RESEARCH ARTICLE

Binding Pattern Elucidation of NNK and NNAL Cigarette Smoke Carcinogens with NER Pathway Enzymes: an Onco-Informatics Study

Qazi Mohammad Sajid Jamal[1], Anupam Dhasmana[2], Mohtashim Lohani[2]*, Sumbul Firdaus[2], Md Yousuf Ansari[3,4], Ganesh Chandra Sahoo[3,4], Shafiul Haque[5]

Abstract

Cigarette smoke derivatives like NNK (4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone) and NNAL (4-(methylnitrosamino)-1-(3-pyridyl)-1-butan-1-ol) are well-known carcinogens. We analyzed the interaction of enzymes involved in the NER (nucleotide excision repair) pathway with ligands (NNK and NNAL). Binding was characterized for the enzymes sharing equivalent or better interaction as compared to +Ve control. The highest obtained docking energy between NNK and enzymes RAD23A, CCNH, CDK7, and CETN2 were -7.13 kcal/mol, -7.27 kcal/mol, -8.05 kcal/mol and -7.58 kcal/mol respectively. Similarly the highest obtained docking energy between NNAL and enzymes RAD23A, CCNH, CDK7, and CETN2 were -7.46 kcal/mol, -7.94 kcal/mol, -7.83 kcal/mol and -7.67 kcal/mol respectively. In order to find out the effect of NNK and NNAL on enzymes involved in the NER pathway applying protein-protein interaction and protein-complex (i.e. enzymes docked with NNK/NNAL) interaction analysis. It was found that carcinogens are well capable to reduce the normal functioning of genes like RAD23A (HR23A), CCNH, CDK7 and CETN2. *In silico* analysis indicated loss of functions of these genes and their corresponding enzymes, which possibly might be a cause for alteration of DNA repair pathways leading to damage buildup and finally contributing to cancer formation.

Keywords: NER pathway enzymes - NNK - NNAL - cigarette smoke carcinogens - docking

Asian Pac J Cancer Prev, **16 (13)**, 5311-5317

Introduction

Lung cancers are powerfully linked with cigarette smoke carcinogens like NNK (4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone) and NNAL (4-(methylnitrosamino)-1-(3-pyridyl)-1-butan-1-ol) (Xue et al., 2014). One of the preliminary critical actions is most likely damage of the hereditary material (DNA) by a cigarette smoke carcinogen. This damage can, beneath the definite status, be repaired by cellular DNA repair mechanisms (Raphael Ceccaldi et al., 2015). Though, if not repaired, cells will try to duplicate their DNA during cell division, but are obstructed by the damage and will do fault duplication progression leading to gene mutations brought onto a trail of uncontrolled cell division leading to a tumor growth. Studies show that in ordinary cells, NER removes numerous types of DNA lesions, defending cell integrity (Rouillon et al., 2011).

However, in cancer cells uncovered to DNA damaging compounds that alter the DNA helix or form unwieldy injuries to the genome, NER take part in the managing the damage, consequently protecting cancer cells from fatality (Nouspikel, 2009). But NNK (4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone) and NNAL (4-(methylnitrosamino)-1-(3-pyridyl)-1-butan-1-ol) can alter the biological activity of NER repair enzymes. Therefore, In order to execute our hypothesis, we have selected 17 enzymes involved in NER pathways and their interaction with cigarette smoke carcinogens NNK (4-(Methylnitrosamino) -1-(3-pyridyl) -1-butanone) and NNAL (4-(methylnitrosamino) -1-(3-pyridyl) -1-Butan-1-ol). Molecular docking analyses were performed using Autodock 4.2 tools. On the basis of obtaining top four docking energies we have selected RAD23A, CCNH, CDK7, and CETN2 genes for the further analysis to know the effect of NNK/NNAL on the corresponding enzymes function. The normal functioning of the selected enzymes describes that RAD23A (UV excision repair protein

*[1]Department of Health Information Management, College of Applied Medical Sciences, Buraydah Colleges, Al-Qassim-Buraydah King Abdulaziz Road, East Qassim University, Buraydah, [5]Research and Scientific Studies Unit, College of Nursing and Allied Health Sciences, Jazan University, Jazan, Saudi Arabia, [2]Environmental Carcinogenesis & Toxico-informatics Laboratory, Department of Biosciences and Bioengineering, Integral University , Uttar Pradesh, [3]Pharmaco-informatics Department, National Institute of Pharmaceutical Education and Research (NIPER), [4]BioMedical Informatics Centre, Rajendra Memorial Research Institute of Medical Sciences (RMRIMS) Agam Kuan,Bihar, India. *For correspondence: mlohani@iul.ac.in*

RAD23 homolog A) involved in nucleotide excision repair and is thought to be functionally equivalent for RAD23B in global genome nucleotide excision repair (GG-NER) by association with XPC. Two human homologs of Rad23 are functionally interchangeable in complex formation and stimulation of XPC repair activity (Sugasawa et al., 1997).

CCNH (cyclin H) regulates CDK7, the catalytic subunit of the CDK-activating kinase (CAK) enzymatic complex. CAK activates the cyclin-associated kinases CDK1, CDK2, CDK4 and CDK6 by threonine phosphorylation. CDK7 (Cyclin-dependent kinase 7) required for DNA-bound peptides-mediated transcription and cellular growth inhibition DNA-Bound peptides control the mRNA transcription through CDK7 (Lu X et al., 2009) and CETN2 (Centrin-2) Involved in global genome nucleotide excision repair (GG-NER) by acting as a component of the XPC complex. Cooperatively with RAD23B appears to stabilize XPC. Centrosome protein centrin2/caltractin1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair (Araki M et al., 2001;S.Matsumoto et al.,2015).

Computational tools such as molecular docking are important to understand the binding capabilities of NNK and NNAL with enzymes involved in NER pathways (Xia et al., 2012). It has never been explored through in silico approaches. Therefore, we used protein-protein docking to know the functional loss of the enzymes due to their interaction with NNK and NNAL. In order to perform protein-protein interaction it is necessary to find out the co-operated functional enzymes encoded by genes using STRING 9.0.5 database (Szklarczyk et al., 2011) and their 3D structures. At last the comparative analysis (Protein-Protein docking *vs* Protein- Complex*) was completed by ZDOCK protocol using Discovery Studio Client 2.5 (Accelrys Software Inc, 2013).

Materials and Methods

Preparation of ligand structures

Ligand file of NNK (4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone) and NNAL (4-(methylnitrosamino)-1-(3-pyridyl)-1-butan-1-ol) were downloaded in .mol format (Figure: 1 and Figure: 2) from ChemSpider Chemical Database (Harry et al., 2010). These files could not directly use by Autodock 4.2 tools (Morris et al., 2009) thus; we have to convert it into .pdb files and also further the ligands were submitted for CHARMm (Brooks et al., 1983; 2009) energy minimization protocol in Discovery Studio Client 2.5.

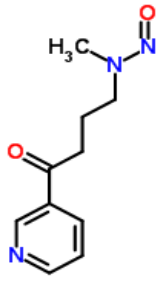

Figure 1. 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone, PubChem Compound ID- 47289, ChemSpider ID 43038

Preparation of protein structures

The structures of enzymes involved in the NER pathways were obtained from Protein Data Bank (Berman et al., 2000) (Table 1). Published structures were edited to remove HETATM and water molecule using Discovery Studio Client 2.5. Energy minimization was performed by the implementation of CHARMm force field (D.T Mirijanian et al., 2014) after addition of hydrogen atoms to the selected enzymes using Accelyrs Discovery studio client 2.5.

Docking studies

Molecular Docking studies were performed to analyze the binding affinity of NNK/NNAL with enzymes involved NER pathways. Autodock (Version 4.2) suite (Morris et al., 1998; 2009) and Cygwin interface was used in the Microsoft Windows 7 professional, operating System on Intel® Xeon® Processor E3-1220 v3 (Quad Core, 3.10GHz Turbo, 8MB) and 256GB 2.5inch Serial ATA Solid State Drive of *Dell Precision T1700* Workstation was used to dock the NNK/NNAL on binding site of the enzymes. Molecular docking methods followed by the searching the best conformation of enzymes and carcinogens complex on the basis of binding energy. Water molecules were removed from the 3D X-ray crystallography structures of enzymes before docking and hydrogen atoms were added to all target enzymes. Kollman united charges and salvation parameters were added to the enzymes. Gasteiger charge was added to the ligands. Grid box was set to cover the maximum part of enzymes and ligands. The values were set to 60×60×60 Å in X, Y and Z axis of a grid point. The default grid points spacing was 0.375 Å. Lamarckian Genetic Algorithm (LGA) (Goodsell et al., 1996; Tsai et al., 2012) was used for enzymes-ligands flexible docking calculations. The LGA parameters like population size (ga_pop_size), energy evaluations (ga_num_generation), mutation rate, crossover rate and step size were set to 150, 2500000, 27000, 0.02, 0.8 and 0.2 Å, respectively. The LGA runs were set at 50 runs. All obtained conformations of enzymes and ligand complex were analyzed the interactions and binding energy of the docked structure using Discovery Studio 2.5 molecular visualization software.

Protein-protein interaction analysis

We found the interacting proteins (used as ligands) of selected enzymes using STRING 9.0.5 database that predict interacting interactions incorporate direct (physical) and indirect (functional) associations derived

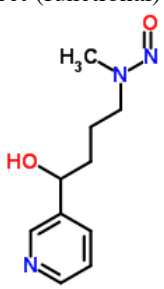

Figure 2. 4-(methylnitrosamino)-1-(3-pyridyl)-1-butan-1-ol, PubChem Compound ID- 104856, ChemSpider ID- 94646dd

from four sources, i.e. genomic context, high throughput experiments (conserved) co-expression and previous knowledge of proteins against your query (Figure:5 A,B,C and D). We used discovery studio Client 2.5 for Zdock (Dock Proteins) Protocol. Zdock scores obtained for both Protein-Protein interactions as well as for Protein-Complex (ligand protein+NNK/NNAL) interaction.

Z dock calculations

Discovery studio Client 2.5 was used to complete protein-protein docking using ZDOCK is an initial stage, rigid body molecular docking algorithm that uses a fast Fourier transform (FFT) algorithm to improve performance for searching in translational space (Chen et al., 2003; Pierce et al., 2014). All of the available structures from PDB were used to calculate the docking poses and the structures obtained were subjected to energy minimization using the smart minimize algorithm (Max steps 200, RMS gradient 0.01) in the program Discovery studio 2.5. The resulting Zdock scores with the highest value were used as appropriate conformational pose (Jamal et al., 2012).

Results and Discussion

In the achievement of the current investigation molecular docking techniques were adopted to explore the binding capabilities of NNK and NNAL with enzymes encoded by respective genes of NER pathways. Primarily the 1IRD (Crystal Structure of Human Carbonmonoxy-Haemoglobin at 1.25 Å Resolution) was used as a positive control and 3CI9 (Human heat shock factor-binding protein 1) as a negative control to validate our docking analysis. Molecular interaction results of these enzymes showed that 1IRD docked with NNK, observed binding energy was -6.68 Kcal/Mol, it docked with NNAL and observed binding energy was -6.31 Kcal/Mol. 3CI9 docked with NNK with the experimental binding energy of -3.91 Kcal/Mol, it interacted with NNAL with binding energy of +2.09 Kcal/Mol. We performed docking analysis between 16 enzymes and NNK/NNAL separately. The observed docking energy between NER pathway enzymes and NNK were ranging from -4.28 kcal/mol to -8.05 kcal/mol (Table 1) similarly between NER pathways enzymes and NNK were ranging from -5.33 kcal/mol to -7.94 kcal/mol. In the completion of next step of our hypothesis, we selected top four NER enzymes encoded by respective genes from Table 1 and Table 2 on the basis of their highest obtained docking energy between NNK and enzymes RAD23A, CCNH, CDK7, and CETN2 were -7.13 kcal/mol, -7.27 kcal/mol, -8.05 kcal/mol and -7.58 kcal/mol respectively. Similarly the highest obtained docking energy between NNAL and enzymes RAD23A, CCNH, CDK7, and CETN2 are -7.46 kcal/mol, -7.94 kcal/mol, -7.83 kcal/mol and -7.67 kcal/mol respectively (Table 1).

The active site characterization analysis of top four enzymes revealed that NNK and NER enzymes shown that RAD23A involved in the building of 2 hydrogen bonds A: GLN21:HE21 - :UNK0:N4 and A: GLN21:HE21 - :UNK0:O2 with the distance of 2.40437 Å and 1.8656 Å respectively with NNK. The hydrophobic pocket characterized by the occurrence of Phe14, Ser17, Leu18, Gln21, Ala22, Phe36, Leu37, Leu38, Gln40, Asn41, Phe42, Asp43 amino acid residues. The estimated inhibition constant of NNK and RAD23A docked complex was 112.03 (Table 2 Figure: 3 A). The CCNH involved in the building of 2 hydrogen bonds A: ASP202: HN - :UNK0:N4 and A: ASP202: HN - :UNK0:O2 with the distance of 2.32473 Å and 1.94374 Å, respectively with NNK. The hydrophobic pocket characterized by the occurrence of Arg23, Met54, Cys57, Lys58, Glu61, Phe87, Lys88, Tyr91, Leu200, Thr201, Asp202, Leu205, Leu258, Lys261, and Tyr262 amino acids residues. The estimated inhibition constant of NNK and CCNH docked complex was 91.24 uM (Table 2 Figure: 3 B). The CDK7 involved in building of 5 hydrogen bonds A: LYS139:HZ3 - :UNK0:O1, A: ASN142:HD22 - :UNK0:O1, A: SER161: HN - :UNK0:O2, A: SER161: HG - :UNK0:N4 and A: SER161: HG - :UNK0:O2 with distance of 1.67265 Å, 2.2816 Å, 2.05534 Å, 1.8268 Å and 2.24771 Å respectively.The hydrophobic pocket characterized by the occurrence of Gly21, Gln22, Phe23, Lys41, His135, Asp137, Lys139, Asn142, Ala154, Asp155, Phe156,

Table 1. NNK and NNAL Docked with EnzymesInvolved in NER Pathway

S. No.	Gene's Name	PDB ID	Accession Code	GenBank	Uniprot	Docking Energy (vdW + Hbond + desolv Energy)	
						Docked with NNK	Docked with NNAL
1.	RPA2	1DPU	NM_002946	NP_002937	P15927	-6.53 kcal/mol	-6.19 kcal/mol
2.	RAD23A (HR23A)	1DV0	NM_005053	NP_005044	P54725	-7.13 kcal/mol	-7.46 kcal/mol
3.	MT1	1G25	NM_002431	NP_002422	P51948	-6.06 kcal/mol	-5.98 kcal/mol
4.	CCNH	1KXU	NM_001239	NP_001230	P51946	-7.27 kcal/mol	-7.94 kcal/mol
5.	RAD23B (HR23B)	1P1A	NM_002874	NP_002865	P54727	-6.46 kcal/mol	-7.01 kcal/mol
6.	CDK7	1UA2	NM_001799	NP_001790	P50613	-8.05 kcal/mol	-7.83 kcal/mol
7.	LIG1	1X9N	NM_000234	NP_000225	P18858	-6.88 kcal/mol	-7.46 kcal/mol
8.	XPA	1XPA	NM_000380	NP_000371	P23025	-6.99 kcal/mol	-6.82 kcal/mol
9.	GTF2H2	1Z60	NM_001515	NP_001506	Q13888	-6.74 kcal/mol	-7.43 kcal/mol
10.	CETN2	1ZMZ	NM_004344	NP_004335	P41208	-7.58 kcal/mol	-7.67 kcal/mol
11.	ERCC1	2A1I	NM_001983	NP_973730	Q7Z7F5	-6.74 kcal/mol	-6.54 kcal/mol
12.	ERCC4 (XPF)	2AQ0	NM_005236	NP_005227	Q92889	-6.41 kcal/mol	-6.23 kcal/mol
13.	GTF2H5 (TTDA)	2JNJ	NM_207118	NP_997001	Q6ZYL4	-6.27 kcal/mol	-5.33 kcal/mol
14.	XPC	2OBH	NM_000380	NP_000371	P23025	-6.25 kcal/mol	-6.08 kcal/mol
15.	GTF2H1	2RNR	NM_005316	NP_005307	P32780	-6.46 kcal/mol	-5.83 kcal/mol
16.	ERCC3 (XPB)	4ERN	NM_000122	NP_000113	P19447	-4.28 kcal/mol	-5.56 kcal/mol

Table 2. Docking Studies of NNK and NER Pathways Enzymes Interaction

S.No.	Enzymes	PDB ID	H-Bonds	H-Bonds Distance (Å)	Residues involved in Hydrophobic region	Docking Energy (vdW + Hbond + desolv Energy) kcal/mol	Inhibition Constant (uM)
1.	RAD23A (HR23A)	1DV0	A:GLN21:HE21 - :UNK0:N4	2.40437		-7.13	112.03
			A:GLN21:HE21 - :UNK0:O2	1.8656	Phe14,Ser17,Leu18,Gln21,Ala22, Phe36,Leu37,Leu38,Gln40,Asn41, Phe42,Asp43		
2.	CCNH	1KXU	A:ASP202:HN - :UNK0:N4	2.32473		-7.27	91.24
				31	Arg23,Met54,Cys57,Lys58,Glu61 ,Phe87,Lys88,Tyr91,Leu200, Thr 201,Asp202,Leu205,Leu258,Lys2 61,Tyr262		
			A:ASP202:HN - :UNK0:O2	1.94374			
3.	CDK7	1UA2	A:LYS139:HZ3 - :UNK0:O1	1.67265		-8.05	26.15
			A:ASN142:HD22 - :UNK0:O1	2.26816	Gly21,Gln22,Phe23,Lys41,His135 ,Asp137,Lys139,Asn142,Ala154,A sp155,Phe156,Gly157,Lys160,Ser 161,Phe162,Thr175		
			A:SER161:HN - :UNK0:O2	2.05534			
			A:SER161:HG - :UNK0:N4	1.8268			
			A:SER161:HG - :UNK0:O2	2.24771			
4.	CETN2	1ZMZ	A:ARG18:HN - :UNK0:N4	2.231		-7.58	49.61
			A:ARG18:HN - :UNK0:O2	1.7428	Gln15,Arg16,Lys17,Arg18,Met19, Leu25,Gln29,Lys30,Gln31,Ile33,A rg34,Phe86,Leu90		
			A:MET19:HN - :UNK0:O2	1.87497			
			A:LYS30:HZ2 - :UNK0:O1	2.09563			

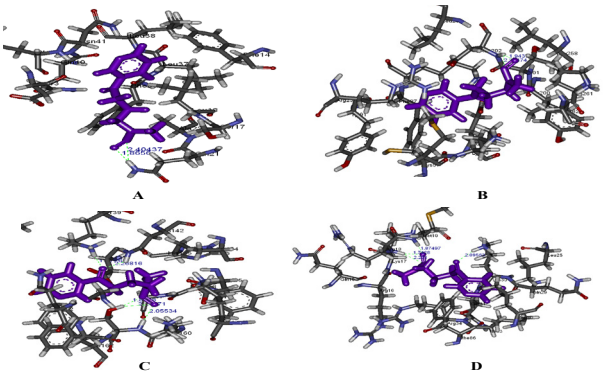

Figure 3. (A) 1DV0 (RAD23A (HR23A)) Docked with NNK (in Purple color) (B) 1KXU (CCNH) Docked with NNK (C) 1UA2 (CDK7) Docked with NNK (D) 1ZMZ (CETN2) Docked with NNK (in Purple color) and the Hydrogen Bonds Shown by Green Dotted Lines. All graphics generated by discovery studio visualizer

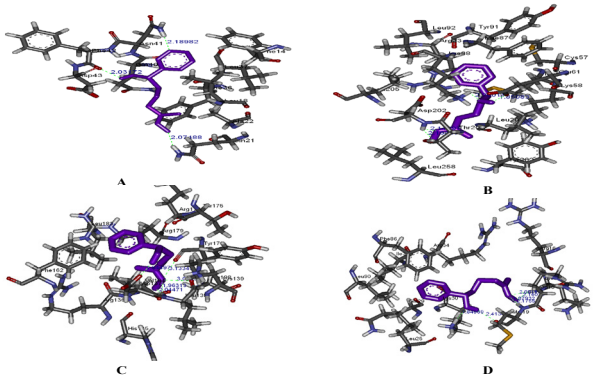

Figure 4. (A) 1DV0 (RAD23A (HR23A)) Docked with NNAL (in Purple color) (B) 1KXU (CCNH) Docked with NNAL C 1UA2 (CDK7) Docked with NNAL (D) 1ZMZ (CETN2) Docked with NNAL (in Purple color) and Hydrogen Bonds Shown by Green Dotted Lines. All graphics generated by discovery studio visualizer

Gly157, Lys160, Ser161, Phe162, and Thr175 amino acids residues. The estimated inhibition constant of NNK and CDK7 docked complex was 26.15 uM (Table 2 Figure:3 C). The CETN2 involved in the building of 4 hydrogen bonds A: ARG18: HN - :UNK0:N4, A: ARG18: HN - :UNK0:O2, A: MET19: HN - :UNK0:O2, and A: LYS30:HZ2 - :UNK0:O1 with the distance of 2.231 Å, 1.7428 Å, 1.87497 Å and 2.09563 Å respectively. The hydrophobic pocket characterized by the occurrence of Gln15, Arg16, Lys17, Arg18, Met19, Leu25, Gln29, Lys30, Gln31, Ile33, Arg34, Phe86, and Leu90 amino acids residues. The estimated inhibition constant of NNK and CETN2 docked complex was 46.61 uM (Table 2 Figure: 3 D).

Furthermore, the active site characterization analysis of top four enzymes also revealed that NNAL and NER enzymes shown that RAD23A involved in the building of 3 hydrogen bonds A: GLN21:HE21 - :UNK0:O2, A: ASN41:HD21 - :UNK0:N4, and: UNK0:H24 - A: PHE42: O with the distance of 2.26011 Å, 2.18982 Å, and 2.03172 Å respectively. The hydrophobic pocket characterized by the occurrence of Phe14, Leu18, Gln21, Ala22, Phe36, Leu37, Gln40, Asn41, Phe42, and Asp43 amino acid residues. The estimated inhibition constant of NNAL and RAD23A docked complex was 67.58 uM (Table 3 Figure: 4 A). The CCNH involved in the building of 4 hydrogen bonds A: LYS88:HZ2 - :UNK0:O1, A: ASP202: HN - :UNK0:N5, A: ASP202: HN - :UNK0:O2, and: UNK0:H24 - A: GLU61:OE1 with the distance of 2.26011 Å, 2.14097 Å, 2.16487 Å and 1.85083 Å, respectively. The hydrophobic pocket characterized by the occurrence of Arg23, Met54, Cys57, Lys58, Glu61,

Table 3. Docking Studies of NNAL and NER Pathways Enzymes Interaction

S.No.	Enzymes	PDB ID	H-Bonds	H-Bonds Distance (Å)	Residues involved in Hydrophobic region	Docking Energy (vdW + Hbond + desolv Energy) kcal/mol	Inhibition Constant (uM)
1.	RAD23A	1DV0	A:GLN21:HE21 - :UNK0:O2	2.07489	Phe14,Leu18,Gln21,Ala22,Phe36,	-7.46 kcal/mol	67.58 uM
			A:ASN41:HD21 - :UNK0:N4	2.18982	Leu37,Gln40,Asn41,Phe42,Asp43		
			: UNK0:H24 - A: PHE42: O	2.03172			
2.	CCNH	1KXU	A:LYS88:HZ2 - :UNK0:O1	2.26011	Arg23,Met54,Cys57,Lys58,Glu61,	-7.94 kcal/mol	29.57 uM
			A:ASP202:HN - :UNK0:N5	2.14097	Phe87,Lys88,Tyr91,Leu92,Leu200,		
			A:ASP202:HN - :UNK0:O2	2.16487	Thr201,Asp202		
			:UNK0:H24 - A:GLU61:OE1	1.85083	Leu205,Leu258,Tyr262		
3.	CDK7	1UA2	A:LEU138:HN - :UNK0:N5	1.96313	His135,Arg136,Asp137,Leu138,Lys139,	-7.83 kcal/mol	39.55 uM
			A:LEU138:HN - :UNK0:O2	2.04471	Phe162,Thr175,Arg176,Tyr178,Arg179		
			:UNK0:H24 - A:ASP137:OD1	1.84497	Leu183,Val194,Ala198		
			:UNK0:N5 - A:LEU138:O	3.00052			
			:UNK0:N5 - A:TYR178:O	3.13345			
4.	CETN2	1ZMZ	A:ARG18:HN - :UNK0:N5	2.0811	Arg16,Lys17, Arg18,Met19,Leu25	-7.67 kcal/mol	43.67 uM
			A:ARG18:HN - :UNK0:O2	2.14078	Gln29,Lys30,Ile33,Arg34,Phe86,		
			A:MET19:HN - :UNK0:N5	1.87932	Leu90		
			A:MET19:HN - :UNK0:O2	2.17551			
			A:LYS30:HZ2 - :UNK0:O1	1.84969			
			:UNK0:H24 - A:MET19:O	2.41349			

Table 4. NNK and NNAL Binding to the Enzymes Reduces their Normal Functions after Analyzing the ZDOCK Scores

S. No.	Selected Genes	PDB ID	Interacted enzymes	PP interaction of NNK with enzymes	ZDOCK Score	PP interaction of NNAL with enzymes	ZDOCK Score
			(Obtained from STRING 9.0.5)		NNK		NNAL
1.	RAD23A (HR23A)	1DV0	2KDE	1DV0 *vs* 2KDE	14.58	1DV0 *vs* 2KDE	14.58
				1DV0+NNK *vs* 2DKE	13.74	1DV0+NNAL *vs* 2DKE	13.76
2.	CCNH	1KXU	1UA2	1KXU *vs* 1UA2	14.24	1KXU *vs* 1UA2	14.24
				1KXU+NNK *vs* 1UA2	13.08	1KXU+NNAL *vs* 1UA2	13.84
3.	CDK7	1UA2	1KXU	1UA2 *vs* 1KXU	14.96	1UA2 *vs* 1KXU	14.96
				1UA2+NNK *vs* 1KXU	13.08	1UA2+NNAL *vs* 1KXU	13.68
4.	CETN2		2GGM	1ZMZ *vs* 2GGM	15.92	1ZMZ *vs* 2GGM	15.92
		1ZMZ		1ZMZ+NNK *vs* 2GMM	15.4	1ZMZ+NNAL *vs* 2GMM	15.55

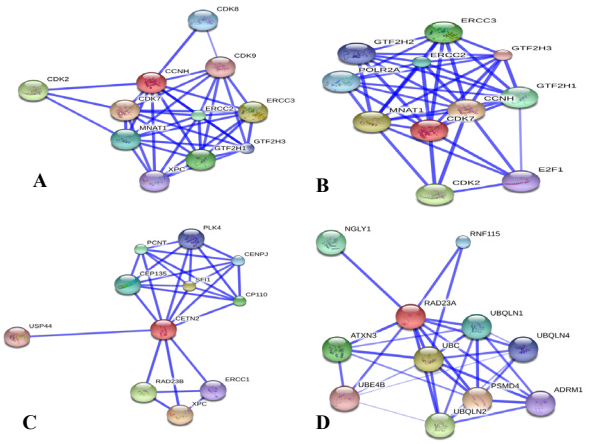

Figure 5. (A) CCNH Interacted with CDK7 (B) CDK7 Interacted with CCCNH (C) CETN2 Interacted with XPC (D) RAD23A Interacted with PSMD4. All interaction networks of selected enzymes obtained from STRING database

Phe87, Lys88, Tyr91, Leu92, Leu200, Thr201, Asp202, Leu205, Leu258, and Tyr262. The estimated inhibition constant of NNAL and CCNH docked complex was 29.57 uM (Table 3 Figure: 4 B). The CDK7 involved in the building of 5 hydrogen bonds A:LEU138:HN - :UNK0:N5, A:LEU138:HN - :UNK0:O2, :UNK0:H24 - A:ASP137:OD1, :UNK0:N5 - A:LEU138:O, and :UNK0:N5 - A:TYR178:O with the distance of 1.96313 Å, 2.04471 Å, 1.84497 Å, 3.00052 Å, and 3.13345 Å respectively. The hydrophobic pocket characterized by the occurrence of His135, Arg136, Asp137, Leu138, Lys139, Phe162, Thr175, Arg176, Tyr178, Arg179, Leu183, Val194, and Ala198 amino acid residues. The estimated inhibition constant of NNAL and CDK7 docked complex was 39.55 uM (Table 3 Figure: 4 C). CETN2 involved in the building of 6 hydrogen bonds A:ARG18:HN - :UNK0:N5, A:ARG18:HN - :UNK0:O2, A:MET19:HN - :UNK0:N5, A:MET19:HN - :UNK0:O2, A:LYS30:HZ2 - :UNK0:O1, and :UNK0:H24 - A:MET19:O with the distance of 2.0811 Å, 2.14078 Å, 1.87932 Å, 2.17551 Å, 1.84969 Å, and 2.41349 Å respectively. the hydrophobic pocket characterized by the occurrence of Arg16, Lys17, Arg18, Met19, Leu25, Gln29, Lys30, Ile33, Arg34, Phe86, and Leu90 amino acids residues. The estimated inhibition constant of NNAL and CETN2 complex was 43.67 uM (Table 3 Figure: 4 D).

In the further analysis, the protein-protein docking was adopted using ZDOCK protocol in Discovery Studio Client 2.5. Initially, we have found out the cooperated the enzymes encoded by genes for four selected enzymes, i.e. RAD23A, CCNH, CDK7, and CETN2 by STRING 9.0.5 database. The found Best closely related enzymes for RAD23A PDB ID: 1DV0 (RAD23 Homolog A) was PDB ID: 2KDE (PSMD4 MCB1, 26S proteasome non-ATPase regulatory subunit 4), for CCNH PDB ID: 1KXU (CDK-Activating Kinase Complex Subunit) was PDB ID: 1UA2 (CDK7,Cell division protein kinase 7), for CDK7 PDB ID: 1UA2 (Cyclin H) was PDB ID: 1KXU (CDK-Activating Kinase Complex Subunit) and for CETN2 PDB ID:1ZMZ (Centrin, EF-Hand Protein, 2) was PDB ID: 2GGM (XPC, DNA-repair protein complementing XP-C cells) (Figure 5 A, B, C and D). Later on we run ZDOCK program for Protein-Protein Docking *vs* Protein-Complex docking analysis.

The obtained Zdock scores 1DV0 *vs* 2KDV was 14.58, 1DV0+NNK *vs* 2KDE was 13.74, 1KXU *vs* 1UA2 was 14.24,1KXU+NNK *vs* 1UA2 was 13.08, 1UA2 *vs* 1KXU was 14.96, 1UA2+NNK *vs* 1KXU was 13.08, 1ZMZ *vs* 2GGM was 15.92 and 1ZMZ+NNK *vs* 2GMM was 15.40 (Table 4).Similarly, obtained Zdock scores 1DV0 *vs* 2KDV was 14.58, 1DV0+NNAL *vs* 2KDE was 13.76, 1KXU *vs* 1UA2 was 14.24,1KXU+NNAL *vs* 1UA2 was 13.84, 1UA2 *vs* 1KXU was 14.96, 1UA2+NNAL *vs* 1KXU was 13.68, 1ZMZ *vs* 2GGM was 15.92 and 1ZMZ+NNAL *vs* 2GMM was 15.55 (Table 4). The results shown that Zdock score of protein complex (contain enzymes and cigarette smoke carcinogens conformation) interaction were higher than protein-protein interaction. Analysis clearly revealed that when NNK/NNAL interacts with NER enzymes their metabolic activity to form complex with its cooperated enzymes reduces significantly. Thus, NNK and NNAL were capable to damage the DNA repair machinery and its will lead to the functional loss of NER enzymes encoded by genes RAD23A, CCNH, CDK7, and CETN2.

Conclusion

This hypothesis able to provide better understanding to explore the molecular interaction of NNK and NNAL with enzymes involved in NER pathways. It is also helpful to understand the biological insights of NNK and NNAL binding efficacy in the progression of cancer. The study revealed that the enzymatic activity of these enzymes RAD23A, CCNH, CDK7, and CETN2 affected by NNK and NNAL. Therefore, the possibility of DNA damage will be increased because these enzymes have an important role in the DNA damage control. Once the DNA repair machinery altered due to interaction of cigarette smoke carcinogens NNK and NNAL the whole biological process will lead to uncontrolled tumor growth and finally cancer will be developed. For the further confirmation of study the *in vivo* and *in vitro* validation needed.

Acknowledgements

The authors are thankful to College of Applied Medical Sciences, Buraydah Colleges, Al Qassim, Saudi Arabia for providing necessary infrastructure facility to complete the study.

References

Accelrys Software Inc., Discovery Studio Modeling Environment, Release 4.0, San Diego: Accelrys Software Inc, 2013.

Araki M, Masutani C, Takemura M, et al (2001). Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair. *J Biol Chem*, **276**, 18665-72.

Berman HM, Westbrook J, Feng Z, et al (2000). The protein data bank. *Nucl Acids Res*, **28**, 235-42.

Brooks BR, Bruccoleri RE, Olafson BD, et al (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem*, **4**, 187-217.

Brooks BR, Brooks CL, Mackerell AD Jr, et al (2009). M. CHARMM: the biomolecular simulation program. *J Comput Chem*, **30**, 1545-614.

Ceccaldi R, WO'Connor K, et al (2015). A unique subset of epithelial ovarian cancers with platinum sensitivity and PARP inhibitor. *Cancer Res*, **75**, 628-34.

Chen R, Li L, Weng Z (2003). ZDOCK: An initial-stage protein-docking algorithm. *Proteins*, **52**, 80-7.

Goodsell DS, Morris GM, Olson AJ, (1996). Automated docking of flexible ligands: applications of AutoDock. *J Mol Recognition*, **9**, 1-5.

Jamal QMS, M Lohani, M H Siddiqui, et al (2012). Molecular interaction analysis of cigarette smoke carcinogens NNK and NNAL with enzymes involved in DNA repair pathways: An in silico approach. *Bioinformation*, **8**, 795-800.

Lv X, Wang J, Dong Z Lv, Qin Y (2009). DNA-Bound peptides control the mRNA transcription through CDK7. *Peptides*, **30**, 681-8.

Mirijanian DT, Mannige RV, Zuckermann RN, et al (2014). Development and use of an atomistic CHARMM-based forcefield for peptoid simulation. *J Comput Chem*, **35**, 360-70.

Morris GM, Goodsell DS, Halliday RS, et al (1998). Automated docking using Lamarckian genetic algorithm and an empirical binding free energy function. *J Comp Chem*, **19**, 1639-62.

Mou F-F, Wang L-W (2013). Predictive role of computer simulation in assessing signaling pathways of crizotinib-treated A549 lung cancer cells. *Asian Pac J Cancer Prev*, **13**, 3119-21.

Morris GM, Huey R, Lindstrom W, et al (2009). Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. *J Comp Chem*, **16**, 2785-91.

Nouspikel T (2009). DNA repair in mammalian cells: Nucleotide excision repair: variations on versatility. *Cell Mol Life Sci*, **66**, 994-1009.

Pence HE, Williams A (2010). Chemspider: an online chemical information resource. *J Chemical Education*, **87**, 1123-4.

Pierce BG, Wiehe K, Hwang H, et al (2014). ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics*, **30**, 1771-3.

Rouillon C, White MF (2011). The evolution and mechanisms of nucleotide excision repair proteins. *Res Microbiol*, **162**, 19-26.

Sugasawa K, Ng JM, Masutani C, et al (1997). Two human homologs of Rad23 are functionally interchangeable in complex formation and stimulation of XPC repair activity. *JH Mol Cell Biol*, **17**, 6924-31.

Syota Matsumoto, Eric S Fischer, Takeshi Yasuda, et al (2015). Functional regulation of the DNA damage-recognition factor DDB2 by ubiquitination and interaction with xeroderma pigmentosum group C protein. *Nucl Acids Res*.

Szklarczyk D, Franceschini A, Kuhn M, et al (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucl Acids Res*, **39**, 561-8.

Tsai CW, Chen JL, Yang CS (2012). An improved LGA for protein ligand docking prediction. Proceedings of IEEE congress on evolutionary computation, pp. 1-6.

Xue J, Yang S, Seng S (2014). Mechanisms of cancer induction by tobacco-specific NNK and NNN. *Cancer*, **6**, 1138-56.