DOI:10.14753/SE.2023.2817

SEMMELWEIS EGYETEM

DOKTORI ISKOLA

Ph.D. értekezések

2817.

JULIANA FERREIRA DE SANTANA

A gyógyszerészeti tudományok korszerű kutatási irányai

című program

Programvezető: Dr. Antal István, egyetemi tanár

Témavezető: Dr. Noszál Béla, professzor emeritus

Társtémavezető: Dr. Arash Mirzahosseini, egyetemi adjunktus

DOI:10.14753/SE.2023.2817

CHARACTERIZATION OF THIOL-DISULFIDE SYSTEMS AS CANDIDATES TO COMBAT OXIDATIVE STRESS

PhD Thesis

Juliana Ferreira de Santana

The Doctoral School of Pharmaceutical Sciences

Semmelweis University

Supervisors: Béla Noszál, D.Sc.

Arash Mirzahosseini, PharmD, Ph.D.

Official reviewers: Ádám Orosz, PharmD, Ph.D.

Fanni Sebák, PharmD, Ph.D.

Head of the Complex Examination Committee: Éva Szökő, D.Sc.

Members of the Complex Examination Committee: Zsolt Szakonyi, D.Sc.

Viola Tomási, Ph.D.

Budapest

2022

DOI:10.14753/SE.2023.2817

TABLE OF CONTENTS

DOI:10.14753/SE.2023.2817

LIST OF ABBREVIATION

Ac	acetylation
ACA	alanine-cysteine-alanine
Ala	alanine
Arg	arginine
Asn	asparagine
CysASSCysA	cystamine
Cys	cysteine
CysASH	cysteamine
CysSSCys	cystine
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DIPEA	N, N-diisopropylethylamine
DMF	N, N-dimethylformamide
DSS	4,4-dimethyl-4-silapentane-1-sulfonate
DTT	dithiothreitol
GRX	glutaredoxin
GSH	glutathione
GSSG	glutathione-disulfide
HATU	hexafluorophosphate azabenzotriazole tetramethyl uronium
Hcy	homocysteine
HcySSHcy	homocystine
HSQC	heteronuclear single quantum coherence

DOI:10.14753/SE.2023.2817

Met	methionine
NADPH	dihydro-nicotinamide adenine dinucleotide phosphate
NCN	asparagine-cysteine-asparagine
NMR	nuclear magnetic resonance
PenA	penicillamine
RCR	arginine-cysteine-arginine
ROS	reactive oxygen species
SAH	S-adenosylhomocysteine
SAM	S-adenosylmethionine
SCS	serine-cysteine-serine
SMC	S-methyl cysteine
Ser	serine
TCT	threonine-cysteine-threonine
Thr	threonine
TIS	triisopropylsilane
TRX	thioredoxin
TRXR	thioredoxin reductase
Val	valine
VCV	valine-cysteine-valine

DOI:10.14753/SE.2023.2817

1. Introduction

1.1 Oxidative stress and redox signaling

The term “oxidative stress” was formulated by Helmut Sies in 1985. [1] First, it was delineated as “a disturbance in the prooxidant-antioxidant balance in favor of the former”. In 2007, to incorporate the term redox signaling the definition was updated to “an imbalance between oxidants and antioxidants in favor of the oxidants, leading to a disruption of redox signaling control and/or molecular damage”. [2] This imbalance in favor of oxidants has been associated with aging, atherosclerosis, carcinogenesis, diabetes, and neurodegeneration. [3]

The oxidants usually referred to as reactive oxygen species (ROS), are produced chiefly within the mitochondria throughout the traditional cellular metabolism. However, within the cytosol and plasma membrane, some enzymes like NADPH oxidase and cytochrome P450 oxidase can originate them as well. [4] ROS have a vital role against infectious agents and in cellular signaling systems. Although their effects are beneficial only if they are present in low or moderate concentrations. [5] In higher concentrations they can be toxic to organisms, nevertheless, through billions of evolution years, cells became capable to evolve antioxidant molecules and detoxifying enzymes to deal with them. [6]

Reactive oxygen species comprise both free radicals and nonradical oxygen derivatives. [4] Free radicals are chemical species with an unpaired electron, which are responsible for the high reactivity of these compounds. [7] Some examples of free radicals include hydroxyl radical ($OH^{\bullet}$) and superoxide anion ($O_2^{\bullet -}$), while nonradical oxygen derivatives can also be chemically stable molecules, such as hydrogen peroxide (H_2O_2) and organic hydroperoxide (ROOH). [8]

In previous researches, the centre of attention was mostly free radicals, which persist as a crucial topic. However, studies showed that hydroperoxides and electrophiles also have fundamental functions in physiologically signaling and regulation of transcription factors. [9] Redox signaling is part of the physiology of cells. It can be defined as a process in which the signal is transported through redox reactions and has a compelling role in pathophysiological responses. [5, 10] For redox signaling to occur, an

unbalanced redox state is necessary by decreasing the activity of antioxidants or increasing ROS generation. [5]

In the signaling pathways, it is important to figure out how ROS can modify a function of proteins. The two principal target residues in the redox regulation are the sulfur-containing cysteine (Cys) and methionine (Met) in proteins. Cys oxidation develops reactive sulfenic acid (– SOH) which can form disulfide bonds (– S – S –) with nearby Cys or go through further oxidation to sulfinic (– SO_2H) or sulfonic (–SO_3H) acid. [3, 11] These types of Cys oxidation are noticed when a protein has its activity modified and may become vulnerable to aggregation and/or degradation. [11] It does not imply that Cys is the only amino acid in redox signaling, but those are the Cys chemical reactions that mainly determine redox signaling containing hydroperoxides and some other electrophiles. [10]

To date there is not, unfortunately, any powerful therapeutic agent to treat chronic diseases generated by oxidative stress. [12] It is currently an unmet medical need to have a therapy based on small molecules such as antioxidants to control oxidative stress. [4]

1.2 Thiol-disulfide systems

The sulfur atom can be found in many different functional groups, in some metabolites (coenzyme A, glutathione - GSH, and mycothiol) and amino acids, such as cysteine, homocysteine, and penicillamine. Due to its versatile reactivity, sulfur is known for having a main role in redox biochemistry. [13, 14] The thiol part (SH) of these compounds has a -2 sulfur oxidation state so they are susceptible to a variety of oxidations. If the thiol is placed in an environment with chemical and physical properties that could bring it into its deprotonated form (e.g. a thiolate anion S^-), which has the ability to react with oxidants and electrophilic molecules. [13] As soon as this occurs, thiolate forms sulfenic acids (RSOH), which is a short-lived species and has the ability to react with other protein thiols or non-protein thiols. When reacting with protein thiols, RSOH can generate disulfide bonds (intramolecular or intermolecular) and when associating with non-protein thiols (e.g., GSH or Cys) it can form mixed (hetero-) disulfide bonds or it can react with amino groups to form sulfenamides (RSNRH). It is possible that a second and usually irreversible oxidation of sulfenic acid and sulfenamides occur to form sulfinic acid (RSO_2H) and sulfonic acid (RSO_3H). [15]

The redox homeostasis of these residues is usually handled by thioredoxin (TRX) and glutaredoxin (GRX) systems, which are conserved oxidoreductases, that use thiol-disulfide reactions to reduce protein disulfides into thiols. Both enzymes keep a sequence pattern (Cys-X-Y-Cys) capable to assist and support the reduction of the target protein. [15, 16]

Regarding the TRX system, the cells carry the cytosolic type TRX1/thioredoxin reductase 1 (TRXR1) and the mitochondrial type TRX2/TRXR2. [16] The reduction by thioredoxin occurs initially with a nucleophilic attack, in which one of its Cys attacks the oxidized thiol of the target protein generating an intermolecular hetero-disulfide bond between the thioredoxin and the target protein. After this, a second nucleophilic attack takes place, now the other Cys attacks the previously made disulfide bond and forms an intramolecular one in thioredoxin, thus completely reducing the target protein. Lastly, the reduction of the disulfide bond in thioredoxin succeeds with thioredoxin reductase. [15]

The GRX system comprehends the GSH/glutathione-disulfide (GSSG), glutathione reductase, and glutaredoxin. GSH, like other molecules that contain Cys, is promptly oxidized and usually, it is the most abundant thiol in cells. [17] This other important thiol-dependent antioxidant is responsible for the removal of ROS by glutathione peroxidase. [18] Primarily, glutaredoxin usually attacks target proteins that are S-glutathionylated, and then an intermolecular disulfide bond with GSH is made. Another GSH attacks this bond, with this GRX is reduced and GSSG is released. This oxidized GSH is formed as ROS donates one electron to a GSH. [17] Finally, GSSG is reduced by glutathione reductase, which is dependent on NADPH. [15] The proportion of GSH/GSSG in the cells is a relevant indicator to quantify the cellular redox potential. [16]

Regardless of their differences, these two systems (thioredoxin and glutaredoxin) can substitute each other in the cells, showing how they are essential for the maintenance of redox homeostasis and their potential of preventing irreversible thiol modifications. [15] The complete microspeciation of thiol-containing amino acids was done using ^{1}H NMR-pH titrations by Mirzahosseini and Noszál, which helped to comprehend their inter-related redox and acid-base processes and may provide a means in the search for antioxidants molecules. [19]

1.2.1 Cysteine

The amino acid Cys contains a thiol moiety, which is ionizable and susceptible to oxidation. It has already been demonstrated that different Cys oxidations are related to specific redox signals. Besides being a substrate of Met and GSH biosynthesis, Cys accomplishes several functions as regulating catalysis, structure, and metal ion trafficking. [13, 20, 21]

Cys is, mainly, a pivotal regulator of redox homeostasis and signaling. Not all residues of Cys present in proteins are likely to be oxidized, it will depend on the solvent accessibility, pK_a (acid dissociation constant), and polarity of the nearby residues. To better understand the biological function of Cys oxidation and how this is connected to some diseases, it is fundamental to be aware of how these modifications occur. [22]

In human plasma, one of the most abundant low molecular weight redox pairs is the Cys/CysSSCys couple, therefore it is possible to determine the plasma redox status by them. Cystine (CysSSCys), the oxidized form of Cys, is formed by the coupling of two Cys via a disulfide bond, which can improve the stability of molecules in adverse environments. This oxidized form has several roles in cellular and metabolic functions. As CysSSCys can be found in biological samples (urine or blood), it may be used as a biomarker for some diseases like cystinosis and metabolic disorders. [23-25]

Another compound also synthesized from Cys and used in this work was S-methyl cysteine (SMC). For its formation, Cys can be methylated in the free state or cysteinyl residues can be methylated in peptides and proteins. [26]

1.2.3 Homocysteine

Homocysteine (Hcy), another amino acid containing a sulfur atom, is not involved directly in the genetic code, so it is known as a nonproteinogenic amino acid. Hcy is formed from the essential amino acid methionine, which forms S-adenosylmethionine (SAM) with ATP. Then, SAM as methyl donor generates S-adenosylhomocysteine (SAH). SAH undergoes hydrolysis and finally, Hcy is produced. [27, 28] Once produced, Hcy can be a substrate for the production of methionine again, by a remethylation by the enzyme N5, N10- methylenetetrahydrofolate reductase or it can be transformed to Cys by cystathionine-synthase. [29]

The mirrorlike disulfide of Hcy is homocystine (HcySSHcy), their names were chosen to point out that each carbon chain of their structure included one more -CH_2 group than those of, Cys and CysSSCys. [30]

1.2.4 Penicillamine

Penicillamine (PenA), which is also considered a nonproteinogenic amino acid, is an analog of Cys, as both present similar chemical characteristics. In its structure, two methyl groups replace the two hydrogens connected to the b-carbon of Cys. PenA basically is composed of a thiol side chain, a carboxyl group, and an a-amino group. [31, 32]

As PenA is resembling to Cys, penicillamine-disulfide, its oxidized form, is analogous to CysSSCys. Penicillamine-disulfide has two carboxylic and two amino groups. [33]

1.2.5 Cysteamine

Cysteamine (CysASH) is an amino-thiol produced from Cys and pantothenate (vitamin B5). Together they synthesize coenzyme A, which is degraded into pantetheine. The enzyme pantetheinase cleaves pantetheine and produces CysASH and pantothenic acid. Then CysASH can be oxidized into hypotaurine, which can be further degraded by hypotaurine dehydrogenase into taurine. [34, 35]

This compound is very reactive, and its activity will be relative to its physiological environment. If oxygen or transition metals are present, CysASH may be oxidized producing its disulfide form, cystamine (CysASSCysA). However, without these factors and under a reducing condition, CysASH can behave as an antioxidant. [36] When CysASH is present in low concentrations, it may impact the cellular redox homeostasis, as it can help to transport Cys into cells, which is a substrate for GSH. [35]

CysASH and CysASSCysA also exist in equilibrium and act as an effective redox sensor. [37, 38]

1.2.6 Glutathione

GSH, also known as y-L-glutamyl-L-cysteinyl-glycine, is the main low molecular weight intracellular antioxidant. This tripeptide is produced by adding Cys to glutamate proceeded by the insertion of glycine. GSH has several functions as it can transport amino

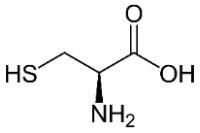
acids through the plasma membrane, sweeps singlet oxygen and hydroxyl radical, acts as a cofactor of some enzymes (e.g., glutathione peroxidase) to remove peroxides, and is also associated with the regulation of the cell cycle. [4, 39]

The participation of GSH in reduction and conjugation reactions is due to the sulfhydryl group (−SH) inherited from Cys. [39] Besides that, GSH does not have the toxicity related to Cys, thereby it favorably fits to sustain a particular thiol-disulfide redox potential. [40] As it has already been said, the oxidation of GSH forms glutathione disulfide (GSSG), these two compounds together are relevant and broadly studied cellular redox couple. To be able to do its function, this system does not demand enzyme catalysis and can be used in aqueous media under different conditions. [41]

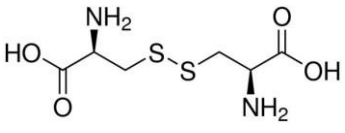
GSSG can be reduced to GSH mainly by glutathione reductase using NADPH. [42] The conversion (oxidation) of GSH into GSSG takes usually place under the conditions of oxidative stress. As GSSG acts as a prooxidant its accumulation can damage the cells, showing why the equilibrium between GSH-GSSG is essential. [43]

The difference between GSH and glutathione ethyl ester is that the latter has the carboxyl group of the glycine residue esterified. Thereby glutathione ethyl ester can be promptly transported into the cells and hydrolysed leading to higher intracellular concentrations of GSH. [44]

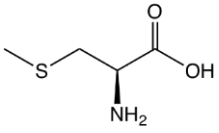
The last compound, S-methyl glutathione, is a thioether synthesized from GSH and S-adenosyl-methionine. The chemotaxis methyltransferase presence is also crucial for its formation. Until today its exact origin is not known, but S-methyl glutathione is commonly detected in yeasts, brain tissues, and *Escherichia coli* and it is a usual product of several methylated drugs and pesticides in animals and plants and also from the metabolism of methyl halides by glutathione S-transferases. [45, 46] All the thiols and disulfides structures mentioned are demonstrated in Figure 1.



Cysteine

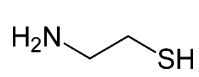


Cystine

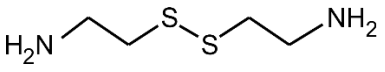


S-methyl cysteine

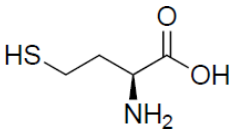
DOI:10.14753/SE.2023.2817



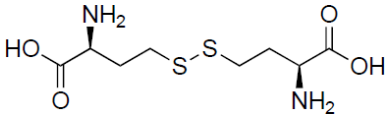
Cysteamine



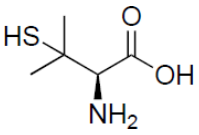
Cystamine



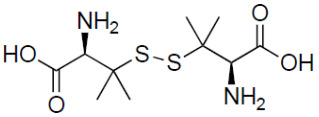
Homocysteine



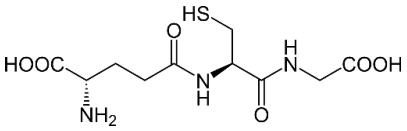
Homocystine



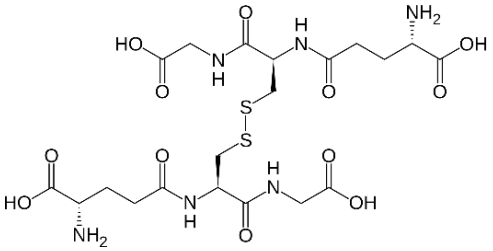
Penicillamine



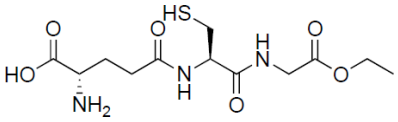
Penicillamine-disulfide



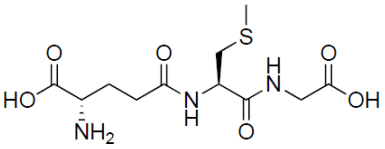
Glutathione



Glutathione-disulfide



Glutathione ethyl ester



S-methyl glutathione

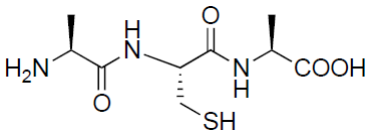
Figure 1. Structural formulae of the thiols and disulfides studied.

1.2.7 Cysteine-containing peptides

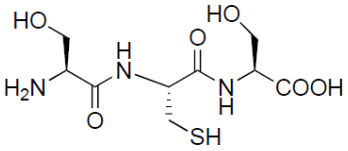
Our studies also included 11 cysteine-containing peptides, divided into two groups. The first group contained 5 peptides with free N- and C- termini, as showed in Figure 2: Ser-Cys-Ser (SCS), Ala-Cys-Ala (ACA), Val-Cys-Val (VCV), Thr-Cys-Thr (TCT), and Asn-Cys-Asn (NCN).

DOI:10.14753/SE.2023.2817

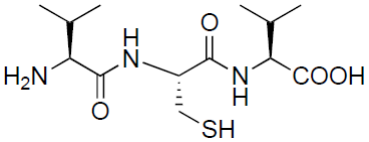
In the second group, there were 6 peptides with N-terminal acetylation and C-terminal amidation, as presented in Figure 3: Arg-Cys-Arg (Ac-RCR-Amide), Thr-Cys-Thr (Ac-TCT-Amide), Ala-Cys-Ala (Ac-ACA-Amide), Val-Cys-Val (Ac-VCV-Amide), Asn-Cys-Asn (Ac-NCN-Amide) and Ser-Cys-Ser (Ac-SCS-Amide). It has been shown that these modifications can protect the peptides from protease degradation [47].



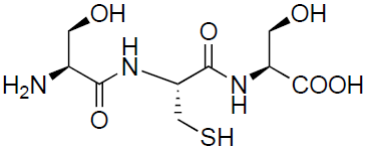
ACA (1)



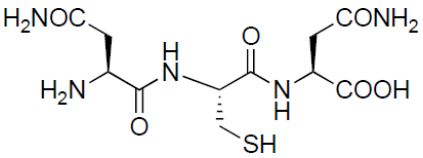
SCS (2)



VCV (3)

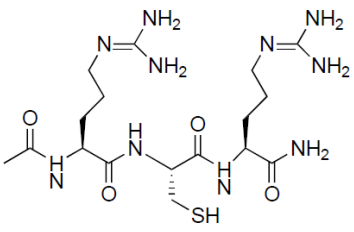


TCT (4)

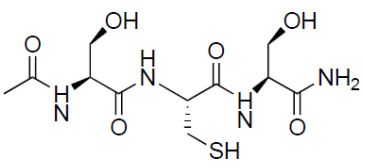


NCN (5)

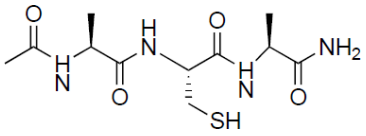
Figure 2. Cysteine-containing peptides with free N- and C- termini.



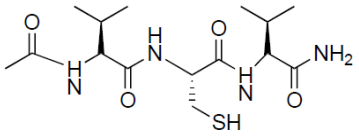
Ac-RCR-Amide (6)



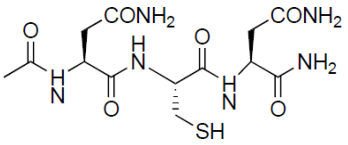
Ac-TCT-Amide (7)



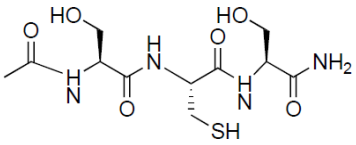
Ac-ACA-Amide (8)



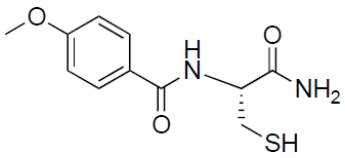
Ac-VCV-Amide (9)



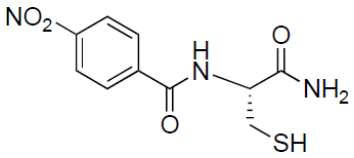
Ac-NCN-Amide (10)



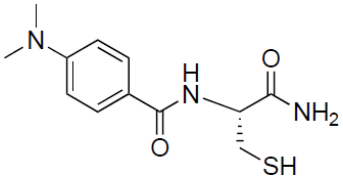
Ac-SCS-Amide (11)



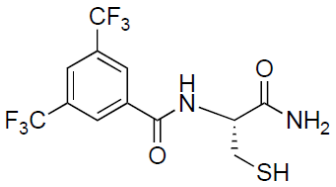
4-methoxybenzoylcysteine amide (12)



4-nitrobenzoylcysteine amide (13)



4-dimethylaminobenzoylcysteine amide (14)



3,5-bis(trifluoromethyl)benzoylcysteine amide (15)

Figure 3. Cysteine-containing peptides with N-terminal acetylation and C- terminal amidation (6-11) and Cys derivatives (12-15).

1.3 Species-specific chemical equilibria

The last years expanded the physicochemical characterization of biomolecules, regarding the biorelevant species-specific parameters, such as experimental determination of species-specific partition coefficients, solubilities, and redox equilibrium constants. [48] The capability to do a correlation between Cys chemical

shifts, redox potential, pK_a, and the molecular structure can bring relevant information about the chemistry of proteins and peptides containing Cys. [49]

It is not possible to measure directly the redox potential for Cys and other thiols through common electrochemical methods since the formation of stable metal-thiolate complexes occur at electrode surfaces. [50] Therefore, the redox potential of thiols can be determined indirectly only, by measurements of equilibrium constants for their reaction with redox systems of known redox potential. [48]

Although macroscopic parameters are used on large scale, they are in principle limited to characterizing the molecule as a whole, not its basic sites. Differently, microscopic parameters can show information and interactions at a submolecular level. The best known submolecular physicochemical parameter is the microscopic protonation constant. It has the capacity to quantify the proton-binding potential of submolecular basic units. [48]

A complete microspeciation of CysASH, Cys, Hcy, and their respective homodisulfides has been elaborated through ^{1}H NMR-pH titrations allowing the comprehension of the detailed acid-base processes of thiol-containing amino acids at a submolecular level. [19, 41] The acid-base microspeciation schemes of CysASH/CysASSCysA, Cys/CysSSCys, and GSH/GSSG were done in previous works by Mirzahosseini et al., they are shown in Figures 4, 5 and 6, respectively. [41, 51] Now the observed correlation is extended between standard redox potentials and thiolate log$K$ [51] to chemical shift values, to highlight the predictive power of NMR parameters available from relatively simple spectroscopic measurements.

DOI:10.14753/SE.2023.2817

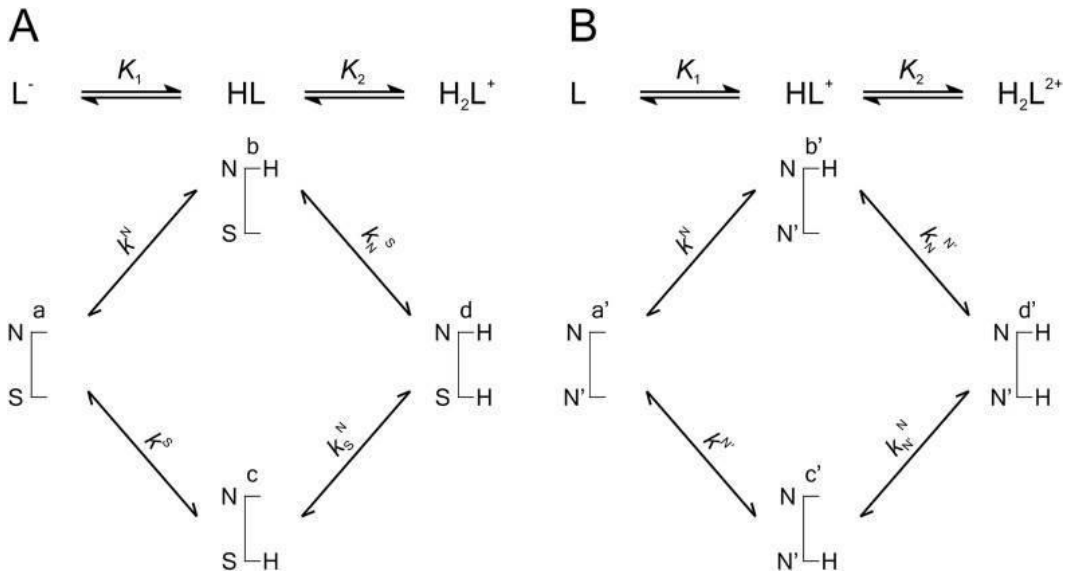


Figure 4. In this figure, the protonation equilibrium schemes of CysASH (A) and CysASSCysA (B) are represented. Stepwise macroscopic protonation constants (K_1 and K_2) characterize the overall basicities of the compounds, where L^- and HL are the successively protonating ligands (top lines). Below in the schematic networks are the microspecies (a, b, c, d and a', b'= c', d') and the microscopic protonation constants (k^N, k^N_S…) characterizing the submolecular localization of the protonation within the molecule. Symbols N and S represent the amino and thiolate, respectively. The figure was reproduced from [51].

DOI:10.14753/SE.2023.2817

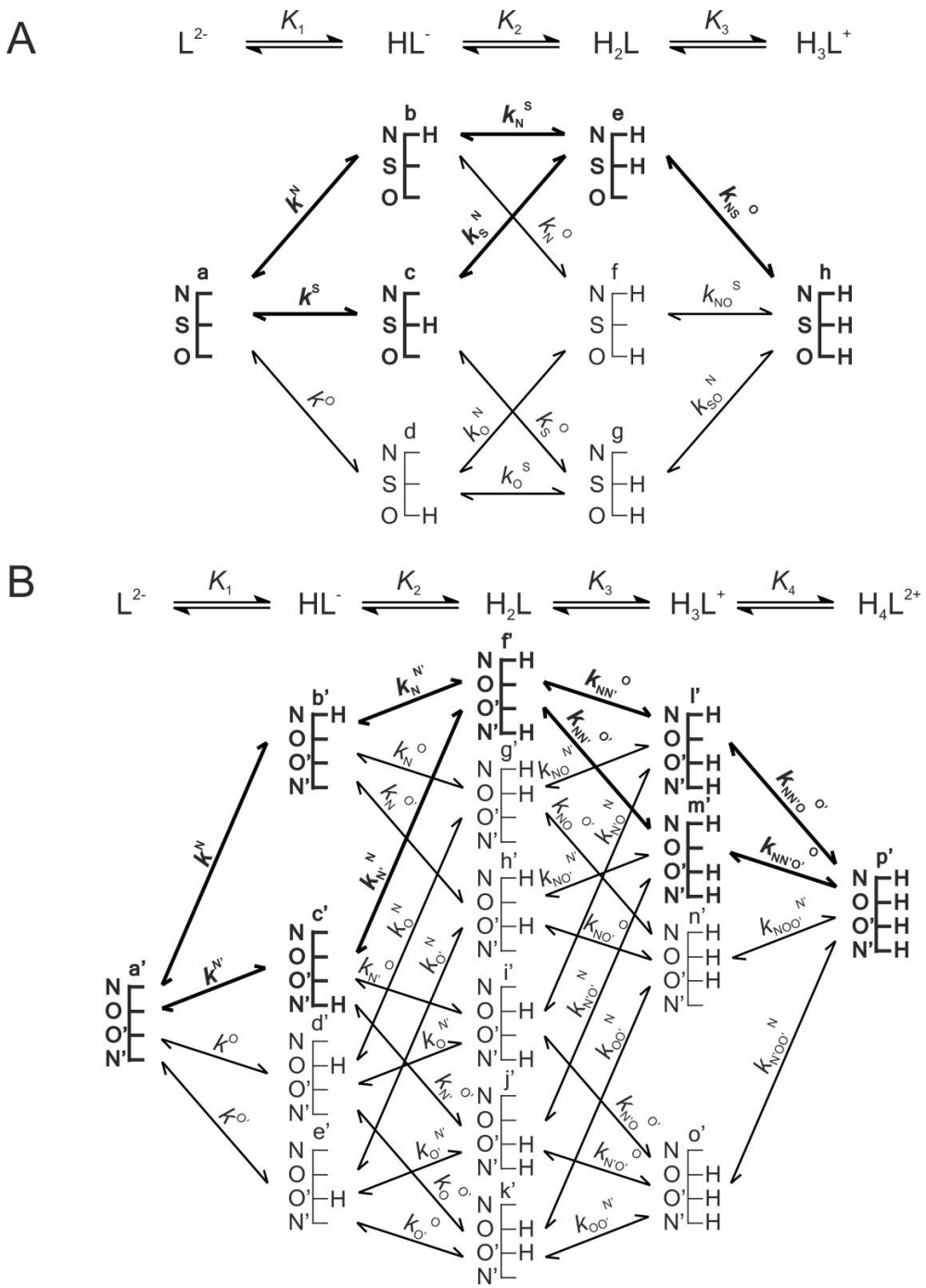


Figure 5. The protonation macro- and microequilibrium schemes of Cys (A) and CysSSCys (B) in terms of stepwise macroscopic protonation constants (K_1, K_2, K_3 …), where L^{2-}, HL^{-}, etc. are the protonating ligands (horizontal, top lines). Below are the species-specific protonation schemes in terms of microspecies (a, b, c …) and microscopic protonation constants (k^N, k^N_S…). The components of the principal

pathways are in bold. The same microspeciation scheme is valid for Hcy, PenA and their homodisulfides. The figure was reproduced from [41, 51].

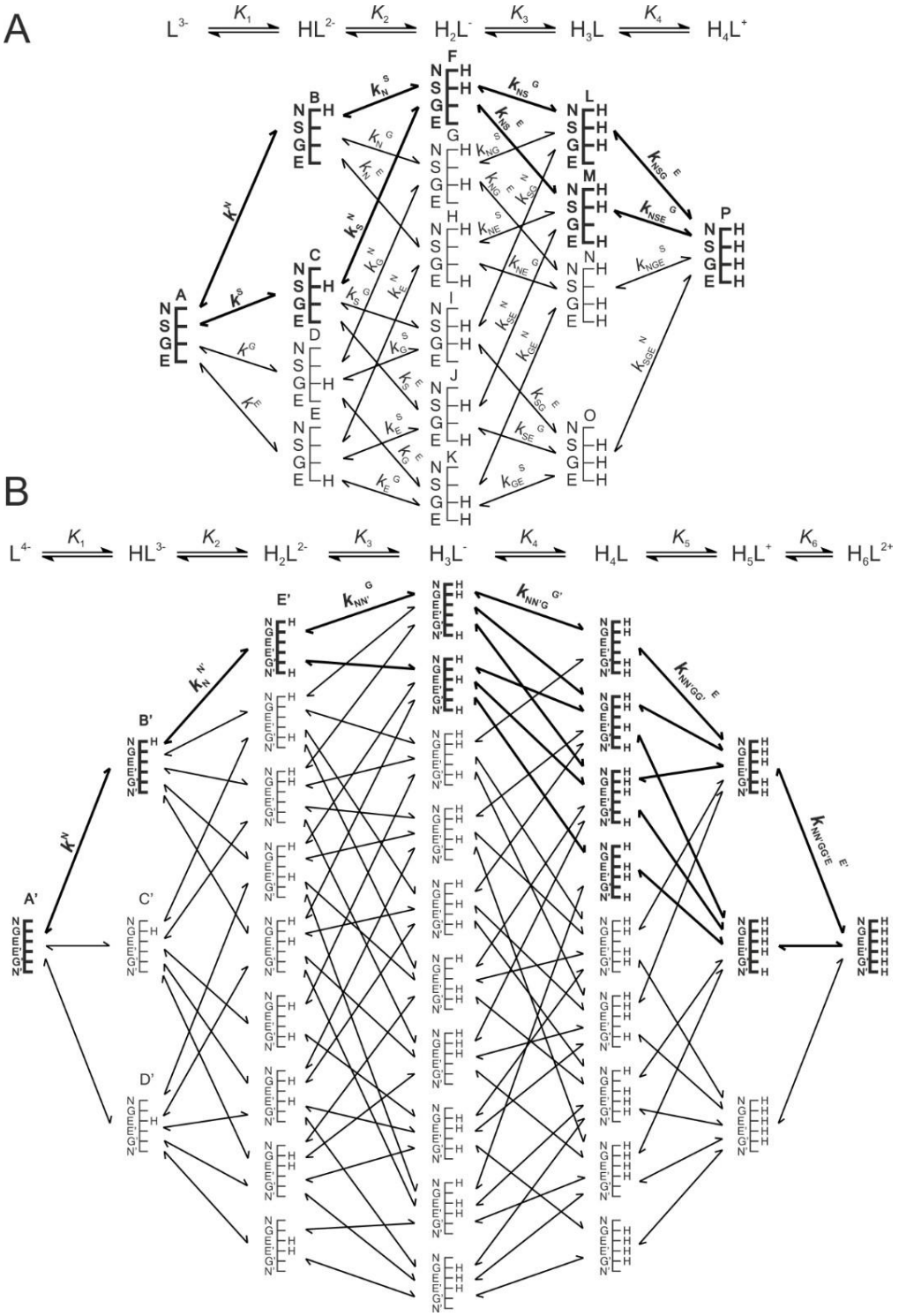


Figure 6. The protonation macro- and microequilibrium schemes of GSH (A) and GSSG (B) in terms of stepwise macroscopic protonation constants (K_1, K_2, K_3 …), where L^{3-}, HL^{2-}, etc. are the successively protonating ligands (top lines). Below are the

species-specific protonation schemes in terms of microspecies (A, B, C …) and microscopic protonation constants (k^{N}, k_{N}^{S} …). The components of the major pathways are in bold. The basic moieties are depicted with their one-letter symbols: N (amino), S (thiolate), G (glycyl carboxylate), and E (glutamyl carboxylate). The figure was reproduced from [41, 51].

DOI:10.14753/SE.2023.2817

2. Objectives

- The goal of our work is to see whether a correlation anticipated in previous work from our group, really exists between the redox potential, and the NMR chemical shifts of nuclei near the sulfur atom in these compounds.

- Determination of limiting ^{1}H and ^{13}C chemical shifts of 13 compounds (thiols and disulfides) and 15 Cys containing peptides.

- The hypothesized correlation is expected on the basis of changing electron densities, which are concomitant influencers of both the redox and NMR parameters.

- The correlations were analysed by plotting microscopic log*k* values pertaining to the thiolate group in various microspecies and the corresponding chemical shift values.

3. Results

3.1 Materials

All the thiols and disulfides studied were obtained from Sigma-Aldrich (Saint Louis, MO, USA) and used without further purification. Except for the compounds (12), (13), (14) and (15), all the peptides containing Cys, (Figures 2 and 3) were purchased from ProteoGenix (Schiltigheim, France). Their homodisulfides were created by adding 1% H_2O_2 to the thiols. Deionized water was produced in our laboratory by a Milli-Q Direct 8 Millipore system.

The Cys derivative compounds (12), (13), (14) and (15) (Figure 3) were synthesized on TentaGel R RAM resin (0.19 mmol/g) with Fmoc-chemistry on a Rink amide linker on a 0.1 mmol scale manually. The coupling of Cys was performed as follows: 3 equivalents of Fmoc-protected amino acid, 3 equivalents of the uronium coupling agent O-(7-azabenzotriazol-1-yl)-N, N, N′, N′-tetramethyluronium hexafluorophosphate (HATU) and 6 equivalents of N, N-diisopropylethylamine (DIPEA) were used in N, N-dimethylformamide (DMF) as solvent with shaking for 3 h. After the coupling steps, the resin was washed 3 times with DMF, once with methanol, and 3 times with dichloromethane. Deprotection was performed with 2% 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) and 2% piperidine in DMF in two steps, with reaction times of 5 and 15 min. After the deprotection of the Fmoc-group, the resin was washed and a further coupling step was carried out with the appropriate benzoic acid derivative and HATU with DIPEA as the coupling agent. The resin was washed with the same solvents as delineated previously. The cleavage was performed with trifluoroacetic acid/water/DL-dithiothreitol (DTT)/triisopropylsilane (TIS) (90:5:2.5:2.5) at 0 °C for 2 h. The cleavage cocktail was evaporated, and the peptide was precipitated with diethyl ether. After the precipitation, the Cys derivatives were dissolved in a 10% acetic acid solution and were lyophilized. As a final purification, the solid residue that remained after lyophilization was digerated with diisopropyl ether. The crystals gained by this step were washed with diisopropyl ether.

3.2 NMR spectroscopy measurements

NMR spectra were recorded on a Varian 600 MHz spectrometer at 298.15 ± 0.1 K. The solvent was H_2O:D_2O 95:5 (V/V), and ionic strength was adjusted to 0.15 mol/L.

DOI:10.14753/SE.2023.2817

The pH values were determined *in situ* by internal indicator molecules (at ca. 1 mmol/L) optimized for ^{1}H NMR. [52, 53] The sample volume was 550 μL and every sample contained ca. 1 mmol/L DSS (3-(trimethylsilyl)propane-1-sulfonate) as a chemical shift reference. The H_2O ^{1}H signal was suppressed with a presaturation sequence; the average acquisition parameters for ^{1}H measurements were: number of transients = 16, number of points = 65536, acquisition time = 3.33 s, relaxation delay = 1.5 s. ^{1}H-^{13}C HSQC measurements were performed with solvent signal presaturation and the following parameters: number of transients = 64, number of increments = 96, number of points = 2884, acquisition time = 149.968 ms, relaxation delay = 1 s.

3.3 Statistical Analysis

Non-linear regression analysis on the titrations were performed using R version 4.0.5 (R Foundation for Statistical Computing, Vienna, Austria) [54] with the function on equation (1):

$$\delta_{\mathrm{obs}}(\mathrm{pH}) = \frac{\delta_{\mathrm{L}} + \delta_{\mathrm{HL}} \times 10^{\log K - \mathrm{pH}}}{1 + 10^{\log K - \mathrm{pH}}} \quad (1)$$

where δ_{L} is the chemical shift of an unprotonated moiety, δ_{HL} is the chemical shift of the protonated moiety, and log$K$ is the base 10 logarithm of the group-specific protonation constant. Linear regression analyses for the chemical shift-log$K$ data were carried out using the R version 4.0.5 (R Foundation for Statistical Computing, Vienna, Austria). [54]

DOI:10.14753/SE.2023.2817

4. Discussion

The species-specific protonation constants of the compounds analysed here were identical to those obtained in previous works of our group. [19, 55] The species-specific NMR chemical shifts of the $^{\alpha}CH$ and $^{\beta}CH_2$ nuclei were determined by measuring ^{1}H and ^{1}H-^{13}C HSQC NMR spectra at limiting pH values (corresponding to the plateaus on the titration curves of the compounds). The species-specific chemical shifts of the microspecies were determined using Sudmeier-Reilley equations; [56] this method was recently elaborated for the analogous selenocysteine/selenocystine pair. [57]

Summarily, the chemical shifts of the compounds registered at limiting pH values afforded the species-specific chemical shift values of the major microspecies corresponding to those pH values, as the contribution of minor microspecies to the mole fraction-weighted observed chemical shifts are insignificant. These data provide the protonation shifts, which in turn allow the determination of the NMR chemical shifts of the minor microspecies. The species-specific chemical shifts of Cys and CysSSCys microspecies are compiled in Table 1, sorted according to thiolate-bearing, thiol-bearing, and corresponding disulfide-bearing microspecies, respectively. The protonation shift is the chemical shift change a nucleus undergoes when the thiolate moiety changes protonation state from unprotonated form to the protonated (thiol) form. The protonation shifts in ppm determined for Cys are as follows: $\Delta\delta^{1}H(^{\beta}CH_2)$ = 0.29 and -0.03; $\Delta\delta^{13}C(^{\beta}CH_2)$ = -2.0; $\Delta\delta^{1}H(^{\alpha}CH)$ = 0.36; $\Delta\delta^{13}C(^{\alpha}CH)$ = -2.6. Analogous data from CysASH, Hcy, and PenA microspecies are compiled in Tables 2 and 3.

Table 1. *The species-specific chemical shifts in ppm were determined for Cys and CysSSCys microspecies (Figure 5). For the thiol- and disulfide-bearing species, the concomitant thiolate protonation constant given in the third column refer to the relevant thiolate-bearing microspecies that give rise to such thiol- or disulfide-bearing species via protonation or oxidation, respectively.*

class of species	microspecies	thiolate log*k*	^{1}H (ppm)			^{13}C (ppm)	
			α	β		α	β
thiolate-bearing (Cys)	Cys a	10.07	3.099	2.460	2.894	63.13	34.75
	Cys b	8.76	3.631	2.746	3.141	61.25	29.61
	Cys d	8.95	3.479	2.566	2.984	61.72	33.87
	Cys f	7.64	4.011	2.853	3.231	59.84	28.73
thiol-bearing (Cys)	Cys c	10.07	3.461	2.745	2.861	60.54	32.74
	Cys e	8.76	3.993	3.032	3.108	58.66	27.60
	Cys g	8.95	3.841	2.852	2.951	59.13	31.86
	Cys h	7.64	4.373	3.138	3.198	57.25	26.72
disulfide-bearing (CysSSCys)	$(Cys)_2$ a’	10.07	3.575	2.899	3.106	57.74	46.27
	$(Cys)_2$ f’	8.76	4.125	3.186	3.387	56.05	40.23
	$(Cys)_2$ k’	8.95	3.971	3.018	3.205	56.40	45.17
	$(Cys)_2$ p’	7.64	4.526	3.335	3.478	54.61	38.76

Multiple linear regression analyses were accomplished on the data found in Tables 1, 2, and 3 using the NMR chemical shifts as independent variables and the log*k* as the dependent variable; this result is depicted in Figures 7 and 8. Note that the assignment of independent and dependent variables is not meant to reflect a causal relationship between the parameters. It is merely designed to establish a model to predict log*K* values from chemical shifts.

Regarding the results from Cys and CysSSCys microspecies, it is demonstrated that for each of the three cases the $^{\alpha}$CH ^{13}C chemical shift had the most reliable contribution to the model. It can also be seen from the regression parameters in Table 4 that this chemical shift has the best fit and predictive potential on the log*K* values.

Although analysing results from CysASH, PenA, and Hcy microspecies in Figure 8 a good fit of the chemical shift data of both ^{1}H and ^{13}C was noticed. The parameters of the multiple linear regression analysis are shown in Tables 4 and 5.

It is important to make some distinctions between the regression parameters and their interpretation; (a) the adjusted R^2 characterizes the vertical dispersion of the data points around the linear fit and quantifies how much the linear model explains the

variability of the data; (b) the slope of the regression line characterizes the degree and direction of response between the dependent and independent variable, i.e. how much a particular thiolate basicity is accompanied by a different chemical shift.

DOI:10.14753/SE.2023.2817

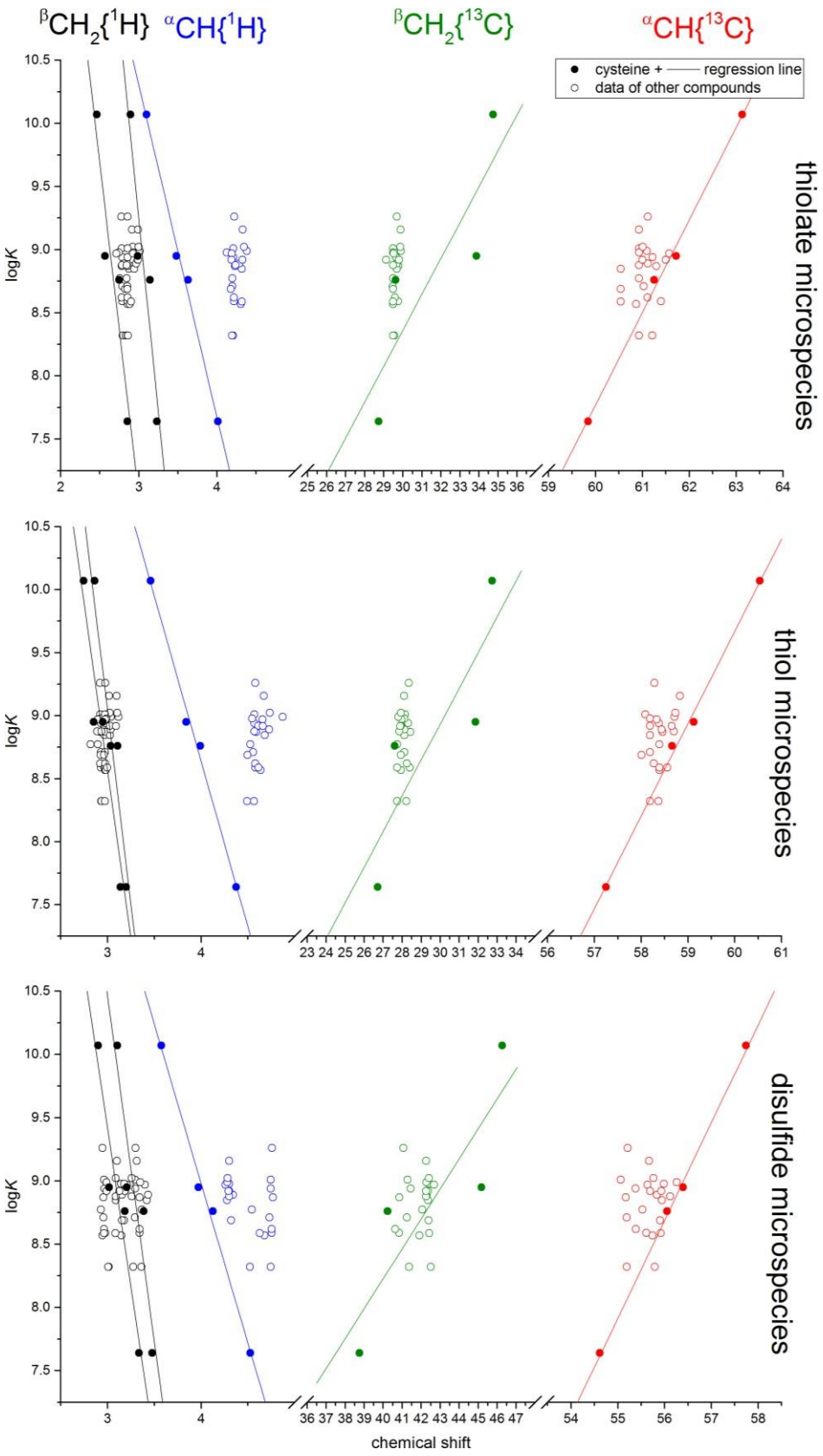


Figure 7. *The multivariate linear regression fits of the chemical shift data vs thiolate basicities.*

DOI:10.14753/SE.2023.2817

Table 2. *The species-specific chemical shifts (on the ppm scale) were determined for CysASH, Hcy, and PenA microspecies (Figures 4 and 5). The thiolate protonation constants were determined previously in [19]. Δδ (Delta delta) indicates how the protonation shift observed on a nucleus, caused by the protonation of the various basic moieties can be quantified in ppm. Note that Δδ is nucleus-dependent; its value is different and distinctive for all types of nuclei and for every atom in the molecule.*

Microspecies	thiolate log*k*	^{1}H (ppm)					^{13}C (ppm)				
		α	**β**	**γ**	**CH_3**		**α**	**β**	**γ**	**CH_3**	
Cysteamine											
CysASH a	9.67	2.646	2.506				48.83	30.57			
CysASH b	8.37	3.076	2.720				46.95	25.43			
Δδ N		*0.431*	*0.213*				*-1.88*	*-5.14*			
Δδ S		*0.129*	*0.118*				*-2.08*	*-1.30*			
Homocysteine											
HCy a	9.94	3.288	1.726	1.805	2.424		58.97	23.71	45.22		
HCy b	8.99	3.807	2.020	2.068	2.503		55.68	23.40	43.58		
HCy d	9.35	3.653	1.793	1.927	2.492		56.95	23.48	44.41		
HCy f	8.4	4.172	2.087	2.190	2.571		51.18	23.05	42.85		
Δδ N		*0.519*	*0.294*	*0.263*	*0.079*		*-3.29*	*-0.31*	*-1.64*		
Δδ S		*0.074*	*0.096*	*0.108*	*0.155*		*0.77*	*-0.81*	*-6.22*		
Δδ O		*0.365*	*0.067*	*0.122*	*0.069*		*-2.02*	*-0.23*	*-0.81*		
Penicillamine											
PenA a	9.34	3.049			1.190	1.435	71.83	49.42		30.91	37.10
PenA b	8.09	3.293			1.291	1.528	69.31	47.81		31.69	35.42
PenA d	8.24	3.464			1.248	1.468	69.94	49.16		30.85	36.66
PenA f	6.99	3.713			1.351	1.566	67.42	47.55		35.36	31.25
Δδ N		*0.249*			*-2.520*	*-5.410*	*0.10*	*0.10*		*-1.61*	*4.51*
Δδ S		*0.404*			*-1.770*	*1.320*	*0.03*	*0.19*		*-1.07*	*-4.95*
Δδ O		*0.415*			*-1.890*	*-0.440*	*0.03*	*0.06*		*-0.26*	*-0.06*

DOI:10.14753/SE.2023.2817

Table 3. *The species-specific chemical shifts (on the ppm scale) were determined for CysASSCysA, HcySSHcy, and penicillamine-disulfide microspecies. The thiolate protonation constants were determined previously in [19]. Δδ (Delta delta) indicates how the protonation shift observed on a nucleus, caused by the protonation of the various basic moieties can be quantified in ppm. Note that Δδ is nucleus-dependent; its value is different and distinctive for all types of nuclei and for every atom in the molecule.*

Microspecies	**log*k***	**^{1}H (ppm)**				
		α	**β**		**γ**	**CH_3**
Cystamine						
CysASSCysA a	9.67	2.928	2.807			
CysASSCysA d	8.37	3.409	3.025			
Δδ NN'		*0.431*	*0.213*			
Homocystine						
HcySSHcy a	9.94	3.347	2.052	1.942	2.783	
HcySSHcy f	8.99	3.860	2.315	2.247	2.826	
HcySSHcy k	9.35	3.727	2.075	2.123	2.859	
HcySSHcy p	8.4	4.240	2.338	2.428	2.902	
Δδ NN'		*0.512*	*0.263*	*0.305*	*0.043*	
Δδ OO'		*0.380*	*0.023*	*0.181*	*0.076*	
Penicillamine disulfide						
Penicillamine-disulfide a	9.34	3.691			1.296	1.321
Penicillamine-disulfide f	8.09	3.982			1.454	1.398
Penicillamine-disulfide k	8.24	3.989			1.340	1.416
Penicillamine-disulfide p	6.99	4.280			1.498	1.493
Δδ NN'		*0.007*			*-0.114*	*0.018*
Δδ OO'		*0.291*			*0.158*	*0.077*

DOI:10.14753/SE.2023.2817

Table 4. *The regression statistics of the Cys and CysSSCys chemical shift data.*

		Intercept		Slope		Statistics	
		Value	Standard error	Value	Standard error	Adj. R-square	P-value
thiolate species	$^{\beta}CH_2\{^1H\}z$	22.92	3.65	-5.294	1.373	0.8222	0.0611
	$^{\beta}CH_2\{^1H\}$	27.66	4.98	-6.141	1.624	0.8159	0.0634
	$^{\beta}CH_2\{^{13}C\}$	-0.16	3.81	0.284	0.120	0.6064	0.1412
	$^{\alpha}CH\{^1H\}$	18.19	0.58	-2.626	0.162	0.9886	0.0038
	$^{\alpha}CH\{^{13}C\}$	-36.21	2.01	0.733	0.033	0.9941	0.0020
thiol species	$^{\beta}CH_2\{^1H\}$	24.43	4.04	-5.294	1.373	0.8222	0.0611
	$^{\beta}CH_2\{^1H\}$	27.46	4.93	-6.141	1.624	0.8159	0.0634
	$^{\beta}CH_2\{^{13}C\}$	0.41	3.57	0.284	0.120	0.6064	0.1412
	$^{\alpha}CH\{^1H\}$	19.14	0.64	-2.626	0.162	0.9886	0.0038
	$^{\alpha}CH\{^{13}C\}$	-34.31	1.92	0.733	0.033	0.9941	0.0020
disulfide species	$^{\beta}CH_2\{^1H\}$	24.38	3.33	-4.992	1.069	0.8739	0.0430
	$^{\beta}CH_2\{^1H\}$	26.86	5.07	-5.465	1.537	0.7952	0.0708
	$^{\beta}CH_2\{^{13}C\}$	-1.25	3.96	0.237	0.093	0.6492	0.1247
	$^{\alpha}CH\{^1H\}$	19.07	0.60	-2.522	0.148	0.9897	0.0034
	$^{\alpha}CH\{^{13}C\}$	-34.60	1.20	0.773	0.021	0.9977	0.0008

DOI:10.14753/SE.2023.2817

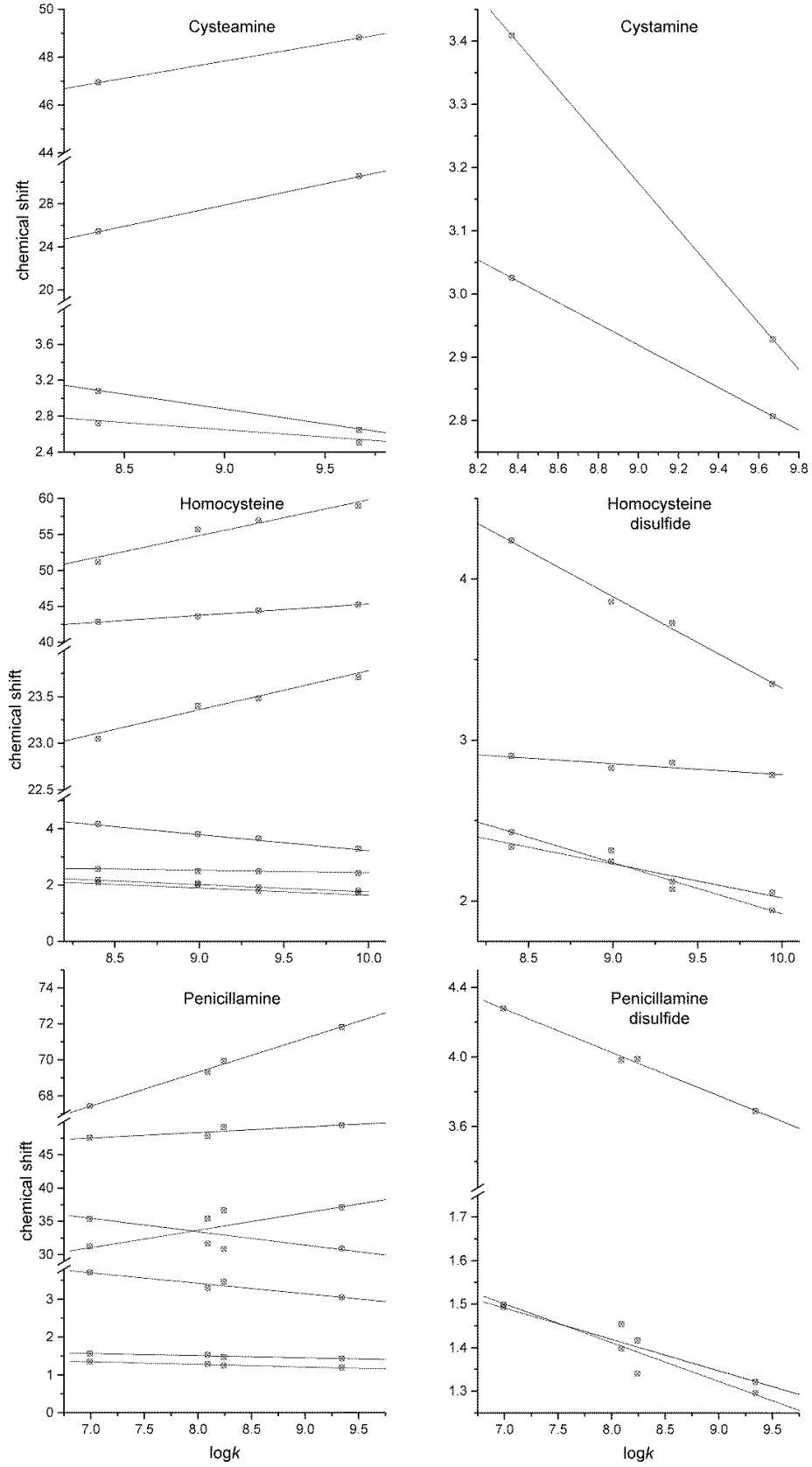


Figure 8. *The multivariate linear regression fits of the chemical shift data vs thiolate basicities.*

Table 5. *The regression statistics of the CysASH, CysASSCysA, Hcy, HcySSHcy, PenA, and penicillamine-disulfide chemical shift data. The best regression models are demonstrated in bold with p-values less than 0.05.*

			Slope		Intercept		Statistics	
			Value	Standard error	Value	Standard error	Adj. R^2	*p*-value
CysASH	^{1}H	α	-0.33	-	5.85	-	-	-
		β	-0.16	-	4.09	-	-	-
	^{13}C	α	1.44	-	34.88	-	-	-
		β	3.95	-	-7.66	-	-	-
CysASSCysA	^{1}H	α	-0.37	-	6.50	-	-	-
		β	-0.17	-	4.43	-	-	-
Hcy	^{1}H	**α**	**-0.57**	***0.02***	**8.92**	***0.21***	**0.9951**	**0.0016**
		βa	-0.26	*0.06*	4.24	*0.57*	0.8415	0.0543
		βa	**-0.26**	***0.02***	**4.36**	***0.20***	**0.9783**	**0.0073**
		γ	**-0.09**	***0.01***	**3.34**	***0.10***	**0.9619**	**0.0128**
	^{13}C	**α**	**4.98**	***0.82***	**10.0**	***7.5***	**0.9228**	**0.0261**
		β	**0.42**	***0.05***	**19.58**	***0.46***	**0.9585**	**0.0139**
		γ	**1.58**	***0.12***	**29.5**	***1.1***	**0.9821**	**0.0060**
HcySSHcy	^{1}H	**α**	**-0.57**	***0.03***	**9.01**	***0.31***	**0.9898**	**0.0034**
		βa	-0.21	*0.08*	4.13	*0.69*	0.6941	0.1078
		βa	**-0.32**	***0.00***	**5.09**	***0.04***	**0.9994**	**0.0002**
		γ	-0.07	*0.03*	3.47	*0.24*	0.6550	0.1225
PenA	^{1}H	**α**	**-0.28**	***0.06***	**5.64**	***0.52***	**0.8553**	**0.0495**
		CH_3a	**-0.07**	***0.01***	**1.84**	***0.08***	**0.9431**	**0.0191**
		CH_3b	-0.06	*0.02*	1.97	*0.13*	0.8093	0.0657
	^{13}C	**α**	**1.89**	***0.10***	**54.23**	***0.86***	**0.9909**	**0.0031**
		β	0.83	*0.37*	41.7	*3.0*	0.5747	0.1536
		CH_3a	-1.91	*0.81*	47.8	*6.6*	0.6033	0.1424
		CH_3b	2.51	*0.83*	14.6	*6.8*	0.7295	0.0947
Penicillamine disulfide	^{1}H	**α**	**-0.25**	***0.01***	**6.02**	***0.11***	**0.9915**	**0.0028**
		CH_3a	-0.09	*0.03*	2.12	*0.25*	0.7183	0.0988
		CH_3b	**-0.07**	***0.01***	**2.00**	***0.07***	**0.9574**	**0.0143**

To broaden the validity of the multiple linear regression model obtained from both groups data (Tables 4 ad 5), other cysteine-derivatives were included in the analysis; notably GSH and other tripeptides were meant to model the varying environments of mid-chain Cys residues. Cys residues with neighboring amino acids of varying electronic effects (compounds **1** – **11**) were assumed to present varying acid-base and NMR characteristics depending on their neighboring residue. Certain non-peptide Cys derivatives with exceptionally electron withdrawing conjugates (compounds **12** – **15**)

were also selected to extend the log*k* range in which data points could be acquired. The typical range of Cys thiolate protonation constants is expected to fall between log*K* 8 – 10; however, since oxidoreductase enzymes must have reactive Cys residues bearing unprotonated thiolate moieties for catalysis, there are plenty of instances in which the Cys thiolate log*k* was found to be much lower than 7. [58-61] Therefore, as an attempt to extend the log*k* range of the linear model well below 7 the compounds listed in Table 6 were selected for further investigation. The species-specific protonation constants of GSH and GSSG were brought from former works. [19, 55] The thiolate protonation constants of the other compounds were determined with ^{1}H NMR-pH titrations by plotting the ^{1}H chemical shift of the Cys $^{\alpha}$CH vs pH. Non-linear regression analyses provided the protonation constants using equation (1). Based on the protonation constants assembled in Table 6, it is evident that the originally anticipated lower thiolate log*k* values were not observed in the Cys derivatives **12** – **15**. Tables 6 – 8 include the NMR chemical shift data determined for these additional compounds as well. In Figure 9 the linear regression fits of only the $^{\alpha}$CH ^{13}C chemical shifts are displayed for the entire data set (Cys/CysSSCys microspecies, GSH/GSSG microspecies and Cys containing peptides).

DOI:10.14753/SE.2023.2817

Table 6. *The species-specific chemical shifts were determined for the compounds studied bearing thiolate moiety (GSH microspecies and compounds presented in Figures 2 and 3).*

compound/ thiolate species		**thiolate log*k***	**$^{\beta}CH_2$**			**$^{\alpha}CH$**	
			$\delta^{1}H$		**$\delta^{13}C$**	**$\delta^{1}H$**	**$\delta^{13}C$**
GSH	GSH A	9.26	2.777	2.856	29.68	4.220	61.12
	GSH B	8.94	2.786	2.855	29.65	4.207	61.22
	GSH D	8.87	2.779	2.858	29.77	4.230	61.30
	GSH E	9.01	2.779	2.855	29.49	4.207	60.93
	GSH G	8.59	2.788	2.857	29.74	4.218	61.40
	GSH H	8.71	2.789	2.854	29.46	4.195	61.03
	GSH K	8.62	2.781	2.857	29.58	4.218	61.11
	GSH N	8.32	2.790	2.856	29.55	4.205	61.21
(**1**)		8.98	2.745	2.930	29.49	4.126	60.93
(**2**)		8.85	2.871	2.942	29.68	4.320	60.54
(**3**)		8.97	2.711	2.923	29.51	4.190	61.57
(**4**)		8.89	2.808	2.964	29.68	4.287	61.12
(**5**)		8.92	2.948	2.986	29.12	4.186	61.50
(**6**)		8.32	2.831	2.860	29.49	4.191	60.93
(**7**)		8.57	2.855	2.885	29.49	4.306	60.87
(**8**)		8.69	2.824	2.852	29.49	4.176	60.54
(**9**)		8.88	2.782	2.844	29.49	4.238	60.93
(**10**)		8.77	2.756	2.845	29.48	4.196	60.93
(**11**)		8.59	2.870	2.903	29.49	4.319	60.54
(**12**)		9.16	2.911	2.990	29.88	4.329	60.93
(**13**)		8.99	2.928	3.015	29.90	4.380	61.10
(**14**)		8.92	2.928	2.991	29.80	4.326	61.00
(**15**)		9.02	2.922	2.999	29.86	4.345	61.01

DOI:10.14753/SE.2023.2817

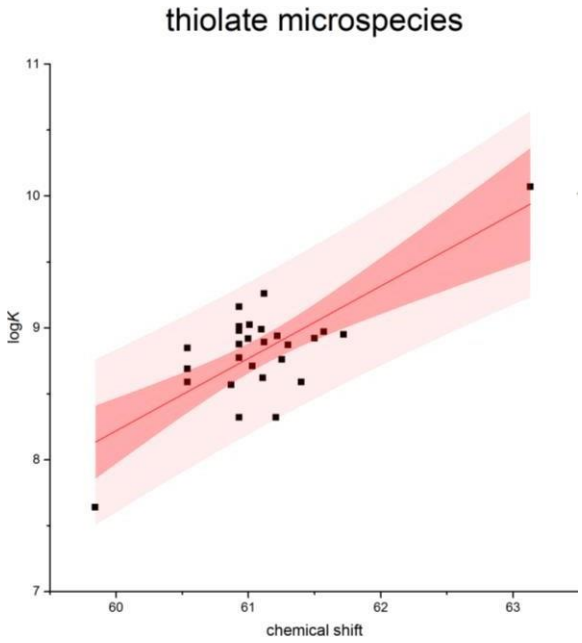
Table 7. *The species-specific chemical shifts together with the protonation shifts determined for all the compounds studied bearing thiol moiety. The mean value and standard deviation of the protonation shifts are 0.14±0.07, 0.08±0.05, -1.62±0.28, 0.37±0.05, -2.64±0.30 for the nuclei found in the table, respectively.*

compound/ thiol species		$^{\beta}CH_2$			$^{\alpha}CH$		$^{\beta}CH_2$			$^{\alpha}CH$	
		δ^1H		$\delta^{13}C$	δ^1H	$\delta^{13}C$	$\Delta\delta^1H$		$\Delta\delta^{13}C$	$\Delta\delta^1H$	$\Delta\delta^{13}C$
GSH	GSH C	2.919	2.975	28.34	4.580	58.28	*0.14*	*0.12*	*-1.3*	*0.36*	*-2.8*
	GSH F	2.929	2.974	28.31	4.568	58.38	*0.14*	*0.12*	*-1.3*	*0.36*	*-2.8*
	GSH I	2.921	2.977	28.43	4.591	58.46	*0.14*	*0.12*	*-1.3*	*0.36*	*-2.8*
	GSH J	2.921	2.973	28.15	4.568	58.09	*0.14*	*0.12*	*-1.3*	*0.36*	*-2.8*
	GSH L	2.930	2.976	28.40	4.578	58.56	*0.14*	*0.12*	*-1.3*	*0.36*	*-2.8*
	GSH M	2.931	2.972	28.12	4.555	58.19	*0.14*	*0.12*	*-1.3*	*0.36*	*-2.8*
	GSH O	2.923	2.975	28.24	4.578	58.27	*0.14*	*0.12*	*-1.3*	*0.36*	*-2.8*
	GSH P	2.933	2.974	28.21	4.566	58.37	*0.14*	*0.12*	*-1.3*	*0.36*	*-2.8*
(**1**)		2.927	2.990	28.12	4.547	58.19	*0.18*	*0.06*	*-1.4*	*0.42*	*-2.7*
(**2**)		2.968	3.026	28.12	4.680	58.19	*0.10*	*0.08*	*-1.6*	*0.36*	*-2.4*
(**3**)		2.896	2.977	27.97	4.658	58.33	*0.18*	*0.05*	*-1.5*	*0.47*	*-3.2*
(**4**)		2.967	3.035	28.12	4.727	58.44	*0.16*	*0.07*	*-1.6*	*0.44*	*-2.7*
(**5**)		2.970	3.028	27.93	4.610	58.19	*0.02*	*0.04*	*-1.2*	*0.42*	*-3.3*
(**6**)		2.936	2.973	27.73	4.494	58.19	*0.10*	*0.11*	*-1.8*	*0.30*	*-2.7*
(**7**)		2.972	2.983	27.93	4.636	58.39	*0.12*	*0.10*	*-1.6*	*0.33*	*-2.5*
(**8**)		2.925	2.962	27.93	4.497	58.01	*0.10*	*0.11*	*-1.6*	*0.32*	*-2.5*
(**9**)		2.888	2.935	27.85	4.570	58.70	*0.11*	*0.09*	*-1.6*	*0.33*	*-2.2*
(**10**)		2.818	2.896	27.73	4.525	58.39	*0.06*	*0.05*	*-1.8*	*0.33*	*-2.5*
(**11**)		2.965	2.996	27.73	4.613	58.39	*0.10*	*0.09*	*-1.8*	*0.29*	*-2.2*
(**12**)		3.018	3.098	28.10	4.672	58.83	*0.11*	*0.11*	*-1.8*	*0.34*	*-2.1*
(**13**)		3.037	3.119	27.80	4.872	58.71	*0.11*	*0.10*	*-2.1*	*0.49*	*-2.4*
(**14**)		3.014	3.092	27.90	4.660	58.65	*0.09*	*0.10*	*-1.9*	*0.33*	*-2.4*
(**15**)		3.023	3.103	27.93	4.734	58.73	*0.10*	*0.10*	*-1.9*	*0.39*	*-2.3*

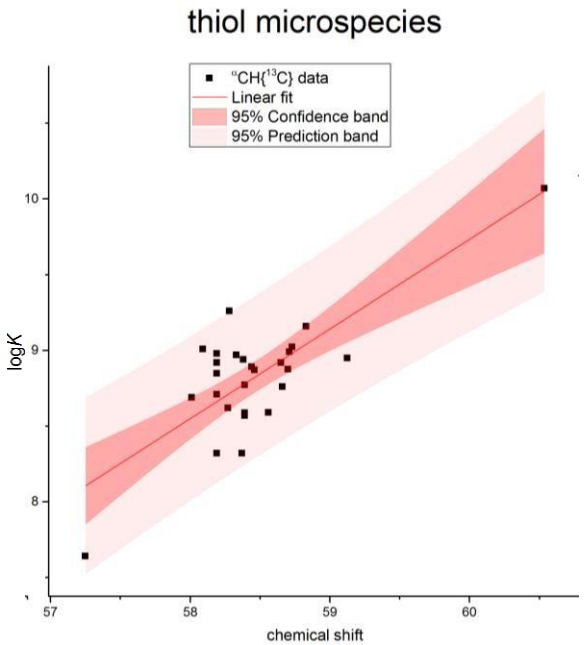
DOI:10.14753/SE.2023.2817

Table 8. *The species-specific chemical shifts of the compounds studied bearing disulfide moiety.*

compound/ disulfide species		**$^{\beta}CH_2$**			**$^{\alpha}CH$**	
		δ^1H		**$\delta^{13}C$**	**δ^1H**	**$\delta^{13}C$**
GSSG	GSSG A’	2.947	3.298	41.06	4.758	55.21
	GSSG E’	2.967	3.297	41.45	4.746	55.38
	GSSG J’	2.956	3.343	40.85	4.769	55.17
	GSSG M’	2.964	3.296	41.28	4.746	55.06
	GSSG X’	2.976	3.343	40.83	4.755	55.61
	GSSG AA’	2.957	3.292	41.24	4.732	55.19
	GSSG AF’	2.963	3.342	40.63	4.755	55.39
	GSSG AJ’	3.015	3.275	41.36	4.744	55.19
homodisulfide of (**1**)		3.146	3.178	42.41	4.300	55.98
homodisulfide of (**2**)		3.088	3.421	42.35	4.278	55.94
homodisulfide of (**3**)		3.207	3.400	42.68	4.261	55.63
homodisulfide of (**4**)		3.153	3.435	42.29	4.341	55.83
homodisulfide of (**5**)		2.989	3.185	42.36	4.300	55.69
homodisulfide of (**6**)		3.007	3.363	42.50	4.522	55.79
homodisulfide of (**7**)		2.946	3.150	41.91	4.677	55.74
homodisulfide of (**8**)		3.177	3.156	42.38	4.320	55.91
homodisulfide of (**9**)		3.088	3.257	42.26	4.299	56.12
homodisulfide of (**10**)		2.931	3.367	42.05	4.542	55.53
homodisulfide of (**11**)		2.962	3.083	42.42	4.628	55.92
homodisulfide of (**12**)		3.100	3.312	42.26	4.299	55.67
homodisulfide of (**13**)		2.988	3.351	42.27	4.281	56.26
homodisulfide of (**14**)		3.189	3.126	42.26	4.293	55.93
homodisulfide of (**15**)		3.089	3.257	42.41	4.285	55.76

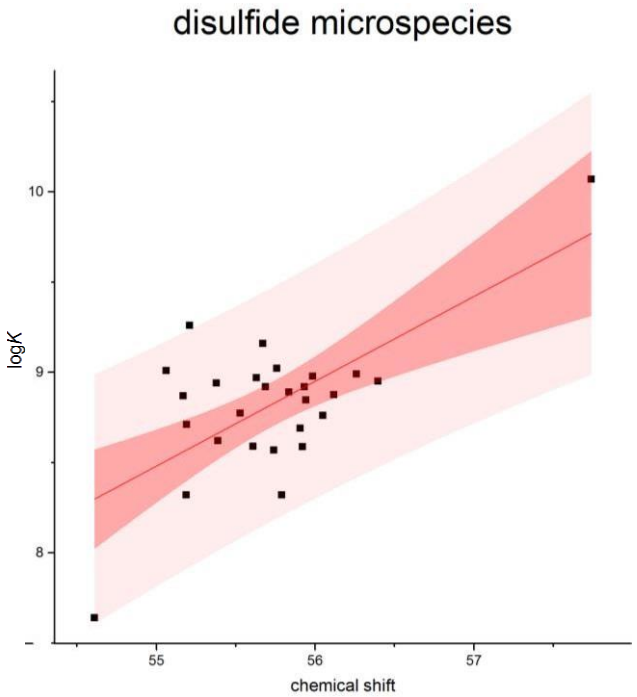


Intercept		Slope		Statistics	
Value	**Standard error**	**Value**	**Standard error**	**Adj. R^2**	***p*-value**
-24.73	**5.96**	**0.549**	**0.097**	**0.5417**	**<0.0001**



Intercept		Slope		Statistics	
Value	**Standard error**	**Value**	**Standard error**	**Adj. R^2**	***p*-value**
-25.82	**5.43**	**0.593**	**0.093**	**0.6043**	**<0.0001**

DOI:10.14753/SE.2023.2817



Intercept		Slope		Statistics	
Value	**Standard. error**	**Value**	**Standard error**	**Adj. R^2**	***p*-value**
-17.38	**5.9**	**0.47**	**0.106**	**0.4189**	**<0.0001**

Figure 9. *The univariate linear regression fits of the $^{\alpha}CH\ ^{13}C$ data together with 95% confidence and prediction bands for Cys/CysSSCys microspecies, GSH/GSSG microspecies and Cys containing peptides (compounds 1 to 15).*

5. Conclusion

In this work, we analysed NMR spectra from Cys, CysSSCys, CysASH, CysASSCysA, Hcy, HcySSHcy, PenA, penicillamine-disulfide, and cysteine-containing peptides to determine their thiolate basicities and the related standard redox potential. The redox potential values were calculated from log*k* in previous works by Mirzahosseini and Noszál. [51] Although Cys as amino acid or moiety is the main regulator of redox homeostasis and signaling, not all Cys residues are expected to be oxidized, it will rely on the solvent accessibility, thiolate basicity, and polarity of the nearby residues. [22] Through the $^{13}C/^{1}H$ chemical shift analysis, a direct and inverse relationship could be identified between them and log*k*, respectively.

Due to the different chemical shift ranges, the ^{1}H protonation shifts are considerably lower compared to the ^{13}C counterparts. Besides that, there are smaller differences between the reduced and oxidized species in terms of ^{1}H chemical shifts as well. In contrast, the ^{13}C chemical shift data of the $^{\alpha}CH$ reveal the redox state of the species and significant physicochemical properties. As shown in Table 5, it is worthy to mention a strong linear relationship between almost every nucleus chemical shifts and thiolate basicities, but particularly in the $^{\alpha}CH$ nuclei, in both ^{1}H and ^{13}C NMR. To support this finding, in a preceding work, Sharma and Rajarathnam already showed that α and β ^{13}C NMR chemical shifts can clearly indicate disulfide bond structure and recognize the reduced and oxidized state of Cys. [49]

It was feasible to notice a better correlation between all chemical shift data and logK in compounds shown in Figure 8. As PenA has lower thiolate basicity, likely due to the shielding effect of the two methyl groups restrained near the thiolate, [26] it is significant to observe that the correlation is maintained. Nevertheless, it is still essential to expand this model to compounds with even lower thiolate basicity.

Regarding the protonation constant results of the cysteine-containing peptides from this work, it is newsworthy to observe that the neighboring residue on a Cys has virtually no bearing on the acid-base properties of the Cys thiolate, implying that the properties of a Cys side-chain can only be perturbed via steric interactions in a peptide. The regression analysis displayed in Figure 9 shows a linear relationship within the data of Cys and CysSSCys microspecies, whereas the correlation data from cysteine-containing peptides only show adherence for the case of $^{\alpha}CH$ ^{13}C. This nucleus reveals

the best conformity to the correlation. Consequently, this is the best option to estimate thiolate properties from NMR data. This is probably because the $^{\alpha}$CH carbon and the sulfur atom are connected via two covalent bonds. The $^{\alpha}$CH ^{1}H chemical shifts of the Cys, however, are presumably perturbed more by the protonation state of adjacent moieties or the presence of a peptide bond, which makes this nucleus a poor indicator on the properties of the status of the sulfur atom. The $^{\beta}CH_2$ nuclei also demonstrate this phenomenon and have a weaker correlation with log*K* altogether. These observations hold for the regression analysis of the thiolate bearing species as well as that of the thiol bearing and the corresponding disulfide bearing species.

It is frequently assumed that chemical shifts of the NMR active nuclei are highly susceptible to changes in the intramolecular microenvironment. Besides that, the correlation from cysteine-containing compound chemical shifts, log*K*, and redox potentials could bring better knowledge about the chemistry and the biological function of its oxidation. [49] Through the gathered chemical shift data set and regression analysis, it is also possible to estimate thiolate basicity/thiol acidity and the concomitant standard redox potential.

Even though, the clear limitation of this method is the window of the regression analysis. Whereas the species-specific thiolate basicities observed in the compounds analysed are limited to a certain range, and the further analysis of derivative compounds did not extend this range (the lowest value analysed was from PenA), in order to extend the scale of the regression more measurements on larger peptides are needed. A detailed literature search for reported thiolate log*K* values in the PKAD Database [62] and the corresponding chemical shift values for the Cys residue using the Biological Magnetic Resonance Data Bank (BMRM) (http://www.bmrb.wisc.edu) was performed. Unfortunately, the literature review generated only a handful of data, that were not reliable enough to incorporate into the model. Thereby our research group is currently investigating larger peptides as we hope that the determination of species-specific chemical shifts of added peptides will extend the regression model for better utility.

NMR spectroscopy has a good capacity to predict some characteristics from simple measurements. Through this technique and analysing the chemical shift of the carbon near the sulfur atom, it was possible to confirm a strong linear relationship within the Cys, CysASH, Hcy, PenA, and their homodisulfides for the chemical shift data vs

DOI:10.14753/SE.2023.2817

thiolate basicity, specifically for the $^{\alpha}CH$ ^{13}C. ^{1}H NMR is another indicator when analysing CysASH, Hcy, and PenA. The forthcoming step in improving this model is to analyse peptides with lower thiolate basicity and extend the correlation that can be used on larger proteins in order to estimate acid-base and redox properties of Cys residues using only the easily accessible NMR chemical shifts. Thereby these results would provide tools for the development of a selective antioxidant compound.

DOI:10.14753/SE.2023.2817

6. Summary

Oxidative stress, known as the imbalance between prooxidants and antioxidants in biological systems in favor of oxidants has been associated with aging, atherosclerosis, carcinogenesis, diabetes, and neurodegeneration. The main moiety targeted by oxidant species in the redox signaling pathways is the thiol (SH) group in the Cys residues, especially in its deprotonated (S^-) form. ^{1}H and ^{13}C NMR measurements were carried out to explore anticipated correlations between chemical shifts versus thiolate basicities and redox potentials of Cys, CysASH, Hcy, PenA, and their homo-disulfides at the submolecular level. Cysteine-containing small peptides were also studied. Regression analysis demonstrated a strong linear relationship for chemical shift vs thiolate log*K* of the microspecies data. The closest correlation was observed for the αCH nuclei in ^{1}H and ^{13}C NMR data. Whereas neither site-specific basicities nor site-specific redox potentials can be directly measured by any means in peptides and proteins containing several thiols and/or disulfide units, these data allow a simple method and predictive power to estimate the above mentioned site-specific physicochemical parameters for analogous sulfur-containing moieties in related biopolymers.

7. References

1. Sies H, Cadenas E. (1985) Oxidative stress: damage to intact cells and organs. Philos Trans R Soc Lond B Biol Sci, 311: 617-631.

2. Sies H. (2015) Oxidative stress: a concept in redox biology and medicine. Redox Biol, 4: 180-183.

3. Ray PD, Huang BW, Tsuji Y. (2012) Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. Cell Signal, 24: 981-990.

4. Sharma GN, Gupta G, Sharma P. (2018) A Comprehensive Review of Free Radicals, Antioxidants, and Their Relationship with Human Ailments. Crit Rev Eukaryot Gene Expr, 28: 139-154.

5. Valko M, Leibfritz D, Moncol J, Cronin MT, Mazur M, Telser J. (2007) Free radicals and antioxidants in normal physiological functions and human disease. Int J Biochem Cell Biol, 39: 44-84.

6. Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoele K, Gollery M, Shulaev V, Van Breusegem F. (2011) ROS signaling: the new wave? Trends Plant Sci, 16: 300-309.

7. Betteridge DJ. (2000) What is oxidative stress? Metabolism, 49:3-8.

8. Sies H, Berndt C, Jones DP. (2017) Oxidative Stress. Annu Rev Biochem, 86: 715-748.

9. Forman HJ. (2016) Redox signaling: An evolution from free radicals to aging. Free Radic Biol Med, 97: 398-407.

10. Forman HJ, Ursini F, Maiorino M. (2014) An overview of mechanisms of redox signaling. J Mol Cell Cardiol, 73: 2-9.

11. Roos G, Messens J. (2011) Protein sulfenic acid formation: from cellular damage to redox regulation. Free Radic Biol Med, 51: 314-326.

12. Guo Q, Li F, Duan Y, Wen C, Wang W, Zhang L, Huang R, Yin Y. (2020) Oxidative stress, nutritional antioxidants and beyond. Sci China Life Sci, 63: 866-874.

13. Comini MA. (2016) Measurement and meaning of cellular thiol:disufhide redox status. Free Radic Res, 50: 246-271.

14. Mutlu H, Ceper EB, Li X, Yang J, Dong W, Ozmen MW, Theato P. (2019) Sulfur Chemistry in Polymer and Materials Science. Macromol Rapid Commun, 40: e1800650:1-51.

15. Groitl B, Jakob U. (2014) Thiol-based redox switches. Biochim Biophys Acta, 1844: 1335-1343.

16. Berndt C, Lillig CH, Holmgren A. (2007) Thiol-based mechanisms of the thioredoxin and glutaredoxin systems: implications for diseases in the cardiovascular system. Am J Physiol Heart Circ Physiol, 292: H1227-1236.

17. Couto N, Wood J, Barber J. (2016) The role of glutathione reductase and related enzymes on cellular redox homoeostasis network. Free Radic Biol Med, 95: 27-42.

18. Lu J, Holmgren A. (2014) The thioredoxin antioxidant system. Free Radic Biol Med, 66: 75-87.

19. Mirzahosseini A, Noszal B. (2014) The species- and site-specific acid-base properties of biological thiols and their homodisulfides. J Pharm Biomed Anal, 95: 184-192.

20. Bak DW, Bechtel TJ, Falco JA, Weerapana E. (2019) Cysteine reactivity across the subcellular universe. Curr Opin Chem Biol, 48: 96-105.

21. Chan KX, Phua SY, Van Breusegem F. (2019) Secondary sulfur metabolism in cellular signalling and oxidative stress responses. J Exp Bot, 70: 4237-4250.

22. Alcock LJ, Perkins MV, Chalker JM. (2018) Chemical methods for mapping cysteine oxidation. Chem Soc Rev, 47: 231-268.

23. Zhang P, Chan W, Ang IL, Wei R, Lam MMT, Lei KMK, Poon TCW. (2019) Gas-Phase Fragmentation Reactions of Protonated Cystine using High-Resolution Tandem Mass Spectrometry. Molecules, 24: 747-755

24. Go YM, Jones DP. (2011) Cysteine/cystine redox signaling in cardiovascular disease. Free Radic Biol Med, 50: 495-509.

25. Salles RCM, Coutinho LH, da Veiga AG, Sant'Anna MM, de Souza GGB. (2018) Surface damage in cystine, an amino acid dimer, induced by keV ions. J Chem Phys, 148: 045107-045108

26. Maw, GA. (1982) Biochemistry of S-Methyl-L-Cysteine and its Principal Derivatives. Sulfur Reports, 2: 1-26.

27. Jakubowski H. (2019) Homocysteine Modification in Protein Structure/Function and Human Disease. Physiol Rev, 99: 555-604.

28. Maron BA, Loscalzo J. (2006) Homocysteine. Clin Lab Med, 26: 591-609.

29. Djuric D, Jakovljevic V, Zivkovic V, Srejovic I. (2018) Homocysteine and homocysteine-related compounds: an overview of the roles in the pathology of the cardiovascular and nervous systems. Can J Physiol Pharmacol, 96: 991-1003.

30. Mudd SH, Finkelstein JD, Refsum H, Ueland PM, Malinow MR, Lentz SR, Jacobsen DW, Brattstrom L, Wilcken B, Wilcken DE, Blom HJ, Stabler SP, Allen RH, Selhub J, Rosenberg IH. (2000) Homocysteine and its disulfide derivatives: a suggested consensus terminology. Arterioscler Thromb Vasc Biol, 20: 1704-1706.

31. Bhushan R, Kumar R. (2010) Enantioresolution of dl-penicillamine. Biomed Chromatogr, 24: 66-82.

32. Ishak R, Abbas O. (2013) Penicillamine revisited: historic overview and review of the clinical uses and cutaneous adverse effects. Am J Clin Dermatol, 14: 223-233.

33. Apruzzese F, Bottari E, Festa MR. (2004) Penicillamine disulfide (PNS) and alkaline cations. Ann Chim, 94: 45-56.

34. Paul BD, Snyder SH. (2019) Therapeutic Applications of Cysteamine and Cystamine in Neurodegenerative and Neuropsychiatric Diseases. Front Neurol, 10: 1315-1324

35. Besouw M, Masereeuw R, van den Heuvel L, Levtchenko E. (2013) Cysteamine: an old drug with new potential. Drug Discov Today, 18: 785-792.

36. Fraser-Pitt DJ, Mercer DK, Smith D, Kowalczuk A, Robertson J, Lovie E, Perenyi P, Cole M, Doumith M, Hill RLR, Hopkins KL, Woodford N, O'Neil DA. (2018) Cysteamine, an Endogenous Aminothiol, and Cystamine, the Disulfide Product of Oxidation, Increase Pseudomonas aeruginosa Sensitivity to Reactive Oxygen and Nitrogen Species and Potentiate Therapeutic Antibiotics against Bacterial Infection. Infect Immun, 86: e00947-00964.

37. Hermanson GT. (2013) Chapter 2 - Functional Targets for Bioconjugation, in Bioconjugate Techniques. Academic Press., 3: 127-228.

38. Ferreira DW, Naquet P, Manautou JE. (2015) Influence of Vanin-1 and Catalytic Products in Liver During Normal and Oxidative Stress Conditions. Curr Med Chem, 22: 2407-2416.

39. Forman HJ, Zhang H, Rinna A. (2009) Glutathione: overview of its protective roles, measurement, and biosynthesis. Mol Aspects Med, 30: 1-12.

40. Sies H. (1999) Glutathione and its role in cellular functions. Free Radic Biol Med, 27: 916-921.

41. Mirzahosseini A, Somlyay M, Noszal B. (2015) Species-Specific Thiol-Disulfide Equilibrium Constant: A Tool To Characterize Redox Transitions of Biological Importance. J Phys Chem B, 119: 10191-10197.

42. Xiao Z, La Fontaine S, Bush AI, Wedd AG. (2019) Molecular Mechanisms of Glutaredoxin Enzymes: Versatile Hubs for Thiol-Disulfide Exchange between Protein Thiols and Glutathione. J Mol Biol, 431: 158-177.

43. Desideri E, Ciccarone F, Ciriolo MR. (2019) Targeting Glutathione Metabolism: Partner in Crime in Anticancer Therapy. Nutrients, 11: 1926-1938.

44. Wellner VP, Anderson ME, Puri RN, Jensen GL, Meister A. (1984) Radioprotection by glutathione ester: transport of glutathione ester into human lymphoid cells and fibroblasts. Proc Natl Acad Sci U S A, 81: 4732-4735.

45. Filipiak P, Hug GL, Bobrowski K, Pedzinski T, Kozubek H, Marciniak B. (2013) Sensitized photooxidation of s-methylglutathione in aqueous solution: intramolecular (S thereforeO) and (S thereforeN) bonded species. J Phys Chem B, 117: 2359-2368.

46. Terwilliger TC, Bollag GE, Sternberg DW, Koshland DE. (1986) S-methyl glutathione synthesis is catalyzed by the cheR methyltransferase in Escherichia coli. J Bacteriol, 1986. 165: 958-963.

47. Arispe N, Diaz JC, Flora M. (2008) Efficiency of histidine-associating compounds for blocking the alzheimer's Abeta channel activity and cytotoxicity. Biophys J, 95: 4879-4889.

48. Mazak K, Noszal B. (2016) Advances in microspeciation of drugs and biomolecules: Species-specific concentrations, acid-base properties and related parameters. J Pharm Biomed Anal, 130: 390-403.

49. Sharma D, Rajarathnam K. (2000) 13C NMR chemical shifts can predict disulfide bond formation. J Biomol NMR, 18: 165-171.

50. Millis KK, Weaver KH, Rabenstein DL. (1993) Oxidation/Reduction Potential of Glutathione. J Org Chem, 58: 4144-4146.

51. Mirzahosseini A, Noszal B. (2016) Species-Specific Standard Redox Potential of Thiol-Disulfide Systems: A Key Parameter to Develop Agents against Oxidative Stress. Sci Rep, 6: 37596-37607

52. Szakacs Z, Hagele G, Tyka R. (2004) H-1/P-31 NMR pH indicator series to eliminate the glass electrode in NMR spectroscopic pK(a) determinations. Analytica Chimica Acta, 522: 247-258.

53. Orgovan G, Noszal B. (2011) Electrodeless, accurate pH determination in highly basic media using a new set of (1)H NMR pH indicators. J Pharm Biomed Anal, 54: 958-964.

54. R Core Team. (2021) R: A language and environment for statistical computing, R Foundation for Statistical Computing.

55. Mirzahosseini A, Somlyay M, Noszál B. (2015) The comprehensive acid–base characterization of glutathione. Chemical Physics Letters, 622: 50-56.

56. Submeier JL, Reilley CN. (1964) Nuclear Magnetic Resonance Studies of Protonation of Polyamine and Aminocarboxylate Compounds in Aqueous Solution. Analytical Chemistry, 36: 1698-1706.

57. Pálla T, Mirzahosseini A, Noszál B. (2020) The species-specific acid-base and multinuclear magnetic resonance properties of selenocysteamine, selenocysteine, and their homodiselenides. Chemical Physics Letters, 741: 137076-137084

58. Nelson JW, Creighton TE. (1994) Reactivity and ionization of the active site cysteine residues of DsbA, a protein required for disulfide bond formation in vivo. Biochemistry, 33: 5974-5983.

59. Zhou DH, Nieuwkoop AJ, Berthold DA, Comellas G, Sperling LJ, Tang M, Shah GJ, Brea EJ, Lemkau LR, Rienstra CM. (2012) Solid-state NMR analysis of membrane proteins and protein aggregates by proton detected spectroscopy. J Biomol NMR, 54: 291-305.

60. Dijkstra K, Karvonen P, Pirneskoski A, Koivunen P, Kivirikko KI, Darby NJ, van Straaten M, Scheek RM, Kemmink J. (1999) Assignment of 1H, 13C and 15N resonances of the a' domain of protein disulfide isomerase. J Biomol NMR, 1999. 14: 195-196.

61. Kortemme T, Darby NJ, Creighton TE. (1996) Electrostatic Interactions in the Active Site of the N-Terminal Thioredoxin-like Domain of Protein Disulfide Isomerase. Biochemistry, 35: 14503-14511.

62. Pahari S, Sun L, Alexov E. (2019) PKAD: a database of experimentally measured pKa values of ionizable groups in proteins. Database (Oxford).

8. Bibliography of candidate's publications

8.1 Publications pertaining to the doctoral thesis

1. **Santana, J. F.;** Mirzahosseini, A.; Noszál, B.. Correlation between the NMR chemical shifts and thiolate protonation constants of cysteamine, homocysteine, and penicillamine. Journal of Spectroscopy, 2022: 9491360-9491368, 2022 - https://doi.org/10.1155/2022/9491360. Impact Factor: 1.750

2. **Santana, J.F.;** Mirzahosseini, A.; Mándity, B.; Bogdán, D.; Mándity, I. and Noszál, B.. Close correlation between thiolate basicity and certain NMR parameters in cysteine and cystine microspecies. PLoS ONE 17(3): e0264866, 2022 - https://doi.org/10.1371/journal.pone.0264866. Impact Factor: 3.752

8.2 Publications pertaining to different subjects

3. Soares, B.A.; Teixeira, K.N.; **de Santana, J.F.;** de Assis, B.L.M.; Zocatelli-Ribeiro, C.; Scandelari, J.P.S.; Thomaz-Soccol, V.; Machado-de-Ávila, R.A.; Alvarenga, L.M.; de Moura, J.. Epitope mapping from *Mycobacterium leprae* proteins: Convergent data from *in silico* and *in vitro* approaches for serodiagnosis of leprosy, Molecular Immunology, Vol. 138, 2021 - https://doi.org/10.1016/j.molimm.2021.07.021

4. **De Santana, J.F.;** da Silva, M.R.B.; Picheth, G.F.; Yamanaka, I.B.; Fogaça, R.L.; Thomaz-Soccol, V.; Machado-de-Avila, R.A.; Chávez-Olórtegui, C.; Sierakowski M.R.; de Freitas, R.A.; Alvarenga, L.M.; de Moura, J..Engineered biomarkers for leprosy diagnosis using labeled and label-free analysis, Talanta, Vol. 187, 2018 - https://doi.org/10.1016/j.talanta.2018.05.027

5. **De Santana, J.F.;** Pilla, V.; Silva, A.C.; Dantas, N.O.; Messias, D.N.; Andrade, A.A.. Optical characterization of core-shell quantum dots embedded in synthetic saliva: Temporal dynamics, Journal of Photochemistry and Photobiology B: Biology, Vol. 151, 2015 - https://doi.org/10.1016/j.jphotobiol.2015.08.016

6. Pilla, V.; Alves, L.P.; **de Santana, J.F**; da Silva, L.G.; Ruggiero, R.; Munin, E.. Fluorescence quantum efficiency of CdSe/ZnS quantum dots embedded in biofluids: pH dependence, Journal of Applied Physics, Vol. 112, 2012 - https://doi.org/10.1063/1.4767470

DOI:10.14753/SE.2023.2817

9. Acknowledgments

Words cannot express my gratitude to Prof. Béla Noszál and Prof. Arash Mirzahosseini for their huge efforts in guiding me through my doctoral work. Without their knowledge, patience, and feedback this work would not be possible.

I am very grateful to Dr. Péter Horváth and every colleague in the Department of Pharmaceutical Chemistry of Semmelweis University for creating a great work environment, especially to Dr. Zoltán Faragó for his friendship and all the help when I arrived in Budapest. I am also grateful for the Stipendium Hungaricum Scholarship for allowing me to pursue my Ph.D. in Hungary.

Lastly, I would like to mention my parents, Roberto and Lucilaine, my sister, Paula and my boyfriend, Marcos. Their belief and support in me have kept my motivation high during this process.