Thakur et al., J Proteomics Bioinform 2019, 12:1
DOI: 10.4172/0974-276X.1000491

Journal of
Proteomics & Bioinformatics

Research Article Open Access

A Multidimensional Approach to Explore the Use of a Small Heterobifunctional Crosslinker based on a Metabolite of the Kynurenine Pathway

Sujeet Kumar Thakur[1], Mitul Srivastava[2], Ankur Kumar[2], Renu Goel[2], Shailendra Asthana[2] and Sambasivan Venkat Eswaran[1]*

[1]Regional Centre for Biotechnology, Faridabad, NCR Biotech Science Cluster, 3rd Milestone, Faridabad-Gurgaon Expressway, Faridabad-121001, Haryana, India
[2]Drug Discovery Research Center, Translational Health Science and Technology Institute, Faridabad, NCR Biotech Science Cluster, 3rd Milestone, Faridabad-Gurgaon Expressway, Faridabad-121001, Haryana, India

Abstract

This study describes the use of a new small heterobifunctional crosslinker for crosslinking of proteins (e.g. lysozyme). This crosslinker is based on 3-hydroxy anthranilic acid (3HAA) that is part of the kynurenine pathway of the degradation of Tryptophan. 3HAA is found in enhanced amounts in disease states in the human body. Small crosslinkers capture interacting protein interfaces better, while the larger ones are more useful for identifying interacting partners. The new crosslinker described here, functions presumably via a 'long lived' transient, leading to enhanced rate of intermolecular crosslinking, which is otherwise difficult to achieve. It contains a photo labile azido group and an amine reactive N-hydroxysuccinimide (NHS) group. Successful crosslinking in two steps (incubation followed by photolysis (366 nm, 6W UV lamp) , has been confirmed using SDS-PAGE, ESI-MS/MS, and bioinformatics analysis via StavroX 3.6.0.1 Docking followed by molecular dynamics simulation studies, have provided detailed structural insights into the 'dimer' formation of lysozyme. Identical conclusions have been obtained, using two different software, and providing a more refined 3D view of the interfaces during protein-protein interactions.

Keywords: 3-Hydroxy anthranilic acid (3HAA); Kynurenine; ESI-MS/MS; Bioinformatics; Molecular Docking; Molecular dynamics simulation; Schrodinger software

Introduction

The method of chemical crosslinking- mass spectrometry-bioinformatics has come of age and is now occupying centre stage, as it contributes to studies on protein-protein interactions (PPIs), antibody drug conjugates (ADCs) and even to Cryo-EM, Zero length, homo and hetero-bifunctional, mass cleavable, and isotope labeled crosslinkers are known previously [1-18]. The most commonly used crosslinkers being homobifunctional crosslinkers, which possess two identical groups at the two ends [often amine reactive N-hydroxysuccinimde (NHS) groups]. A very popular such reagent is BS^2G [Bis (sulphosuccinimidyl) 2,2,4,4, glutarate]. Heterobifunctional crosslinkers, on the other hand, contain two different groups at the two ends (e.g. amine reactive NHS group and the photo reactive azide group). Both small and large crosslinkers are available, the former providing better information about interfaces while the latter are more useful for identifying interacting partners. The protocol commonly used in such studies involves crosslinking, SDS-PAGE, MALDI-MS, ESI-MS/MS of the 'dimeric' band and analysis using different bioinformatics tools. As uncrossed linked fragments dominate and one is trying to identify the intra- and inter- molecular crosslinking, the real difficulty lies in creating and identifying inter-molecularly crosslinked fragments. This is often referred to finding 'a needle in a haystack' problem. Crosslinked fragments are invariably charged (+1, + 2, +3, +4.), though separating them selectively is often referred to as finding 'two needles in a haystack' problem In recent years, there have been great improvements in mass spectrometric techniques and simultaneously in bioinformatics tools for specially identifying intermolecular crosslinked fragments. StavroX 3.6.0.1 [19] is one such software, which helps in identifying intermolecular crosslinked fragments. The difficulty lies in increasing the rate of intermolecular reaction. In this regard, heterobifunctional crosslinkers, based on fluorinated [20] and those based on 'long lived transients', have proved to be very useful for this purpose. This has been highlighted in recent papers from our laboratory [21-23]. The role of 3-hydroxykynurenine in aggregation of α-Crystallin in cataractogenesis has been previously highlighted [24]. The homobifunctional crosslinker BS^2G and mass spectrometry has been used to compare the aggregation tendency of the mutant and the wild type α-Crystallin and its importance in cataract formation [25]. Here, we have discussed the use of a small heterobifunctional crosslinker. It is based on 3-hydroxy anthranilic acid (3HAA), whose levels are found to be enhanced in patients suffering from different diseases [26a-c]. 3HAA is a part of the alternative kynurenine pathway of Tryptophan metabolism (different from the better-known Serotonin pathway). The ratio of 3HAA to anthranilic acid has been shown to be crucial in diseased states [27a]. 3HAA is further converted into the toxin quinolinic acid. Even small amounts of this toxin in the human brain can lead to death. Considering the importance of crosslinkers in disease biology and therapeutics, we aimed to developed the small heterobifunctional crosslinker based on 3HAA that could be particularly useful for studying interacting interfaces during PPIs, and especially those leading to disease states in the human body. It may be noted that the new crosslinker is based on a critical metabolite of the essential amino acid Tryptophan. We also further wished to combine and integrate by combining and integrating the results from chemical crosslinking-mass spectrometry-

***Corresponding author:** Sambasivan Venkat Eswaran, Regional Centre for Biotechnology, Faridabad, NCR Biotech Science Cluster, 3rd Milestone, Faridabad-Gurgaon Expressway, Faridabad-121001, Haryana, India, Tel: + +001292848888; E-mail: samba.eswaran@rcb.res.in

Received December 17, 2018; **Accepted** January 22, 2018; **Published** January 29, 2019

Citation: Thakur SK, Srivastava M, Kumar A, Goel R, Asthana S, et al. (2019) A Multidimensional Approach to Explore the Use of a Small Heterobifunctional Crosslinker based on a Metabolite of the Kynurenine Pathway. J Proteomics Bioinform 12: 010-0017. doi: 10.4172/0974-276X.1000491

bioinformatics and all atom molecular dynamics studies, and try to validate and further refine 3D-representation of the interfaces during the protein-protein interactions with and without the crosslinker.

Materials and Methods

Synthesis of the new crosslinker & crosslinking of lysozyme

The new heterobifunctional crosslinker, 2-Azido-3-methoxy-benzoic acid-2, 5-dioxo-pyrrolidin-1-yl ester (135 mg), (II) was prepared from 2-Amino-3-methoxy-benzoic acid (I) 200 mg (1.19 mmol) was taken and dissolved in 8ml of concentrated hydrochloric acid and 2 ml of water and cooled at 0°C, this was diazotized by slow addition of sodium nitrate (140 mg, 1.2 mmol) in minimum amount of water required. A solution of sodium azide NaN_3 (120 mg, 2 mmol) and sodium acetate (3.36 g,40 mmol) in minimum water was taken and slowly added to the diazotized solution, when an off white solid settled down on the bottom of the vessel. This compound was filtered and thus compound (II) was obtained (yield, 175 mg). 175 mg of compound (II) was dissolved in 10 ml of dichloromethane (DCM) and adding N-hydroxysucinimide NHS, 59.8 mg, and 0.52 mmol, dicyclohexyl carbodiimide (DCC) (107 mg, 0.52 mmol) were added, and the reaction mixture was stirred at room temperature overnight. The solution was filtered to separate the urea side product. The filtrate was distilled and put in a desiccator with P_2O_5 when a pale white compound (III) was obtained. The scheme for synthesis of (III) is shown in Supplementary Figure 1 and MALDI-MS spectrum of (III) is shown in Supplementary Figure 2; m. p.- 192°C; MS, m/z- 290.5 [27b]. Lysozyme (14 kDa) was incubated overnight with the new crosslinker and then subjected to photolysis for 366 nm for 30 minutes, when intermolecular crosslinking occurs, a 'dimeric' band can be observed in the SDS-PAGE at around 28 kDa Supplementary Figure 3.

Lysozyme

The protein chosen for this study is lysozyme as it is a well-studied protein. It is made up of 129 amino acids and is a globular protein, which is found in human tears and the white of the chicken egg. It contains 6 lysine, 11 arginine and 8 cysteine residues. This protein has a molecular weight of 14.4 kDa (Figures 1 and 2). After incubation with the new crosslinker in PBS Buffer at 7.4 pH value with (1: 200) protein and crosslinker are incubated and the sample was photolyzed at 366 nm with 6W UV lamp for 30 minutes.

SDS-PAGE and in-gel digestion

10 micrograms of lysozyme was incubated with the crosslinker,

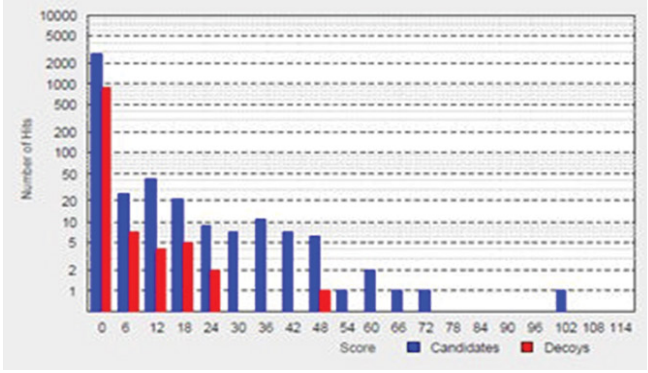

Figure 1: Screen shot of the decoy Analysis for m/z 1904.912.

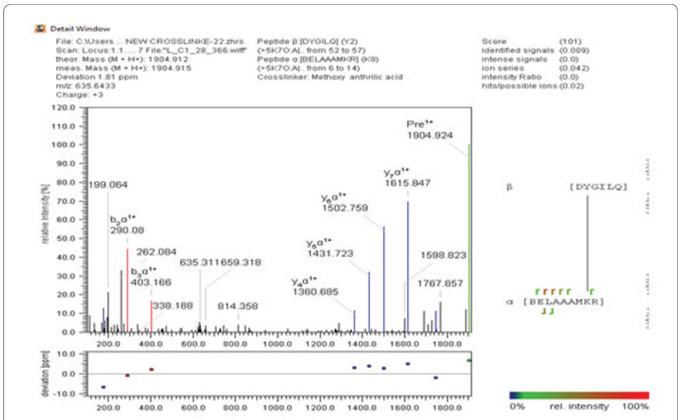

Figure 2: The annotation with the extent of deviation and identified peak m/z 1904.912.

photolyzed and then resolved by SDS-PAGE. The gel was stained with Coomassie brilliant blue r-250 and destained with water. Gel pieces were excised and in-gel digestion was carried out [28]. The excised bands were destained with 40mM ammonium bicarbonate (ABC) in 40% acetonitrile (ACN). The gel bands were subjected to reduction and alkylation using 5mM dithiothreitol (DTT) (60°C for 45 min) and alkylation using 10mM iodoacetamide (IAA). The gel sections were dehydrated with 100% ACN, followed by digestion with trypsin (Gold mass-spectrometry trypsin; Promega, Madison, WI) at 37°C for 10-12 h. The peptides were removed from the gel pieces with 0.4% formic acid in 50% ACN solution and finally with 100% ACN. The extracted peptides were vacuum-dried and stored at 80°C until LC-MS/MS analysis.

LC-MS/ MS analysis

All fractions were evaluated by 5600 Triple-TOF mass spectrometer which is directly linked to reverse-phase high-pressure liquid chromatography Ekspert-nanoLC 415 system (Eksigent; Dublin, CA). The trap column (200 µm × 0.5 mm) and the analytical column (75 µm × 15 cm) were both from Eksigent, packed with 3 µm ChromXP C-18 (120 Å) used for reverse phase elution by Ekspert-nanoLC 415 system. 0.1% formic acid in water was used as mobile phase A and mobile phase B is 0.1% formic acid in ACN. All fractions were eluted from the analytical column at a flow rate of 250 ml/min using an initial gradient elution of 10% B from 0 to 5 min, transitioned to 40% over 120 min, ramping up to 90% B for 5 min, holding 90% B for 10 min, followed by re-equilibration of 5% B at 10 min with a total run time of 150 min. Peptides were injected into the mass spectrometer using 10 µm Silica Tip electrospray Pico Tip emitter (New Objective Cat. No. FS360-20-10-N-5-C7-CT), and the ion source was operated with the following parameters: ISVF = 1950; GS1 = 20; CUR = 12. The data dependent acquisition (DDA) experiments was set to obtain a high resolution TOF-MS scan over a mass range 100–1250 m/z, followed by MS/MS scans of 20 ion candidates per cycle with activated rolling collision energy, operating the instrument in high sensitivity mode using the Analyst TF 1.7 where each 1 second MS survey scan was followed by 3 MS/MS scans of 3 seconds. The selection criteria for the parent ions included the intensity, where ions have to be greater than 150 cps, with a charge state between +2 to +5, mass tolerance of 50 mDa and on a dynamic exclusion list were present. Mass spectra (MS) and tandem mass spectra (MS/MS) were recorded in positive-ion and high-sensitivity mode with a resolution of ~35,000 full-width half-maximum.

Before running samples in the mass spectrometer, calibration of spectra was done after acquisition of every sample using dynamic LC–MS and MS/MS acquisitions of 100-fmol β-galactosidase. MS/MS, its mass, had fragmented once an ion and isotopes were excluded from further MS/MS fragmentation for 12s. The ion accumulation time was set to 250 ms (MS) and to 70 ms (MS/MS). The collected raw files spectra were stored in (dot) .wiff format.

Data analysis

All raw mass spectrometry files were searched in Protein Pilot software v. 5.0.1 (SCIEX) with the Paragon algorithm for relative protein quantification and identification. For Paragon searches, the following settings were used: Sample type: Identification; Cysteine Alkylation: Iodoacetamide, Digestion: Trypsin; Instrument: TripleTOF5600; Species: homosapiens; maximum allowed missed cleavages 1, Search effort: Thorough ID; Results Quality: Correction was automatically applied. The search was conducted using a through identification effort of a Ref-seq. database from the National Center for Biotechnology Information (NCBI) website (https://www.ncbi.nlm.nih.gov/refseq/). False discovery rate analysis was also performed through decoy database (mention database). Carbamidomethylation (C) was used as a fixed modification. The peptide and product ion tolerance of 0.05 Da was used for searches. The output of this search is a group file and this file contains the following information that is required for targeted data extraction: protein name and accession, cleaved peptide sequence, modified peptide sequence, relative intensity, precursor charge, unused Protscore, confidence, and decoy result.

Data submitted

The raw data has been submitted to public data repositories Proteome Exchange via Massive and Pride. The raw files obtained from mass-spectrometer, files from database search and result files can be downloaded with the database identifier. As mentioned, the data was downloaded from Pride through Proteome Exchange with the dataset identifier.

Identification of intermolecular crosslinked peptides

StavroX 3.6.0.1 was used to identify the intermolecularly cryosslinked peptides. It enables quick and efficient identification of the intermolecular crosslinked peptides. This software calculates the theoretical crosslinks and estimates them to the precursors of the MS/MS data stored in the form of (dot) .mgf file. This further leads to the identification of the 'hits' and 'scores' and given accordingly. For analysis, we need to load the original FASTA sequence of our protein (Supplementary Figures 3 and 4) along with the MS/MS data. It provides options to select the desired cross-linker along with the scope to add new cross-linkers. The crosslinker used in this experiment has a chemical composition of $C_{12}H_{10}N_4O_5$ with the molecular mass of 290.2316. No changes were made in the amino acid sequence section. Missed cleaveges were taken to be equal to three during this analysis. An unspecific digest option was selected along with minimum peptide length as 5 and maximum peptide length as 10. The precursor precision was 3.0 ppm and fragment ion precision was 10.0 ppm, the lower mass limit as 200.0 Da and upper mass limit as 3500 Da (our mass-spec range). The S/N ratio was selected to be 2. Only 'b' and 'y' ions were selected with the score cut-off of -1 and pre score intensity greater than 10%. FASTA sequence of lysozyme was loaded and appropriate settings were selected to run the process.

Structural Bioinformatics

The X-ray crystallographic structure of lysozyme (PDB ID: 5K7O, resolution 1.8 Å) [29] was prepared using Protein Preparation wizard of Maestro. The molecule III was drawn in 2D sketcher of Schrodinger (version 2017-1) [30] and then it was prepared an optimized using LIGPREP [31] module, which generates tautomers, and possible ionization states at the pH range 7 ± 2 using Epik [32] and also generates all the stereoisomers of the cross-linker if necessary. The optimization was done using the OPLS3 force field [33]. To check the activity of III and to find its binding site as the co-crystal is absent, the blind docking studies were performed [34-37]. It was done using AUTODOCK4.2 [38]. The most favorable pose was picked based on high docking energy and number of cluster. To be more confident about the binding site of 5K7O, we also performed Sitemap [39] module of Schrodinger. The complex so formed after docking was subjected to Molecular Dynamics Simulation for 100ns. MD simulations were carried out in DESMOND [40] module of Schrodinger Suite using the OPLS3 force field and the system was solvated with TIP3P solvent model. Orthorhombic box shape was chosen, as it suits best for the globular proteins, with the edge length of 10 Å ensuring the minimal distance between atoms of protein and edge of the box. Counter ions were added to neutralize the systems. For docked complexes, the above parameters were same and counter ions were added at least 20 Å from III. The system relaxed before the actual simulation by a series of energy minimization and short MD simulations. There are mainly six relaxation steps in this process, where minimization of solute restrained and without restraints are carried in first two steps. Step three through step six are short MD simulations of 12 ps, 12ps, 12ps and 24ps each using NPT ensemble each using the NPT ensemble at 10, 10, 300, and 300k, respectively. In between at Step 5 the pocket is solvated as well. Velocity resampling is carried in steps three to five, while at step six it is not done. The NPT ensemble was employed for the simulations with Nose-Hover chain thermostat and the Martyna-Tobias-Klein barostat. RESPA integrator was used with a time step of 0.002 ps. For short-range Columbic interactions, a 9.0 Å cut off radius was considered. Bonds to hydrogen were constrained using the M-SHAKE algorithm of DESMOND. The simulation was carried out for total 100ns for each system and the coordinates were saved at intervals of 20 ps that are referred to as "frames" in this study. The whole protocol was implemented in order to observe stability of III at the binding site. After the validation of III at predicted site, the most stable complex was generated (hereafter COM). This COM was subjected to protein-protein docking [41] in order to understand the cross-linking activity of III between Y53 of one chain and K13 of its replica chain. Protein-protein docking was accomplished in two different ways: 1) APO-APO docking, to observe the distance between Y53 and K13 and 2) in presence of III (COM-APO docking) to see the behavior of III that, in its presence whether, the distance between these two residues get reduced or not ? In addition, separate experiments were setup in order to understand the energy difference of complexes in presence/absence of III. The protein-protein docking was done using PIPER [42] and Swarmdock [43]. Both different programs were used for obtaining the consensus pose and also to increase the efficiency of our result. The docked complexes energy (APO-APO and COM-APO) was obtained from Swarmdock.

Results and Discussion

The new small hetero-bifunctional crosslinker, 2-Azido-3-methoxy-benzoic acid-2, 5-dioxo-pyrrolidin-1-yl ester, (III) has been prepared and characterized spectroscopically. It's MALDI-MS spectrum showed the M+ to be m/z 290.5. Lysozyme was chosen for this study, as it is a well-studied protein. Lysozyme (14 kDa), was incubated overnight with the new crosslinker and then subjected to photolysis at (366 nm, 6W UV lamp) for 30 minutes. This mixture

was subjected to SDS -PAGE, the 'dimeric' band excised, subjected to trypsinization, enrichment using 'zip-tip', followed by ESI-MS, MS/MS. The data thus obtained was fed into the StavroX 3.6.0.1 software as a.mgf file. Thus, 7532 spectra out of 7553 spectra were compared by the software to 4223745 theoretical candidates. The software identified 40135 possible cross-links were identified within 1 minute and 32 seconds of the run. Major fragments identified by StavroX 3.6.0.1 are shown in Table 1. Thus, the Software StavroX 3.6.0.1 gave the Decoy analysis which shown in Figure 1, which shows the candidate scores along with the decoy data. This analysis shows the highest score to be 107, out of all the calculated possible theoretical candidates. The red bars represent the false positives obtained from the reverse FASTA sequence; the blue bars represent real positive peaks. The larger the number of blue bands more successful is the crosslinking. Figure 2 shows the annotation with the extent of deviation and the identified 'b' and 'y' ions. This analysis was repeated for each of other intermolecular crosslinked fragment identified by the software. The intermolecular crosslinked peaks decorated/ modified peaks are identified by the software are included in the (Supplementary Figures 5-11). The analysis of one peak m/z value is shown below. The crosslinked candidate detail spectrum (Figure 2) gives us details about the peptides that have been involved in the process of crosslinking and shows the annotation with the extent of deviation and the identified 'b' and 'y' ions. The details of modified fragments ion along with 'b' and 'y' ions of highest score 107 of peak m/z value 1904.915 is shown in Table 2. Peptide 1 contains "K13" sequence of "BELAAAMKR" and peptide 2 contains "L53" of "DYGILQ" sequence, with the major crosslinking site suggested as a link between K13 and L53. According to the results obtained, it implies that intermolecular cross-linking has occurred between K13 (Lysine) of alpha peptide from the sequence 10-20 of one Lysozyme and L53 (Leucine) of beta peptide from the sequence 50-60 of another moiety of Lysozyme. Some most intense inter and intra crosslinked peptides according to observed mass as well specified sequence identified by StavroX 3.6.0.1 are shown in Table 3. Previous crosslinking studies on lysozyme lead to the identification of very few crosslinks and that to just one intramolecularly cosslinked fragments. On the other hand, use of our new heterobifunctional crosslinkec using the two steps protocol has led to the identification of a very large number of both intra and inter molecular crosslinks emphasizing the efficient nature of our crosslinker.

Residue level structural insights as to how the crosslinker facilitates dimerization

The study aims at finding the cross-linking activity of compound III between two homogenous chains of lysozyme that involves K13 and Y53 residues of respective chains. The blind docking (BD) data suggests that major active site is found near the active cleft of lysozyme (Figure 1), which is well supported with Site Map results showing only one prominent site with considerable Site Score of 0.81 (SI-X). We picked the lowest energy conformers from all the 89 conformations out of 200 found at this site (Figure 1). In addition, the (groove-binding cavity) shows that compound III fits well at this site (Figure 1). There are 12 residues E35, N46, D52, Q57, I58, N59, W63, I98, A107, W108, V109 and L156 that are lining the cavity under 3.5 Å (Figure 2) and contains most of the residues recognized previously as important for hen-egg white lysozyme original ligand binding which approves our blind docking data [44-48]. However, our intention was to observe the crosslinking activity of III between residues K13 and Y53 and for that it was important to understand the dynamic nature of III at this site, a 100ns simulation was performed which justifies the stability of III at this pocket. However, it changes the orientation after first half (50 ns) of the simulation with respect to its initial position but at the point of convergence, it seems that III has gained its stable state at RMSD average value of ~6.0 Å (Figure 3A-C). Since, the pocket is not completely buried, this fluctuation from Docked to MD pose can be easily understood, and it is very important that III remained at this site. To understand the complete activity of III, a protein-protein setup was very important to observe how it is performing the crosslinking activity between K13 and Y53 in terms of interaction with both sets of proteins. Since it is documented that the crosslinker must hold both the interacting protein partners with significant binding affinity, therefore, we performed an APO-APO docking (establishment of dimer), and found that the residues (K13: chain B and Y53: chain A) that are important for crosslinking activity as shown by our previous study [49], are falling apart with 20.1 Å. From the distance analysis between two chains gave us information about the space localized at dimer state (APO-APO) and possibly this space might be occupied by the crosslinker (Figure 4A). At the next step, we performed COM-APO (COM is APO +III; dimer formation in presence of III) docking and we found that the distance of these two residues (K13: chain B and Y53: chain A) has been significantly reduced from 20.1 to 14.65 Å (Figure 4B). Since, III has a spacer length of 9.94 Å, it can bridge the distance of 14.65 Å between Cα-atom of K13: chain B and oxygen atom of Y53: chain A. It has well documented in previous studies that crosslinker with spacer length of 7.7 Å can bridge distance upto 25 Å [2,50-52] which clearly justifies our prediction of III functioning as a crosslinker which corroborates well with our experimental findings. Additionally, we found that in presence of III, chain B has flipped its orientation and fits more tightly (in terms of docking energy) than APO-APO state. The III at the junction of both the chains is found to be 6.2 Å away from Y53 (chain A) while from K13 (chain B) it is 8.0 Å (Figure 4B). In addition, the space in APO-APO is occupied well by III bringing closer both the chains in COM-APO and forming interactions with the key residues.

Intens.	rel. Intens.	m/z	calc.	Dev(pp…	Type	z	Peptide	Loss
1765.0	100.0	1904.924	1904.912	6.56	P0	1	α	0
1229.1	69.6	1615.839	1615.839	4.912	y7	1	α	0
991.2	56.2	1502.759	1502.759	2.72	y6	1	α	0
782.9	44.4	290.08	290.081	-0.75	b2	1	α	0
564.7	32	1413.723	1431.717	3.782	y5	1	α	0
291.7	16.5	403.166	403.165	2.278	b3	1	α	0
224.2	12.7	175.118	175.119	-6.578	y1	1	α	0
214.1	12.1	1887.896	1887.885	5.585	P0	1	α	17
202.6	11.5	1360.685	1360.68	3.034	y4	1	α	0
197.1	11.2	1744.878	1744.878	-1.923	y8	1	α	0

Table 1: Major fragments identified by StavroX 3.6.0.1.

B	b-H_2O	b-NH_3	AA	Y	Y-H_2O	Y-NH_3
Peptide: a						
Charge: +1						
161.038	143.027	144.011	B			
290.081	272.07	273.054	E	1744.881	1726.871	1727.855
403.165	385.154	386.138	L	1615.839	1529.828	1598.812
474.202	456.191	457.175	A	1502.755	1484.744	1485.728
545.239	527.228	528.212	A	1431.717	1413.707	1414.691
616.276	598.265	599.249	A	1360.68	1342.67	1343.654
747.316	729.306	730.29	M	1289.643	1271.633	1272.617
1730.8	1712.79	1713.774	K	1158.603	1140.592	1141.576
			R	175.119	157.108	158.092
Charge: +2						
81.023	72.017	72.509	B			
145.544	136.539	137.031	E	872.944	863.939	864.431
202.086	193.081	193.573	L	808.423	799.418	799.91
237.604	228.599	229.091	A	751.881	742.876	743.368
273.123	264.118	264.61	A	716.362	707.357	707.849
308.642	299.636	300.128	A	680.844	671.839	672.331
374.162	365.157	365.649	M	645.325	636.320	636.812
865.904	856.898	857.39	K	579.805	570.800	571.292
			R	88.063	79.058	79.55
Charge: +3						
54.351	48.347	48.675	B			
97.365	91.362	91.69	E	582.299	576.295	576.623
135.06	129.056	129.383	L	539.284	533.281	533.609
158.739	152.735	153.063	A	501.59	495.586	495.914
182.418	176.414	176.742	A	477.911	471.907	472.235
206.097	200.093	200.421	A	454.232	448.228	448.556
249.777	243.773	244.101	M	430.553	424.549	424.877
577.605	571.601	571.929	K	386.872	380.869	381.197
			R	59.045	53.041	53.369
Peptide: β						
Charge: +1						
116.034	98.024	99.008	D			
1475.653	1457.643	1458.627	Y	1789.885	1771.874	1772.858
1532.675	1514.664	1515.648	G	430.266	412.255	413.239
1645.759	1627.748	1628.732	I	373.245	355.234	356.218
1758.843	1740.832	1741.816	L	260.160	242.15	243.134
			Q	147.076	129.066	130.050
Charge: +2						
58.521	49.515	50.007	D			
738.33	729.325	729.817	Y	895.446	886.441	886.933
766.841	757.836	758.328	G	215.637	206.631	207.123
823.383	814.378	814.87	I	187.126	178.121	178.613
879.925	870.92	871.412	L	130.584	121.579	122.071
			Q	74.042	65.037	65.529
Charge: +3						
39.35	33.346	33.674	D			
492.556	486.552	486.88	Y	597.300	591.296	591.624
511.563	505.56	505.888	G	144.094	138.09	138.418
549.258	543.254	543.582	I	125.086	119.083	119.411
586.952	580.949	581.277	L	87.392	81.388	81.716
			Q	49.697	43.693	44.021

Table 2: Modified fragment ions obtained from the StavroX 3.6.0.1 analysis.

We also quantified our results based on docking energy after the APO-APO and COM-APO complexes were formed. A considerable ~7.0 kcal/mole energy difference was observed between two systems (SI-XI). After the manual introspection from the results obtained from Swarmdock and PIPER, we found around 100 complexes out of 250 generated from each system that are showing our proposed orientation. This new and small heterobifunctional crosslinker based on 3HAA is expected to be useful for other such similar studies. Being small it can help in identifing interfaces of protein-protein interaction. As it is a heterobifunctional crosslinker, it can be used in a two-step protocol involving an initial incubation followed by the subsequent step of photolysis. It is based on 3HAA, which itself is a part of the alternative kynurenine pathway of Tryptophan metabolism. Being a natural metabolite it could have implications in studies on cataractogenisis and for phototherapy of karatoconus, a disease of the human cornea (Figures 5 and 6).

Cross-linking reagent	Cross-linking product	Observed mass	Sequence (amino acids)
Intermolecular CXL-290	N-Terminal of K-Y Y 53-K 13	1904.915	10-20, 50-60 +XL
	K13-T43	1864.888	10-20, 30-50 +XL
	K13-S24	1655.859	10-20, 20-30 +XL
	K13-S86	1650.842	10-20, 80-90 +XL
Intramolecular CXL	K116-K33	2092.038	110-120, 30-40 +XL

Table 3: Inter and intra molecularly crosslinked peptides identified by StavroX.3.6.0.1.

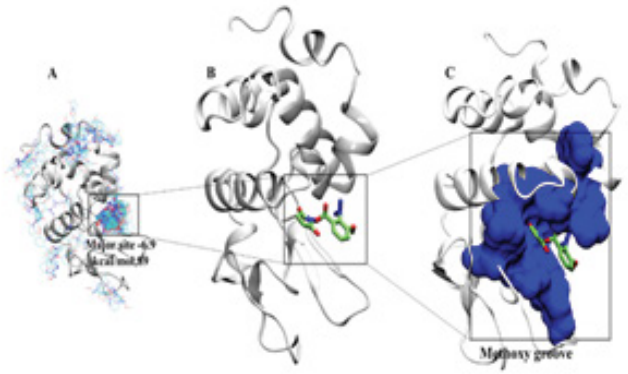

Figure 3: Blind docking on 5K7O: (A) Conformations distributed throughout protein surface and major cluster/site found. (B) Lowest energy conformer of III bound at active cleft (rendered in licorice and atom wise, C: green, N: blue, O: red. (c) The molecular surface of residues lining the binding site is highlighted in blue and III in green.

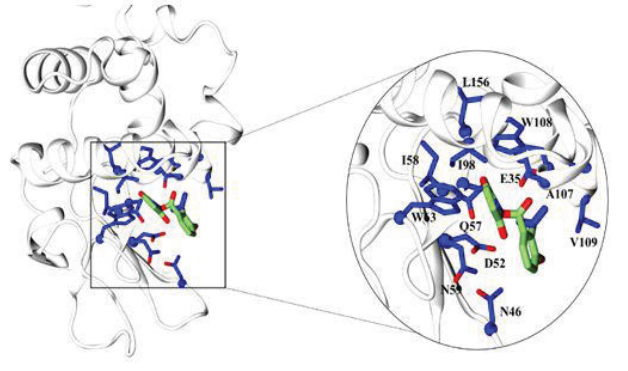

Figure 4: Binding mode and interaction map of III cross-linker.

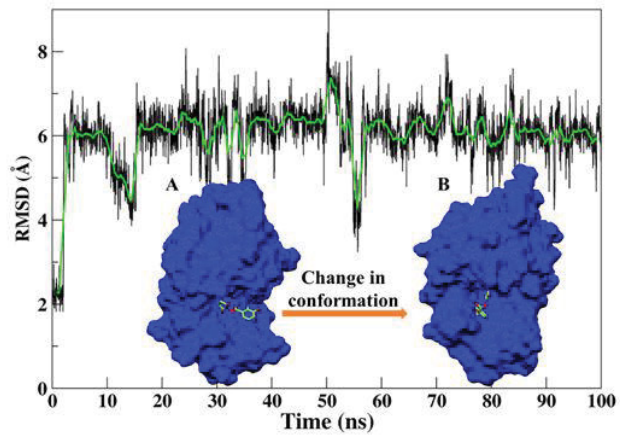

Figure 5: The monitoring of MD trajectory over 100 ns of COM. (A) III is stable in docked pose till first half (50 ns). (B) During converge of simulation, III shifts its position but remains stable in pocket.

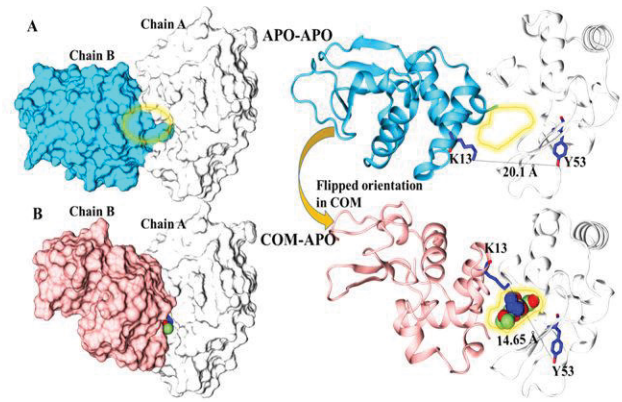

Figure 6: Protein-Protein docking in presence/absence of III. (A) Molecular surface and respective cartoon view of APO-APO docked model highlights the space (in yellow) between two chains (chain A: white and chain B: cyan) and 20.1 Å distance between K13 and Y53 (B) COM-APO docked model highlights that the space is occupied by III and chain B is found in different orientation (chain A: white and chain B: pink), the respective image shows that the distance between K13 and Y53 significantly reduces to 14.6 Å.

Conclusion

Proteomewise crosslinker designing is very challenging specially identifying crosslinkers of the right length and selectivity that could capture key interactions. The chemical crosslinkers are able to provide the identities of interactors, interacting residues and interacting interface site. Here, multi-faceted tools have been used to integrate and validate the data from chemical crosslinking-MS of lysozyme. Structural bioinformatics have been used extensively to explore the residue wise key information with and without small molecules, and StavroX 3.6.0.1 is particularly useful for identifying intermolecular crosslinking of proteins. Such intermolecular crosslinking has been carried out using a small heterobifunctional crosslinker [53,54] based on an essential amino acid Tryptophan a metabolite of the kynurenine pathway. This was done using the protocol Incubation, Photolysis, SDS-PAGE, excision of 'dimeric'band, trypsinization, ESI-MS, MS/MS. Work described here shows that identical results are obtained by the use of two different bioinformatics software and a more refined 3D

representation of the interfaces is obtained. This paves the way for the use of this analysis for understanding protein-protein interaction for designing of potential crosslinkers [55,56].

Acknowledgement

Authors thank Dr Sudhanshu Vrati, Executive Director, Regional Centre for Biotechnology (RCB) Faridabad, Haryana & Dr Gagandeep Kang, Executive Director, Translational Health Science & Technology Institute (THSTI), Faridabad, Haryana for research facilities and AcSIR for Emertius Professoship (Hony.) (to SVE). We thank Dr. Dinesh Mahajan, Drug Discovery Research Center (DDRC), Translational Health Science & Technology Institute (THSTI), Faridabad, and Haryana for the mass spectral studies. We thank the Advanced Technology Platform Center (ATPC, RCB) & Dr. Nirpender Singh for his advice and for supervising for ESI-MS/MS studies. We thank Dr Andrea Sinz, University of Halle-Salle, Germany for helpful discussions & Dr Michael Goetze from the same group for providing the STAVROX 3. 6.0.1 Software and advising us in its use. MS, RG and SA are thankful for THSTI intramural core grant for funding.

References

1. Rappsilber J (2011) The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. J Struct Biol 173: 530-545.
2. Sinz A (2014) The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks. Expert Rev Proteomics 11: 733-743.
3. Stengel F, Aebersold R, Robinson CV (2012) Joining Forces: Integrating Proteomics and Cross-linking with the Mass Spectrometry of Intact Complexes. Mol Cell Proteomics 11: R111.014027.
4. Hermanson GT (2008) Bioconjugate Techniques IInd Edn, Acadmic Press, New York.
5. Schmidt C, Urlaub H (2017) Combining cryo-electron microscopy (cryo-EM) and cross-linking mass spectrometry (CX-MS) for structural elucidation of large protein assemblies. Current opinion in Struc Biol 46: 157-168.
6. Ewens CA, Panico S, Kloppesteck P, McKeown C, Ebong IO, et al. (2014) The p97-FAF1 protein complex reveals a common mode of p97 adaptor binding. J Biol Chem 289: 12077-12084.
7. Leitner A, Faini M, Stengel F, Abersold R (2016) Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. Trends in Biochem Sci 41: 20-32.
8. Fischer L, Chen ZA, Rappsilber J, Cox J (2013) Quantitative cross-linking/mass spectrometry using isotope-labelled cross-linkers and MaxQuant. J Proteomics 88: 120-128.
9. Walzthoeni T, Joachimaik LA, Rosenberger G, Röst HL, Malmström L, et al. (2015) xTract: software for characterizing conformational changes of protein complexes by quantitative cross-linking mass spectrometry. Nature Methods 12: 1185-1190.
10. Russel D, Lasker K, Webb B, Tjioe E, Peterson B, et al. (2012) Putting the Pieces Together: Integrative Modeling Platform Software for Structure Determination of Macromolecular Assemblies. PLOS Biol 10: e1001244.
11. Dihazi GH, Sinz A (2003) Mapping low-resolution three-dimensional protein structures using chemical cross-linking and Fourier transform ion-cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrom 17: 2005-2014.
12. Back JW, Sanz MA, de Jong L, Grivell LA, Muijsers AO, et al. (2002) A structure for the yeast-prohibitin complex: Structure prediction and evidence from chemical crosslinking and mass spectrometry. Protein Science 11: 2471-2478.
13. Holding AN (2015) XL-MS: Protein cross-linking coupled with mass spectrometry. Methods 8: 54-63.
14. Ehresmann C, Moine H, Mougel M, Dondon J, Manago MG, et al. (1986) Cross-linking of initiation factor IF3 to Escherichia coli 30S ribosomal subunit by trans-dlamminedichloroplatinum (II): characterization of two cross-linking sites in 16S rRNA; a possible way of functioning for IF3. Nucleic Acids Res 14: 4803-4821.
15. Sutherland BW, Toews J, Kast J (2008) Utility of formaldehyde cross-linking and mass spectrometry in the study of protein–protein interactions. J Mass Spectrom 43: 699 -715.
16. Fritzsche R, Ihling Ch, Gotze M, Sinz A (2012) Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis. Rapid Commun Mass Spectrom 26: 653-658.
17. Kao A, Chiu CL, Vellucci D, Yang Y, Patel VR, et al. (2010) Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. Mol Cell Proteomics 10: M110.002212.
18. Cravatt BF, Gabriel MS, Yates III JR (2007) The biological impact of mass-spectrometry-based proteomics. Nature 450: 991-1000.
19. Goetze M, Pettelkau J, Fritzsche R, Ihling CH, Schafer M, et al. (2015) Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. J Am Soc Mass Spectrom 26: 83-97.
20. Soundarajan N, Liu SH, Soundarajan S, Platz MS (1993) Synthesis and binding of new polyfluorinated aryl azide to alpha chymotrypsin, New reagent for photoaffinity labelling. Boiconj Chem 4: 256-261.
21. Kaur D, Luk HL, Coldren W, Srinivas PM, Sridhar L, et al. (2014) Concomitant Nitrene and Carbene Insertion Accompanying Ring Expansion: Spectroscopic, X-ray and Computational Studies. J Org Chem 79: 1199-1205.
22. Xue J, Luk LH, Eswaran SV, Platz MS, Hadad CM, et al. (2012) Ultrafast Infrared and UV-vis Studies of the Photochemistry of 2-Methoxy-6-Methoxycarbonylphenyl Azide in Solution. J Phys Chem 116: 5325-5334.
23. Eswaran, SV, Kaur D, Jana K, Khamaru K, Prabhakar S, et al. (2017) Nitrene insertion into an adjacent o-methoxy group followed by nucleophilic addition to the heterocumulene intermediate: Experimental and computational studies. Tetrahedron 73: 5280-5288.
24. Goldstein LE, Leopold MC, Huang X, Atwood CS, Saunders AJ, et al. (2000) 3-Hydroxy kynurenine and 3-Hydroxy anthranilic acid generate hydrogen peroxide and permute alpha-Crystallin crosslinking by metal ion reduction. Biochemistry 39: 7266-7275.
25. Kannan R, Santoshkumar P, Mooney B, Sharma KK (2013) Identification of Subunit-Subunit Interaction Sites in αA-WT Crystallin and Mutant αA-G98R Crystallin Using Isotope-Labeled Cross-Linker and Mass Spectrometry. PLoS ONE 8: e65610.

26a. Lim CK, BruceJ, Sundaram G, Gullemin JU (2010) Understanding the rate of the kynurenine pathway in multiple scelerosis progression. Int J Tryptophan Res 3: 51-59.

26b. Moroni F (1999) Tryptophan metabolism and brain function: focus on kynurenine and other indole metabolites. Eur J Pharmacol 375: 87-100.

26c. Schwarcz R, Pellicciari R (2002) Manipulation of brain kynurenines: Glial target , neuronal effects and clinical opportunities . J Pharmacol Exp Therap 303: 1-10.

27a. Darlington LG, Forrest CM, Mackay GM, Smith RA, Smith AJ, et al.(2010) On the Biological importance of the 3- Hydroxyanthranilic acid :Anthranilic acid ratio. Int J Tryptophan Res 3: 51-59.

27b. Kaur D (2013) Unusual products based on Aryl Nitrenes and Cateractogenesis, Ph. D. thesis, University of Delhi, India.

28. Goel R, Murthy KR, Srikanth SM, Pinto SM, Bhattacharjee M, et al. (2013) Characterizing the normal proteome of human ciliary body. Cilin Proteomics 10: 9.

29. de la Cruz MJ, Hattne J, Shi D, Seidler P, Rodriguez J, et al. (2017) Atomic-resolution structures from fragmented protein crystals with the cryoEM method MicroED. Nat Methods 14: 399-402.

30. Maestro, version (2017) Schrödinger, LLC, New York.

31. LigPrep, version 4.2 (2017) Schrödinger, LLC, New York.

32. Epik Version 4.0 (2017): Schrödinger, LLC, New York, NY.

33. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. J Am Chem Soc 118: 11225-11236.

34. Srivastava M, Suri C, Singh M, Mathur R, Asthana S, et al. (2018) Molecular dynamics simulation reveals the possible druggable hotspots of USP7. Oncotarget 9.

35. Kanwal A, Kasetti S, Putcha UK, Asthana S, Banerjee SK, et al. (2016) Protein kinase C-mediated sodium glucose transporter 1 activation in precondition-induced cardioprotection. Drug Des Devel Ther 10: 2929-2938.

36. Asthana S, Shukla S, Vargiu AV, Ceccarelli M, Ruggerone P, et al. (2013) Different molecular mechanisms of inhibition of bovine viral diarrhea virus and hepatitis C virus RNA-dependent RNA polymerases by a novel benzimidazole. Biochemistry 52: 3752-3764.

37. Asthana S, Shukla S, Ruggerone P, Vargiu AV (2014) Molecular mechanism of viral resistance to a potent non-nucleoside inhibitor unveiled by molecular simulations. Biochemistry 53: 6941-6953.

38. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, et al. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem 30: 2785-2791.

39. SiteMap, Version 4.2 (2017) Schrödinger, LLC, New York, NY.

40. Desmond Molecular Dynamics System, Version 2.3 (2017) D. E. Shaw Research, New York, NY and Maestro-Desmond Interoperability Tools, Schrödinger (2017) New York, NY.

41. Mattapally S, Singh M, Murthy KS, Asthana S, Banerjee SK, et al. (2018) Computational modeling suggests impaired interactions between NKX2.5 and GATA4 in individuals carrying a novel pathogenic D16N NKX2.5 mutation. Oncotarget 9: 13713-13732.

42. Kozakov D, Brenke R, Comeau SR, Vajda S (2006) PIPER: An FFT-based protein docking program with pairwise potential. Proteins 65: 392-406.

43. Torchala M, Moal IH, Chaleil RAG, Fernandez-Recio J, Bates PA, et al. (2013) SwarmDock: a server for flexible protein-protein docking. Bioinformatics 29: 807-809.

44. Yang S-T, Wang H, Guo L, Gao Y, Liu Y, et al. (2008) Interaction of fullerenol with lysozyme investigated by experimental and computational approaches. Nanotechnology 19: 395101.

45. Strynadka NCJ, James MNG (1991) Lysozyme revisited: Crystallographic evidence for distortion of an N-acetylmuramic acid residue bound in site D. J Mol Biol 220: 401-424.

46. Vocadlo DJ, Davies GJ, Laine R, Withers SG (2001) Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. Nature 412: 835-838.

47. Harata K, Muraki M, Jigami Y (1993) Role of Arg115 in the catalytic action of human lysozyme. J Mol Biol 233: 524-535.

48. Calvaresi M, Bottoni A, Zerbetto F (2015) Thermodynamics of binding between proteins and carbon nanoparticles: the case of C60 @ lysozyme. J Phys Chem 119: 28077-28082.

49. Thakur S, Eswaran SV (2017) A New Heterobifunctional Crosslinker Based on an "Introverted" Acid: Mass Spectrometric and Bioinformatics Studies, Analysis of Intermolecular Crosslinking of Proteins. J Anal Bioanal Tech 8.

50. Chen ZA, Jawhari A, Fischer L, Buchen C, Tahir S, et al. (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. EMBO J 29: 717-726.

51. Dimova K, Kalkhof S, Pottratz I, Ihling C, Liepold T, et al. (2009) Structural insights into the calmodulin-Munc13 interaction obtained by cross-linking and mass spectrometry. Biochemistry 48: 5908-5921.

52. Green NS, Reisler E, Houk KN (2001) Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers. Protein Sci 10: 1293-1304.

53. Thakur KS, Eswaran SV (2018) ESI-MS and Stavrox 3.6.0.1 Investigations of Crosslinking by an Aryl-Azido-NHS Heterobifunctional Crosslinker. J Anal Bioanal Tech 9: 2.

54. Thakur KS, Pal S, Kumar A, Goel R, Eswaran SV, et al. (2018) ESI-MS-Bioinformatics Studies on crosslinking of αA-Crystallin and Lysozyme Using a new Aryl Azido- N-Hydroxy Succinimidyl Heterobifuntional crosslinker based on a metabolite of the alternative kynurenine Pathway. J Proteomics Bioinform 11: 192-200.

55. Dihazi GH, Sinz A (2003) Mapping low –resolution three dimensional protein structures using chemical crosslinking and furrier-transform ion-cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrum 17: 2005-2014.

56. Crosslinker for keratoconus. https://youtu.be/sQHfDWgEhQs