12-2020

Degree of Conservation of Methionines Found to be Oxidized in the Human Urinary Proteome

Alexis Hall
University of Arkansas, Fayetteville

Degree of Conservation of Methionines Found to be
Oxidized in the Human Urinary Proteome

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Chemistry

by

Alexis Hall
John Brown University
Bachelor of Science in Biochemistry, 2017

December 2020
University of Arkansas

This thesis is approved for recommendation to the Graduate Council.

_____
Wesley Stites, Ph.D.
Thesis Director

_____
Josh Sakon, Ph.D.
Committee Member

_____
Chenguang Fan, Ph.D.
Committee Member

_____
Jackson Lay Jr., Ph.D.
Committee Member

Abstract

In previous work from this laboratory, methionine containing peptides from the human urinary proteome were examined by mass spectrometry for the degree of methionine oxidation to the sulfoxide form. While this demonstrated that many of the methionines detected were capable of being oxidized, the question of whether these methionines are important in the structure and/or function of the parent proteins came about. In some proteins, methionine oxidation has been linked to conformational changes and alteration of function and thus can serve as a mechanism for reversible regulation of activity. It is hypothesized that methionines which might serve a regulatory purpose when cycled between the thioether and sulfoxide form will be conserved throughout many species. A list of species whose genomes have been sequenced was designed to span a wide distance across the phylogenetic tree, so that conservation of specific methionines could be analyzed. Some methionines that were susceptible to oxidation were found to be much more conserved than other methionines and also more conserved than their surrounding residues. It is hypothesized that upon further study, these methionines will prove to be more important to the structure and/or function of their respective proteins than the other methionines that were less conserved than their surroundings or those with a low overall conservation. Focusing future work to examine the impact upon protein structure and function of oxidation and reduction of this group of conserved methionines would seem more likely to be fruitful than a broad examination of the redox effect of all methionines.

Table of Contents

Introduction

Post-translational modifications of proteins serve a variety of purposes, but in many ways, an important underlying theme is that they affect the activity of proteins. They, for example, may stabilize a protein, change it from one structure to another, reveal or conceal an active site, or alter a binding affinity. One of the various types of post-translational modifications which has received comparatively little attention is oxidation of methionine residues to the sulfoxide form.

While other amino acids can be oxidized, cysteine and methionine are most susceptible.[1] The oxidation of methionine to methionine sulfoxide is particularly notable as a post-translational modification as it is readily reversible, as reviewed briefly below. The utility of a reversible modification such as phosphorylation in regulation of protein activity is obvious from any introductory biochemistry course and, recently, it has been shown that the oxidation and reduction of methionine is important for the regulation of certain signaling proteins.[2,3] Interestingly, it has also been found that the oxidation of methionine often occurs simultaneously with other post-translational modifications. One study found that methionine residues that are more readily oxidized were located near amino acids that are readily phosphorylated.[4]

While there is some evidence that methionine oxidation and reduction can serve a regulatory function, there are not a large number of examples. In part this is likely because detection of the modification is much more difficult than determining if a protein is phosphorylated or cleaved. There is no good spectroscopic signal, no charges are introduced or removed, and radioactive labeling is not practical. The main way to determine if a specific methionine is oxidized is to examine it using mass spectrometry. Further complicating matters is

the fact that the modification can occur fairly easily and a brief discussion of this process is in order.

Reactive oxygen species (ROS), such as superoxide, singlet oxygen, ozone, peroxides, and hydroxyl and peroxyl radicals, are produced in cells as part of normal cellular functions, such as from NADPH oxidases.[5-8] ROS can be introduced in the body through other means, such as smoking, as well. In the presence of reactive oxygen species, a 2-electron oxidation of methionine forms a double bond between the sulfur of a methionine residue and an oxygen atom from a ROS. Figure 1 shows the structures of both methionine and methionine sulfoxide.[10] Methionine sulfoxide (MetSO) exists in 2 epimers, an S configuration and an R configuration.[1] This oxidation changes the side chain of methionine from a smaller flexible hydrophobic chain to a larger, less flexible, hydrophilic chain.[9] The sulfur of MetSO holds a partial positive charge and the oxygen carries a partial negative charge. These changes affect various protein-protein interactions and protein-substrate interactions.[10] Even the reactivity of the side chain is altered as methionine can be alkylated while MetSO is resistant to alkylation.[11]

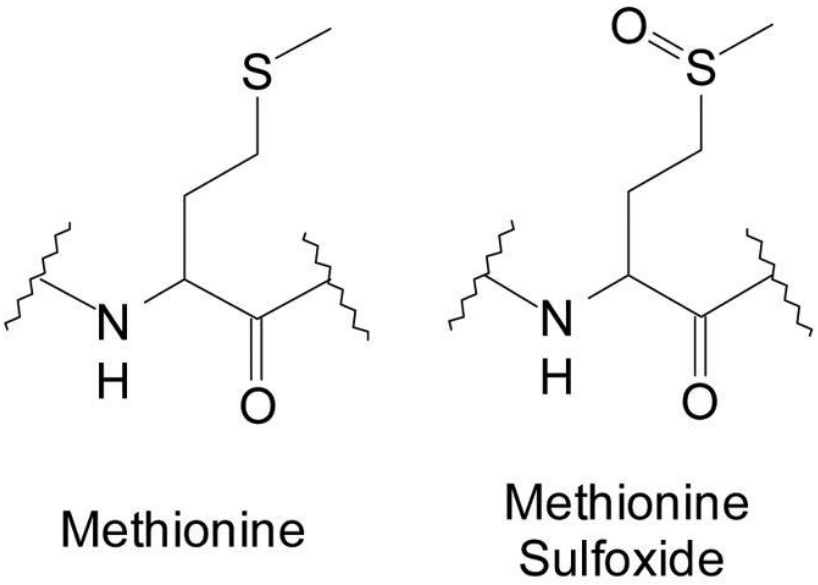

Figure 1. Structures of methionine and methionine sulfoxide

Methionine sulfoxide reductase (MSR) returns methionine sulfoxide in peptides and proteins back to the thioether form of methionine. In order to reverse oxidation of both epimers of MetSO, two methionine sulfoxide reductases exist, MSRA and MSRB, each one specific for an epimer. MSRA reduces S-MetSO while MSRB reduces R-MetSO.[12] Figure 2 shows the stereochemistry of each epimer of MetSO. MSRs contain cysteine residues in the active site that interact with the oxygen of MetSO in order to reduce it back to methionine. Figure 3 shows the mechanism of MSRA to reduce MetSO in bacteria, a similar mechanism to that of mammals.[13]

The ability of methionine to undergo oxidation and reduction readily allows it to guard cells from oxidative damage.[14,15] This is a general ability of any methionine that is accessible to ROS and to MSR and can be thought of as analogous to the role of the glutathione (GSH/GSSG) redox buffering system. Our interest here is not in this general capacity to react with and effectively neutralize ROS and protect the cell from damage, but rather in the possible effect of specific methionines to alter the function of the protein that they are a part of. The oxidation might be caused by a general increase in the concentration of oxidative species or it might be targeted at a specific methionine in a specific protein.

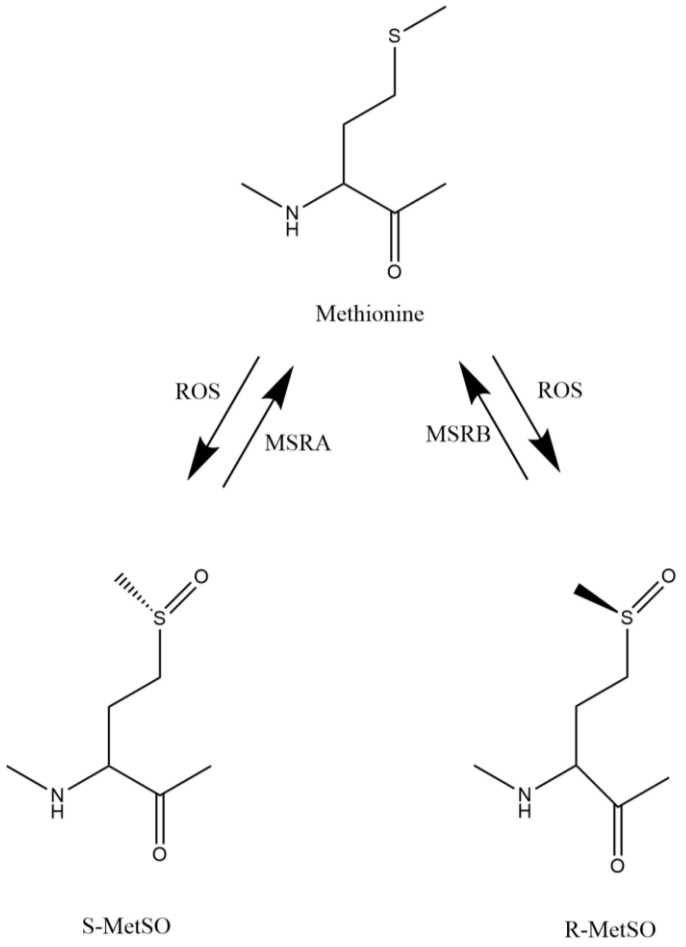

Figure 2. Formation of 2 epimers of methionine sulfoxide by ROS. The S configuration is reduced by MSRA while the R configuration is reduced by MSRB.

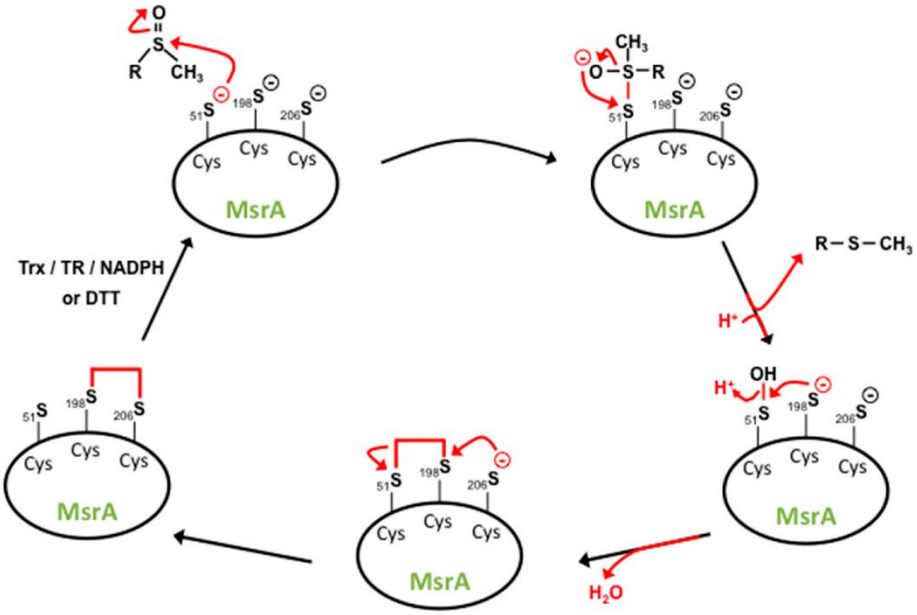

Figure 3. Mechanism of reduction of methionine sulfoxide by MSRA in bacteria. The same mechanism is used in mammalian species to reduce methionine sulfoxide from dos Santos et. al.[13]

One example of methionine oxidation as a role in cell signaling is the role it can play in oxidation-induced ion channels such as the thermosensitive transient receptor potential ion channels. This receptor has been studied in mice (though also present in humans) and was found to lose its oxidation-induced gating function almost completely when two specific methionine residues were substituted. This lost function is a necessary function for macrophages to perform phagocytic activity. Oxidation of these residues is critical for the ion channel to function.[16]

In calcium regulatory proteins such as calmodulin, the oxidation of methionine to methionine sulfoxide leads to a conformational change in the protein. This change prevents activity between proteins and their targets which then signals energy production to slow, leading to a slowed production of other reactive oxygen species.[17,18] Another study performed in *Aspergillus nidulans*, a fungus, found that when MetSO is present at position 169, NirA is

inactive. The presence of nitrogen reduces the MetSO back to methionine, activating NirA, a regulatory gene for nitrogen assimilation. It is believed that the reduction of MetSO to methionine exposes the activation domain of NirA due to a conformational change.[19]

Methionine modification to methionine sulfoxide can also lead to the progression of various diseases. An example of this includes the decreased rate of the thrombomodulin-thrombin complex to activate protein C, shifting hemostasis toward clot formation when the methionine at position 388 in thrombomodulin is oxidized to methionine sulfoxide.[10] In contrast, the kinase CaMKII is activated by the oxidation of methionine residues M281 and M282 which causes association of the regulatory and catalytic domains and keeps CaMKII in the active state until oxidation is reversed by methionine sulfoxide reductase. When studied in mice, it was found that an increase in oxidation of CaMKII can lead to heart arrhythmias. When this same kinase is highly oxidized in lung tissues it can contribute to the development of asthma.[20] In Apolipoprotein A-I, oxidation of Met148 prevents the activation of the enzyme lecithin-cholesterol acyltransferase (LCAT) which is necessary for the conversion of cholesterol to cholesterol ester. This can lead to fatty plaque formation in arteries.[20] Along with these 2 examples, the same study also noted that oxidation of Met 388 in transmembrane protein TM, Met1606 in VWF (linked to von Willebrand disease), Met 46 and Met49 in actin, and Met78, Met367, and Met476 in fibrinogen have also been linked to the advancement of thrombosis and vascular disease.

Methionine oxidation has also been linked to other diseases such as Parkinson's Disease. Formation of MetSO in alpha-synuclein can lead to cytotoxicity connected to Parkinson's Disease.[21] In this case, there are 4 methionine residues that can be oxidized to MetSO. Oxidation of the methionine residues prevents fibrillation in the non-oxidized form of alpha-synuclein. It

was noted that the greater the number of oxidized methionine residues, the greater the inhibition of fibrillation in the non-oxidized protein.

In Alzheimer's disease, aggregation of amyloid beta proteins leads to cell death in neuronal cells. Studies have shown that methionine 35 in this protein can be readily oxidized to MetSO. One study found that this oxidation actually slowed the aggregation of the beta amyloid proteins which in turn slowed cell death. It was noted however that aggregation does still occur although at a slower rate therefore cell death still occurs.[22] One study noted that the presence of MetSO in this protein may signal the activation of MSRA which could also help slow the rate of toxicity however, as age and oxidative stress increase, the activity of MSRA decreases.[23] Another study stated that it is possible this same methionine could donate just one electron and form a radical cation, which would be considered an irreversible oxidation of methionine. This oxidation could also lead to damage within the beta amyloid protein.[24]

There are several clear examples of the use of methionine redox chemistry to regulate protein function. But every solvent exposed methionine is susceptible to oxidation.[25] How can we identify methionine residues that might be playing a regulatory role? Previously, our laboratory conducted a proteomic study of human urine and identified some specific methionines that seem susceptible to oxidation. A general trend has been identified to show that an increase in the degree of conservation correlates to a higher role of importance that amino acids play in the structure or function of its protein.[26-28] We hypothesize that methionines that are playing a redox regulatory role are more likely to be conserved than those that are not. By aligning human proteins with methionines with the sequences of homologous proteins from various species, one can determine the degree of conservation. It is expected that as distance across the phylogenetic tree increases, the degree of overall conservation decreases. However, if an amino acid is vital

for the structure or function of its protein, it is more likely to be conserved across a greater distance than a residue that is not as vital to the protein. A comparison of well conserved and less well conserved methionines is interesting.

Procedure

In a previous study from this lab, human urine samples were collected from 40 individuals. These samples were processed and analyzed using mass spectrometry which identified methionine containing peptides.[29] An Excel sheet containing 308 methionine containing peptides, some of which contain more than one methionine, from 156 proteins was obtained from this previous study. The rate of oxidation of methionine to methionine sulfoxide in these peptides was calculated and included in this Excel sheet.

A conservation study was then performed on these proteins in order to identify which methionine residues were highly conserved and which residues were not conserved. To perform this study, the proteins identified from mass spectrometry were searched using the BLAST tool by inserting their protein accession number. The BLAST protocol was used to search for homologs of the identified proteins across a set list of 39 species, including humans. These species were selected to provide a broad range across the phylogenetic tree. The chosen species had also had their genomes sequenced, increasing the likelihood of finding homologs. This species list contains 23 mammals, 13 non-mammal chordates, and 3 non-chordates. The species list can be viewed in Table 1. A total of 26 proteins were not considered in the remainder of this study due to difficulties when using the BLAST search tool to find sequence homologs. Of these 26 proteins identified using mass spectrometry, 11 resulted in only human sequences when searched using the BLAST tool, 2 consisted of a repeat sequence throughout making it unclear which methionine residue was detected, and 3 proteins contained multiple oxidized methionines

in a single peptide, which greatly complicated analysis. In addition, one detected peptide was matched to be a fragment of La-related protein 4. The accession code for this protein failed to run using the BLAST tool. A total of 9 proteins detected did not have calculated oxidation values and were therefore not used for this study on oxidation and conservation. The other 130 proteins were used in the remainder of this study.

Table 1. Set species list used for setting BLAST parameters when search for protein sequences for a multiple sequence alignment.

Species	Common Name	Taxonomic ID
Mammals		
Homo sapiens	Human	9606
Gorilla gorilla	Gorilla	9593
Pongo abelii	Sumatran orangutan	9601
Otolemur garnettii	Bushbaby	30611
Mus musculus	Mouse	10090
Oryctolagus cuniculus	European rabbit	9986
Sus scrofa	Wild boar	9823
Bos primigenius taurus	Cow	9913
Equus caballus	Horse	9796

Table 1 (Cont.)

Species	Common Name	Taxonomic ID
Delphinapterus leucas	Beluga whale	9749
Neophocaena phocaenoides	Finless porpoise	34892
Felis catus	Domestic cat	9685
Canis lupus familiaris	Domestic dog	9615
Ailuropoda melanoleuca	Giant panda	9646
Odobenus rosmarus	Walrus	9707
Megaderma lyra	Greater false vampire bat	9413
Erinaceus europaeus	Western European hedgehog	9365
Elephas maximus	Asian elephant	9783
Trichechus manatus	West Indian manatee	9778
Ornithorhynchus anatinus	Platypus	9258
Phascolarctos cinereus	Koala	38626
Monodelphis domestica	Gray short-tailed opossum	13616
Sarcophilus harrisii	Tasmanian devil	9305
Non-mammal Chordates		
Anas platyrhynchos	Mallard	8839

Table 1 (Cont.)

Species	Common Name	Taxonomic ID
Gallus gallus	Chicken	9031
Haliaeetus albicilla	White-tailed eagle	8969
Pogona vitticeps	Central bearded dragon	103695
Chelonia mydas	Green sea turtle	8469
Aliigator mississippiensis	American alligator	8496
Nanorana parkeri	High Himalaya frog	125878
Ambystoma mexicanum	Axolotl	8296
Gadus morhua	Atlantic cod	8049
Tetraodon nigroviridis	Green spotted puffer	99883
Oncorhynchus mykiss	Rainbow trout	8022
Petromyzon marinus	Lamprey	7757
Ciona savignyi	Sea squirt	51511
Non-chordates		
Saccoglossus kowalevskii	Acorn worm	10224
Acanthaster planci	Crown of thorns starfish	133434
Bacillus cereus		1396

The homolog searches were completed from January 2019 through September 2019 using the Basic Local Alignment Search Tool (BLAST) on the National Center for Biotechnology Information website.[30] Due to the limitation of the number of species that can be specifically searched in one BLAST run, three links were created in order to obtain homolog sequences from all of the chosen species. All of the links opened the BLAST tool, each with the tax IDs of a portion of the species list from above uploaded into the search set parameters. The first link contained the species from *Homo sapiens* through *Ornithorhynchus anatinus*. The second link was set to search for *Homo sapiens* along with the species from *Phascolarctos cinereus* to *Acanthaster planci*. The third and final BLAST search was set to search for sequences belonging to *Homo sapiens* and *Bacillus cereus*. BLASTP was used to run a protein-protein BLAST search. The general parameter values were set as follows: max target sequences equal to 100, automatic parameter adjustments for short input sequences, expect threshold set to 10, word size set to 6, and max matches in a query range equal to 0. The scoring parameters were set as follows: matrix set to BLOSUM62, gap costs set to existence: 11 and extension: 1, and the compositional adjustments set to conditional compositional score matrix adjustment. Masking and filters were not selected.

After using the BLAST search tool to find homologous proteins, sequence results were saved for the species from the list that had a matching protein sequence. All sequences were then saved into a Word document and aligned using the Clustal Omega online multiple sequence aligner (https://www.ebi.ac.uk/Tools/msa/clustalo/).[31] Alignments were then saved in a Word document to be analyzed. Methionines in peptides that were detected by mass spectrometry were found in the human sequence within the alignment. Methionines conserved in homologous proteins were counted in order to calculate the percentage of conservation of the specific residues

across all species. When counting conserved residues, the counts were taken for three groups: mammals, chordates (including mammals), and total conservation across the other 38 species. Humans were not included in these counts. Percent conservation was calculated for each of the three groups by dividing the number of conserved side chains at each position by the number of homologous proteins found. For example, if a protein BLAST search resulted in sequences for 25 species, but only 15 of those sequences conserved the residue found in humans within those 25, the percent conservation for the group of all species would be 15/25 or 60%. Conservation counts were also taken for three surrounding residues on each side of the methionine. The surrounding counts were averaged to give an average conservation count for the left (N-terminus) and right (C-terminus) sides. These average counts were then divided by the total possible counts as before to give an average percent conservation for the N and C flanking sequences.

In addition to calculating percent conservation in the above manner, ConSurf was used to find a conservation score for the methionines within the proteins. ConSurf is an online program that estimates conservation scores based on algorithms that take evolutionary models and phylogenetic distance into account. ConSurf also predicts whether a residue is more likely to be buried or exposed when calculating scores.[26] ConSurf allows users to search for homologs and calculate conservation scores by either uploading a protein sequence or a PDB code. Both of these methods were used initially. When using ConSurf and uploading a PDB structure code for the protein, the sequence provided belonging to structure given with the uploaded PDB code did not identically match the sequence that was retrieved when running the BLAST search of the protein. In order to run the exact sequences downloaded from the BLAST search, the protein sequence was directly uploaded into ConSurf, rather than uploading a PDB code. In order to do

this, the parameters had to be set to select that a structure did not exist. While there are structures of many of the proteins detected in this study, we did not use these structures for ConSurf. However, ConSurf did match our uploaded sequence to known structures that were very similar in sequence to our proteins. The parameters in ConSurf were set to select up to 150 sequences that sampled a list of homologs that fell within a specified percent identity range. For this study, three ranges of percent identity were created to allow for a sequence search that spanned more broadly across the phylogenetic tree. These ranges were: 35% minimal to 95% maximum (the default setting which most often gave results that were 80% or greater identity match and thus relatively close in the phylogenetic tree), 40% to 65%, and 10% to 35%. By using these 3 ranges, results were obtained not only from various classes of animals but also from some plants and fungi.

Results and Discussion

After all the proteins from which one or more methionine containing peptides detected in a proteomic study of human urinary proteins were aligned and conservation of the methionine and surrounding residues was calculated, a PIVOT table was created with the data. The percentage of methionine conservation among the 39 species is plotted in Figure 4 against the percent oxidation found in the proteomics data for that methionine. Note that some or even all of the oxidation might have occurred during processing of the samples and it is equally possible that a methionine might have been detected for unknown reasons in only the oxidized or reduced forms. In other words, the percentage of oxidation we observed is only a crude indicator that such an oxidation event is possible and by no means definitive evidence that it occurs under physiological conditions. However, this is the only broadly sampled data we are aware indicating what methionines in what proteins are susceptible to oxidation and a wide range of the

percentage of oxidation was found in this work. As expected, even poorly conserved methionines are susceptible to oxidation and some highly conserved methionines showed only a slight propensity to found in the oxidized state.

Again, reasoning that the methionines that are most likely to have possible redox regulatory role are those that can be oxidized readily and those that are well conserved, we somewhat arbitrarily selected for more detailed investigation of proteins with one or more methionines with 80% or greater conservation as well as 50% or greater oxidation, falling in the upper right of Figure 4. A list of these methionine containing peptides and their corresponding proteins can be found in Table 2. A second group of peptides containing methionines with a low methionine oxidation percentage and high percentage of methionine conservation was also selected to compare to the high conservation, high oxidation proteins. The range for this list of methionines was greater than 80% conservation but less than 50% oxidation (lower right of Figure 4) and can be viewed in Table 3.

A plot was also created to determine which proteins on average fell into these same two groups as described above. Figure 5 was created by averaging the percentage of oxidation and percentage of conservation of all methionines within a protein. Each data point represents a different protein. Those proteins that had an average of 80% or greater conservation and an average of 50% or greater oxidation, which fall in the upper right of Figure 5, are listed in Table 4. The proteins that had an average of 80% or greater conservation and an average of less than 50% oxidation, which fall in the lower left of Figure 5, are listed in Table 5. The proteins in both Table 4 and Table 5 were used for further study with ConSurf.

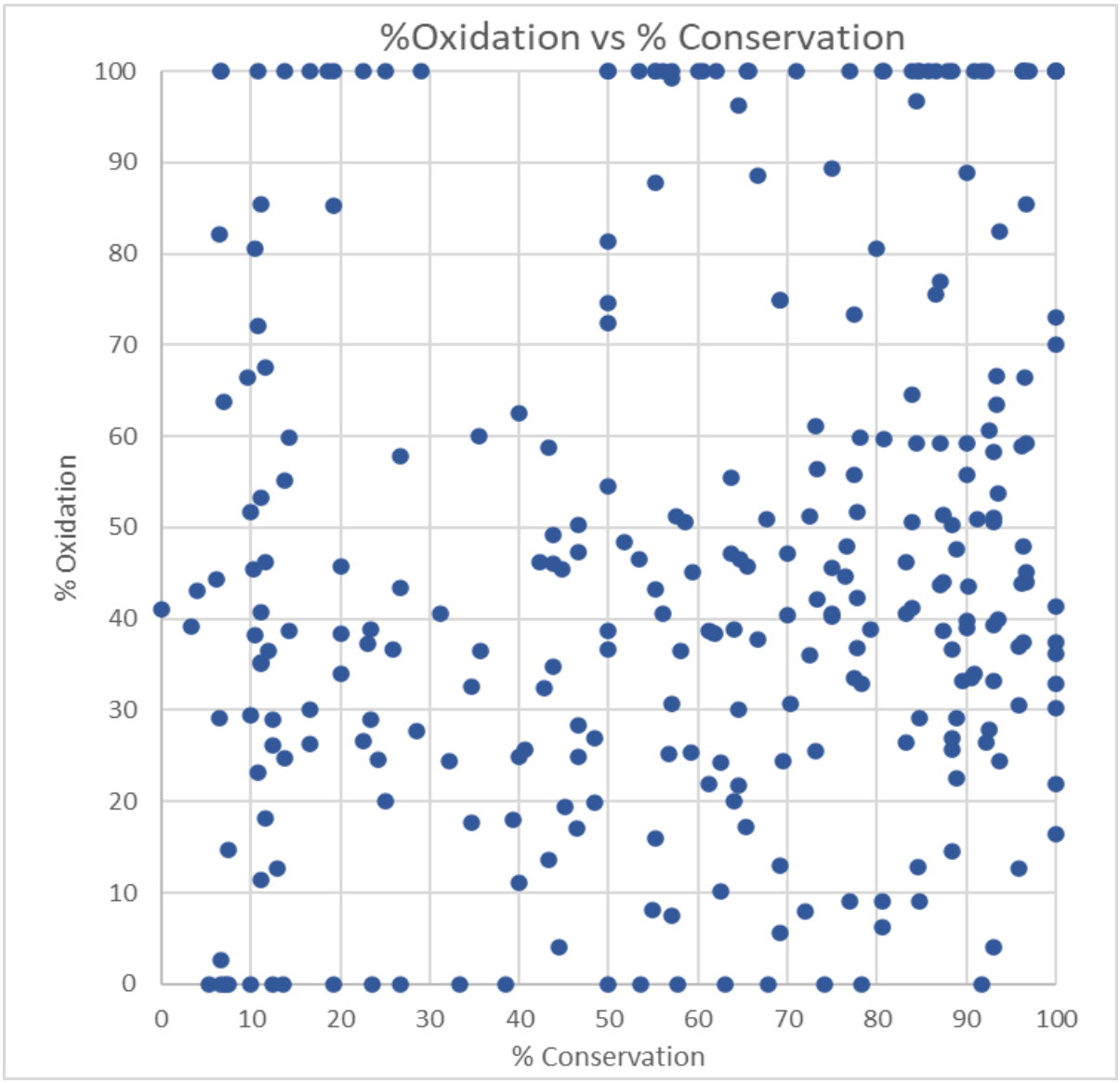

Figure 4. Scatter plot of % oxidation vs. % conservation of methionine residues.

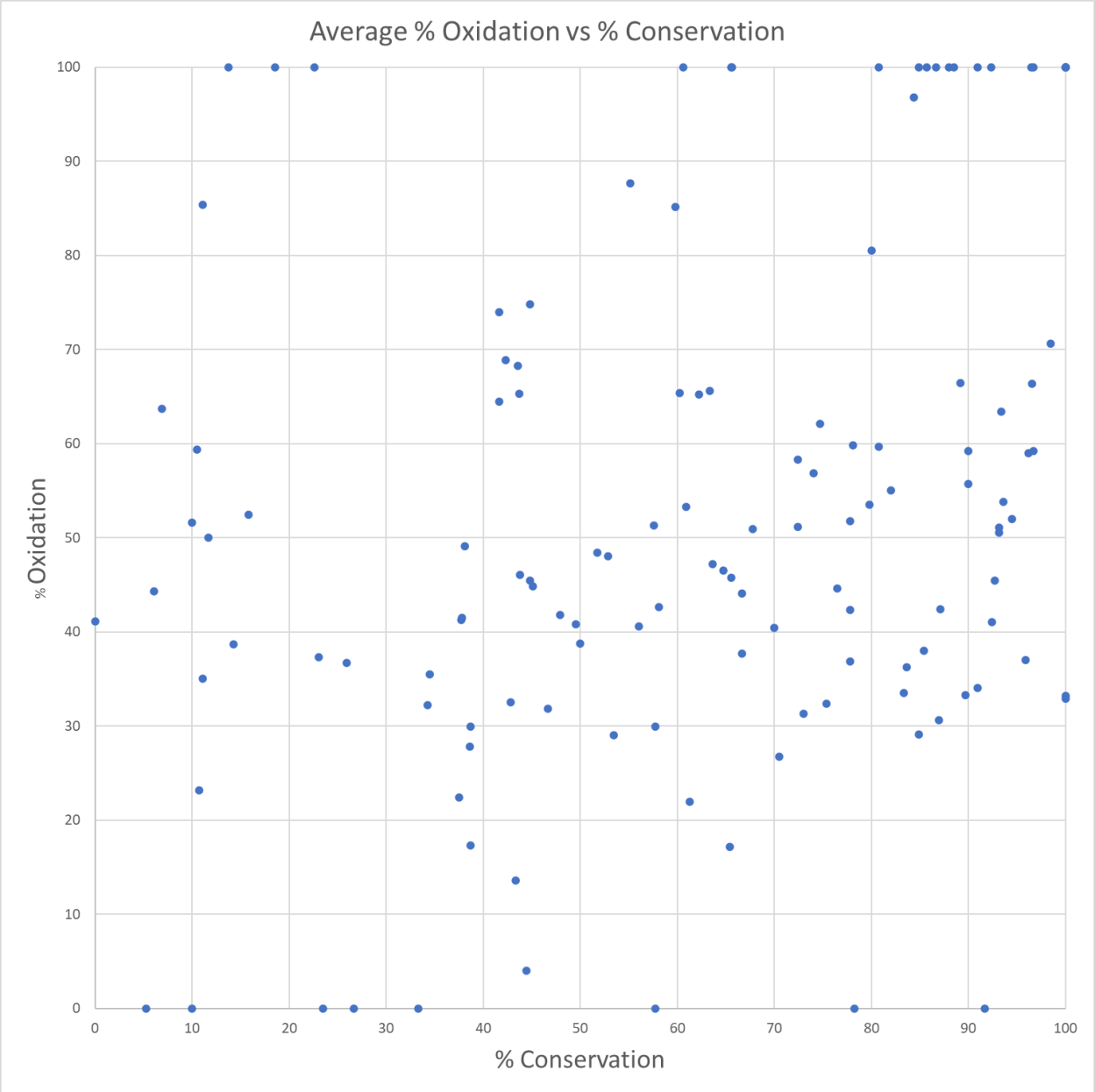

Figure 5. Scatter plot of average % oxidation vs. average % conservation of methionine residues within a protein.

Table 2. Methionines of interest selected with greater than 80% conservation and greater than 50% oxidation.

Protein Accession	Protein Description	Methionine Position #	% Total Conservation	%Ox
P02533	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4	272	100	100
P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4	274	100	100
P0C0L4	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=2	1128	100	100
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6	306	100	100
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3	457	100	70.07
P15311	Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4	200	100	100
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3	326	100	100
P60709	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1	305	100	73.07
P68032	Actin, alpha cardiac muscle 1 OS=Homo sapiens GN=ACTC1 PE=1 SV=1	229	96.9697	100
P98164	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3	1905	96.77419	85.41
P13797	Plastin-3 OS=Homo sapiens GN=PLS3 PE=1 SV=4	169	96.66667	100
Q9NZH0	G-protein coupled receptor family C group 5 member B OS=Homo sapiens GN=GPRC5B PE=2 SV=2	381	96.66667	59.2
P54802	Alpha-N-acetylglucosaminidase OS=Homo sapiens GN=NAGLU PE=1 SV=2	238	96.55172	66.39
P06753	Tropomyosin alpha-3 chain OS=Homo sapiens GN=TPM3 PE=1 SV=2	11	96.42857	100
Q02413	Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2	208	96.42857	100
O43707	Alpha-actinin-4 OS=Homo sapiens GN=ACTN4 PE=1 SV=2	679	96.2963	100
P29508	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2	40	96.2963	100
Q96P63	Serpin B12 OS=Homo sapiens GN=SERPINB12 PE=1 SV=1	41	96.15385	58.98
P04745	Alpha-amylase 1 OS=Homo sapiens GN=AMY1A PE=1 SV=2	338	93.75	82.43
Q562Z4	Actin-like protein (Fragment) OS=Homo sapiens GN=ACT PE=3 SV=1	70	93.54839	53.82
O00560	Syntenin-1 OS=Homo sapiens GN=SDCBP PE=1 SV=1	57	93.33333	63.43
P07339	Cathepsin D OS=Homo sapiens GN=CTSD PE=1 SV=1	201	93.33333	66.68
P07384	Calpain-1 catalytic subunit OS=Homo sapiens GN=CAPN1 PE=1 SV=1	438	93.10345	51.1
P30740	Leukocyte elastase inhibitor OS=Homo sapiens GN=SERPINB1 PE=1 SV=1	228	93.10345	58.3

Table 2. (Cont.)

Protein Accession	Protein Description	Methionine Position #	% Total Conservation	%Ox
P51654	Glypican-3 OS=Homo sapiens GN=GPC3 PE=1 SV=1	316	93.10345	50.56
P02787	Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=3	483	92.59259	60.68
P13646	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=4	340	92.30769	100
P02751	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=4	1548	91.66667	100
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2	467	91.30435	50.93
P11142	Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1	61	90.90909	100
P05155	Plasma protease C1 inhibitor OS=Homo sapiens GN=SERPING1 PE=1 SV=2	324	90	88.91
P68431	Histone H3.1 OS=Homo sapiens GN=HIST1H3A PE=1 SV=2	94	90	59.26
Q8N4F0	BPI fold-containing family B member 2 OS=Homo sapiens GN=BPIFB2 PE=1 SV=2	254	90	55.76
P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3	242	88.46154	100
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	296	88.46154	50.35
Q8WZ42	Titin OS=Homo sapiens GN=TTN PE=1 SV=4	16670	88	100
P01011	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2	284	87.5	51.42
P10253	Lysosomal alpha-glucosidase OS=Homo sapiens GN=GAA PE=1 SV=4	671	87.09677	59.31
P15144	Aminopeptidase N OS=Homo sapiens GN=ANPEP PE=1 SV=4	212	87.09677	77.03
P07602	Prosaposin OS=Homo sapiens GN=PSAP PE=1 SV=2	255	86.66667	75.62
P07737	Profilin-1 OS=Homo sapiens GN=PFN1 PE=1 SV=2	86	86.66667	100
P05067	Amyloid beta A4 protein OS=Homo sapiens GN=APP PE=1 SV=3	303	85.71429	100
Q92616	eIF-2-alpha kinase activator GCN1 OS=Homo sapiens GN=GCN1 PE=1 SV=6	1377	84.84848	100
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3	247	84.61538	100
P04745	Alpha-amylase 1 OS=Homo sapiens GN=AMY1A PE=1 SV=2	302	84.375	59.19
Q2M2H8	Probable maltase-glucoamylase 2 OS=Homo sapiens GN=MGAM2 PE=2 SV=3	623	84.375	96.77

Table 2. (Cont.)

Protein Accession	Protein Description	Methionine Position #	% Total Conservation	%Ox
A8K2U0	Alpha-2-macroglobulin-like protein 1 OS=Homo sapiens GN=A2ML1 PE=1 SV=3	1371	84	50.59
P10253	Lysosomal alpha-glucosidase OS=Homo sapiens GN=GAA PE=1 SV=4	427	83.87097	64.54
P15144	Aminopeptidase N OS=Homo sapiens GN=ANPEP PE=1 SV=4	249	83.87097	100
P08185	Corticosteroid-binding globulin OS=Homo sapiens GN=SERPINA6 PE=1 SV=1	76	80.76923	100
Q99456	Keratin, type I cytoskeletal 12 OS=Homo sapiens GN=KRT12 PE=1 SV=1	251	80.76923	59.66
P98164	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3	2629	80.64516	100
P01040	Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1	1	80	80.55

Table 3. Methionines of interest selected with greater than 80% conservation and less than 50% oxidation.

Protein Accession	Protein Description	Methionine Position #	% Total Conservation	%Ox
P06870	Kallikrein-1 OS=Homo sapiens GN=KLK1 PE=1 SV=2	122	100	30.24
P06870	Kallikrein-1 OS=Homo sapiens GN=KLK1 PE=1 SV=2	197	100	36.17
P15144	Aminopeptidase N OS=Homo sapiens GN=ANPEP PE=1 SV=4	486	100	37.44
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3	157	100	21.9
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2	337	100	16.44
P61626	Lysozyme C OS=Homo sapiens GN=LYZ PE=1 SV=1	35	100	32.87
P68032	Actin, alpha cardiac muscle 1 OS=Homo sapiens GN=ACTC1 PE=1 SV=1	192	100	41.33
P00450	Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1	259	96.77419	44.08
P07339	Cathepsin D OS=Homo sapiens GN=CTSD PE=1 SV=1	219	96.66667	45.11
O43707	Alpha-actinin-4 OS=Homo sapiens GN=ACTN4 PE=1 SV=2	328	96.2963	37.5
O43707	Alpha-actinin-4 OS=Homo sapiens GN=ACTN4 PE=1 SV=2	883	96.2963	48.02
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3	284	96.15385	43.95
P02751	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=4	1283	95.83333	30.51
P02751	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=4	1783	95.83333	12.62
Q7Z3Y9	Keratin, type I cytoskeletal 26 OS=Homo sapiens GN=KRT26 PE=1 SV=2	87	95.83333	37.04
Q16769	Glutaminyl-peptide cyclotransferase OS=Homo sapiens GN=QPCT PE=1 SV=1	223	93.75	24.4
P98164	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3	1004	93.54839	39.97
P30740	Leukocyte elastase inhibitor OS=Homo sapiens GN=SERPINB1 PE=1 SV=1	228	93.10345	4.08
P30740	Leukocyte elastase inhibitor OS=Homo sapiens GN=SERPINB1 PE=1 SV=1	307	93.10345	39.37
P54107	Cysteine-rich secretory protein 1 OS=Homo sapiens GN=CRISP1 PE=1 SV=1	65	93.10345	33.25
P29508	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2	243	92.59259	27.83
P41222	Prostaglandin-H2 D-isomerase OS=Homo sapiens GN=PTGDS PE=1 SV=1	145	92.30769	26.49
Q9UNN8	Endothelial protein C receptor OS=Homo sapiens GN=PROCR PE=1 SV=1	30	91.66667	0

Table 3. (Cont.)

Protein Accession	Protein Description	Methionine Position #	% Total Conservation	%Ox
Q8WUM4	Programmed cell death 6-interacting protein OS=Homo sapiens GN=PDCD6IP PE=1 SV=1	358	90.90909	34.05
P04745	Alpha-amylase 1 OS=Homo sapiens GN=AMY1A PE=1 SV=2	117	90.625	33.54
Q12907	Vesicular integral-membrane protein VIP36 OS=Homo sapiens GN=LMAN2 PE=1 SV=1	277	90.32258	43.58
P15309	Prostatic acid phosphatase OS=Homo sapiens GN=ACPP PE=1 SV=3	75	90	39.04
P15309	Prostatic acid phosphatase OS=Homo sapiens GN=ACPP PE=1 SV=3	75	90	39.75
P55290	Cadherin-13 OS=Homo sapiens GN=CDH13 PE=1 SV=1	267	89.65517	33.29
O43707	Alpha-actinin-4 OS=Homo sapiens GN=ACTN4 PE=1 SV=2	767	88.88889	22.58
P29508	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2	275	88.88889	29.19
P29508	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2	317	88.88889	47.57
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	296	88.46154	36.66
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6	469	88.46154	25.73
P41222	Prostaglandin-H2 D-isomerase OS=Homo sapiens GN=PTGDS PE=1 SV=1	64	88.46154	14.58
P41222	Prostaglandin-H2 D-isomerase OS=Homo sapiens GN=PTGDS PE=1 SV=1	64	88.46154	27
P02751	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=4	926	87.5	38.75
Q16769	Glutaminyl-peptide cyclotransferase OS=Homo sapiens GN=QPCT PE=1 SV=1	65	87.5	43.96
P10253	Lysosomal alpha-glucosidase OS=Homo sapiens GN=GAA PE=1 SV=4	695	87.09677	43.73
P09668	Pro-cathepsin H OS=Homo sapiens GN=CTSH PE=1 SV=4	200	84.84848	29.13
P60709	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1	153	84.84848	9
P05154	Plasma serine protease inhibitor OS=Homo sapiens GN=SERPINA5 PE=1 SV=3	275	84.61538	12.89
Q12907	Vesicular integral-membrane protein VIP36 OS=Homo sapiens GN=LMAN2 PE=1 SV=1	271	83.87097	41.22
O60494	Cubilin OS=Homo sapiens GN=CUBN PE=1 SV=5	2069	83.33333	46.15
P05090	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1	177	83.33333	26.45
P05090	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1	177	83.33333	40.55

Table 3. (Cont.)

Protein Accession	Protein Description	Methionine Position #	% Total Conservation	%Ox
P98164	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3	1299	80.64516	9.06
P98164	Low-density lipoprotein receptor-related protein 2 OS=Homo sapiens GN=LRP2 PE=1 SV=3	2629	80.64516	6.19

Table 4. Proteins of interest selected with greater than 80% conservation and greater than 50% oxidation.

Protein Accession Number	Protein	Average % Conservation	Average % Oxidation
P15311	Ezrin	100.0	100.0
P0C0L4	Complement C4-A	100.0	100.0
P08779	Keratin, type I cytoskeletal 16	100.0	100.0
P02533	Keratin, type I cytoskeletal 14	100.0	100.0
P68032	Actin, alpha cardiac muscle 1	98.5	70.7
Q9NZH0	G-protein coupled receptor family C group 5 member B	96.7	59.2
P13797	Plastin-3	96.7	100.0
P54802	Alpha-N-acetylglucosaminidase	96.6	66.4
P06753	Tropomyosin alpha-3 chain	96.4	100.0
Q96P63	Serpin B12	96.2	59.0
O43707	Alpha-actinin-4	94.4	52.0

Table 4. (Cont.)

Protein Accession Number	Protein	Average % Conservation	Average % Oxidation
Q562Z4	Actin-like protein (Fragment)	93.5	53.8
O00560	Syntenin-1	93.3	63.4
P51654	Glypican-3	93.1	50.6
P07384	Calpain-1 catalytic subunit	93.1	51.1
P13646	Keratin, type I cytoskeletal 13	92.3	100.0
P11142	Heat shock cognate 71 kDa protein	90.9	100.0
Q8N4F0	BPI fold-containing family B member 2	90.0	55.8
P68431	Histone H3.1	90.0	59.3
P13645	Keratin, type I cytoskeletal 10	89.1	66.4
P02538	Keratin, type II cytoskeletal 6A	88.5	100.0
Q8WZ42	Titin	88.0	100.0

Table 4. (Cont.)

Protein Accession Number	Protein	Average % Conservation	Average % Oxidation
P07737	Profilin-1	86.7	100.0
P05067	Amyloid beta A4 protein	85.7	100.0
Q92616	eIF-2-alpha kinase activator GCN1	84.8	100.0
Q2M2H8	Probable maltase-glucoamylase 2	84.4	96.8
P04745	Alpha-amylase 1	82.0	55.1
Q99456	Keratin, type I cytoskeletal 12	80.8	59.7
P08185	Corticosteroid-binding globulin	80.8	100.0

Table 5. Proteins of interest selected with greater than 80% conservation and less than 50% oxidation.

Protein Accession Number	Protein	Average % Conservation	Average % Oxidation
P06870	Kallikrein-1	100.0	33.2
P61626	Lysozyme C	100.0	32.9
Q7Z3Y9	Keratin, type I cytoskeletal 26	95.8	37.0
P02751	Fibronectin	92.7	45.5
P60709	Actin, cytoplasmic 1	92.4	41.0
Q9UNN8	Endothelial protein C receptor	91.7	0.0
Q8WUM4	Programmed cell death 6-interacting protein	90.9	34.1
P55290	Cadherin-13	89.7	33.3
Q12907	Vesicular integral-membrane protein VIP36	87.1	42.4
P35908	Keratin, type II cytoskeletal 2 epidermal	87.0	30.6
Q16769	Glutaminyl-peptide cyclotransferase	85.4	38.0

Table 5 (Cont.)

Protein Accession Number	Protein	Average % Conservation	Average % Oxidation
P09668	Pro-cathepsin H	84.8	29.1
P30740	Leukocyte elastase inhibitor	83.6	36.3
P05090	Apolipoprotein D	83.3	33.5

Scatter plots were also created for the proteins to examine how the percent conservation of methionine compared to the percent conservation of the surrounding residues. The first scatter plot, Figure 6, shows the average percent conservation of the three residues to the left, or on the N-terminus side, of the detected methionine compared to the percent conservation of the methionine residues within a protein. Figure 7 shows the average percent conservation of the three residues to the right, or on the C-terminus side, of the detected methionine compared to the percent conservation of the methionine residues detected within a protein. The overall average percent conservation of the surroundings was calculated by averaging the C-terminus and N-terminus percent conservations together. This average was plotted with the percent conservation of the methionine in Figure 8. One protein, Cystatin-A, was not included in these graphs because the detected methionine was at position 1, therefore conservation of surroundings could only be counted for the C-terminus side of the methionine.

Using Figures 6, 7, and 8 it was determined that in general, the degree of methionine conservation correlated with the degree of conservation of the surrounding residues. However, methionine is generally less conserved than the surrounding residues. Because there is only one codon that corresponds to methionine, it is expected that methionine would be less conserved than any other amino acid except for tryptophan, also with only a single codon. A change in any of the three positions in the codon would result in a mutation from methionine. In Figures 6, 7, and 8, a line runs through the plot. Any points below this line represent detected methionines that are more conserved than the average of its surroundings. This greater than expected conservation may indicate a greater degree of selective pressure on these residues, preserving them as methionines.Therefore, the methionine residues that fell below the line and were above 70% conserved were selected as residues of interest. These residues should be examined further for

the impact of oxidation and reduction on the protein. These selected methionines are represented in Tables 6-10.

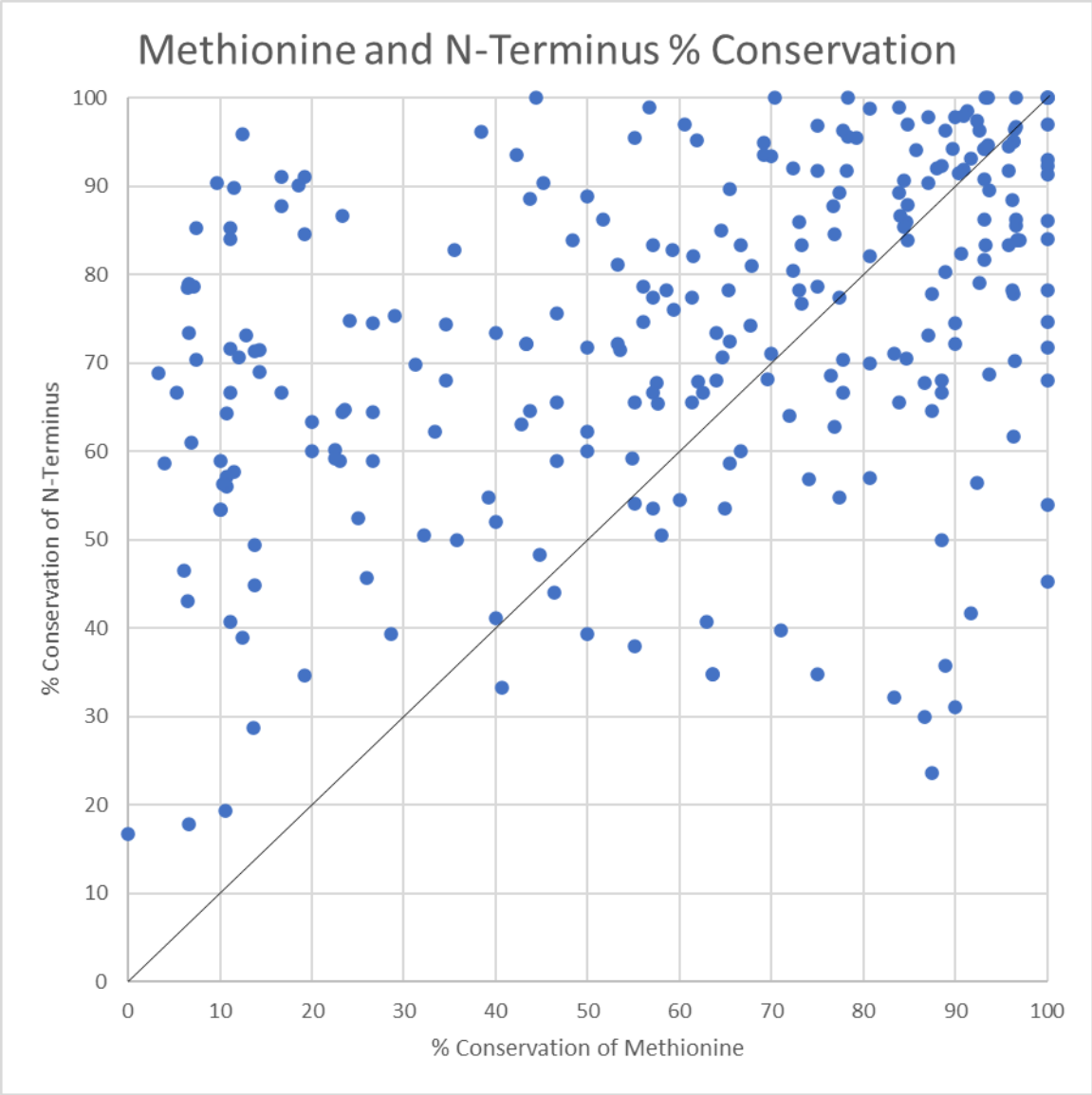

Figure 6. Average percentage of conservation of residues on the N-terminus of the detected methionine and percentage of conservation of the detected methionine residues.

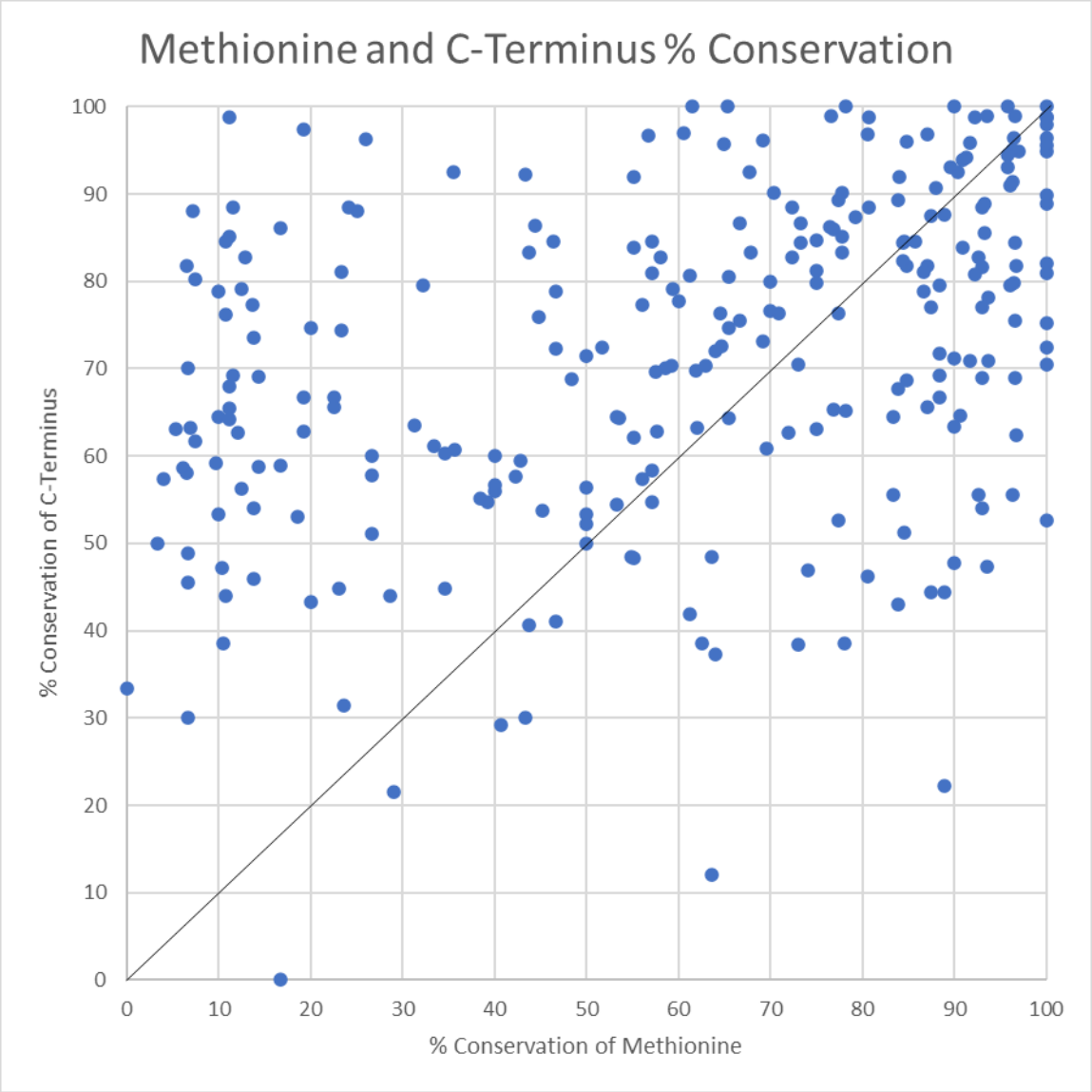

Figure 7. Average percentage of conservation of residues on the C-terminus of the detected methionine and percentage of conservation of the detected methionine residues.

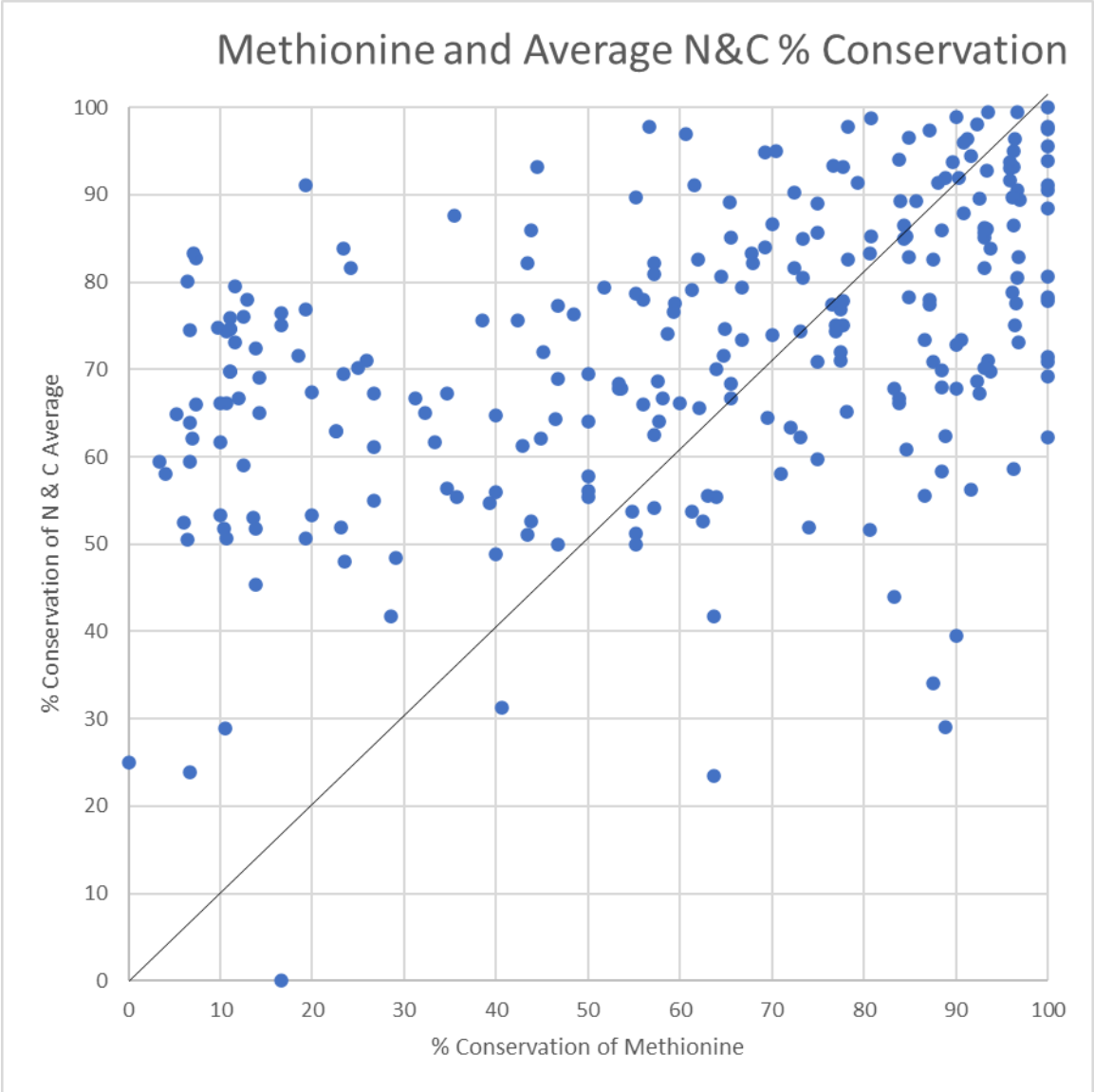

Figure 8. Average percentage of conservation of residues on the N and C-terminus of the detected methionine and percentage of conservation of the detected methionine residues.

For the proteins with conserved methionines with high or low oxidation levels, ConSurf was also used in order to calculate conservation.[26] Using ConSurf, conservation scores were generated over the three identity ranges described in methods (the default 35-95% conservation which typically resulted in alignment of only proteins with overall high levels of overall conservation close to humans in the phylogenetic tree, a mid-range of overall conservation of 40-65% and the most divergent range 10-35% range of overall conservation, aligning species far from humans on the phylogenetic tree). In addition to the conservation data for these three ranges, the multiple sequence alignments between the set list of species determined as described earlier generated by Clustal Omega for the proteins of interest were uploaded into ConSurf in order to compare the ConSurf conservation scores for these alignments with the scores calculated using the sequences matched within the three percent identity ranges set in ConSurf. The % total conservation scores calculated previously were also compared. These results are summarized in Table 11, which compares the conservation scores for those proteins determined to have an average methionine conservation greater than 80% and an average methionine oxidation of greater than 50% and Table 12, which compares the conservation scores for those proteins determined to have an average methionine conservation greater than 80% but an average methionine oxidation of less than 50%.

In Tables 11 and 12, the “score (normalized)” row is the normalized conservation score produced by ConSurf. The more negative scores represent residues that are more highly conserved while the more positive scores represent the residues that are less conserved throughout various homologues. The “color” value is assigned by ConSurf as a number 1-9 matched to the conservation scores, where 9 is the most highly conserved and 1 is the least conserved. If a known structure of the protein exists, this color identification matches each

number to a specific color in order to color code residues based on their conservation scores. The "confidence interval" row is representative of the conservation interval determined by ConSurf for the normalized score. The "confidence interval color" row is the confidence interval for the color number assigned to the residue. The "percent" column represents the conservation percentages calculated with all possible sequences for the set species list described in Table 1 obtained before using ConSurf. Tables 13 and 14 summarize these findings by only focusing on the color value assigned and the percentage of conservation calculated. In some cases, the ConSurf program failed to run the sequence given. In most of these instances, there were not enough homologues matched to properly align the sequences and calculate a conservation score. Titin failed to run in ConSurf because the sequences uploaded were too large for the program. Lysozyme C failed to run using the Clustal upload because less than 5 of the species from the created list had sequence results from using the BLAST search tool. ConSurf requires a minimum of 5 matches in order to calculate a conservation score. Reproducibility of ConSurf was verified by rerunning several proteins at each of the specified percent identity ranges 7-10 days after the first ConSurf trials were run. The results were identical for all proteins at all ranges.

Table 13. Summary of ConSurf results for proteins with an average percentage of conservation greater than 80% and an average percentage of oxidation greater than 50%. "Close" represents the ConSurf search with up to a 95% identity. "Mid" represents the ConSurf search with a 40-65% identity match. "Far" represents the ConSurf search using 10-35% identity match. "Clustal" represents the ConSurf search using the multiple sequence alignment downloaded from Clustal using the set species list. Each number in the "Close", "Mid", "Far", and "Clustal" column represents a number, 1-9, assigned by ConSurf where 1 is the least conserved and 9 is the most conserved.

Protein	Methionine Position	Close	Mid	Far	Clustal	% Conservation
Actin, alpha cardiac muscle 1	192	4	2	4	8	100.0%
	229	3	2	3	5	97.0%
eIF-2-alpha kinase activator GCN1	1377	6	5	8	4	84.9%
Keratin, type I cytoskeletal 14	272	8	8	6	9	100.0%
Syntenin-1	57	9	9	5	9	93.3%
Alpha-actinin-4	328	4	3	7	8	96.3%
	679	1	1	2	8	96.3%
	767	5	4	5	1	88.9%
	883	9	9	8	8	96.3%
Ezrin	200	9	9	8	9	100.0%
Complement C4-A	1128	9	9	8	9	100.0%
Keratin, type I cytoskeletal 16	274	8	8	5	9	100.0%
G-protein coupled receptor family C group 5 member B	381	8	8	6	9	96.7%
Plastin-3	169	7	7	6	7	97.0%

Table 13(Cont.)

Protein	Methionine Position	Close	Mid	Far	Clustal	% Conservation
Alpha-N-acetylglucosaminidase	238	9	9	8	9	96.6%
Tropomyosin alpha-3 chain	11	8	4	4	9	96.4%
Serpin B12	41	8	9	8	9	96.15%
Actin-like protein (fragment)	70	5	7	6	1	93.6%
Glypican-3	316	3	2	6	8	93.1%
Calpain-1 catalytic subunit	438	4	3	1	9	93.1%
Keratin, type I cytoskeletal 13	341	6	6	7	7	92.3%
Heat shock cognate 71 kDa protein	61	6	1	6	4	90.9%
BPI fold-containing family B member 2	254	9	Error	9	8	90.0%
Histone H3.1	94	1	1	4	1	90.0%
Keratin, type I cytoskeletal 10	291	7	6	6	8	78.3%
	306	8	8	9	9	100.0%
Keratin, type II cytoskeletal 6A	242	6	6	6	7	88.5%
Titin	16670	Error	Error	Error	Error	88.0%
Profilin-1	86	8	7	7	7	86.7%
Amyloid beta A4 protein	303	6	1	Error	2	85.7%
Probable maltase-glucoamylase 2	623	9	9	Error	8	84.4%

Table 13(Cont.)

Protein	Methionine Position	Close	Mid	Far	Clustal	% Conservation
Alpha-amylase 1	117	9	9	8	9	90.6%
	198	6	5	8	3	59.4%
	302	4	4	4	6	84.4%
	338	8	7	6	8	93.75
Keratin, type I cytoskeletal 12	251	5	5	5	5	80.8%
Corticosteroid-binding globulin	76	7	7	8	7	80.8%

Table 14. Summary of ConSurf results for proteins with an average percentage of conservation greater than 80% and an average percentage of oxidation less than 50%. "Close" represents the ConSurf search with up to a 95% identity. "Mid" represents the ConSurf search with a 40-65% identity match. "Far" represents the ConSurf search using 10-35% identity match. "Clustal" represents the ConSurf search using the multiple sequence alignment downloaded from Clustal using the set species list. Each number in the "Close", "Mid", "Far", and "Clustal" column represents a number, 1-9, assigned by ConSurf where 1 is the least conserved and 9 is the most conserved.

Protein	Methionine Position	Close	Mid	Far	Clustal	% Conservation
Kallikrein-1	122	Error	9	8	9	100.0%
	197	Error	9	7	9	100.0%
Lysozyme C	35	6	6	6	Error	100.0%
Keratin, type I cytoskeletal 26	87	9	9	8	8	95.9%
Fibronectin	926	6	7	Error	3	87.5%
	1283	3	6	Error	9	95.8%
	1548	7	7	Error	5	91.7%
	1783	9	9	Error	9	95.8%
Actin, cytoplasmic 1	153	4	5	7	1	84.8%
	305	8	7	7	9	100.0%
Endothelial protein C receptor	30	8	8	4	8	91.7%
Programmed cell death 6-interacting protein	358	9	9	8	7	90.9%
Cadherin-13	267	8	8	7	6	89.7%

Table 14 (Cont.)

Keratin, type II cytoskeletal 2 epidermal	294	6	Error	5	4	69.6%
	337	9	Error	7	9	100.0%
	467	8	Error	9	8	91.3%
Glutaminyl-peptide cyclotransferase	65	6	4	1	7	87.5%
	223	7	7	4	8	93.8%
	332	2	1	2	5	75.0%
Pro-cathepsin H	200	5	5	5	8	84.8%
Leukocyte elastase inhibitor	156	1	1	3	1	55.2%
	228	9	9	9	8	93.1%
	307	7	8	7	8	93.1%
Apolipoprotein D	177	8	7	7	7	83.3%
Vesicular integral-membrane protein VIP36	271	5	5	7	4	83.9%
	277	5	5	4	6	90.3%

When using ConSurf, as expected, conservation scores generally decreased as phylogenetic distance increased although there were some residues within the tested proteins that had higher conservation scores in the low % match category than in the high % match category. It was determined that when uploading alignments generated by Clustal using the set species list as described in Table 1, the conservation scores tended to be higher and of more significance than the conservation scores obtained when ConSurf chose which homologs to align. This indicates that while some residues may not seem too highly conserved or of a large importance when only using ConSurf, those same residues can be of significance in ConSurf when the species list is restricted to those in Table 1.

In addition to running the protein sequences through ConSurf in order to verify the conservation numbers calculated in this study, individual research was done to see if any of the proteins of interest had already been studied for the effect of oxidation of methionine residues on the protein's structure and function. Several of the proteins of interest had been previously studied for the effect of the formation of MetSO on the protein. Below are summaries of the findings of some of the proteins of interest.

Profilin-1 was found to be a protein of interest with 100% oxidation and over 85% conservation in all 3 groups of species. In this study it was found that the methionine at position 86 was oxidized. Profilin-1 is an isoform of profilin which is a protein that binds to actin for structural support of the cytoskeleton. It was found that profilin binds to actin at residues 128-134 of profilin.[32] It was also discovered that when profilin is present at high concentrations, actin will not polymerize. When profilin is present at low concentrations, actin polymerization increases.[33] Profilin has also been found to bind to phosphatidylinositol-4,5-bisphosphate.[32] The methionine at position 114 in profilin-1 was determined to be of importance in the progression of

ALS (amyotrophic lateral sclerosis). When a mutation occurs at this position, changing methionine to threonine, the structure of profilin-1 becomes destabilized due to the creation of a cleft in the protein.[34] Possible treatment options for ALS were studied by targeting this cleft. Although this specific methionine residue was not detected in the current study, it is interesting to note that the normalized conservation score for this residue was highly conserved with a value of -0.724 and a confidence interval ranging from -0.969 to -0.602 when the percent identity range was set from 35% to 95% identity.

As mentioned in the introduction, beta amyloid proteins are of interest in the study of Alzheimer's Disease (AD). It has been found that amyloid beta aggregates in the cells of Alzheimer's patients leading to cell death. In previous studies, it was found that methionine 35 of beta amyloid readily oxidizes to methionine sulfoxide and this formation of MetSO slows the aggregation of the proteins.[22,35] One study in particular noted that this oxidation leads to a conformational change by unwinding the helix of bound proteins at the C-terminus. Following this unwinding, it was noted that interactions become unstable and the C-terminus of this protein begins to interact more strongly with lipids of the cell bilayer. This interaction prevents the rest of the protein from binding with the bilayer, reducing the overall affinity that amyloid beta has for the bilayer.[22] This previous study was performed on peptides made of up to 40 residues that are cleaved from a precursor protein such as the beta amyloid A4 protein studied here. In this current study, a small peptide was detected with an oxidized methionine. The peptide identified was matched to have originated from the amyloid beta A4 protein. When this protein was sequenced, the identified methionine was from position 303 of the amyloid beta A4 precursor. When the beta amyloid peptide studied in AD was aligned with the amyloid beta A4 from this study, Met35 of the peptide was equivalent to Met706 of amyloid beta A4. While the same

methionine residue was not detected, it is clear that methionine oxidation plays a major role in the progression of AD.

Actin has 6 different isoforms. These isoforms are very similar, only differing by a few residues primarily at the N terminus.[36] In this study, peptides from actin, cytoplasmic 1 and actin, alpha cardiac muscle 1 were detected. In actin, cytoplasmic 1, methionine oxidation to MetSO has been noted at residues 44 and 47. It was found that the formation of MetSO at positions 44 and 47 promotes actin disassembly while the reduction of MetSO back to methionine promotes actin assembly. In the original study on these two residues, they were identified as residues 46 and 49. In the PDB, these residues were matched to residues 44 and 47 of human actin, cytoplasmic 1.[37] In this current study, MetSO formation was found at residues 153 and 305 of actin, cytoplasmic 1. Residues of actin, cytoplasmic 1 were not detected at positions 44 or 47. Another study found that many methionine residues can be oxidized. These residues were found to be located at positions 44, 47, 176, 190, 227, 269, and 355 in actin. It was also noted that Met82 can form a sulfone.[38] In the current study methionine residues at positions 192 and 229 were detected in the sulfoxide form in actin, alpha cardiac muscle 1. These residues when aligned with actin, cytoplasmic 1 align at positions 190 and 227, two of the residues previously identified as sites of methionine sulfoxide formation. Met190 was found to be one of two methionine residues that when oxidized completely prevents polymerization. It was noted that in F-actin, Met227 is less accessible than in G-actin.[38]

Alpha-1-antitrypsin has also previously been studied for the effect of methionine oxidation on protein structure. In our study, methionines at positions 250, 375, and 382 were detected however none of these methionines were determined to be of high interest because of their lower percent conservation. M250 was found to be 64% conserved and 20.1% oxidized.

M375 was found to be only 12% conserved and 36.5% oxidized. M382 was also determined to be 64% conserved and 38.8% oxidized. Interestingly, these same methionines when studied previously, although not chosen as targets of interest here, were found to have a significant effect on the structure of alpha-1-antitrypsin[39]. M250, M375, and M382 align to positions 226, 351, and 358 of the previous study. In that particular study, it was determined that oxidation of M351 or M358, both of which lie in the reactive center loop of alpha-1-antitrypsin, leads to a loss of inhibitory activity against neutrophil elastase. It was also discovered that oxidation of M358 can lead to a loss of inhibitory activity against pancreatic elastase. M226, which lies near and is hydrogen bonded to the reactive center loop, was also oxidized and it was noted that this oxidation could occur due to a conformational change induced by oxidation of M351 and M358, making M226 more solvent exposed. Similar to alpha-1-antitrypsin, we detected methionines of interest in alpha-1-antichymotrypsin. Of the four detected methionines, two were considered to be potentially interesting due to their higher percentage of conservation. M205 of alpha-1-antichymotrypsin was 75% conserved and 89.4% oxidized. M284 was determined to be 87.5% conserved and 51.4% oxidized. The sequence of alpha-1-antichymotrypsin was aligned with that of alpha-1-antitrypsin. It was determined that these methionines were not conserved in alpha-1-antitrypsin. The other two detected methionines, M204 and M219, were also not conserved between alpha-1-antichymotrypsin and alpha-1-antitrypsin. These two methionines were determined to have conservation percentages of 12.5% and 16.7%, respectively. Previous studies on these methionines or other methionines of alpha-1-antichymotrypsin have not yet been found.

While not regulatory, oxidation of methionines in apolipoprotein D was found to be involved in reduction of hydroperoxyeicosatetraenoic acid (HpETE) to hydroxyeicosatetraenoic acid (HETE), a step in the arachidonic acid pathway of the inflammatory response. In the

previous study, M93 was found to be the primary methionine involved in this reduction, but M157 was also determined to be involved in this reduction of HpETE to HETE.[40] These methionines align to positions 113 and 177, respectively, in our study. While M113 was not detected, M177 was detected and determined to be of potential interest because it was found to have a conservation percentage of 83.3%.

A previous study on complement C4 found that 52.6% of methionine residues were oxidized.[41] It was suggested that because complement C4 belongs to the alpha-2-macroglobulin family, this oxidation of methionine residues could lead to activation of the protein. In our study of complement C4-A, one methionine residue was detected at position 1128 and was determined to be potentially interesting due to a conservation percentage of 100%. Our study also detected several methionines in alpha-2-macroglobulin-like protein 1. Of seven detected methionines, only one was determined to be of interest. M1116 had a conservation percentage of 72%. While it is not clear whether this particular methionine may play an important role in alpha-2-macroglobulin, oxidation of four other residues has been shown to suppress the function of the protein. One residue in particular, M358 (not detected in our study), can lead to a complete loss in protein function when oxidized.[42]

A final example is that of fibronectin. A previous study on fibronectin found that of the 14 methionines that they detected, only 3 were determined to be targets of oxidation.[43] Those methionines were at positions 926, 1783, ad 2050. Of those 3 methionines, 2 were detected in our study. M926 and M1783, along with M1283 and M1548, were detected in our study. Of those, all except for M1548 were determined to be of interest. M1548 should not be ignored however. The percentage of conservation for M1548 was high at 91.7% and was found to be 100% oxidized. It was not previously included as being of interest because the surrounding

residues were more highly conserved than the methionine itself. M926, M1283, and M1783 had conservation percentages of 87.5%, 95.8%, and 95.8% while their surrounding residues were all less conserved than the methionine. It is not yet understood what role these particular methionines carry out in the function of fibronectin.

Conclusions

This study identified a list of potentially important methionine residues within proteins that may serve a critical role in determining a protein's structure or carrying out its function. Within this study it is clear that some methionine residues are much more conserved than others indicating that they may be of greater importance. Some of the identified residues of importance within the selected proteins of interest have already been identified as being highly involved in structural or functional roles. While several of the identified proteins of interest have already been studied for the effect of MetSO formation, this list could potentially provide target proteins to be further studied to determine the effect of MetSO formation. Those methionines that were not chosen as methionines of interest should not be ignored however. As revealed in the examination of alpha-1-antitrypsin, residues not considered to be of high interest here could still be vital for a protein's function. This list of residues and proteins of interest, while they should not be considered as the only residues of interest, could serve as a "map" of where to focus future studies on these proteins. Future studies on these proteins could provide a greater understanding into cell signaling processes, disease progression, and even possible treatments for diseases that advance due to the modification of methionine to MetSO.

Table 6. Methionines that were more conserved than the surroundings and were at least 70% conserved in all 3 graphs (Fig. 6-8)

Actin, alpha cardiac muscle 1
192
229
Actin, cytoplasmic 1
305
Alpha-1-antichymotrypsin
284
Alpha-2-macroglobulin-like protein 1
1116
Alpha-actinin-4
328
679
883
Alpha-amylase 1
117
338
Alpha-N-acetylglucosaminidase
238
Aminopeptidase N
486
Apolipoprotein D
177
BPI fold-containing family B member 2
254
Calpain-1 catalytic subunit
438

Table 6 (Cont.)

Cathepsin D
219
Ceruloplasmin
259
Complement C4-A
1128
Cubilin
2069
Cysteine-rich secretory protein 1
65
Desmoglein-1
208
Endothelial protein C receptor
30
Fibronectin
1283
1783
Glutaminyl-peptide cyclotransferase
65
223
Glypican-3
316
Kallikrein-1
122
197

Table 6 (Cont.)

Keratin, type I cytoskeletal 10
306
Keratin, type I cytoskeletal 14
272
Keratin, type I cytoskeletal 16
274
Keratin, type I cytoskeletal 9
157
326
Keratin, type I cytoskeletal 14
272
Keratin, type I cytoskeletal 16
274
Keratin, type I cytoskeletal 9
157
326
Keratin, type II cytoskeletal 1
296
Keratin, type II cytoskeletal 2 epidermal
337
Keratin, type II cytoskeletal 5
284
Keratin, type II cytoskeletal 6A
242

Table 6 (Cont.)

Low-density lipoprotein receptor-related protein 2
1905
2629
Lysosomal alpha-glucosidase
427
695
Plasma protease C1 inhibitor
324
Plasma serine protease inhibitor
275
Pro-cathepsin H
200
Profilin-1
86
Prosaposin
255
Prostaglandin-H2 D-isomerase
64
145
Prostatic acid phosphatase
75
Serpin B12
41

Table 6 (Cont.)

Serpin B3
40
243
275
304
317
Syntenin-1
57

Table 7. Methionines that were more conserved than the surroundings and were at least 70% conserved in only the graph comparing the N-terminus conservation (Fig. 6)

Beta-globin
56
Hemoglobin subunit beta
56
Low-density lipoprotein receptor-related protein 2
1299

Table 8. Methionines that were more conserved than the surroundings and were at least 70% conserved in the graph comparing the N-terminus conservation and the graph comparing the average of both the N-terminus and C-terminus (Fig. 6 and 8)

Alpha-1-antichymotrypsin
205
Aminopeptidase N
693
Fibronectin
926

Table 8 (Cont.)

Hemoglobin subunit delta
56
Keratin, type I cytoskeletal 26
87
Lysosomal alpha-glucosidase
440
Plasma serine protease inhibitor
78

Table 9. Methionines that were more conserved than the surroundings and were at least 70% conserved in only the graph comparing the C-terminus conservation (Fig. 7)

Alpha-actinin-4
767
Alpha-amylase 1
302
Amyloid beta A4 protein
303
Keratin, type I cytoskeletal 10
291
Keratin, type I cytoskeletal 9
234

Table 10. Methionines that were more conserved than the surroundings and were at least 70% conserved in the graph comparing the C-terminus conservation and the graph comparing the average of both the N-terminus and C-terminus (Fig. 7 and 8)

Aminopeptidase N
212
249
Cathepsin D
201
Ceruloplasmin
599
Desmoglein-1
425
eIF-2-alpha kinase activator GCN1
1377
Ezrin
200
G-protein coupled receptor family C group 5 member B
381
Keratin, type II cytoskeletal 1
259
469
Keratin, type II cytoskeletal 5
457
Leukocyte elastase inhibitor
228
307

Table 10 (Cont.)

Low-density lipoprotein receptor-related protein 2
1004
Lysosomal alpha-glucosidase
408
Maltase-glucoamylase, intestinal SV=2
1623
Programmed cell death 6-interacting protein
358
Prostaglandin-H2 D-isomerase
187
Serotransferrin
483

Table 11. Conservation chart for proteins of interest with at least 80% conservation and at least 50% oxidation. Comparisons are shown between ConSurf chosen homologues at various distances on the phylogenetic tree (shown by close, mid, and far ranges), ConSurf calculations using a Clustal alignment of sequences from the set list of species, and the calculated % conservation without using ConSurf.

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Actin, alpha cardiac muscle 1	192	**Score (normalized)**	0.444	0.83	0.28	-0.646	100%
		Color	4	2	4	8	
		Confidence Interval	0.085, 0.679	0.417, 1.145	-0.078, 0.438	-1.377, -0.328	
		Confidence Interval Colors	5, 3	4, 2	5, 4	9, 6	
	229	**Score (normalized)**	0.745	1.012	0.876	0.103	96.97%
		Color	3	2	3	5	
		Confidence Interval	0.34, 0.907	0.604, 1.145	0.438, 1.058	-0.883, 0.583	
		Confidence Interval Colors	4, 2	3, 2	4, 2	9,3	

Table 11 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
eIF-2-alpha kinase activator GCN1	1377	**Score (normalized)**	-0.27	-0.012	-0.828	0.222	84.85%
		Color	6	5	8	4	
		Confidence Interval	-0.5, -0.176	-0.286, 0.172	-0.963, -0.738	-0.313, 0.532	
		Confidence Interval Colors	7, 6	6, 4	8, 8	6, 2	
Keratin, type I cytoskeletal 14	272	**Score (normalized)**	-0.975	-1.001	-0.476	-0.833	100%
		Color	8	8	6	9	
		Confidence Interval	-1.107, -0.914	-1.129, -0.937	-0.747, -0.313	-0.993, -0.772	
		Confidence Interval Colors	8, 8	8, 8	7, 6	9, 9	
Syntenin-1	57	**Score (normalized)**	-1.038	-0.971	0.022	-0.914	93.33%
		Color	9	9	5	9	
		Confidence Interval	-1.099, -1.013	-1.026, -0.93	-0.974, 0.74	-1.135, -0.835	
		Confidence Interval Colors	9, 9	9, 9	7, 3	9, 9	

Table 11 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Alpha-actinin-4	328	**Score (normalized)**	0.125	0.515	-0.752	-0.649	96.30%
		Color	4	3	7	8	
		Confidence Interval	-0.091, 0.253	0.076, 0.781	-1.195, -0.43	-0.935, -0.578	
		Confidence Interval Colors	5, 4	5, 2	8, 6	9, 8	
	679	**Score (normalized)**	1.229	1.774	1.336	-0.649	96.30%
		Color	1	1	2	8	
		Confidence Interval	0.82, 1.614	0.781, 2.856	0.242, 2.06	-0.935, -0.578	
		Confidence Interval Colors	2, 1	2, 1	4, 1	9, 8	
	767	**Score (normalized)**	0.11	0.145	-0.11	0.77	88.89%
		Color	5	4	5	1	
		Confidence Interval	-0.181, 0.253	-0.236, 0.361	-0.784, 0.385	-0.202, 1.117	
		Confidence Interval Colors	6, 4	6, 4	7, 4	6, 1	

Table 11 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Alpha-actinin-4 (Cont.)	883	**Score (normalized)**	-0.988	-0.991	-1.548	-0.649	96.30%
		Color	9	9	8	8	
		Confidence Interval	-1.067, -0.952	-1.125, -0.94	-2.119, -1.279	-0.935, -0.578	
		Confidence Interval Colors	9, 9	9, 8	9, 8	9, 8	
Ezrin	200	**Score (normalized)**	-1.213	-1.179	-0.865	-0.740	100%
		Color	9	9	8	9	
		Confidence Interval	-1.238, -1.196	-1.252, -1.16	-1.018, -0.774	-0.861, -0.702	
		Confidence Interval Colors	9,9	9,9	8,7	9,9	
Complement C4-A	1128	**Score (normalized)**	-1.193	-1.432	-1.137	-1.232	100%
		Color	9	9	8	9	
		Confidence Interval	-1.293, -1.159	-1.843, -1.252	-1.255, -1.092	-1.391, -1.173	
		Confidence Interval Colors	9,9	9,9	9,8	9,9	

Table 11 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Keratin, type I cytoskeletal 16	274	**Score (normalized)**	-0.961	-0.871	-0.097	-0.879	100%
		Color	8	8	5	9	
		Confidence Interval	-1.082, -0.88	-1.016, -0.797	-0.401, 0.033	-1.024, -0.830	
		Confidence Interval Colors	8,8	8,7	6,5	9,9	
G-protein coupled receptor family C group 5 member B	381	**Score (normalized)**	-0.981	-1.108	-0.349	-1.158	96.67%
		Color	8	8	6	9	
		Confidence Interval	-1.148, -0.862	-1.378, -0.976	-1.225, 0.244	-1.400, -1.046	
		Confidence Interval Colors	9,8	9,8	8,4	9,9	
Plastin-3	169	**Score (normalized)**	-0.701	-0.57	-0.531	-0.521	96.67%
		Color	7	7	6	7	
		Confidence Interval	-0.856, -0.572	-0.781, -0.463	-0.947, -0.283	-0.931, -0.301	
		Confidence Interval Colors	8,7	7,6	7,6	9,6	

Table 11 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Alpha-N-acetylglucosa-minidase	238	**Score (normalized)**	-1.181	-1.123	-1.141	-0.852	96.55%
		Color	9	9	8	9	
		Confidence Interval	-1.247, -1.125	-1.211, -1.106	-1.254, -1.106	-1.042, -0.752	
		Confidence Interval Colors	9,9	9,9	9,8	9,8	
Tropomyosin alpha-3 chain	11	**Score (normalized)**	-0.732	0.222	0.3	-0.545	96.43%
		Color	8	4	4	9	
		Confidence Interval	-0.937, -0.593	-0.22, 0.574	-0.712, 1.1015	-0.689, -0.547	
		Confidence Interval Colors	9,7	6,3	7,3	9,9	
Serpin B12	41	**Score (normalized)**	-1.091	-0.966	-1.14	-1.008	96.15%
		Color	8	9	8	9	
		Confidence Interval	-1.254, 0.975	-1.074, -0.93	-1.353, -1.036	-1.225, -0.903	
		Confidence Interval Colors	9,8	9,8	8,7	9,8	

Table 11 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Actin-like protein (fragment)	70	**Score (normalized)**	-0.042	-0.54	-0.534	1.817	93.55%
		Color	5	7	6	1	
		Confidence Interval	-0.383, 0.117	-0.726, -0.427	-0.805, -0.356	-0.214, 3.139	
		Confidence Interval Colors	6,5	7,6	7,6	7,1	
Glypican-3	316	**Score (normalized)**	0.497	0.927	-0.238	-0.676	93.10%
		Color	3	2	6	8	
		Confidence Interval	0.148, 0.64	0.19, 1.330	-0.472, -0.102	-1.015, -0.504	
		Confidence Interval Colors	4,3	4,1	6,5	9,7	
Calpain-1 catalytic subunit	438	**Score (normalized)**	0.421	0.756	1.521	-0.839	93.10%
		Color	4	3	1	9	
		Confidence Interval	0.069, 0.553	0.373, 0.813	1.017, 1.906	-0.952, -0.803	
		Confidence Interval Colors	5,3	4,2	2,1	9,9	

Table 11 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Keratin, type I cytoskeletal 13	341	**Score (normalized)**	-0.341	-0.176	-0.803	-0.386	92.31%
		Color	6	6	7	7	
		Confidence Interval	-0.519, -0.253	-0.373, -0.052	-0.996, -0.667	-0.691, -0.198	
		Confidence Interval Colors	7,6	6,5	8,7	9,6	
Heat shock cognate 71 kDa protein	61	**Score (normalized)**	-0.277	1.407	-0.605	0.225	90.91%
		Color	6	1	6	4	
		Confidence Interval	-0.451, -0.193	0.904, 1.367	-0.862, -0.434	-0.538, 0.678	
		Confidence Interval Colors	7,6	2,1	7,6	8,2	
BPI fold-containing family B member 2	254	**Score (normalized)**	-1.136	ERROR	-1.67	-1.435	90%
		Color	9	ERROR	9	8	
		Confidence Interval	-1.322, -1.024	ERROR	-2.025, -1.537	-1.965, -1.227	
		Confidence Interval Colors	9,8	ERROR	9,8	9,8	

Table 11 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Histone H3.1	94	**Score (normalized)**	2.621	2.537	0.523	2.121	90%
		Color	1	1	4	1	
		Confidence Interval	1.544, 2.797	1.442, 2.69	0.083, 0.758	0.704,2.559	
		Confidence Interval Colors	1,1	1,1	5,3	1,1	
Keratin, type I cytoskeletal 10	291	**Score (normalized)**	-0.583	-0.17	-0.382	-0.673	78.26%
		Color	7	6	6	8	
		Confidence Interval	-0.793, -0.409	-0.442, -0.013	-0.629, -0.265	-1.009, -0.496	
		Confidence Interval Colors	7,6	6,5	7,6	9,7	
	306	**Score (normalized)**	-1.154	-1.02	-1.124	-1.013	100%
		Color	8	8	9	9	
		Confidence Interval	-1.267, -1.078	-1.138, -0.931	-1.245, -1.087	-1.192, -0.933	
		Confidence Interval Colors	9,8	8,8	9,8	9,9	

Table 11 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Keratin, type II cytoskeletal 6A	242	**Score (normalized)**	-0.286	-0.158	-0.507	-0.442	88.46%
		Color	6	6	6	7	
		Confidence Interval	-0.473, -0.2	-0.378, -0.057	-0.708, -0.37	-0.726, -0.260	
		Confidence Interval Colors	7,6	6,5	7,6	9,6	
Titin	16670	**Score (normalized)**	ERROR	ERROR	ERROR	ERROR	88%
		Color	ERROR	ERROR	ERROR	ERROR	
		Confidence Interval	ERROR	ERROR	ERROR	ERROR	
		Confidence Interval Colors	ERROR	ERROR	ERROR	ERROR	
Profilin-1	86	**Score (normalized)**	-1.027	-0.773	-0.901	-0.764	86.67%
		Color	8	7	7	7	
		Confidence Interval	-1.211, -0.909	-1.11, -0.536	-1.245, -0.697	-1.242, -0.461	
		Confidence Interval Colors	9,8	8,6	8,7	9,6	

Table 11 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Amyloid beta A4 protein	303	**Score (normalized)**	-0.123	2.215	ERROR	0.728	85.71%
		Color	6	1	ERROR	2	
		Confidence Interval	-0.373, 0.001	0.984, 3.863	ERROR	-0.175, 1.264	
		Confidence Interval Colors	7,5	1,1	ERROR	6,1	
Probable maltase-glucoamylase 2	623	**Score (normalized)**	-1.047	-1.198	ERROR	-1.006	84.38%
		Color	9	9	ERROR	8	
		Confidence Interval	-1.11, -1.013	-1.325, -1.16	ERROR	-1.219, -0.877	
		Confidence Interval Colors	9,9	9,9	ERROR	9,8	
Keratin, type I cytoskeletal 12	251	**Score (normalized)**	0.068	0.144	-0.178	0.116	80.77%
		Color	5	5	5	5	
		Confidence Interval	-0.159, 0.251	-0.167, 0.253	-0.426, -0.001	-0.450, 0.397	
		Confidence Interval Colors	6,4	6,4	6,5	7,3	

Table 11 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Corticosteroid-binding globulin	76	**Score (normalized)**	-0.572	-0.517	-1.077	-0.697	80.77%
		Color	7	7	8	7	
		Confidence Interval	-0.774, -0.422	-0.67, -0.389	-1.231, -0.975	-1.040, -0.442	
		Confidence Interval Colors	7,6	7,6	8,7	8,6	
Alpha-amylase 1	117	**Score (normalized)**	-1.161	-1.067	-1.188	-0.820	90.63%
		Color	9	9	8	9	
		Confidence Interval	-1.232, -1.128	-1.1, -1.057	-1.338, -1.112	-0.981, -0.716	
		Confidence Interval Colors	9,9	9,9	9,8	9,8	
	198	**Score (normalized)**	-0.167	-0.063	-0.901	0.419	59.375%
		Color	6	5	8	3	
		Confidence Interval	-0.388, 0.011	-0.268, 0.057	-1.112, -0.753	-0.084, 0.855	
		Confidence Interval Colors	6,5	6,5	8,7	5,1	

Table 11 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Alpha-amylase 1 (Cont.)	302	**Score (normalized)**	0.374	0.279	0.296	-0.264	84.38%
		Color	4	4	4	6	
		Confidence Interval	0.011, 0.487	0.057, 0.378	-0.106, 0.429	-0.575, -0.084	
		Confidence Interval Colors	5,3	5,4	5,4	8,5	
	338	**Score (normalized)**	-0.865	-0.587	-0.466	-0.696	93.75%
		Color	8	7	6	8	
		Confidence Interval	-0.969, -0.778	-0.717, -0.558	-0.687, -0.3	-0.898, -0.575	
		Confidence Interval Colors	8,8	8,7	7,6	9,8	

Table 12. Conservation chart for proteins of interest with at least 80% conservation and less than 50% oxidation showing comparisons between ConSurf chosen homologs at various distances on the phylogenetic tree (shown by close, mid, and far ranges), ConSurf calculations using a Clustal alignment of sequences from the set list of species, and the calculated % total conservation without using ConSurf

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Kallikrein-1	122	**Score (normalized)**	ERROR	-1.101	-1.127	-1.172	100%
		Color	ERROR	9	8	9	
		Confidence Interval	ERROR	-1.171, -1.049	-1.203, -1.058	-1.312, -1.129	
		Confidence Interval Colors	ERROR	9,9	9,8	9,9	
	197	**Score (normalized)**	ERROR	-1.201	-0.834	-1.172	100%
		Color	ERROR	9	7	9	
		Confidence Interval	ERROR	-1.290, -1.171	-0.953, -0.777	-1.312, -1.129	
		Confidence Interval Colors	ERROR	9,9	8,7	9,9	

Table 12 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Lysozyme C	35	**Score (normalized)**	-0.190	-0.279	-0.230	ERROR	100%
		Color	6	6	6	ERROR	
		Confidence Interval	-0.443, -0.085	-0.543, -0.103	-0.813, 0.059	ERROR	
		Confidence Interval Colors	6,5	7,5	7,5	ERROR	
Keratin type I cytoskeletal 26	87	**Score (normalized)**	-1.208	-1.313	-1.096	-0.905	95.833%
		Color	9	9	8	8	
		Confidence Interval	-1.284, -1.187	-1.370, -1.308	-1.246, -1.006	-1.272, -0.729	
		Confidence Interval Colors	9,9	9,9	8,8	9,8	
Actin, cytoplasmic 1	153	**Score (normalized)**	0.465	0.004	-0.839	1.579	84.848%
		Color	4	5	7	1	
		Confidence Interval	0.054, 0.624	-0.291, 0.137	-1.020, -0.714	0.062, 2.365	
		Confidence Interval Colors	5,3	6,5	7,7	5,1	

Table 12 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Actin, cytoplasmic 1 (Cont.)	305	**Score (normalized)**	-0.831	-0.660	-0.827	-1.136	100%
		Color	8	7	7	9	
		Confidence Interval	-0.998, -0.717	-0.794, -0.602	-1.095, -0.714	-1.532, -0.968	
		Confidence Interval Colors	8,7	7,7	8,7	9,9	
Endothelial protein C receptor	30	**Score (normalized)**	-0.840	-0.935	0.327	-0.751	91.667%
		Color	8	8	4	8	
		Confidence Interval	-0.960, -0.761	-1.305, -0.688	-0.300, 0.668	-1.041, -0.579	
		Confidence Interval Colors	9,8	9,7	6,3	9,8	
Programmed cell death 6-interacting protein	358	**Score (normalized)**	-0.974	-1.048	-1.324	-0.713	90.909%
		Color	9	9	8	7	
		Confidence Interval	-1.094, -0.895	-1.188, -0.984	-1.471, -1.236	-1.109, -0.477	
		Confidence Interval Colors	9,8	9,8	9,8	9,7	

Table 12 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Cadherin-13	267	**Score (normalized)**	-0.989	-1.035	-0.852	-0.184	89.655%
		Color	8	8	7	6	
		Confidence Interval	-1.124, -0.930	-1.193, -0.910	-1.042, -0.749	-0.613, 0.072	
		Confidence Interval Colors	8,8	9,8	8,7	8,5	
Keratin, type II cytoskeletal 2 epidermal	294	**Score (normalized)**	-0.299	-0.423	0.035	0.379	69.565%
		Color	6	6	5	4	
		Confidence Interval	-0.542, -0.182	-0.587, -0.328	-0.306, 0.227	-0.220, 0.803	
		Confidence Interval Colors	7,6	7,6	6,4	6,2	
	337	**Score (normalized)**	-1.170	-1.184	-0.551	-1.110	100%
		Color	9	9	7	9	
		Confidence Interval	-1.246, -1.113	-1.268, -1.126	-0.730, -0.406	-1.316, -1.042	
		Confidence Interval Colors	9,9	9,9	7,6	9,9	

Table 12 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Keratin, type II cytoskeletal 2 epidermal (Cont.)	467	**Score (normalized)**	-0.896	-0.814	-1.329	-0.641	91.304%
		Color	8	8	9	8	
		Confidence Interval	-1.035, -0.797	-0.950, -0.725	-1.457, -1.282	-0.997, -0.394	
		Confidence Interval Colors	8,8	8,7	9,9	9,7	
Glutaminyl-peptide cyclotransferase	65	**Score (normalized)**	-0.341	0.290	2.572	-0.430	87.5%
		Color	6	4	1	7	
		Confidence Interval	-0.567, 0.202	-0.080, 0.485	1.255, 3.117	-0.730, -0.202	
		Confidence Interval Colors	7,6	5,3	2,1	8,6	
	223	**Score (normalized)**	-0.706	-0.615	0.244	-0.819	93.75%
		Color	7	7	4	8	
		Confidence Interval	-0.872, -0.637	-0.770, -0.500	-0.124, 0.451	-1.016, -0.679	
		Confidence Interval Colors	8,7	8,7	5,4	9,8	

Table 12 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Glutaminyl-peptide cyclotransferase	332	**Score (normalized)**	0.794	1.149	1.196	-0.089	75%
		Color	2	1	2	5	
		Confidence Interval	0.422, 1.013	0.485, 1.294	0.657, 1.255	-0.505, 0.118	
		Confidence Interval Colors	4,2	3,1	3,2	7,5	
Pro-cathepsin H	200	**Score (normalized)**	-0.122	-0.076	0.010	-0.675	84.848%
		Color	5	5	5	8	
		Confidence Interval	-0.328, -0.022	-0.242, 0.004	-0.217, 0.149	-0.894, -0.538	
		Confidence Interval Colors	6,5	6,5	6,5	9,7	
Leukocyte elastase inhibitor	156	**Score (normalized)**	1.652	1.080	0.690	1.131	55.172%
		Color	1	1	3	1	
		Confidence Interval	1.215, 1.742	0.614, 1.216	0.226, 0.847	0.530, 1.336	
		Confidence Interval Colors	1,1	3,1	4,3	3,1	

Table 12 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Leukocyte elastase inhibitor (Cont.)	228	**Score (normalized)**	-1.336	-1.234	-1.456	-0.749	93.103%
		Color	9	9	9	8	
		Confidence Interval	-1.431, -1.302	-1.305, -1.196	-1.554, -1.428	-0.987, -0.586	
		Confidence Interval Colors	9,9	9,9	9,9	9,7	
	307	**Score (normalized)**	-0.706	-0.808	-0.865	-0.844	93.103%
		Color	7	8	7	8	
		Confidence Interval	-0.856, -0.595	-0.934, -0.773	-0.997, -0.791	-1.067, -0.702	
		Confidence Interval Colors	8,7	8,8	7,7	9,8	
Apolipoprotein D	177	**Score (normalized)**	-0.942	-0.766	-0.764	-0.462	83.333%
		Color	8	7	7	7	
		Confidence Interval	-1.149, -0.848	-0.968, -0.650	-1.041, -0.572	-0.810, -0.222	
		Confidence Interval Colors	8,7	8,7	8,6	8,6	

Table 12 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Vesicular integral-membrane protein VIP36	271	**Score (normalized)**	-0.069	0.083	-0.271	0.261	83.871%
		Color	5	5	7	4	
		Confidence Interval	-0.288, 0.073	-0.230, 0.273	-0.919, -0.596	-0.270, 0.563	
		Confidence Interval Colors	6,5	6,4	8,7	6,2	
	277	**Score (normalized)**	-0.054	0.127	0.305	-0.252	90.323%
		Color	5	5	4	6	
		Confidence Interval	-0.383, 0.073	-0.230, 0.273	-0.118, 0.586	-0.635, -0.076	
		Confidence Interval Colors	6,5	6,4	5,3	8,5	
Fibronectin	926	**Score (normalized)**	-0.125	-0.602	ERROR	0.330	87.5%
		Color	6	7	ERROR	3	
		Confidence Interval	-0.415, 0.021	-1.366, -0.145	ERROR	-0.311, 0.829	
		Confidence Interval Colors	7,5	9,6	ERROR	7,1	

Table 12 (Cont.)

Protein	Methionine Position		Close	Mid	Far	Clustal	% Conservation
Fibronectin (Cont.)	1283	**Score (normalized)**	0.610	-0.144	ERROR	-0.730	95.833%
		Color	3	6	ERROR	9	
		Confidence Interval	0.136, 0.805	-0.949, 0.424	ERROR	-0.946, -0.668	
		Confidence Interval Colors	4,2	8,4	ERROR	9,8	
	1548	**Score (normalized)**	-0.439	-0.704	ERROR	-0.088	91.667%
		Color	7	7	ERROR	5	
		Confidence Interval	-0.715, -0.263	-1.366, -0.264	ERROR	-0.572, 0.233	
		Confidence Interval Colors	8,6	9,6	ERROR	8,4	
	1783	**Score (normalized)**	-0.929	-1.250	ERROR	-0.730	95.833%
		Color	9	9	ERROR	9	
		Confidence Interval	-1.077, -0.855	-1.818, -0.949	ERROR	-0.946, -0.668	
		Confidence Interval Colors	9,8	9,8	ERROR	9,8	

References

1. Kim, G.; Weiss, S. J.; Levine, R. L., Methionine oxidation and reduction in proteins. *Biochim Biophys Acta* **2014,** *1840*, 901-5.

2. Hoshi, T.; Heinemann, S., Regulation of cell function by methionine oxidation and reduction. *J Physiol* **2001,** *531*, 1-11.

3. Lim, J. M.; Kim, G.; Levine, R. L., Methionine in proteins: it's not Just for protein initiation anymore. *Neurochem Res* **2019,** *44*, 247-57.

4. Veredas, F. J.; Cantón, F. R.; Aledo, J. C., Methionine residues around phosphorylation sites are preferentially oxidized in vivo under stress conditions. *Sci Rep* **2017,** *7*, 40403.

5. Bedard, K.; Krause, K. H., The NOX family of ROS-generating NADPH oxidases: physiology and pathophysiology. *Physiol Rev* **2007,** *87*, 245-313.

6. Brieger, K.; Schiavone, S.; Miller, F. J.; Krause, K. H., Reactive oxygen species: from health to disease. *Swiss Med Wkly* **2012,** *142*, w13659.

7. Forrester, S. J.; Kikuchi, D. S.; Hernandes, M. S.; Xu, Q.; Griendling, K. K., Reactive oxygen species in metabolic and inflammatory signaling. *Circ Res* **2018,** *122*, 877-902.

8. Thannickal, V. J.; Fanburg, B. L., Reactive oxygen species in cell signaling. *Am J Physiol Lung Cell Mol Physiol* **2000,** *279*, L1005-28.

9. Aledo, J. C., Methionine in proteins: The Cinderella of the proteinogenic amino acids. *Protein Sci* **2019,** *28*, 1785-96.

10. Stites, W. E.; Froude, J. W., Does the oxidation of methionine in thrombomodulin contribute to the hypercoaguable state of smokers and diabetics? *Med Hypotheses* **2007,** *68*, 811-21.

11. Saunders, C. C.; Stites, W. E., An electrophoretic mobility shift assay for methionine sulfoxide in proteins. *Anal Biochem* **2012,** *421*, 767-9.

12. Maupin-Furlow, J. A., Methionine sulfoxide reductases of archaea. *Antioxidants* **2018**, 7, 124.

13. dos Santos, S.L.; Petropoulos, I.; Friguet, B., The oxidized protein repair enzymes methionine sulfoxide reductases and their roles in protecting against oxidative stress, in ageing and in regulating protein function. *Antioxidants* **2018**, 7, 191.

14. Bin, P.; Huang, R.; Zhou, X., Oxidation resistance of the sulfur amino acids: Methionine and Cysteine. *Biomed Res Int* **2017,** *2017*, 9584932.

15. Weissbach, H.; Resnick, L.; Brot, N., Methionine sulfoxide reductases: history and cellular role in protecting against oxidative damage. *Biochim Biophys Acta* **2005,** *1703*, 203-12.

16. Fricke, T. C.; Echtermeyer, F.; Zielke, J.; de la Roche, J.; Filipovic, M. R.; Claverol, S.; Herzog, C.; Tominaga, M.; Pumroy, R. A.; Moiseenkova-Bell, V. Y.; Zygmunt, P. M.; Leffler, A.; Eberhardt, M. J., Oxidation of methionine residues activates the high-threshold heat-sensitive ion channel TRPV2. *Proc Natl Acad Sci U S A* **2019,** *116*, 24359-65.

17. Bigelow, D. J.; Squier, T. C., Redox modulation of cellular signaling and metabolism through reversible oxidation of methionine sensors in calcium regulatory proteins. *Biochim Biophys Acta* **2005,** *1703*, 121-34.

18. Bigelow, D. J.; Squier, T. C., Thioredoxin-dependent redox regulation of cellular signaling and stress response through reversible oxidation of methionines. *Mol Biosyst* **2011,** *7*, 2101-9.

19. Gallmetzer, A.; Silvestrini, L.; Schinko, T.; Gesslbauer, B.; Hortschansky, P.; Dattenböck, C.; Muro-Pastor, M. I.; Kungl, A.; Brakhage, A. A.; Scazzocchio, C.; Strauss, J., Reversible oxidation of a conserved methionine in the nuclear export sequence determines subcellular distribution and activity of the fungal nitrate regulator NirA. *PLoS Genet* **2015,** *11*, e1005297.

20. Gu, S. X.; Stevens, J. W.; Lentz, S. R., Regulation of thrombosis and vascular function by protein methionine oxidation. *Blood* **2015,** *125*, 3851-9.

21. Glaser, C. B.; Yamin, G.; Uversky, V. N.; Fink, A. L., Methionine oxidation, alpha-synuclein and Parkinson's disease. *Biochim Biophys Acta* **2005,** *1703*, 157-69.

22. Lockhart, C.; Smith, A. K.; Klimov, D. K., Methionine oxidation changes the mechanism of Aβ peptide binding to the DMPC bilayer. *Sci Rep* **2019,** *9*, 5947.

23. Jiang, B.; Moskovitz, J., The functions of the mammalian methionine sulfoxide reductase system and related diseases. *Antioxidants (Basel)* **2018,** *7*.

24. Schöneich, C., Methionine oxidation by reactive oxygen species: reaction mechanisms and relevance to Alzheimer's disease. *Biochim Biophys Acta* **2005,** *1703*, 111-9.

25. Requena, J. R.; Dimitrova, M. N.; Legname, G.; Teijeira, S.; Prusiner, S. B.; Levine, R. L., Oxidation of methionine residues in the prion protein by hydrogen peroxide. *Arch Biochem Biophys* **2004,** *432* (2), 188-95.

26. Ashkenazy, H.; Erez, E.; Martz, E.; Pupko, T.; Ben-Tal, N., ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* **2010,** *38*, W529-33.

27. Capra, J. A.; Singh, M., Predicting functionally important residues from sequence conservation. *Bioinformatics* **2007,** *23*, 1875-82.

28. Valdar, W. S., Scoring residue conservation. *Proteins* **2002,** *48*, 227-41.

29. Qassab, A., The First In Vivo Human Methionine Sulfide Proteome and the Impact of Smoking, 2018.

30. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic local alignment search tool. *J Mol Biol* **1990,** *215*, 403-10.

31. Sievers, F.; Higgins, D. G., Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci* **2018,** *27*, 135-145.

32. Metzler, W. J.; Bell, A. J.; Ernst, E.; Lavoie, T. B.; Mueller, L., Identification of the poly-L-proline-binding site on human profilin. *J Biol Chem* **1994,** *269*, 4620-5.

33. Shao, J.; Welch, W. J.; Diprospero, N. A.; Diamond, M. I., Phosphorylation of profilin by ROCK1 regulates polyglutamine aggregation. *Mol Cell Biol* **2008,** *28*, 5196-208.

34. Boopathy, S.; Silvas, T. V.; Tischbein, M.; Jansen, S.; Shandilya, S. M.; Zitzewitz, J. A.; Landers, J. E.; Goode, B. L.; Schiffer, C. A.; Bosco, D. A., Structural basis for mutation-induced destabilization of profilin 1 in ALS. *Proc Natl Acad Sci U S A* **2015,** *112*, 7984-9.

35. Nakamura, M.; Shishido, N.; Nunomura, A.; Smith, M. A.; Perry, G.; Akutsu, H.; Hayashi, T., Specific reaction of Met 35 in amyloid beta peptide with hypochlorous acid. *Free Radic Res* **2010,** *44*, 734-41.

36. Perrin, B. J.; Ervasti, J. M., The actin gene family: function follows isoform. *Cytoskeleton (Hoboken)* **2010,** *67*, 630-4.

37. Lee, B. C.; Péterfi, Z.; Hoffmann, F. W.; Moore, R. E.; Kaya, A.; Avanesov, A.; Tarrago, L.; Zhou, Y.; Weerapana, E.; Fomenko, D. E.; Hoffmann, P. R.; Gladyshev, V. N., MsrB1 and MICALs regulate actin assembly and macrophage function via reversible stereoselective methionine oxidation. *Mol Cell* **2013,** *51*, 397-404.

38. Dalle-Donne, I.; Rossi, R.; Giustarini, D.; Gagliano, N.; Di Simplicio, P.; Colombo, R.; Milzani, A., Methionine oxidation as a major cause of the functional impairment of oxidized actin. *Free Radic Biol Med* **2002,** *32*, 927-37.

39. Griffiths, S. W.; Cooney, C. L., Relationship between protein structure and methionine oxidation in recombinant human alpha 1-antitrypsin. *Biochemistry* **2002,** *41*, 6245-52.

40. Bhatia, S.; Knoch, B.; Wong, J.; Kim, W. S.; Else, P. L.; Oakley, A. J.; Garner, B., Selective reduction of hydroperoxyeicosatetraenoic acids to their hydroxy derivatives by apolipoprotein D: implications for lipid antioxidant activity and Alzheimer's disease. *Biochem J* **2012,** *442*, 713-21.

41. Blanchin, S.; Estienne, V.; Durand-Gorde, J. M.; Carayon, P.; Ruf, J., Complement activation by direct C4 binding to thyroperoxidase in Hashimoto's thyroiditis. *Endocrinology* **2003,** *144*, 5422-9.

42. Reddy, V. Y.; Desorchers, P. E.; Pizzo, S. V.; Gonias, S. L.; Sahakian, J. A.; Levine, R. L.; Weiss, S. J., Oxidative dissociation of human alpha 2-macroglobulin tetramers into dysfunctional dimers. *J Biol Chem* **1994,** *269*, 4683-91.

43. Nybo, T.; Cai, H.; Chuang, C. Y.; Gamon, L. F.; Rogowska-Wrzesinska, A.; Davies, M. J., Chlorination and oxidation of human plasma fibronectin by myeloperoxidase-derived oxidants, and its consequences for smooth muscle cell function. *Redox Biol* **2018,** *19*, 388-400.