of C114 (Kooistra et al., 2020), this indicates that formation of the ACR covalent adduct will hinder ligand binding and/or impair receptor activation. Moreover, Cys at this position (3.36, following the Ballesteros-Weinstein generalized numbering for class A GPCRs) is conserved across dopamine receptors D2, D3 and D4. Given the role of these receptors in dopaminergic neurotransmission, ACR modification of Cys(3.36) might be one of the molecular mechanisms by which ACR intoxication mimics Parkinsonian symptoms.

### 3.2.4 Dopamine transporter (DAT)

Dopamine transporters are integral membrane proteins responsible for regulating dopamine neurotransmitter concentrations at the synaptic cleft (Giros and Caron, 1993). Chemicals such as peroxynitrite and 2-aminoethyl methanethiosulfonate (MTSEA), which have in common the potential to modify cysteine sulphydryls, are known to inhibit DAT (Park et al., 2002). Mutagenesis data has also shown that oxidation of C342 causes a decrease in DAT activity (Park et al., 2002). Furthermore, Cys modification is enhanced if the transporter is in the inward-facing state (Chen et al., 2000). To understand this differential reactivity of the two conformational states of DAT, we performed two covalent dockings for C342, using DAT structures in either outward- and inward-facing conformations (hereafter, OF and IF). Since experimental structural information for human DAT is missing, we generated homology models of the two transporter conformations. The templates used for the OF and IF models were the *Drosophila melanogaster* DAT (PDB code 6M2R) (Pidathala et al., 2021) and the human serotonin transporter (PDB code 6DZZ) (Coleman et al., 2019), respectively. The target-template sequence identities are 56.2% (OF) and 52.4% (IF); thus, the models are expected to be medium-to-high quality (Chothia and Lesk, 1986; Olivella et al., 2013; Piccoli et al., 2013). We further assessed the quality of the models by calculating their Ramachandran plots (Supplementary Figure S1) and QMEANbrane local quality values (Supplementary Figure S2). The percentage of residues in favored/allowed regions is 93.3%/98.7% (OF) and 95.6%/99.4% (IF), whereas the predicted local quality scores are above 0.7 (except for loop regions or not resolved in the template structures). Thus these

two quality assessments support the reliability of the DAT homology models used here. SASA calculations show that C342 is more solvent exposed in the IF model, with SASA values four-fold larger than the OF model (see Supplementary Table S1). Therefore, cysteine accessibility seems to play a role in the observed higher reactivity of ACR with the IF state (Chen et al., 2000). Our covalent docking results (see Table 1) further support the enhanced ACR modification in the IF state. The top cluster for the IF model (Figure 4D) had a more favorable score of  $-15.1$  a.u. than the one ( $-2.0$  a.u.) for the OF model (Figure 4E). Such preferential binding of acrylamide to C342 in the IF state could alter the conformational transition between the two states crucial for dopamine transport, resulting in altered neurotransmitter concentrations. Considering the link between DAT and Parkinson's disease, it is tempting to suggest that this might be responsible, at least in part, for the PD-like symptoms of acrylamide neurotoxicity (McHugh and Buckley, 2015; Jayaramayya et al., 2020).

Besides C342, we also performed covalent docking for C315. Mutagenesis experiments and transport assays upon treatment with sulphydryl reagents have shown that C342 is the main modification site responsible for transport inhibition in wild-type DAT (Chen et al., 2000). However, C315 can also be modified and have a minor contribution to transport inhibition in a DAT C90A/C306A/C319F/C342A mutant construct lacking C342 (Whitehead et al., 2001), with higher C135 accessibility in the IF state (Chen et al., 2000; Whitehead et al., 2001). The docking score for this alternative C135 site ( $-11.9$  a.u.) in the IF state is less favorable than for the main C342 site ( $-15.1$  a.u.), further supporting our proposal that the docking score can help discriminate the most reactive Cys within a given protein target.

### 3.2.5 Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)

GAPDH is a housekeeping enzyme involved in both glycolysis, as well as apoptotic cell signaling. Multiple studies, both experimental and computational, have shown that acrylamide can covalently modify GAPDH, inhibiting enzymatic activity (Tanii and Hashimoto, 1985). Moreover, such enzymatic inactivation is concentration- and time-dependent, as well as pH sensitive (Martyniuk et al., 2011). C152 has been identified as the most reactive cysteine compared to two other solvent exposed cysteine residues, C156 and C247 (Martyniuk et al., 2011). At small concentrations of acrylamide, almost only C152 is modified. As C152 is essential for GAPDH catalysis by acting as nucleophile, formation of the Michael adduct will result in enzyme inhibition (Martyniuk et al., 2011). However, at higher concentrations, ACR adducts with C156 and C247 are also formed and have been shown to further contribute to enzyme inhibition. To identify features that could explain this differential reactivity, we performed covalent docking for each of the aforementioned Cys residues (see Table 1; Figure 4F; Supplementary Figure S17–20). The top docking cluster for residue C152 had score of  $-12.3$  a.u., significantly more favourable than those for C156 and C247 ( $46.1$  and  $9.9$  a.u., respectively). Moreover, the last two dockings showed poses with C-S distances outside the covalent bond range. Therefore, modeling of the Michael adduct indicates that C152 is the primary binding site of ACR in GAPDH, in agreement with experiments (Martyniuk et al., 2011). In addition, both C156 and

C247 had a calculated  $pK_a$  above 12, and thus are less likely to become deprotonated. Instead, the calculated  $pK_a$  value of 6.6 for C152 suggests that the sulfur atom can be present, at least partially, as a thiolate anion. Besides, C152 forms a hydrogen bond with residue H179 (Figure 4F); this could help deprotonate C152. Indeed, H179 activates the thiol group during enzymatic catalysis (Soukri et al., 1989). Moreover, the resulting doubly protonated H179 and its respective positive charge could electrostatically stabilize the enolate-type intermediate of the Michael addition reaction. Furthermore, the SASA values further support C152 as the most reactive residue, since its predicted accessibility is two orders of magnitude higher than C156 and C247 (see Table 1). Altogether, the computational results are in agreement with the experimental evidence that C152 is the primary site of acrylamide modification of GAPDH. Moreover, the higher reactivity of C152 with respect to C156 and C247 seems to correlate with the more favorable score obtained for the first cysteine.

### 3.2.6 Hemoglobin (Hb)

Hemoglobin is a heme-containing protein responsible for oxygen transport from lungs to other tissues. Structurally, Hb is a heterotetramer (Ahmed et al., 2020) formed by two  $\alpha$  and two  $\beta$  subunits that assemble as dimer of dimers ( $\alpha_1\beta_1$  and  $\alpha_2\beta_2$ , respectively). Mass spectrometry showed that C93 within the  $\beta$  chains and C104 in the  $\alpha$  chains are modified by acrylamide, with C93 being the most reactive site (Basile et al., 2008). Hence, measuring the levels of ACR-modified Hb in plasma can be used to monitor acrylamide exposure (Barber et al., 2001; Basile et al., 2008). The two aforementioned reactive cysteines have a predicted  $pK_a$  value above 12. However, C93 is more solvent-exposed compared to C104 (see Supplementary Table S1), suggesting that C93 is more accessible to acrylamide. This is in line with the covalent docking results obtained here (Table 1). The top docking cluster for C93 shows a significantly better docking score ( $-13.9$  a.u.) compared to  $-3.3$  a.u. for C104. Indeed, docking simulations with C104 did not result in a properly formed covalent bond between Cys and acrylamide, which could explain the less favorable docking scores compared to C93. Moreover, C93 has three nearby potential proton acceptors, H346, H97 and D294 (Figure 4G), belonging to the same  $\beta_2$  subunit. C93 is also located near K40 of the adjacent  $\alpha_1$  chain, which stabilizes the adduct by forming a hydrogen bond with the ligand oxygen atom (see Supplementary Figures S21–S23). In addition, the nearby positively charged side chain could help stabilize the transient negative charge developed on the ligand oxygen atom during the nucleophilic attack. Instead, C104 only has a single nearby residue, H103, which could deprotonate the thiol group or stabilize the adduct. Taken together, our computational analysis suggests that C93 in the  $\beta$  subunit should be the primary site for acrylamide adduct formation in Hb, whereas binding to C104 in the  $\alpha$  subunit is likely to occur only at higher acrylamide concentrations or longer times, in line with the experimentally observed reactivity (Basile et al., 2008). Moreover, the location of C93 at the interface between the  $\alpha_1$  and  $\beta_2$  suggests that covalent modification of this cysteine by acrylamide could affect Hb function. Oxygen binding to Hb induces changes within this quaternary structure, i.e. a conformational transition from deoxyhemoglobin (T-state) to oxyhemoglobin (R-state). The largest movement occurs between the  $\alpha_1$ C-helix and the  $\beta_2$ FG corner and a smaller change

takes place between the  $\alpha_1$ FG corner and the  $\beta_2$ C-helix (Perutz, 1979). Thus, the aforementioned  $\alpha_1\beta_2$  intersubunit location of C93 might alter the transition from the T to R state triggered by oxygen binding to Hb and/or its cooperativity mechanism. We propose here that the effect of ACR modification on C93 could be tested experimentally, for instance, by measuring oxygen saturation at different oxygen partial pressures, after incubation with acrylamide. In the absence of such experimental validation, the functional impact of ACR modification is partially supported by a previous experimental study showing that covalent modification of C93 and C104 by other (larger) organic compounds prevented formation of the Hb tetramer (Hwang and Greer, 1980).

### 3.2.7 NEM-sensitive factor (NSF)

N-ethylmaleimide(NEM)-sensitive factor is a homohexameric ATPase (Hoyle et al., 1996). In the presynaptic neuron, NSF, together with SNARE proteins, is involved in fusion of neurotransmitter-loaded vesicles with the cell membrane and vesicle recycling and thus is key for synaptic neurotransmission (May et al., 2001). Previous studies indicated that thiol reagents (e.g., NEM or NO) inhibit NSF (Matsushita et al., 2003). Mass spectrometry data showed that acrylamide modifies cysteine sulphydryls, thereby altering the ATPase activity of NSF (Barber and LoPachin, 2004). Moreover, experimental evidence indicates that C264 is a critical residue for NSF function (Zhao et al., 2007). This cysteine is located in a so-called Walker A motif, important to ATP binding and thus for NSF ATPase function.

Due to the lack of experimental structural information for human NSF, we generated a homology model based on a *Cricetulus griseus* template (PDB code 3J94) (Zhao et al., 2015), which has sequence identity of 98.4%. Therefore, the model is expected to be high quality (Chothia and Lesk, 1986; Olivella et al., 2013; Piccoli et al., 2013). Indeed, its Ramachandran plot (Supplementary Figure S1) shows 90.4%/96.5% residues in favored/allowed regions (comparable to the 93.0%/98.0% values, respectively, for the template structure, solved at 4.20Å resolution). Additionally, the QMEANDisCo local quality values (Supplementary Figure S2) are mostly above 0.7 (except for loop regions or not resolved in the template structures), further supporting the quality of the model. Our computational analysis using this homology model showed that, among the potential attachment points of ACR, C264 is both the most susceptible to deprotonation and the most accessible to acrylamide, with a  $pK_a$  value of 7.7 and a SASA value of 85.2Å<sup>2</sup>. The docking results further support the modification of C264 by acrylamide (Figure 4H). Structural inspection of the docking poses revealed that ACR would partially block access to the ATP binding site, thus hindering ATP binding and decreasing NSF activity.

### 3.2.8 Vesicular proton ATPase (v-ATPase)

Filling of the synaptic vesicles with neurotransmitters relies on the proton gradient created by the vesicular proton ATPases (v-ATPases) using ATP. N-ethylmaleimide (NEM), a sulphydryl reagent similar to ACR, reduces H<sup>+</sup> uptake, as well as decreases v-ATPase activity (Shulamit and Sihra, 1989). The modification site responsible for v-ATPase inhibition by sulphydryl reagents (Feng and Forgac, 1992; LoPachin and Barber, 2006) has been proposed to be C254, which is located in a loop segment of the v-ATPase catalytic

subunit, corresponding to the so-called Walker A (GAFGCGKT) motif coordinating ATP binding and hydrolysis. Physicochemical characterization of C254 revealed that this cysteine has a  $pK_a$  value of 10.3 and a SASA value of 11.0 Å, further supporting this particular cysteine as ACR target site. Thus, we performed the corresponding Haddock calculation for C254. The covalent docking poses obtained here (Figure 4I) show that the ligand is placed at the entrance of the active site and thus can hinder ATP binding. Nonetheless, the loop where C254 is located exhibits large rearrangements during the conformational cycle of  $v$ -ATPase (see Supplementary Figure S28) and such structural changes cannot be modeled with covalent docking. Hence, we integrated additional experimental data for validation. Our hypothesis is indirectly supported by the experimental observation that NEM inactivation of  $v$ -ATPase is associated with exposure of a single cysteine residue that can be protected by incubation with nucleotides (Hunt and Sanders, 1996). Moreover, another experimental study showed that C254 modification, either through formation of a disulfide bridge with C532 or through adduct formation with NEM, causes inactivation of the  $v$ -ATPase (Feng and Forgac, 1994). Therefore, we surmise that modification of C254 by ACR might have a similar inhibitory effect.

### 3.3 Prediction of ACR-modifiable Cys candidates with limited experimental information

Besides performing dockings for ACR protein targets with experimentally validated reactive Cys (see previous Sections 3.2.1–3.2.8), we also considered protein targets that are known to be modified by ACR or NEM, but for which the specific cysteine(s) forming the covalent adduct is not known (see Supplementary Table S1). In order to pinpoint the most likely reactive Cys candidates, we first analyzed cysteine properties (SASA and  $pK_a$  values), inspected their microenvironment (since nearby residues can favor Cys deprotonation) and checked for post-translational modifications and Cys conservation (see Section 2.2). Similar criteria have been previously used to predict cysteines potentially involved in disulfide bridges, metal binding, post-translational modifications or catalysis (as nucleophile) (Soylu and Marino, 2015, 2017; Li et al., 2021). For most proteins, this first filtering rendered multiple Cys residues potentially targeted by ACR. Hence, covalent docking was performed for each candidate Cys. Based on the observation that more favorable docking scores appear to correlate with Cys reactivity against ACR within a given protein (see e.g., Sections 3.2.2 and 3.2.5), we used the docking scores as second filter to suggest the most likely Cys to be modified by ACR (see Supplementary Table S1). We would like to emphasize that the Haddock scoring function used here (Kurkcuoglu et al., 2018; Koukos et al., 2019) is not normalized and thus the docking score is used only to predict ACR reactivity for Cys residues belonging to the same protein, but not to compare different proteins. Further details for this subset of proteins with no experimental information on the reactive cysteine(s) are provided in the Supplementary Material S1. The protein-ligand interaction fingerprints of the obtained docking poses are shown in Supplementary Material S1, whereas Supplementary Material S2

includes the average Haddock score and size of the docking clusters, as well as the individual scores for the top four structures of each cluster.

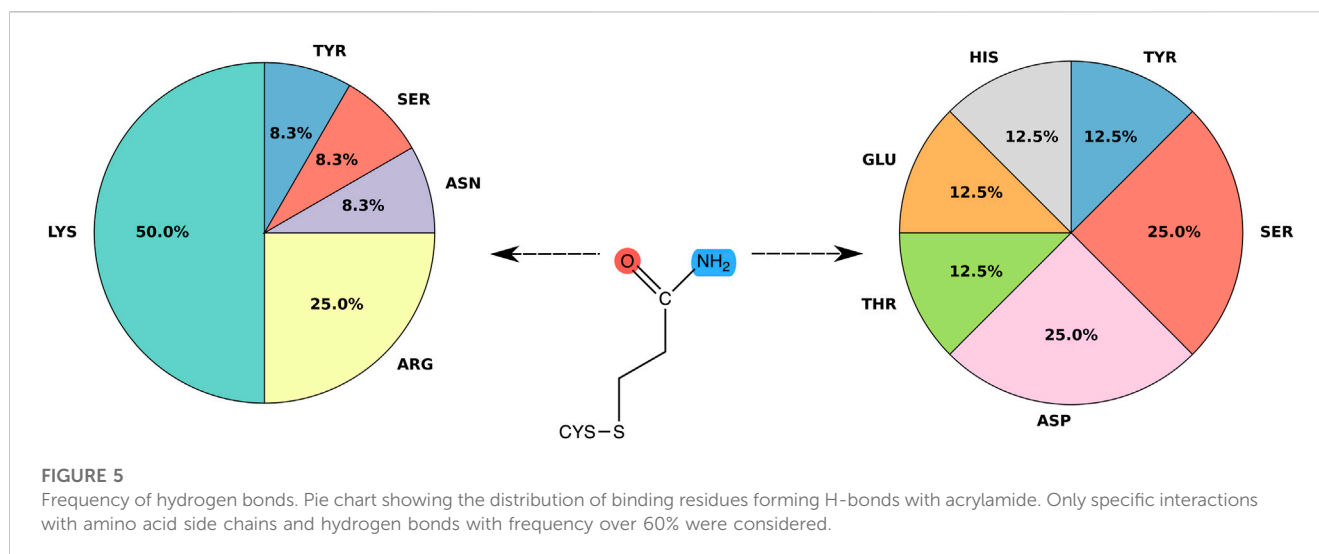
### 3.4 Hydrogen bonds

For each of the considered protein targets in Supplementary Table S1, we analyzed the hydrogen bond (H-bond) network between the ligand and its surrounding binding site residues. The representative clusters of each docking were pooled together and hydrogen bond frequencies were calculated as explained in the Methods Section 2.

Figure 5 shows ligand-protein H-bonds classified by type of amino acid. Only H-bonds with amino acid sidechains and frequency over 60% are displayed; the use of other thresholds turned out not to significantly change the amino acid ranking. Interactions were grouped based on whether the H-bond was formed with either the carbonyl or the amino group of the ligand. Among the H-bonds formed with the carbonyl group, lysine is the residue with the highest frequency (50.0%), followed by arginine (25.0%). After these positively charged residues, asparagine, serine and tyrosine are next in the ranking, with a frequency of 8.3% each. The ligand amino group formed instead H-bonds with polar residues, such as serine (20%), as well as histidine, tyrosine, aspartic acid, threonine and glutamic acid (14.3% each). By grouping amino acids of similar chemical characteristics a more clear picture emerges. In particular, positively charged amino acids, i.e. lysine and arginine, act as main H-bond donors to the carbonyl group of the covalent adduct and have a combined frequency of 75%. Therefore, our structure-based analysis suggests a preference of the ligand carbonyl group to interact with positively charged amino acids, in line with a previous sequence only-based analysis with other thiol-reactive electrophiles (Dennehy et al., 2006). The ligand amino group shows a more diverse picture, in that we did not observe any amino acid preference to interact with the amino group. In particular, the ligand nitrogen does not prefer to interact with negatively charged amino acids (aspartate and glutamate, see Figure 5). We surmise that the specificity of the H-bonds with the carbonyl group with respect to the amino group is a remnant of the role of the H-bond donors in the mechanism of the Michael addition reaction. H-bonding to the carbonyl group does not only contribute to stabilize the resulting Michael adduct, but can also help to stabilize the negative charge developed in the enolate-type reaction intermediate.

## 4 Discussion

Cysteine residues are one of the least frequent (3.3%) proteinogenic amino acids (King and Jukes, 1969). Nevertheless, this residue is disproportionately involved in a variety of important protein functions due to the nucleophilic and redox properties of the thiol group (e.g. catalytic or regulatory activities), as well as its ability to form disulfide bridges and bind metal ions. In this regard, proteins within the synaptic vesicle cycle are of interest, since they are considered as cysteine-rich proteins (Calakos and Scheller, 1996).



Acrylamide (ACR) is a toxicant that has been shown to affect protein function by reacting with Cys residues of several protein targets (LoPachin and Barber, 2006; LoPachin et al., 2007, 2008; LoPachin and Gavin, 2012). Such covalent adduct formation proceeds through a Michael addition mechanism and requires the Cys thiol group to be accessible to acrylamide, as well as deprotonated (i.e. thiolate, see Figure 1, step 2), so that it can act as nucleophile (Figure 1, step 3).

Here, we first calculated Cys physicochemical properties to assess whether these intrinsic values would help pinpoint the most reactive cysteine(s) in a given ACR protein target. The solvent accessible surface area (SASA) is used as proxy of the Cys exposure to ligands. However, the SASA values can significantly vary depending on the resolution of the crystal structure (which can affect the accuracy of the position of the Cys side chain) and/or the functional state of the structure chosen (for the analysis for proteins undergoing large conformational changes during their functional cycle). For instance, the accessibility of C342 in DAT is larger in the IF state compared to OF (see Supplementary Table S1). As for the prediction of  $pK_a$  values, popular software packages, such as H++ (Anandakrishnan et al., 2012), PROPKA (Rostkowski et al., 2011) or MCCE2 (Song et al., 2009), are quite successful in predicting the acidity of aspartic or glutamic acid, but their performance for Cys residues is significantly lower (Awoonor-Williams and Rowley, 2016). For instance, C283 in CK has an experimentally validated  $pK_a$  value of  $\sim 5.6$  (Pahari et al., 2019), yet both H++ and PROPKA estimated values of around 9 (Wang et al., 2006). Unfortunately, both experimental and computational approaches show limitations at estimating  $pK_a$  values. Experimentally, NMR is the most commonly used technique for  $pK_a$  estimation; however, such experiments are quite demanding, in that they require recording multidimensional ( $^{13}\text{C}$  and  $^{15}\text{N}$ ) spectra at different pH values (Stivers et al., 1996; Czerwinski et al., 2001; Hass and Mulder, 2015). Hence, PKAD (a database of experimentally measured  $pK_a$  values of ionizable groups in proteins) (Pahari et al., 2019) contains data only for 1,350 residues in 157 wild-type proteins and for 232 residues in 45 mutant proteins; out of these, only 20 values correspond to Cys residues. Computationally, besides implicit

solvent methods such as PROPKA (Rostkowski et al., 2011) and H++ (Anandakrishnan et al., 2012) mentioned here, explicit solvent methods include, among others, constant pH (Harris and Shen, 2019; Aho et al., 2022) or quantum mechanics/molecular mechanics (QM/MM) (Riccardi et al., 2005). Such molecular dynamics-based approaches have been shown to perform better at predicting  $pK_a$  values; however, their high computational cost makes them prohibitive for systematic applications, such as the one presented here. Moreover, even such advanced approaches may have limitations. For instance, convergence problems might arise due to limited sampling of the titratable residue conformations and of the reorganization of the surrounding protein environment and water molecules upon protonation state change. Moreover, in the case of QM/MM, estimating the energetic contribution of proton dissociation from the protein to the bulk solvent is far from trivial (Mangold et al., 2011; Borstnar et al., 2012; Wang et al., 2020). Therefore, in this work we decided to take advantage of a low computational cost approach such as the H++ webserver (Anandakrishnan et al., 2012) and then use the calculated  $pK_a$  values only to rank Cys residues within a given target protein by acidity, rather than considering the predicted absolute values. Moreover, as thiolate formation is favored in the presence of H-bonding and positively charged residues (Roos et al., 2013), we further inspected nearby residues in order to identify possible candidates to increase Cys acidity (such as Lys and Arg) or act as proton acceptors (such as His, Asp and Glu). However, this initial filtering of the candidate Cys residues based on physicochemical properties and microenvironment effects is not able to pinpoint a single ACR target site, but rather helps to discard the least likely Cys sites, as shown in Section 3.3. This is not surprising, because such analysis has been used to predict Cys reactivity in general (see e.g., Soylu and Marino, 2015, 2017; Li et al., 2021)), but does not take into account specific features related to acrylamide reactivity.

In this work, we have compiled a list of protein targets associated to ACR toxicity and biomonitoring of ACR exposure and used a covalent docking approach to model the adduct formed upon Michael addition reaction of acrylamide with Cys residues of these proteins (Supplementary Tables S1–S3). First, we modeled

the Cys-ACR covalent adduct for a set of protein targets for which the most reactive Cys is experimentally verified (i.e. those target proteins with a Cys highlighted in bold font in [Supplementary Table S1](#)). This is the case for (i) C34 of albumin, (ii) C283 of creatine kinase, (iii) C114 of dopamine D3 receptor, (iv) C342 of dopamine transporter, (v) C152 of GAPDH, (vi) C93 of hemoglobin, (vii) C264 of NSF and (viii) C254 of v-ATPase. The covalent docking approach used here, based on scaling down the van der Waals parameters of the Cys sulfur atom and defining two distance restraints, allows to streamline the generation of structural models of the protein-ACR adducts, compared to more computationally intensive approaches, such as QM/MM ([Mondal and Warshel, 2020](#); [Luo, 2021](#); [Mihalovits et al., 2022](#)). Moreover, in the case of creatine kinase, DAT, GAPDH and hemoglobin, we also applied covalent docking to secondary Cys sites shown experimentally to be modified at increasing ACR concentrations or longer incubation times (see [Sections 3.2.2–3.2.6](#)). We found that the docking score was able to discriminate between primary and secondary ACR sites of the aforementioned proteins, as well as the higher reactivity of C342 of the DAT in the IF state compared to the OF one (see [Section 3.2.4](#)). Therefore, we surmised that covalent docking scores can help identify the main Cys reacting with ACR within a given protein and proceeded to apply the same approach to other protein targets associated to ACR toxicity for which the reactive Cys is unknown (see [Supplementary Table S1](#)). Although the aforementioned traditional approaches ([Soylu and Marino, 2015, 2017](#)), based on solvent accessibility, pKa prediction and H-bonding environment, can help pinpoint possible reactive Cys, the combination with covalent docking, as proposed here, can help better discriminate between different cysteines within the protein. Based on our observation that higher Cys reactivity against ACR turned out to correlate with more favorable docking scores, we suggest that the following Cys are modified by ACR (marked with an asterisk in [Supplementary Table S1](#)): (i) C240 of alcohol dehydrogenase, (ii) C134, C239, C268 and C289 of aldolase, (iii), C134 of immunoglobulin kappa light chain, (iv) C398 of enolase, (v) C381 and C530 of estrogen receptor, (vi) C663 of kinesin KIF1C, (vii) C260 and C287 of kinesin KIF2C, (viii) C395 of immunoglobulin G1 H Nie, (ix) C997 of topoisomerase IIa, and (x) C164 and C188 of sex hormone-binding globulin. Nevertheless, further experiments are needed to validate our computational predictions.

Analysis of the residues surrounding the ACR covalent adducts modeled in the present study shows that acrylamide binding sites are enriched in positively charged Arg and Lys residues (see [Figure 5](#)). Thus, our structure-based study confirms the hypothesis put forward in a previous sequence only-based study with other thiol-reactive electrophiles ([Dennehy et al., 2006](#)). In addition, our work shows that residues such as His, Asp or Glu are often found in close proximity of ACR-modified Cys sites. We surmise that the particular amino acid composition of acrylamide binding sites may have catalytic effects on covalent adduct formation, in line with a previous study on model peptide systems ([Lutolf et al., 2001](#); [Nair et al., 2014](#)). Cys deprotonation (step 2 in [Figure 1](#)) may be favored in the presence of positively charged Lys and Arg (which lower the Cys pKa) and His/Asp/Glu residues (which can either further decrease the Cys pKa by H-bonding or act as proton acceptors). Moreover, the Michael addition reaction (step 3 in

[Figure 1](#)) proceeds *via* an enolate-type intermediate, which may be stabilized in the presence of Lys/Arg/His interacting with the negatively charged oxygen atom, thus decreasing the reaction energy barrier.

We would like to note here that, although the docking protocol used here (see [Section 2.3.2](#)) includes a final refinement step in the presence of explicit water molecules, our analysis of the protein-ligand H-bonds was focused on direct interactions only and thus does not include water-mediated interactions. However, water molecules could also play a catalytic role in acrylamide adduct formation. Water can participate in Cys deprotonation (step 2 in [Figure 1](#)), either as proton shuttle between the reactive Cys and a nearby proton acceptor residue or as a base itself. Unfortunately, such role is also difficult to predict on the basis of the experimental protein structures alone, as many show only a limited number of structured water molecules or lack water molecules altogether ([Gnesia and Carugo, 2017](#); [Caldararu et al., 2020](#)). In addition, water can help further stabilize the covalent adduct (formed upon step 3 in [Figure 1](#)) by mediating H-bonds between the adduct and nearby residues, as observed in other protein-ligand complexes ([Breiten et al., 2013](#)). However, here we were interested in predicting the protein binding site determinants of acrylamide reactivity, rather than quantifying the binding energetics.

The covalent protocol used here has two potential limitations. As for the SASA and pK<sub>a</sub> calculations, the docking results might depend on the input protein structures. To minimize this dependency, we chose the highest resolution structure available for each protein target (to minimize possible errors in the accuracy of the position of the Cys side chain) and employed a fully flexible docking approach (to allow the protein environment to adjust to the presence of the ACR adduct). Nonetheless, docking approaches cannot model large conformational changes. Therefore, analysis of proteins undergoing large structural rearrangements during their functional cycle might require previous knowledge on the conformational state preferentially targeted by ACR (as done here for the dopamine transporter, see [Section 3.2.4](#)). Moreover, classical docking scoring functions, such as the one used here (see [Section 2.3.2](#)), have been parameterized to describe noncovalent protein-ligand interactions and thus can have limited accuracy at describing covalent ligands ([Scarpino et al., 2018](#); [Sottriffer, 2018](#)). Nonetheless, such approaches still exhibit a high success rate for certain warhead chemistry and ligand features, in particular the Michael addition reaction and small size ligands ([Scarpino et al., 2018](#)), as it is the case for the Cys-ACR adducts considered here.

In conclusion, the application of covalent docking to ACR protein targets has provided molecular insights into the binding site where the covalent adduct is formed upon Michael addition. Such sites are enriched in Lys and Arg residues and additionally contain H-bonding residues that stabilize the covalent adduct. Docking scores emerge as a predictive tool to pinpoint Cys residues most likely to be modified by ACR within a given protein. Therefore, the computational workflow presented here ([Figure 2](#)) could serve to filter putative ACR protein targets and candidate reactive Cys resulting from mass spectrometry-based proteomics studies and prioritize those that are more likely to be true positives. However, given the limitations of docking, such ranking should be used to guide follow-up validation studies. Mutagenesis and biochemical experiments would help to assess the impact of ACR on protein function and eventually (neuro)toxicity and computational simulations would

provide further insights into the reaction mechanism of ACR modification, as done for other covalent inhibitors (Mondal and Warshel, 2020; Luo, 2021; Mihalovits et al., 2022).

The computational workflow presented here is based on experimental structures from the Protein Data Bank (Berman et al., 2000; Burley et al., 2021). However, recently developed machine learning-based protein structure prediction algorithms (Baek et al., 2021; Jumper et al., 2021; Ahdritz et al., 2022; Lin et al., 2022) could also be used to generate input protein structures. Moreover, here we performed the covalent docking calculations with HADDOCK (Dominguez et al., 2003), because its availability as a webserver (De Vries et al., 2010; van Zundert et al., 2016) and the minimal preparation of the protein structures required makes our workflow accessible to both new and experienced docking users. Nonetheless, processing the large number of possible candidate ACR protein targets and reactive Cys sites emerging from mass spectrometry-based proteomics will require automated covalent docking workflows, which could integrate either HADDOCK or other docking programs (Scarpino et al., 2018), such as GOLD (Borisek et al., 2015) or Schrödinger (Zhu et al., 2014). Upon such covalent docking-based initial screening, the most promising protein targets and reactive Cys sites could be further filtered and analyzed using more computationally intensive QM/MM methods (Luo, 2021; Mihalovits et al., 2022), such as empirical valence bond (Warshel and Levitt, 1976; Warshel, 1991).

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

## Author contributions

NM and MA-P performed research and analyzed data. PC and MA-P designed research. NM and MA-P drafted the initial version of the manuscript. All authors edited and approved the manuscript.

## References

- Ahdritz, G., Bouatta, N., Kadyan, S., Xia, Q., Gerecke, W., O'Donnell, T. J., et al. (2022). OpenFold: Retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization. *bioRxiv*. doi:10.1101/2022.11.20.517210
- Ahmed, M., Ghatge, M., and Safo, M. (2020). Hemoglobin: Structure, function and allostery. *Subcell. Biochem.* 94, 345–382. doi:10.1007/978-3-030-41769-7\_14
- Aho, N., Buslaev, P., Jansen, A., Bauer, P., Groenhof, G., and Hess, B. (2022). Scalable constant pH molecular dynamics in GROMACS. *J. Chem. Theory Comput.* 18, 6148–6160. doi:10.1021/acs.jctc.2c00516
- Alberts, G. L., Pregenzer, J. F., and Im, W. B. (2009). Contributions of cysteine 114 of the human D3 dopamine receptor to ligand binding and sensitivity to external oxidizing agents. *Brit. J. Pharmacol.* 125, 705–710. doi:10.1038/sj.bjp.0702136
- Anandkrishnan, R., Aguilar, B., and Onufriev, A. V. (2012). H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* 40, W537–W541. doi:10.1093/nar/gks375
- Awoonor-Williams, E., and Rowley, C. N. (2016). Evaluation of methods for the calculation of the pK<sub>a</sub> of cysteine residues in proteins. *J. Chem. Theory Comput.* 12, 4662–4673. doi:10.1021/acs.jctc.6b00631
- Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., et al. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 373, 871–876. doi:10.1126/science.abb8754
- Barber, D. S., Hunt, J. R., Ehrlich, M. F., Lehning, E., and LoPachin, R. M. (2001). Metabolism, toxicokinetics and hemoglobin adduct formation in rats following subacute and subchronic acrylamide dosing. *Neurotoxicology* 22, 341–353. doi:10.1016/s0161-813x(01)00024-9
- Barber, D. S., and LoPachin, R. M. (2004). Proteomic analysis of acrylamide-protein adduct formation in rat brain synaptosomes. *Toxicol. Appl. Pharmacol.* 201, 120–136. doi:10.1016/j.taap.2004.05.008
- Barber, D. S., Stevens, S., and LoPachin, R. M. (2007). Proteomic analysis of rat striatal synaptosomes during acrylamide intoxication at a low dose rate. *Toxicol. Sci.* 100, 156–167. doi:10.1093/toxsci/kfm210
- Basile, A., Ferranti, P., Moccaldi, R., Spagnoli, G., and Sannolo, N. (2008). Proteomic approach for the analysis of acrylamide-hemoglobin adducts: Perspectives for biological monitoring. *J. Chromatogr. A* 1215, 74–81. doi:10.1016/j.chroma.2008.10.093
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T., Weissig, H., et al. (2000). The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242. doi:10.1093/nar/28.1.235

## Funding

This work was supported by the “Ernesto Illy Foundation” (Trieste, Italy). Open-access publication was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)-491111487.

## Acknowledgments

This project was proposed by Dr. Ernesto Illy. We are grateful to Mrs. Anna Illy (Ernesto Illy Foundation, Trieste, Italy) and Dr. Luciano Navarini (illy caffè S.p.A, Trieste, Italy) for useful discussions. We thank the Central Library of the Forschungszentrum Jülich for making the open access publication possible.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphar.2023.1125871/full#supplementary-material>

- Bienert, S., Waterhouse, A., de Beer, T. A. P., Tauriello, G., Studer, G., Bordoli, L., et al. (2017). The SWISS-model repository — New features and functionality. *Nucleic Acids Res.* 45, D313–D319. doi:10.1093/nar/gkw1132
- Borisek, J., Vizovisek, M., Sosnowski, P., Turk, B., Turk, D., Mohar, B., et al. (2015). Development of N-(functionalized benzoyl)-homocycloleucyl-glycinonitriles as potent cathepsin K inhibitors. *J. Med. Chem.* 58, 6928–6937. doi:10.1021/acs.jmedchem.5b00746
- Borstnar, R., Repic, M., Kamerlin, S. C. L., Vianello, R., and Mavri, J. (2012). Computational study of the pK<sub>a</sub> values of potential catalytic residues in the active site of monoamine oxidase B. *J. Chem. Theory Comput.* 8, 3864–3870. doi:10.1021/ct300119u
- Bouysset, C., and Fiorucci, S. (2021). ProLIF: A library to encode molecular interactions as fingerprints. *J. Cheminform.* 13, 72–79. doi:10.1186/s13321-021-00548-6
- Breiten, B., Lockett, M. R., Sherman, W., Fujita, S., Al-Sayah, M., Lange, H., et al. (2013). Water networks contribute to enthalpy/entropy compensation in protein–ligand binding. *J. Amer. Chem. Soc.* 135, 15579–15584. doi:10.1021/ja4075776
- Burley, S. K., Bhikadiya, C., Bi, C., Bittrich, S., Chen, L., Crichtow, G. V., et al. (2021). RCSB Protein Data Bank: Powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res.* 49, D437–D451. doi:10.1093/nar/gkaa1038
- Bušová, M., Bencko, V., Veszelits Laktičová, K., Holcátová, I., and Vargová, M. (2020). Risk of exposure to acrylamide. *Cent. Eur. J. Public Health* 28, S43–S46. doi:10.21101/cejph.a6177
- Calakos, N., and Scheller, R. H. (1996). Synaptic vesicle biogenesis, docking, and fusion: A molecular description. *Physiol. Rev.* 76, 1–29. doi:10.1152/physrev.1996.76.1.1
- Caldararu, O., Ignjatovic, M. M., Ryde, E. O. U., and Ryde, U. (2020). Water structure in solution and crystal molecular dynamics simulations compared to protein crystal structures. *RSC Adv.* 10, 8435–8443. doi:10.1039/c9ra09601a
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: Architecture and applications. *BMC Bioinform.* 10, 421. doi:10.1186/1471-2105-10-421
- Carter, D. C., and Ho, J. X. (1994). “Structure of serum albumin,” in *Advances in Protein Chemistry* (Elsevier), 45, 153–203.
- Catalgol, B., Özhan, G., and Alpertunga, B. (2009). Acrylamide-induced oxidative stress in human erythrocytes. *Hum. Exp. Toxicol.* 28, 611–617. doi:10.1177/0960327109350664
- Chen, N., Ferrer, J. V., Javitch, J. A., and Justice, J. B., Jr. (2000). Transport-dependent accessibility of a cytoplasmic loop cysteine in the human dopamine transporter. *J. Biol. Chem.* 275, 1608–1614. doi:10.1074/jbc.275.3.1608
- Chen, V. B., Arendall, W. B., III, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., et al. (2010). MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D.* 66, 12–21. doi:10.1107/S0907444909042073
- Chothia, C., and Lesk, A. (1986). The relation between the divergence of sequence and structure in proteins. *EMBO J.* 5, 823–826. doi:10.1002/j.1460-2075.1986.tb04288.x
- Coleman, J. A., Yang, D., Zhao, Z., Wen, P.-C., Yoshioka, C., Tajkhorshid, E., et al. (2019). Serotonin transporter–ibogaine complexes illuminate mechanisms of inhibition and transport. *Nature* 569, 141–145. doi:10.1038/s41586-019-1135-1
- Czerwinski, R. M., Harris, T. K., Massiah, M. A., Mildvan, A. S., and Whitman, C. P. (2001). The structural basis for the perturbed pK<sub>a</sub> of the catalytic base in 4-oxalocrotonate tautomerase: Kinetic and structural effects of mutations of Phe-50. *Biochemistry* 40, 1984–1995. doi:10.1021/bi0024714
- Davies, M., Nowotka, M., Papadatos, G., Dedman, N., Gaulton, A., Atkinson, F., et al. (2015). ChEMBL web services: Streamlining access to drug discovery data and utilities. *Nucleic Acids Res.* 43, 612–W620. doi:10.1093/nar/gkv352
- De Vries, S. J., Van Dijk, M., and Bonvin, A. M. (2010). The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* 5, 883–897. doi:10.1038/nprot.2010.32
- Dennehy, M. K., Richards, K. A., Wernke, G. R., Shyr, Y., and Liebler, D. C. (2006). Cytosolic and nuclear protein targets of thiol-reactive electrophiles. *Chem. Res. Toxicol.* 19, 20–29. doi:10.1021/tx050312l
- Dominguez, C., Boelens, R., and Bonvin, A. M. (2003). HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J. Amer. Chem. Soc.* 125, 1731–1737. doi:10.1021/ja026939x
- Erkekoglu, P., and Baydar, T. (2014). Acrylamide neurotoxicity. *Nutr. Neurosci.* 17, 49–57. doi:10.1179/1476830513Y.0000000065
- EU Commission (2017). *Commission Regulation (EU) 2017/2158 of 20 November 2017 establishing mitigation measures and benchmark levels for the reduction of the presence of acrylamide in food*. Available at: <http://data.europa.eu/eli/reg/2017/2158/oj> (Accessed December 5th, 2022).
- European Food Safety Authority Benford, D., Bignami, M., Chipman, J. K., and Ramos Bordajandi, L. (2022). Assessment of the genotoxicity of acrylamide. *EFSA J.* 20, e07293.
- Faria, M., Valls, A., Prats, E., Bedrossiantz, J., Orozco, M., Porta, J. M., et al. (2019). Further characterization of the zebrafish model of acrylamide acute neurotoxicity: Gait abnormalities and oxidative stress. *Sci. Rep.* 9, 7075. doi:10.1038/s41598-019-43647-z
- Faria, M., Ziv, T., Gómez-Canela, C., Ben-Lulu, S., Prats, E., Novoa-Luna, K. A., et al. (2018). Acrylamide acute neurotoxicity in adult zebrafish. *Sci. Rep.* 8, 7918. doi:10.1038/s41598-018-26343-2
- Fasano, M., Curry, S., Terreno, E., Galliano, M., Fanali, G., Narciso, P., et al. (2005). The extraordinary ligand binding properties of human serum albumin. *IUBMB Life* 57, 787–796. doi:10.1080/15216540500404093
- Feng, C.-H., and Lu, C.-Y. (2011). Modification of major plasma proteins by acrylamide and glycidamide: Preliminary screening by nano liquid chromatography with tandem mass spectrometry. *Anal. Chim. Acta* 684, 80–86. doi:10.1016/j.aca.2010.10.042
- Feng, Y., and Forgac, M. (1992). Cysteine 254 of the 73-kDa A subunit is responsible for inhibition of the coated vesicle (H<sup>+</sup>)-ATPase upon modification by sulfhydryl reagents. *J. Biol. Chem.* 267, 5817–5822. doi:10.1016/s0021-9258(18)42626-9
- Feng, Y., and Forgac, M. (1994). Inhibition of vacuolar H<sup>+</sup>-ATPase by disulfide bond formation between cysteine 254 and cysteine 532 in subunit A. *J. Biol. Chem.* 269, 13224–13230. doi:10.1016/s0021-9258(17)36822-9
- Ferreira de Lima, E., and Carloni, P. (2011). “Acrylamide binding to its cellular targets: Insights from computational studies,” in *Computational biology and applied bioinformatics*. Editor H. Lopez (Rijeka, Croatia: InTech Europe), 431–442.
- Foloppe, N., Sagemark, J., Nordstrand, K., Berndt, K. D., and Nilsson, L. (2001). Structure, dynamics and electrostatics of the active site of glutaredoxin 3 from *Escherichia coli*: Comparison with functionally related proteins. *J. Mol. Biol.* 310, 449–470. doi:10.1006/jmbi.2001.4767
- Galesa, K., Bren, U., Kranjc, A., and Mavri, J. (2008). Carcinogenicity of acrylamide: A computational study. *J. Agric. Food Chem.* 56, 8720–8727. doi:10.1021/jf800965y
- Giros, B., and Caron, M. G. (1993). Molecular characterization of the dopamine transporter. *Trends Pharmacol. Sci.* 14, 43–49. doi:10.1016/0165-6147(93)90029-j
- Gnesia, M., and Carugo, O. (2017). How many water molecules are detected in X-ray protein crystal structures? *J. Appl. Crystallogr.* 50, 96–101. doi:10.1107/s1600576716018719
- Guenther, H., Anklam, E., Wenzl, T., and Stadler, R. H. (2007). Acrylamide in coffee: Review of progress in analysis, formation and level reduction. *Food Addit. Contam.* 24, 60–70. doi:10.1080/02652030701243119
- Ha, H.-J., Yoon, D.-H., Park, S., and Kim, H.-J. (2011). Fluorescence turn-on probe for biothiols: Intramolecular hydrogen bonding effect on the Michael reaction. *Tetrahedron* 67, 7759–7762. doi:10.1016/j.tet.2011.08.002
- HADDOCK developer team (2018). *HADDOCK covalent docking tutorial*. Available at: [https://www.bonvinlab.org/education/biomolecular-simulations-2018/HADDOCK\\_tutorial/](https://www.bonvinlab.org/education/biomolecular-simulations-2018/HADDOCK_tutorial/) (Accessed December 6th, 2022).
- Harris, R., and Shen, J. (2019). GPU-Accelerated implementation of continuous constant pH molecular dynamics in Amber: pK<sub>a</sub> predictions with single-pH simulations. *J. Chem. Inf. Model.* 59, 4821–4832. doi:10.1021/acs.jcim.9b00754
- Hass, M. A. S., and Mulder, F. A. A. (2015). Contemporary NMR studies of protein electrostatics. *Annu. Rev. Biophys.* 44, 53–75. doi:10.1146/annurev-biophys-083012-130351
- Hoyle, J., Phelan, J. P., Bermingham, N., and Fisher, E. M. C. (1996). Localization of human and mouse N-ethylmaleimide-sensitive factor (NSF) gene: A two-domain member of the AAA family that is involved in membrane fusion. *Mamm. Genome* 7, 850–852. doi:10.1007/s003359900249
- Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: Visual molecular dynamics. *J. Mol. Graph.* 14, 33–38, 27–28. doi:10.1016/0263-7855(96)00018-5
- Hunt, I. E., and Sanders, D. (1996). The kinetics of N-ethylmaleimide inhibition of a vacuolar H<sup>+</sup>-ATPase and determination of nucleotide dissociation constants. *Plant Physiol.* 110, 97–103. doi:10.1104/pp.110.1.97
- Hwang, P. K., and Greer, J. (1980). Interaction between hemoglobin subunits in the hemoglobin-haptoglobin complex. *J. Biol. Chem.* 255, 3038–3041. doi:10.1016/s0021-9258(19)85848-9
- Jayaramayya, K., Iyer, M., Venkatesan, D., Balasubramanian, V., Narayanasamy, A., Subramaniam, M. D., et al. (2020). Unraveling correlative roles of dopamine transporter (DAT) and Parkin in Parkinson's disease (PD) – a road to discovery? *Brain Res. Bull.* 157, 169–179. doi:10.1016/j.brainresbull.2020.02.001
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi:10.1038/s41586-021-03819-2
- Karaca, E., Melquiond, A. S., De Vries, S. J., Kastritis, P. L., and Bonvin, A. M. (2010). Building macromolecular assemblies by information-driven docking: Introducing the HADDOCK multibody docking server. *Mol. Cell. Proteomics* 9, 1784–1794. doi:10.1074/mcp.M000051-MCP201
- King, J. L., and Jukes, T. H. (1969). Non-Darwinian evolution. *Science* 164, 788–798. doi:10.1126/science.164.3881.788

- Kooistra, A. J., Mordalski, S., Pándy-Szekeres, G., Esguerra, M., Mamyrbekov, A., Munk, C., et al. (2020). GPCRdb in 2021: Integrating GPCR sequence, structure and function. *Nucleic Acids Res.* 49, D335–D343. doi:10.1093/nar/gkaa1080
- Kopanska, M., Lukac, N., Kapusta, E., and Formicki, G. (2015). Acrylamide influence on activity of acetylcholinesterase, thiol groups, and malondialdehyde content in the brain of Swiss mice. *J. Biochem. Mol. Toxicol.* 29, 472–478. doi:10.1002/jbt.21717
- Koukos, P. I., Xue, L. C., and Bonvin, A. M. (2019). Protein-ligand pose and affinity prediction: Lessons from D3R grand challenge 3. *J. Comput. Aided Mol. Des.* 33, 83–91. doi:10.1007/s10822-018-0148-4
- Koutsidis, G., Simons, S. P. J., Thong, Y. H., Haldoupis, Y., Mojica-Lazaro, J., Wedzicha, B. L., et al. (2009). Investigations on the effect of amino acids on acrylamide, pyrazines, and Michael addition products in model systems. *J. Agric. Food Chem.* 57, 9011–9015. doi:10.1021/jf9014763
- Kumar, J., Das, S., and Teoh, S. L. (2018). Dietary acrylamide and the risks of developing cancer: Facts to ponder. *Front. Nutr.* 5, 14. doi:10.3389/fnut.2018.00014
- Kurkcuoglu, Z., Koukos, P. I., Citro, N., Trellet, M. E., Rodrigues, J., Moreira, I. S., et al. (2018). Performance of HADDOCK and a simple contact-based protein–ligand binding affinity predictor in the D3R Grand Challenge 2. *J. Comput. Aided Mol. Des.* 32, 175–185. doi:10.1007/s10822-017-0049-y
- Li, J., Li, D., Yan, Y., Xu, T., Li, P., and He, D. (2015). Acrylamide induces locomotor defects and degeneration of dopamine neurons in *Caenorhabditis elegans*. *J. Appl. Toxicol.* 36, 60–67. doi:10.1002/jat.3144
- Li, S., Yu, K., Wu, G., Zhang, Q., Wang, P., Zheng, J., et al. (2021). pCysMod: Prediction of multiple cysteine modifications based on deep learning framework. *Front. Cell Dev. Biol.* 9, 617366. doi:10.3389/fcell.2021.617366
- Lim, E., Pon, A., Djombou, Y., Knox, C., Shrivastava, S., Chi Guo, A., et al. (2009). T3DB: A comprehensively annotated database of common toxins and their targets. *Nucleic Acids Res.* 38, 781–786. doi:10.1093/nar/gkp934
- Lin, L., Perryman, B., Friedman, D., Roberts, R., and Ma, T. S. (1994). Determination of the catalytic site of creatine kinase by site-directed mutagenesis. *Biochem. Biophys. Acta* 1206, 97–104. doi:10.1016/0167-4838(94)90077-9
- Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., et al. (2022). Evolutionary-scale prediction of atomic level protein structure with a language model. *bioRxiv*. doi:10.1101/2022.07.20.500902
- Liu, H., and May, K. (2012). Disulfide bond structures of IgG molecules: Structural variations, chemical modifications and possible impacts to stability and biological function. *mAbs* 4, 17–23. doi:10.4161/mabs.4.1.18347
- LoPachin, R. M., Barber, D. S., and Gavin, T. (2008). Molecular mechanisms of the conjugated  $\alpha$ ,  $\beta$ -unsaturated carbonyl derivatives: Relevance to neurotoxicity and neurodegenerative diseases. *Toxicol. Sci.* 104, 235–249. doi:10.1093/toxsci/kfm301
- LoPachin, R. M., Barber, D. S., Geohagen, B. C., Gavin, T., He, D., and Das, S. (2007). Structure-toxicity analysis of type-2 alkenes: *In vitro* neurotoxicity. *Toxicol. Sci.* 95, 136–146. doi:10.1093/toxsci/kfl127
- LoPachin, R. M., and Barber, D. S. (2006). Synaptic cysteine sulfhydryl groups as targets of electrophilic neurotoxins. *Toxicol. Sci.* 94, 240–255. doi:10.1093/toxsci/kfl066
- LoPachin, R. M., and Gavin, T. (2012). Molecular mechanism of acrylamide neurotoxicity: Lessons learned from organic chemistry. *Environ. Health Perspect.* 120, 1650–1657. doi:10.1289/ehp.1205432
- Lü, Z.-R., Zou, H.-C., Park, S. J., Park, D., Shi, L., Oh, S. H., et al. (2009). The effects of acrylamide on brain creatine kinase: Inhibition kinetics and computational docking simulation. *Int. J. Biol. Macromol.* 44, 128–132. doi:10.1016/j.ijbiomac.2008.11.003
- Luo, Y. L. (2021). Mechanism-based and computational-driven covalent drug design. *J. Chem. Inf. Model.* 61, 5307–5311. doi:10.1021/acs.jcim.1c01278
- Lutolf, M., Tirelli, N., Cerritelli, S., Cavalli, L., and Hubbell, J. (2001). Systematic modulation of Michael-type reactivity of thiols through the use of charged amino acids. *Bioconj. Chem.* 12, 1051–1056. doi:10.1021/bc015519e
- Mangold, M., Rolland, L., Costanzo, F., Sprik, M., Sulpizi, M., and Blumberger, J. (2011). Absolute  $pK_a$  values and solvation structure of amino acids from Density Functional based molecular dynamics simulation. *J. Chem. Theory Comput.* 7, 1951–1961. doi:10.1021/ct100715x
- Marino, S. M., and Gladyshev, V. N. (2011). Redox biology: Computational approaches to the investigation of functional cysteine residues. *Antioxid. Redox Signal* 15, 135–146. doi:10.1089/ars.2010.3561
- Marino, S. M., Roos, G., and Gladyshev, V. N. (2013). “Computational redox biology: Methods and applications,” in *Oxidative stress and redox regulation* (Springer), 187–211.
- Martyniuk, C. J., Fang, B., Koomen, J. M., Gavin, T., Zhang, L., Barber, D. S., et al. (2011). Molecular mechanism of glyceraldehyde-3-phosphate dehydrogenase inactivation by  $\alpha$ ,  $\beta$ -unsaturated carbonyl derivatives. *Chem. Res. Toxicol.* 24, 2302–2311. doi:10.1021/tx200437y
- Matsushita, K., Morrell, C., Cambien, B., Yang, S.-X., Yamakuchi, M., Boa, C., et al. (2003). Nitric oxide regulates exocytosis by S-nitrosylation of N-ethylmaleimide-sensitive factor. *Cell* 115, 139–150. doi:10.1016/s0092-8674(03)00803-1
- May, A. P., Whiteheart, S. W., and Weis, W. I. (2001). Unraveling the mechanism of the vesicle transport ATPase NSF, the N-ethylmaleimide-sensitive factor. *J. Biol. Chem.* 276, 21991–21994. doi:10.1074/jbc.R100013200
- Mazmanian, K., Sargsyan, K., Grauffel, C., Dudev, T., and Lim, C. (2016). Preferred hydrogen-bonding partners of cysteine: Implications for regulating Cys functions. *J. Phys. Chem. B* 120, 10288–10296. doi:10.1021/acs.jpcc.6b08109
- McHugh, P. C., and Buckley, D. A. (2015). The structure and function of the dopamine transporter and its role in CNS diseases. *Vitam. Horm.* 98, 339–369. doi:10.1016/bs.vh.2014.12.009
- Mendez, D., Gaulton, A., Bento, A. P., Chambers, J., De Veij, M., Félix, E., et al. (2018). ChEMBL: Towards direct deposition of bioassay data. *Nucleic Acids Res.* 47, 930–940. doi:10.1093/nar/gky1075
- Meng, F.-G., Zhou, H.-W., and Zhou, H.-M. (2001). Effects of acrylamide on creatine kinase from rabbit muscle. *Int. J. Biochem. Cell. Biol.* 33, 1064–1070. doi:10.1016/s1357-2725(01)00079-6
- Michino, M., Beuming, T., Donthamsetti, P., Newman, A. H., Javitch, J. A., and Shi, L. (2015). What can crystal structures of aminergic receptors tell us about designing subtype-selective ligands? *Pharmacol. Rev.* 67, 198–213. doi:10.1124/pr.114.009944
- Mihalovits, L. M., Ferenczy, G. G., and u, G. M. K. (2022). The role of quantum chemistry in covalent inhibitor design. *Int. J. Quantum Chem.* 122, e26768. doi:10.1002/qua.26768
- Mondal, D., and Warshel, A. (2020). Exploring the mechanism of covalent inhibition: Simulating the binding free energy of  $\alpha$ -ketoamide inhibitors of the main protease of SARS-CoV-2. *J. Chem. Inf. Model.* 59, 4601–4608. doi:10.1021/acs.biochem.0c00782
- Mottram, D. S., Wedzicha, B. L., and Dodson, A. T. (2002). Acrylamide is formed in the Maillard reaction. *Nature* 419, 448–449. doi:10.1038/419448a
- Murray, S. M., Addell, B. M., and Wu, C.-W. (2020). Neuron-specific toxicity of chronic acrylamide exposure in *C. elegans*. *Neurotoxicol. Teratol.* 77, 106848. doi:10.1016/j.ntt.2019.106848
- Nagashima, D., Zhang, L., Kitamura, Y., Ichihara, S., Watanabe, E., Zong, C., et al. (2019). Proteomic analysis of hippocampal proteins in acrylamide-exposed Wistar rats. *Arch. Toxicol.* 93, 1993–2006. doi:10.1007/s00204-019-02484-9
- Nair, D. P., Podgorski, M., Chatani, S., Gong, T., Xi, W., Fenoli, C. R., et al. (2014). The thiol-michael addition click reaction: A powerful and widely used tool in materials chemistry. *Chem. Mat.* 26, 724–744. doi:10.1021/cm402180t
- Noort, D., and Hulst, A. (2003). Modification of human serum albumin by acrylamide at cysteine-34: A basis for a rapid biomonitoring procedure. *Arch. Toxicol.* 77, 543–545. doi:10.1007/s00204-003-0484-5
- Olivella, M., González, Ángel, Pardo, L., and Deupí, X. (2013). Relation between sequence and structure in membrane proteins. *Bioinformatics* 29, 1589–1592. doi:10.1093/bioinformatics/btt249
- Pahari, S., Sun, L., and Alexov, E. (2019). PKAD: A database of experimentally measured pKa values of ionizable groups in proteins. *Database* 2019, baz024. doi:10.1093/database/baz024
- Papamokos, G., Dreyer, J., Navarini, Luciano, and Carloni, P. (2014). Trapping acrylamide by a Michael addition: A computational study of the reaction between acrylamide and niacin. *Int. J. Quantum Chem.* 114, 553–559. doi:10.1002/qua.24610
- Park, S. U., Ferrer, J. V., Javitch, J. A., and Kuhn, D. M. (2002). Peroxynitrite inactivates the human dopamine transporter by modification of cysteine 342: Potential mechanism of neurotoxicity in dopamine neurons. *J. Neurosci. Res.* 22, 4399–4405. doi:10.1523/JNEUROSCI.22-11-04399.2002
- Pedersen, A. O., and Jacobsen, J. (1980). Reactivity of the thiol group in human and bovine albumin at pH 3–9, as measured by exchange with 2,2'-dithiodipyridine. *Eur. J. Biochem.* 106, 291–295. doi:10.1111/j.1432-1033.1980.tb06022.x
- Pennisi, M., Malaguarnera, G., Puglisi, V., Vinciguerra, L., Vacante, M., and Malaguarnera, M. (2013). Neurotoxicity of acrylamide in exposed workers. *Int. J. Environ. Res. Public Health* 10, 3843–3854. doi:10.3390/ijerph10093843
- Perutz, M. F. (1979). Regulation of oxygen affinity of hemoglobin: Influence of structure of the globin on the heme iron. *Annu. Rev. Biochem.* 48, 327–386. doi:10.1146/annurev.bi.48.070179.001551
- Philippe, L. M., Jerven, B., Edouard, d. C., Elisabeth, G., Parit, B., Andrea, H. A., et al. (2022). SwissBioPics — An interactive library of cell images for the visualization of subcellular location data. *Database*, 1–5.
- Piccoli, S., Suku, E., Garonzi, M., and Giorgetti, A. (2013). Genome-wide membrane protein structure prediction. *Curr. Genomics* 14, 324–329. doi:10.2174/13892029113149990009
- Pidathala, S., Mallela, A. K., Deepthi, J., and Penmatsa, A. (2021). Structural basis of norepinephrine recognition and transport inhibition in neurotransmitter transporters. *Nat. Commun.* 12, 2199. doi:10.1038/s41467-021-22385-9
- Pruser, K. N., and Flynn, N. E. (2011). Acrylamide in health and disease. *Front. Biosci. (Sch. Ed.)* 3, 41–51. doi:10.2741/s130
- Raldúa, D., Casado, M., Prats, E., Faria, M., Puig-Castellví, F., Pérez, Y., et al. (2020). Targeting redox metabolism: The perfect storm induced by acrylamide poisoning in the brain. *Sci. Rep.* 10, 312. doi:10.1038/s41598-019-57142-y



- Reynolds, T. (2002). Acrylamide and cancer: Tunnel leak in Sweden prompted studies. *J. Natl. Cancer Inst.* 94, 876–878. doi:10.1093/jnci/94.12.876
- Riccardi, D., Schaefer, P., and Cui, Q. (2005). pK<sub>a</sub> Calculations in solution and proteins with QM/MM free energy perturbation simulations: A quantitative yeast of QM/MM protocols. *J. Phys. Chem. B* 109, 17715–17733. doi:10.1021/jp0517192
- Roos, G., Foloppe, N., and Messens, J. (2013). Understanding the pK<sub>a</sub> of redox cysteines: The key role of hydrogen bonding. *Antioxid. Redox Signal* 18, 94–127. doi:10.1089/ars.2012.4521
- Rostkowski, M., Olsson, M. H., Sondergaard, C. R., and Jensen, J. H. (2011). Graphical analysis of pH-dependent properties of proteins predicted using PROPKA. *BMC Struct. Biol.* 11, 6. doi:10.1186/1472-6807-11-6
- Sampath, V., Zhao, X., and Caughey, W. (2001). Anesthetic-like interactions of nitric oxide with albumin and hemoproteins. A mechanism for control of protein function. *J. Biol. Chem.* 276, 13635–13643. doi:10.1074/jbc.M006588200
- Scarpino, A., Ferenczy, G. G., and Keseru, G. M. (2018). Comparative evaluation of covalent docking tools. *J. Chem. Inf. Model.* 58, 1441–1458. doi:10.1021/acs.jcim.8b00228
- Schouten, M. A., Tappi, S., and Romani, S. (2020). Acrylamide in coffee: Formation and possible mitigation strategies – a review. *Crit. Rev. Food Sci. Nutr.* 60, 3807–3821. doi:10.1080/10408398.2019.1708264
- Semla, M., Goc, Z., Martiniaková, M., Omelka, R., and Formicki, G. (2017). Acrylamide: A common food toxin related to physiological functions and health. *Physiol. Res.* 66, 205–217. doi:10.33549/physiolres.933381
- Sheng, Q., Zou, H.-C., Lü, Z.-R., Park, Y.-D., Yao, S.-J., Yan, Y. B., et al. (2009). Effects of acrylamide on the activity and structure of human brain creatine kinase. *Int. J. Mol. Sci.* 10, 4210–4222. doi:10.3390/ijms10104210
- Shulamit, C., and Sihra, T. S. (1989). Characterization of a H<sup>+</sup>-ATPase in rat brain synaptic vesicle. *J. Biol. Chem.* 264, 8281–8288.
- Song, Y., Mao, J., and Gunner, M. R. (2009). MCCE2: Improving protein pK<sub>a</sub> calculations with extensive side chain rotamer sampling. *J. Comput. Chem.* 30, 2231–2247. doi:10.1002/jcc.21222
- Sottriffer, C. (2018). Docking of covalent ligands: Challenges and approaches. *Mol. Inf.* 37, 1800062. doi:10.1002/minf.201800062
- Soukri, A., Mougín, A., Corbier, C., and Wonacott, A. (1989). Role of the histidine 176 residue in glyceraldehyde-3-phosphate dehydrogenase as probed by site-directed mutagenesis. *Biochemistry* 28, 2586–2592. doi:10.1021/bi00432a036
- Soylu, I., and Marino, S. M. (2017). Cppe: A comprehensive computational platform for sequence and structure-based analyses of cysteine residues. *Bioinformatics* 33, 2395–2396. doi:10.1093/bioinformatics/btx181
- Soylu, I., and Marino, S. M. (2015). Cy-preds: An algorithm and a web service for the analysis and prediction of cysteine reactivity. *Bioinformatics* 31, 278–291. doi:10.1093/bioinformatics/btt497
- Steinegger, M., Meier, M., Mirdita, M., Harald, V., Haunsberger, S. J., and Johannes, S. (2019). HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinform.* 20, 473. doi:10.1186/s12859-019-3019-7
- Stivers, J. T., Abeygunawardana, C., Mildvan, A. S., Hajipour, G., and Whitman, C. P. (1996). 4-Oxalocrotonate tautomerase: pH dependence of catalysis and pK<sub>a</sub> values of active site residues. *Biochemistry* 35, 814–823. doi:10.1021/bi9510789
- Sunghwan, K., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., et al. (2019). PubChem 2019 update: Improved access to chemical data. *Nucleic Acids Res.* 47, D1102–D1109. doi:10.1093/nar/gky1033
- Swaen, G. M. H., Haidar, S., Burns, C. J., Bodner, K., Parsons, T., Collins, J. J., et al. (2007). Mortality study update of acrylamide workers. *J. Occup. Environ. Med.* 49, 396–401. doi:10.1136/oem.2006.030130
- Tanii, H., and Hashimoto, K. (1985). Effect of acrylamide and related compounds on glycolytic enzymes of rat brain. *Toxicol. Lett.* 26, 79–84. doi:10.1016/0378-4274(85)90188-2
- The UniProt Consortium (2019). UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* 47, D506–D515. doi:10.1093/nar/gky1049
- Tong, G. C., Cornwell, W. K., and Means, G. E. (2004). Reactions of acrylamide with glutathione and serum albumin. *Toxicol. Lett.* 147, 127–131. doi:10.1016/j.toxlet.2003.10.021
- van Zundert, G., Rodrigues, J., Trellet, M., Schmitz, C., Kastiris, P., Karaca, E., et al. (2016). The HADDOCK2.2 web server: User-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* 428, 720–725. doi:10.1016/j.jmb.2015.09.014
- Wang, P.-F., Flynn, A. J., Naor, M. M., Jensen, J. H., Cui, G., Merz, K. M., Jr., et al. (2006). Exploring the role of the active site cysteine in human muscle creatine kinase. *Biochemistry* 45, 11464–11472. doi:10.1021/bi0607002
- Wang, R., Carnevale, V., Bourguet, E., and Klein, M. L. (2020). First-principles calculation of water pK<sub>a</sub> using the newly developed SCAN functional. *J. Phys. Chem. Lett.* 11, 54–59. doi:10.1021/acs.jpclett.9b02913
- Warshel, A. (1991). *Computer modeling of chemical reactions in enzymes and solutions*. Wiley.
- Warshel, A., and Levitt, M. (1976). Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* 103, 227–249. doi:10.1016/0022-2836(76)90311-9
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. doi:10.1093/nar/gky427
- Weber, B. W., Kimani, S. W., Varsani, A., Cowan, D. A., Hunter, R., Venter, G. A., et al. (2013). The mechanism of the amidases: mutating the glutamate adjacent to the catalytic triad inactivates the enzyme due to substrate mispositioning. *J. Biol. Chem.* 288, 28512–28523.
- Whitehead, R. E., Ferrer, J. V., Javitch, J. A., and Justice, J. B. (2001). Reaction of oxidized dopamine with endogenous cysteine residues in the human dopamine transporter. *J. Neurochem.* 76, 1242–1251. doi:10.1046/j.1471-4159.2001.00125.x
- Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., et al. (2018). MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* 27, 293–315. doi:10.1002/pro.3330
- Wishart, D., Arndt, D., Pon, A., Sajed, T., Chi Guo, A., Djoumbou, Y., et al. (2015). T3db: The toxic exposome database. *Nucleic Acids Res.* 43, 928–934. doi:10.1093/nar/gku1004
- Zhao, C., Selvin, J. T., and Whiteheart, S. W. (2007). Cellular functions of NSF: Not just SNAPs and SNAREs. *FEBS Lett.* 581, 2140–2149. doi:10.1016/j.febslet.2007.03.032
- Zhao, M., Dong, L., Zhu, C., Hu, X., Zhao, L., Chen, F., et al. (2019). Proteomic profiling of primary astrocytes and co-cultured astrocytes/microglia exposed to acrylamide. *Neurotoxicology* 75, 78–88. doi:10.1016/j.neuro.2019.09.005
- Zhao, M., Wu, S., Zhou, Q., Vivona, S., Cipriano, D., Cheng, Y., et al. (2015). Mechanistic insights into the recycling machine of the SNARE complex. *Nature* 518, 61–67. doi:10.1038/nature14148
- Zhu, K., Borrelli, K. W., Greenwood, J. R., Day, T., Abel, R., Farid, R. S., et al. (2014). Docking covalent inhibitors: A parameter free approach to pose prediction and scoring. *J. Chem. Inf. Model.* 54, 1932–1940. doi:10.1021/ci500118s