VFAST Transactions on Software Engineering
http://vfast.org/journals/index.php/VTSE@ 2021, ISSN(e): 2309-3978, ISSN(p): 2411-6246
Volume 9, Number 2, April-June, 2021 *pp: 20-29*

Protein Carbonylation Sites Prediction using Biomarkers of Oxidative Stress in Various Human Diseases: A Systematic Literature Review

Anam Shahzadi [1]*, Muhammad Adeel Ashraf[1], Muhammad Sohaib Akram [2]

[1]Department of Computer Sciences, *University of Management and Technology, Lahore,Pakistan*
[2]Govt Graduate College Mailsi, Pakistan
Corresponding Author: adeel.ashraf@umt.edu.pk

ABSTRACT
Protein carbonylation is a non-enzymatic, irreversible, post translational modification (PTM). Carbonylation basically occurs due to the ROS, these species cause the oxidation of proteins and it will lead towards post translational modification of proteins known as carbonylation. In this ROS induces the carbonyl groups into the side chain of amino acid lysine (K), Proline (P), Arginine (R), Threonine (T). Carbonylation is known as a major hall mark or oxidative stress and leads to various diseases like age and age-related diseases. Different techniques and tools have been presented for detection of protein carbonylation, yet still there is no accurate result. In this systematic literature review I try to provide deep understanding of protein carbonylation sites, various techniques, comparison of tools relative to the techniques and its role in different diseases.

KEYWORDS
Protein carbonylation sites; ROS; PTM; Oxidation; Oxidative stress.

JOURNAL INFO
HISTORY: Received: April 06, 2021 Accepted: September 15, 2021 Published: September 30, 2021

1. INTRODUCTION

The creation of Oxygen reactive species (ROS) is an unstoppable consequence of an aerobic cellular respiration. Oxygen reactive species are extensively active like (OH radical). Almost all the types of molecules could be oxidatively modifies or damaged by these ROS, like proteins, sugars, nucleic acid and lipids.

When we see the oxidation of proteins, it could be intensively harmful because major functions in living cells are carried out with the help of proteins [1]. Non-enzymatic, free radical-mediated oxidation of protein is normal in biological processes.

The oxidative damage caused by the specified radicals on lysozyme, apolipoprotein B in LDL, and ∝-1-antitrypsin. There was related loss of functionality and zero impact in each of these instances.

However, the consequences of oxidation are majorly negative. More often oxidation is dependent on (i) oxidizing organisms (ii) primary structure (iii) either oxidants have access to targeted sequences of amino acid inside that protein [2].

ROS could be major facilitator and intermediator of cancer, heart diseases, endothelial abnormality, atherosclerosis, and intestinal tract disorders, diabetes, brain degenerative disease, and inflammation, as well as ischemic and post ischemic pathologies Usually,

Protein oxidation take place as a consequence of exposure to ROS, but it is managed with the aid of different antioxidants.

When this balance is affected and the level of pro-oxidant increases the level of anti-oxidants then oxidative stress take place.

Different kinds of post translational modifications (PTM) are because of oxidative stress such as nitration, hydroxylation, carbonylation, sulfhydrylation, & glutathionylation.

Between them the carbonylation is commonly known as irreversible protein damage. The Binding of a chemical group with a particular protein residue is known as post translational modifications (PTM), that performs an important role in initiation of various cellular reactions like cell proliferation, protein denaturation, monitoring of gene expression, and interactions among proteins.

Binding and elimination of chemical groups of proteins are facilitated by enzymes [4]. PTM sites of proteins determine the state of activity of protein, distribution, turnover, and it's reaction with other proteins. For instance, when kinase cascades are activated and deactivated by reversible including and excluding the phosphate group in the signaling process One of the most commonly studied forms of oxidation is the protein carbonylation, which is an introduction of an aldehyde and ketone groups into the side chain of amino acid and known as a biomarker or hallmark of oxidative stress.

Carbonyls commonly made via interaction with ROS like hydrogen peroxide (H_2O_2), hydroxyl radical (OH), and superoxide anion (O_2), in order to cause oxidation and also made lipid peroxidation final products through reacting with lipid peroxidation products like 4-hydroxy-trans-2-nonenal (HNE), malondialdehyde (MDA), acrolein (ACR) [6].

Protein carbonylation which produces a reactive-carbonyls in protein, like ketone, aldehyde and lactam is commonly known as PTM irreversible damage [1].

Protein carbonylation role in tissue, organ age, and cell has gain special attention. And it is also a causative agent of various diseases like Alzheimer's disorder, lung disorder, fetal renal failure, diabetes etc.

Different processes in the protein carbonylation altered the side groups of k, R, T, and pro residues also metal-catalyzed oxidation (MCO) [4]. Various experiments have been conducted for the identification of protein carbonylation sites such as Mass spectrometry (MS) based proteomics (useful for site specific detection of carbonylation sites), spectrophotometric, slot blotting, and enzyme-linked immunosorbent.

Technique which is used for detection carbonylated sites and measure carbonylation level is Mass spectrometry (MS) it is the one of the most popular technique.

Mainly, four kinds of amino acid residues are found more susceptible to carbonylation: lysine (k), proline (p), arginine (R), and threonine (T). Additionally, in RKPT areas, large amount of carbonylated sites are identifies and they have strong propensity regarding clusters.

But, to utilize the traditional experimental methods these would require considerably longer time and work to identify the protein carbonylation sites [7]. In this situation several tools are already made for protein carbonylation sites detection.

The aim of this study is to analyze the major concepts of protein carbonylation sites and conduct a detail systematic literature review on previously used methods and computational tools for carbonylation sites and comparison of already existing computational tools.

We also see protein carbonylation sites role in ageing and age-related diseases.

Section II describes the related work of this study, Section III describes the methodology, Section IV describes about how different techniques and approaches (Mass spectrometry and proteomics) are used for detecting protein carbonylation sites, Section V describes about what are different tools and models for detecting carbonylation sites and their comparison, describes about how protein carbonylation sites play role in different human diseases, and in Section VII describes the conclusion.

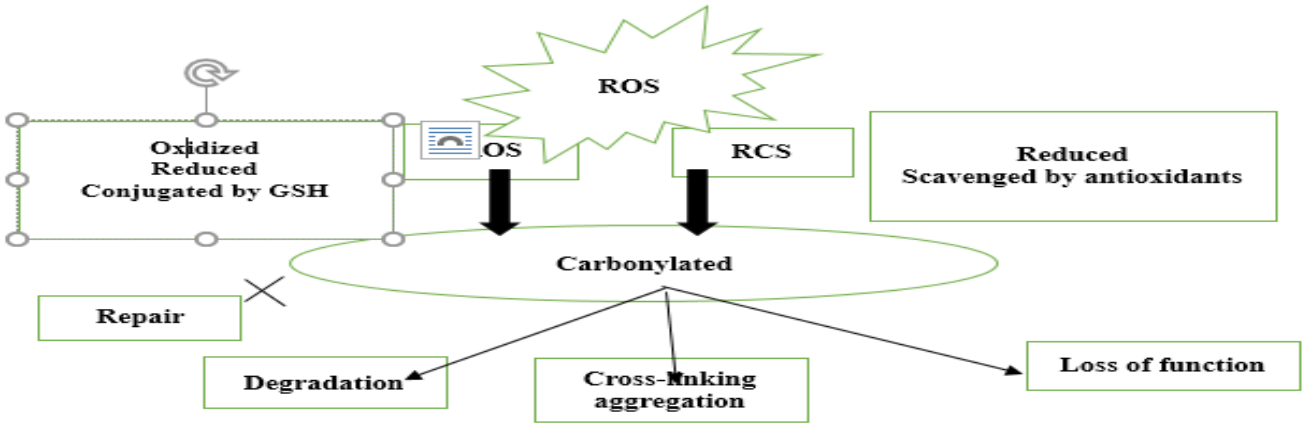

Figure 1: Carbonylation and its effects

2. RELATED WORK

A large number of experiments have been carried out for the identification of protein carbonylation sites in previous years. There are several other methods which were utilized for carbonylation identification and some of them stated that R, K, P, T site of amino acids have major susceptibility towards carbonylation. However, many of these approaches are time-consuming and costly that's why various computational tools and models invented, those who have high capability for identification of protein carbonylation sites. In this related work some previous work has been represented, which basically include invention of tools and techniques.

In this paper [8], a common approaches for the detection of protein carbonylation sites and major potential and pitfalls has been presented. The major objective was to provide guidance for the best option about which approach must be choosed and utilized and its advantages and limitations. This article [9], presents the knowledge gap which was presented in trastuzumab through detection and concisely analysing the metal catalysed corbonylation sites. Carbonylation sites in originators and 4 biosimilars of an immunoglobulin G (IgG1) clinical monoclonal antigen, trastuzumab, a very first line or effective therapy for HER2 positive breast cancer have been investigated in this.

The framework and the static areas of the antigen have MCO carbonylation sites enrichment, although no carbonylated amino acid have been found for any of the products in the CH_3 and CDR areas shown in this study.

In this article [10] to detect the protein carbonyl sites in human proteins a computational tool CarSPred has been

presented. In this author [11] presents a new bioinformatic tool called iCar-PseCp indicator for the protein carbonylation sites. For the classification of carbonylated and non-carbonylated potien sites, this article [12] analyze the informative features, multiple characterisics including twenty amino acids composition (AAC), amino acid pairs composition (AAPC), positional weighted matrics (PWM), and position specific scoring matrics (PSSM) were studied. MDD-Carb, an online tool to promote the analysis of carbonylated large-scale proteomics. In this a new bioinformatics tool, CarSite was suggested in order to find out the human proteins carbonylation sites [13]. This article predicts a new tool PredCar-site for carbonylation identification [14]. In this article [15], they invent a computational tool iCarPS to detect the protein carbonylation sites dependent on the data sequence.
In this [16] new tools was predicted CarSPred 2.0 based on the inovative technique called max-significance and min-redundancy. In this [17] they create a database manually for identification of protein carbonylation sites which contain 1495 carbonylated sites nad 3781 carbonylated sites of rats, yeasts, and humans.

The basic goal of this systematic literature review is to provide the detail analysis about the protein carbonylation sites, their origin, previously used approaches and techniques for carbonylation sites detection, how these techniques replaced with the help of new computational tools, and protein carbonylation role in different diseases like in aging and diabetes etc.
We have done following steps in research methodology.

3.1 RESEARCH OBJECTIVE

This SLR has following objective.
1-How different approaches are used to find out carbonylation sites. And how they are related with one another.
2-How different tools and models are used for protein carbonylation sites and how they are better as compared to previous techniques and approaches.

All the papers which I download was not related to my area of interest. That's why I perform inclusion and exclusioncriteria so in this way I exclude all data which was not related to my interest. In first phase there were 150 articles and after removal of duplication just 100 articles left. After that 40 articles were rejected when I try to select articles based on title, abstract, and keywords. And after further elaboration and analysis 20 articles were excluded because they didn't have information related to my objectives. And in last I perform quantitative evaluation and after this, 27 articles were left which I choose for my systematic literature review.

3.6 RESULT

In this SLR a complete research process has been descried. And after this research objectives also describes. I include articles which were related to my topic.
3-How protein carbonylation sites are hallmark of oxidative stress and this oxidative stress leads towards ageing and various other diseases.

3.2 Research process

Figure 1: Research process

3.4 SEARCH STRING

I conduct this SLR using following search terms and gather data which was related to my topic and objectives. I try to gather data from popular electronic sources like Science Direct, open biology, IEEE, Journal of proteome research, EMBO journal, Cross mark, Original research, Analytical biochemistry, Journal of molecular science, MDPI, Oxford, Cross mark, Informa healthcare, Analytical chemistry, Biomolecules. Mainly I try to use popular electronic resources

3.5 CRITERIA FOR INCLUSION AND EXCLUSION STUDY

And after complete process I choose best articles and conduct systematic literature review. And in search strategy I used popular electronic sources and I try to include those articles which were of good quality and give major information of my topic and conduct research in between 2010 to 2020 years. All steps are described in detail in this section.

There are different approaches and techniques used for identification of protein carbonylation sites. Among them most common are mass spectrometry and proteomics.

4.2 DETECTION OF PROTEIN CARBONYLATION AS A GROUP

Second technique is to targeting the native protein. Carbonylated proteins will be biotinylated first and after that choosed by avidin affinity chromatography in this approach. Mass spectrometry aids in detection of

Table 1: Three questions related to the existing approaches

.Research questions	Motivation
Q#1: How different techniques and approaches (Mass spectrometry and proteomics) are used for detecting protein carbonylation sites?	To find out at which extent these approaches are successful. And how they are related with one another.
Q#2 What are the different tools and models for detecting carbonylation sites and how their comparison?	To find out how different tools and models are better as compared to previous techniques and approaches.
Q#3 How protein carbonylation sites play role in different human diseases?	To find out how carbonylation sites are hallmark of oxidative stress and this oxidative stress leads towards ageing and various diseases.

unoxidized peptides portions of these proteins, by using proteolysis. This approach has several benefits. One of them is the unchanged and changed PTM carrying peptides are available for detection. When there is no ionization of PTM carrying peptides then other peptide fragments will be available for detection at that time. And various carbonylated proteins pass through sulfhydryl oxidation, tyrosine nitrosylation, methionine oxidation, is another benefit.

3.3 RESEARCH QUESTIONS

This SLR consists of three questions.

Table 2. Searching Techniques used

Electronic sources	Search terms
IEEE	(“Carbonylation and their role in oxidative stress” “Carbonylation sites in aging and age- related diseases” “different tools for detecting carbonylation sites”)
Science Direct	(“Mass spectrometry-based techniques for detecting carbonylation sites” “How different approaches are used for detecting carbonylation sites” “Role of carbonylation sites in diabetes”)
Cross mark	(“Computational tools for detection of carbonylation sites” “Tools and models for carbonylation detection” “Protein carbonylation sites cause breast cancer”)
journal of proteome research	**(“What are proteomic based approaches used for detection of carbonylation sites, “All types of techniques used for detecting protein carbonylation sites”)**
Springer	(“Role of carbonylation sites in aging and age-related diseases” Tools and models for detection of carbonylation sites” “Role of Carbonylation in oxidative stress”)
Oxford	(“Detail analysis of protein carbonylation sites” “How protein carbonylation occur” “iCar-PseCp for detecting protein carbonylation sites”)
MDPI	(“Protein carbonylation sites major hallmark of oxidative stress” “ How different diseases associated with protein carbonylation sites”)
Google scholar	(“Carsite for carbonylation detection” “CarsPred computational tools for protein carbonylation detection” “MDD-Crab tool for protein carbonylation detection” “PredCar-Site” “iCarPs” “CarSPred 2.0 for identification of carbonylation”)

4.3 MULTIDIMENSIONAL FRACTIONATION

In third approach, the fractionate affinity mark proteins via liquid chromatography before detection of peptides through mass spectrometry and proteolysis. In the second, dimension.

Table 2. Table 1: Different approaches for identification of protein carbonylation sites [8]

Table 3: Different approaches for identification of protein carbonylation sites [8]

Approaches	Pros	Cons
Mass spectrometry	Permits identification of oxidized proteins and their oxidation sites, absolute quantification could be carried out.	A very tough technique, require advanced equipment, selective oxidized enrichment, peptides proteins are required.
Spectrometry	Antibody independent, signal improved and speedy.	With TCA precipitation, inactivate protein.
Dot blot	Having larger throughput and sensitive.	Nill
ELISA	Sensitive and large throughput	Standardization differs among available sets and individual sets in lab.
2-DE	In conjunction with mass spectrometry is capable of detect the oxidized proteins in complicated mixture.	Just Absolute quantitation is possible. Derivatization can impact on protein PI before electrophoresis.
GC-MS	Having sensitivity for non-purified sample when you have a SIM.	Hydrolysis of a specimen is required. The commercially accessible markers synthesis is not required.
LC-Fluorescence or MS	Sensitive and for non-purified sample while utilizing MS.	Derivatization is required. the commercially accessible markers synthesis is not required.
Carbonyl Western Blot	From complicated samples provides data about proteins.	There is only possibility of relative quantitation.

fractionation is usually get via RPC besides the affinity selection in the first dimension of chromatography. Trypsin digested and peptide fragments are protein fractions which are choose from RPC column which are detected via mass spectrometr.For example, 87 carbonylation sites of yeast have been detected in the same way. From protein, aging process oxidatively changed and unchanged the tryptic peptides, promoting protein detection depend upon the peptide sequence survey. To detect the parent protein the unchanged peptides are utilized. However, for the detection of oxidation sites the targeted carbonylated peptides are utilized. The two-dimensional electrophoresis could also be utilized, this technique is not limited to the liquid chromatography. But 2-4 Dinitrophenylhydrazine, durining Fractionation will cause the sample loss via modifying the isoelectric point of protei

4.4 PROTEOMICS BASED ON 2 GEL ELECTROPHORESIS

In the study of different diseases in humans as well as in animals for example Alzheimer disease, chronic bronchitis, and diabetes and so on, this method [19] has been utilized extensively when 2 gel electrophoresis is paired with carbonyl specific probes, it makes it easier not only for the separation and detecting the carbonylated proteins but also for estimating the amount of carbonylation of every protein in comparison to its whole amount. The most frequently utilized carbonylated detection system is dependent on Dinitrophenylhydrazine derivatization and immune recognition with anti-Dinitrophenylhydrazine antigen on 2D gel. First of all, it was written by Shacter and colleagues in 1944 and now it is available globally with the trade name OxyBlot.

4.5 QUANTIFICATION OF CARBONYLATED SITES BASED ON MASS SPECTROMETRY

For detailed understanding about the processes which are involve in protein oxidation of complicated biological processes we must have to use computational tools to detect the protein damage in that time and within various cellular compartments. For quantitative scaling the methods of choice are by combining immune-chemistry and MS both with 2-DE but, the mass spectrometry mainly utilized to detect carbonylated sites in proteins. However, the basic objective of proteolysis is to identify and quantify the oxidative amino acids in proteins.

6 MASS SPECTROMETRY-BASED APPROACHES

Mass spectrometry could be utilized to monitor any protein modification [8] without a priori hypotheses about which kind of alternation it is. Depend on the mass shift among the peptide masses and genome deduced sequence experimentally detected each protein alternation could be detected. Conversely, this method is repetitive and not appropriate for large throughput study of complicated protein sites due to the limitation in the adequate database search algorithms which could deal with such type of information. This is accomplished basically through particular method of enrichment and chemical derivatization which targets the specific modification class. Mass spectrometry of proteins is an analytical technique which is utilized to find out the protein masses and peptides masses and clear chemical structure of them. Mass spectrometry is an ideal technique for studying the proteins modifications, the reason for this is that covalent addition and loss of an amino acid chemical moiety cause to increase and decrease of that residue's molecular mass. For instance, methionine residue's oxidation increases to 147Da addition with single oxygen atom. By detecting a distinct mass rise and decrease of the intact protein or peptide it is possible to allocate a corresponding modification. In a recent years, Mass spectrometry methods for the study of proteins, both native and oxidized have been improved significantly and could basically categorized into top-down that entails the study of intact proteins and their fragmentation inside the Mass spectrometer, and bottom-up study that contains the enzymatically absorption of proteins into peptide mixture before being incorporated into the instrument.

5. AGING

Aging is known as an individual's gradual deterioration in its biochemical activities and physiological activities. In different diseases aging is a major risk factor like cancer, diabetes, and obesity and so on. The molecular and cellular processes of aging and the relation among aging and emergence of aging are not very well known. One of the main reason of aging is loss in cellular proteomic homeostatic processes which enables the accumulation of defective proteins. This fact is noted in both eukaryotic and prokaryotes showing that evolutionary conservation of molecular processes [22]. Proteins help in sustaining cell functional integrity and tissue's homeostasis shown in different studies, that's why they are important in the process of aging and longevity [23]. Studies have also revealed that radiations that caused morbidity and mortality are damaged oxidatively impaired proteins. Oxidized proteins gathered with the age and induce phenotype of reversible aging with fixed result [24].

5.1 PROTEIN CARBONYLATION SITES IN AGING

Proteomic techniques are responsible for glucose processes which were considered as carbonylated proteins and also detect the enzymes from glycolytic processes. Highly enrich proteins endorse their recognition, in spite of this aging, in various researches those proteins usually appeared as carbonylated sites. Within the mitochondria, pyruvate acquired from the glucose oxidation is metabolized into the acetyl-CoA, that in response oxidized in the Krebs cycle. In term of their alternation in ageing and in many diseases, the mitochondrial process is always answerable for energy operations, also having the complications which form the OXPHOS have possibly the most studied form. The other enzyme is also crucial in aging research the reason for this is that it appears to be oxidized repeatedly and also responsible for ATP production (creatine kinase) [25].

Commonly, oxidation of protein interconnect with the physiological PTM and then pretend as a physiological change. These conceptual physiological post translational modifications operated by ROS and effected by a sequence polymorphism, that could play the crucial operations in cell signaling process and in differentiation mechanism determination. But the elongation and chaperones factors form bacteria to human cells are exquisitely efficient proteomic sites for carbonylation and it is not surprising. The obtained phenotype of exclusive oxidative modification to bacterial proteins are due to the lack of global macromolecule biosynthesis, higher mutation rate and susceptibility to modification caused by radiation. There is also typical aging phenotype in humans, in which protein

carbonylation promote quasi-exponentially with the age of person, equal to the rise in ARD and death rate.

5.2 EXPERIMENT TO FIND THE RELATIONSHIP BETWEEN CARBONYLATION AND ITS ROLE IN AGING:

There are large number of experiments related to the relating between protein carbonylation sites and aging and age-related disorders. So, I will survey one of the study [26] which was done on 10,001 subsamples of Dalmatians research, that is a vast research based biobank in Croatia. To study genetics was the main goal of this study, social and environmental situation of health and diseases with a focus on fetal diseases, that in Croatia and other counties are responsible for death.

They selected 50 healthy subjects for this study, from the island of Korcula. Where there would no evidence of any form of underlying disorder which could have make impact on the result a stable subject status was analyzed by the detailed analysis of its medical data. Their related biological plasma specimens were obtained from the 10.001nDalmatians freezer and operated in the study. To examine the information, descriptive and analytical methods have been utilized. These methods majorly consist of means and standard deviations, also with the correlation investigation. The built-in functions of MS excel was the base for the pattern investigation and data fitting, that gives the ability to quantify the model's explain variance. The method of cross validation and the study of imputation was the bases for the final analytical stage. The datasets were distributed into the two equals parts, then the remaining half was utilized to acquire the data related to the chronological age. The remaining age information to be impute based on the level of carbonylation and gender is enabled by this. The same procedure was repeated producing the total of five full sets of the expected age data, for every of the two sub-sets [28-51]. In order to give an idea of probability of utilizing carbonylation as an aging indicator in large datasets, this information was then utilized in the division study. In IBM SPSS the imputations and correlation were carried out.

5.3.CONCLUSION RELATED TO CARBONYLATION SITES AND AGING

This carbonylation is linked with the chronological age concluded by this study, but the one-dimensional carbonylation has restricted capacity in the age prediction. That is based on a few numbers of changes which they find out, but also on the significant variance found in analyzed datasets. The extent of carbonylation variance was quite large, whereas variation percentage of the estimated age was close to just 15% in the best case scenario, stated by carbonylation.

The research concluded that carbonylation values are more common in elderly men than the values in younger men. The age is attenuated by biological regulation operation known as antioxidant plausible description for this reality, that function as ROS scavenger mechanism. The result of the study carried out in 2006 is similar with the result of this study, that concluded the plasma antioxidant potential that cause decrease as an age function ability and to nullify oxidative stress.

So on, this study predicts that in women carbonylation process, shows the lower values in younger women, but in elderly women a stronger uncertain tendency of carbonylation. This causes due to the consequence of the minimum of the estrogen level to a degree equivalent to that of age-matched men in women following puberty. This study indicated that there is a crucial link in between pulse wave velocity and large carbonylation level and augmentation index (AI). Both augmentation index and higher level of PWV are associated with the vast carbonylation. Maximum PWV and augmentation index values predict reduce atrial compliance that in turn is corelated with increased aging. Standard progressive matrices (SPM) are helpful methods for analytical appraisal by abstract reasoning activities and intelligence diagnosis. They find out that an increased degree of carbonylation is correlated with worst results indicated in the Raven Standard Progressive Matrices test.

6. DIABETES

As decreased state of 3 respiration of mitochondria obtained from obese hearts and diabetic type 2 db/db mice was first demonstrated a 30 years ago, huge quantity of data of the mitochondrial damage in diabetic hearts was gathered. Mechanisms related to the impaired mitochondrial oxidative capacity and modified morphology also including oxidative demolition, transcriptional and translational alternations in the subunit expression of oxidative phosphorylation, damaged handling of mitochondrial calcium and modifications in cardiac insulin signaling. The regulation cause aberrant signaling pathways compromising various signaling substances or activities of enzymes and maybe impacted via a chronic hyperglycemia and ROS development of metabolic pathway. Thioredoxin interaction protein has been shown to linked with oxidative stress as an example and to be high glucose upregulated. The high glucose ROS signaling pathway is because of thioredoxin interaction protein which is a key factor and also required by the NADPH oxidase isoform in mesangial cells to cause the production of mitochondrial and overall cellular hydrogen peroxide in mesangial cells, and could therefore be a possible objective to overcome diabetic nephropathy.

In additionally, the high-level cellular glucose metabolism is the main reason of the higher amount of MG plasma level, which have been observed in the patients with newly identified diabetes type 2, relating increased MG levels and its metabolism with glyoxalase 1 to medical microvascular complications for example retinopathy and nephronpathy [27]. It has recently been suggested that in diabetic nephropathy the GST activity could play a crucial part because, it is exceedingly active in nephropathy patients, as compared to the healthy individuals. It is essential to

create the GST isoform liable for this discovery, although there is proof that many GST isoforms have both transferase and selenium-independent peroxidase mechanism.

7. PROTEIN CARBONYLATION IN OTHER RELATED DISEASES [27]

Because of the 3-hydroxy-3-methylglutaris acid-CoA lyase shortage, 3-hydroxy-3-methylglutaric aciduria is an autosomal recessive situation, causing in the accumulation of 3-hydroxy-3-methylglutaric (HMG) and 3-methylglutaric acid (MGA) in an effected individual's cells and body fluids. It is clear that the 3-hydroxy-3-methylglutaric and 3-methylglutaaric acid increase the activity of glutathione peroxidase and catalase but reduce the cellular levels of reduced glutathione, superoxide dismutase and glutathione reductase activities. Because the carbonylation pathways are assumed to be the cause of protein confirmation, behaviour and function that's why it is assumed that oxidative stress causing protein carbonylation of the aforementioned redox enzymes could also play role in the generation of 3-hydroxy-3-methylglutaric acid.

In patients having diabetes, the chances of occurrence of cardiovascular disorders such as atherosclerosis are higher and linked with the degree of glycemia regulation without providing knowledge into causality, demonstrated by the epidemiological trials. The pathological results of age accumulation are pathways for vascular disorders in diabetes like damaged vasodilatory reaction due to the inhibition of nitric oxide, muscle cell malfunction, oversupply of endothelial growth elements, fetal inflammation, impaired fibrinolytic ability, and increase platelet estimation. In type 2 diabetes patients and cardiovascular disorders,the tissue resistor metalloproteinase-3 expression in carotid atherosclerotic plaque is decreased, especially in regions with enrichment of monocyte type cells. Over expression of metalloproteinase in transgenic mice isolated macrophages cause minimize in strength of oxidative stress, as demonstrated by the lower lipid peroxidation, protein carbonylation and atheromas nitration thereby provding further link among diabetes, vascular cells oxidative carbonylation and atherosclerosis.

8. CONCLUSION

In this article a detail analysis of protein carbonylation sites has been presented. Introduction to carbonylation its origin is also discussed. This systematic literature provides deep understanding about the various approaches and techniques used in carbonylation identification. Among them the Mass spectrometry is the most common on. R, K, P, T site are more susceptible for carbonylation indicated by these techniques. Besides this, different computational tools are also described and their comparison table is also provided. In each tool they used the same basic datasets for evaluation. Each tool has its own accuracy and capability for detection of carbonylation sites. These tools are much better as compared to the existing approaches and techniques because they are time-consuming and costly. And these tools provide accurate results. We know that carbonylation is a hallmark of oxidative stress and this cause various other diseases and that's why I try to also provide the brief detail about the role of carbon

ylation sites in different human diseases. However, there is still research gap in identification of carbonylated sites. And many other researches are still going for their identification.

REFRENCES

[1] Rao, R. S. P, Møller and I.M, "Pattern of occurrence and occupancy of carbonylation sites in proteins. Proteomics," Proteomics, vol. 11, no. 21, pp. 4166-4173, 2011.

[2] Augustyniak, E, Adam, A, Wojdyla, K, Rogowska-Wrzesinska, A, Willetts, R, Korkmaz, A, Griffiths and H. R, "Validation of protein carbonyl measurement: a multi-centre study," Redox biology, vol. 4, pp. 149-157, 2015.

[3] Fedorova, M, Bollineni, R.C, Hoffmann and R, "Protein carbonylation as a major hallmark of oxidative damage: update of analytical strategieS," Mass spectrometry reviews, vol. 33, no. 2, pp. 79-97, 2014.

[4] Weng, S.L, Huang, K.Y, Kaunang, F.J, Huang, C.H, Kao, H.J, Chang, T.H, Lee and T.Y, "Investigation and identification of protein carbonylation sites based on position-specific amino acid composition and physicochemical features," BMC bioinformatics, vol. 18, no. 3, pp. 125-141, 2017.

[5] Xu, Y, Wang, X, Wang, Y, Tian, Y, Shao, X, Wu, L.Y, Deng and N, "Prediction of posttranslational modification sites from amino acid sequences with kernel methods," Journal of theoretical biology, vol. 344, pp. 78-87, 2014.

[6] Coffey, C.M, Gronert and S, "A cleavable biotin tagging reagent that enables the enrichment and identification of carbonylation sites in proteins," Analytical and bioanalytical chemistry, vol. 408, no. 3, pp. 865-874, 2016.

[7] sJia, J, Liu, Z, Xiao, X, Liu, B, Chou and K.C, "iCar-PseCp: identify carbonylation sites in proteins by Monte Carlo sampling and incorporating sequence coupled effects into general PseAAC," Oncotarget, vol. 7, no. 23, pp. 34558-34570, 2016.

[8] Rogowska-Wrzesinska, A, Wojdyla, K, Nedić, O, Baron, C, P, Griffiths and H.R, "Analysis of protein carbonylation—pitfalls and promise in commonly used

methods," Free radical research, vol. 48, no. 10, pp. 1145-1162, 2014.

[9] Joshi, S, Kumari, S, Rathore and A.S, "Identification and characterization of carbonylation sites in trastuzumab biosimilars," International Journal of Biological Macromolecules, 2020.

[10] Lv, H, Han, J, Liu, J, Zheng, J, Liu, R, Zhong and D, "CarSPred: a computational tool for predicting carbonylation sites of human proteins," PloS one, vol. 9, no. 10, p. e111478, 2014.

[11] Jia, J, Liu, Z, Xiao, X, Liu, B, Chou and K.C, "iCar-PseCp: identify carbonylation sites in proteins by Monte Carlo sampling and incorporating sequence coupled effects into general PseAAC," Oncotarget, vol. 7, no. 23, pp. 34558-34570, 2017.

[12] Kao, H.J, Weng, S.L, Huang, K.Y, Kaunang, F.J, Hsu, J, B. K, Huang, C.H, Lee and T.Y, "MDD-carb: a combinatorial model for the identification of protein carbonylation sites with substrate motifs," BMC systems biology, vol. 11, no. 7, pp. 127-140, 2017.

[13] Zuo, Y, Jia and C.Z, " CarSite: identifying carbonylated sites of human proteins based on a one-sided selection resampling method," Molecular BioSystems, vol. 13, no. 11, pp. 2362-2369, 2017.

[14] Hasan, M.A.M, Li, J, Ahmad, S, Molla and M.K.I, "predCar-site: Carbonylation sites prediction in proteins using support vector machine with resolving data imbalanced issue," Analytical biochemistry, vol. 525, pp. 107-113, 2017.

[15] Zhang, D, Xu, Z.C, Su, W, Yang, Y.H, Lv, H, Yang, H, Lin and H, "iCarPS: a computational tool for identifying protein carbonylation sites by novel encoded features," Bioinformatics, 2020.

[16] Lyu, H, Hao, L, Zheng, J, Liu, C, Liu, Y, Shang and Y, "Predicting Carbonylation Sites of Human Proteins with a New Max-Significance and Min-Redundancy Feature Selection Criterion," In 2018 9th International Conference on Information Technology in Medicine and Education (ITME), pp. 265-272, 2018.

[17] Rao, R.S.P, Zhang, N, Xu, D, Møller and I.M, "CarbonylDB: a curated data-resource of protein carbonylation sites," Bioinformatics, vol. 34, no. 14, pp. 2518-2520, 2018.

[18] Madian, A.G, Regnier and F.E, "Proteomic identification of carbonylated proteins and their oxidation sites," Journal of proteome research, vol. 9, no. 8, pp. 3766-3780, 2010.

[19] Møller, I.M, Rogowska-Wrzesinska, A, Rao and R.S.P, "Protein carbonylation and metal-catalyzed protein oxidation in a cellular perspective," Journal of proteomics, vol. 74, no. 11, pp. 2228-2242, 2011.

[20] Verrastro, I, Pasha, S, Jensen, K.T, Pitt, A.R, Spickett and C.M, "Mass spectrometry-based methods for identifying oxidized proteins in disease: advances and challenges," Biomolecules, vol. 5, no. 2, pp. 378-411, 2015.

[21] Artemenko, K, Mi, J, Bergquist and J, "Mass-spectrometry-based characterization of oxidations in proteins," Free radical research, vol. 49, no. 5, pp. 477-493, 2015.

[22] Santos, A.L, Lindner and A.B, "Protein posttranslational modifications: roles in aging and age-related disease," Oxidative Medicine and Cellular Longevity, vol. 2017, 2017.

[23] Mittal, A, Rath and P.C, "Protein Structure and Function in Aging and Age-Related Diseases. In Models," Molecules and Mechanisms in Biogerontology, pp. 3-26, 2020.

[24] Krisko, A, Radman and M, "Protein damage, ageing and age-related diseases," Open biology, vol. 9, no. 3, p. 180249, 2019.

[25] Cabiscol, E, Tamarit, J, R. and J, "Protein carbonylation: proteomics, specificity and relevance to aging. Mass spectrometry reviews," Mass spectrometry reviews, vol. 33, no. 1, pp. 21-48, 2014.

[26] Langert and L, "PLASMA PROTEIN CARBONYLATION AS BIOMARKER FOR ESTIMATE OF AGING," Doctoral dissertation, University of Split. School of Medicine. Public health, 2018.

[27] Hecker, M, Wagner and A.H, "Role of protein carbonylation in diabetes," Journal of inherited metabolic disease, vol. 41, no. 1, pp. 29-38, 2018.

[28] Saeed, S.; Mahmood, M. K.; Khan, Y. D., An exposition of facial expression recognition techniques. Neural Computing and Applications 2018, 29 (9), 425-443.

[29] Butt, A. H.; Khan, Y. D., CanLect-Pred: A cancer therapeutics tool for prediction of target cancerlectins using experiential annotated proteomic sequences. IEEE Access 2019, 8, 9520-9531.

[30] Amanat, S.; Ashraf, A.; Hussain, W.; Rasool, N.; Khan, Y. D., Identification of lysine carboxylation sites in proteins by integrating statistical moments and position relative features via general PseAAC. Current Bioinformatics 2020, 15 (5), 396-407.

[31] Ilyas, S., Hussain, W., Ashraf, A., Khan, Y. D., Khan, S. A., & Chou, K. C. (2019). iMethylK-PseAAC: Improving accuracy of lysine methylation sites identification by incorporating statistical moments and position relative features into general PseAAC via Chou's 5-steps rule. Current Genomics, 20(4), 275-292.

[32] Hussain, W.; Rasool, N.; Khan, Y. D., A Sequence-Based Predictor of Zika Virus Proteins Developed by Integration of PseAAC and Statistical Moments. Combinatorial chemistry & high throughput screening 2020, 23 (8), 797-804.
[33] Khan, Y. D.; Alzahrani, E.; Alghamdi, W.; Ullah, M. Z., Sequence-based Identification of Allergen Proteins Developed by Integration of PseAAC and Statistical Moments via 5-Step Rule. Current Bioinformatics 2020, 15 (9), 1046-1055.
[34] Mahmood, M. K.; Ehsan, A.; Khan, Y. D.; Chou, K.-C., iHyd-LysSite (EPSV): Identifying Hydroxylysine Sites in Protein Using Statistical Formulation by Extracting Enhanced Position and Sequence Variant Feature Technique. Current Genomics 2020, 21 (7), 536-545.
[35] Naseer, S.; Hussain, W.; Khan, Y. D.; Rasool, N., IPhosS (Deep)-PseAAC: Identify phosphoserine sites in proteins using deep learning on general pseudo amino acid compositions via modified 5-Steps rule. IEEE/ACM Transactions on Computational Biology and Bioinformatics 2020.
[36] Naseer, S.; Hussain, W.; Khan, Y. D.; Rasool, N., Sequence-based identification of arginine amidation sites in proteins using deep representations of proteins and PseAAC. Current Bioinformatics 2020, 15 (8), 937-948.
[37] Shah, A. A.; Khan, Y. D., Identification of 4-carboxyglutamate residue sites based on position based statistical feature and multiple classification. Scientific Reports 2020, 10 (1), 1-10.
[38] Awais, M.; Hussain, W.; Rasool, N.; Khan, Y. D., iTSP-PseAAC: Identifying Tumor Suppressor Proteins by Using Fully Connected Neural Network and PseAAC. Current Bioinformatics 2021, 16 (5), 700-709.
[39] Hussain, W.; Rasool, N.; Khan, Y. D., Insights into Machine Learning-based approaches for Virtual Screening in Drug Discovery: Existing strategies and streamlining through FP-CADD. Current Drug Discovery Technologies 2021, 18 (4), 463-472.
[40] Khan, Y. D.; Khan, N. S.; Naseer, S.; Butt, A. H., iSUMOK-PseAAC: prediction of lysine sumoylation sites using statistical moments and Chou's PseAAC. PeerJ 2021, 9, e11581.
[41] Malebary, S. J.; Khan, R.; Khan, Y. D., ProtoPred: Advancing Oncological Research Through Identification of Proto-Oncogene Proteins. IEEE Access 2021, 9, 68788-68797.
[42] Malebary, S. J.; Khan, Y. D., Evaluating machine learning methodologies for identification of cancer driver genes. Scientific reports 2021, 11 (1), 1-13.
[43] Malebary, S. J.; Khan, Y. D., Identification of Antimicrobial Peptides Using Chou's 5 Step Rule. CMC-COMPUTERS MATERIALS & CONTINUA 2021, 67 (3), 2863-2881.
[44] Naseer, S.; Ali, R. F.; Khan, Y. D.; Dominic, P., iGluK-Deep: computational identification of lysine glutarylation sites using deep neural networks with general pseudo amino acid compositions. Journal of Biomolecular Structure and Dynamics 2021, 1-14.
[45] Naseer, S.; Hussain, W.; Khan, Y. D.; Rasool, N., NPalmitoylDeep-PseAAC: A Predictor of N-Palmitoylation Sites in Proteins Using Deep Representations of Proteins and PseAAC via Modified 5-Steps Rule. Current Bioinformatics 2021, 16 (2), 294-305.
[46] Naseer, S.; Hussain, W.; Khan, Y. D.; Rasool, N., Optimization of serine phosphorylation prediction in proteins by comparing human engineered features and deep representations. Analytical Biochemistry 2021, 615, 114069.
[47] Khanum, S., Ashraf, M. A., Karim, A., Shoaib, B., Khan, M. A., Naqvi, R. A., ... & Alswaitti, M. Gly-LysPred: Identification of Lysine Glycation Sites in Protein Using Position Relative Features and Statistical Moments via Chou's 5 Step Rule.
[48] Lv, H., Dao, F. Y., Zhang, D., Yang, H., & Lin, H. (2021). Advances in mapping the epigenetic modifications of 5-methylcytosine (5mC), N6-methyladenine (6mA), and N4-methylcytosine (4mC). Biotechnology and Bioengineering.
[49] Zulfiqar, H., Sun, Z. J., Huang, Q. L., Yuan, S. S., Lv, H., Dao, F. Y., ... & Li, Y. W. (2021). Deep-4mCW2V: A sequence-based predictor to identify N4-methylcytosine sites in Escherichia coli. Methods.
[50] Liu, Y., Wang, X., & Liu, B. (2019). A comprehensive review and comparison of existing computational methods for intrinsically disordered protein and region prediction. Briefings in bioinformatics, 20(1), 330-346.
[51] Zhang, D., Xu, Z. C., Su, W., Yang, Y. H., Lv, H., Yang, H., & Lin, H. (2021). iCarPS: a computational tool for identifying protein carbonylation sites by novel encoded features. Bioinformatics, 37(2), 171-177.