

# Proteome-wide screening for the analysis of protein targeting of *Chlamydia pneumoniae* in endoplasmic reticulum of host cells and their possible implication in lung cancer development

YANYAN LI<sup>1</sup>; SHAHANAVAJ KHAN<sup>2,3,4,\*</sup>; ANIS AHMAD CHAUDHARY<sup>5</sup>; HASSAN AHMED RUDAYNI<sup>5</sup>; ABDUL MALIK<sup>2</sup>; ASHWAG SHAMI<sup>6,\*</sup>

<sup>1</sup> Department of Oncology, Zhengzhou Central Hospital Affiliated to Zhengzhou University, Zhengzhou, 450007, China

<sup>2</sup> Department of Pharmaceutics, College of Pharmacy, King Saud University, Riyadh, 11451, Saudi Arabia

<sup>3</sup> Department of Biosciences, Shri Ram Group of College (SRGC), Muzaffarnagar, 251001, India

<sup>4</sup> Department of Health Sciences, Novel Global Community Educational Foundation, Sydney, Australia

<sup>5</sup> Department of Biology, College of Science, Imam Mohammad Ibn Saud Islamic University (IMSUI), Riyadh, 11623, Saudi Arabia

<sup>6</sup> Department of Biology, College of Sciences, Princess Nourah bint Abdulrahman University, Riyadh, 11617, Saudi Arabia

**Key words:** Bacterial infection, *C. pneumoniae*, Proteomes, ER, Lung cancer

**Abstract:** Available reports have confirmed a link between bacterial infection and the progression of different types of cancers, including colon, lungs, and prostate cancer. Here we report the *Chlamydia pneumoniae* proteins targeting in endoplasmic reticulum (ER) using *in-silico* approaches and their possible role in lung cancer etiology. We predicted 48 proteins that target human ER, which may be associated with protein folding and protein-protein interactions during infection. The results showed *C. pneumoniae* proteins targeting human ER and their implications in lung cancer growth. These targeted proteins may be involved in competitive interactions between host and bacterial proteins, which may change the usual pathway functions and trigger the development of lung cancer. Moreover, *C. pneumoniae* unfolded protein accumulation in the human ER possibly induces ER stress, consequently activating the unfolded protein response (UPR), and providing a favorable microenvironment for cancer growth. The current study showed the *C. pneumoniae* protein targeting in ER of host cell and their implication in lung cancer growth. These results may help researchers better manage lung cancer and establish a molecular mechanism for *C. pneumoniae* lung cancer association.

## List of abbreviations

ER:	Endoplasmic reticulum
IL:	Interleukin
MW:	molecular weight
kDa:	kiloDalton
pI:	isoelectric point
UPR:	unfolded protein response

## Introduction

Lung cancer is one of the very common cancer, with 1.8 million (approximately 13% of the total cancer) new cases estimated

globally in 2012 (Ferlay *et al.*, 2015). The most common cancer is lung carcinoma and leading cause of death, after breast cancer in female and prostate cancer in men (Bray *et al.*, 2018). The association of infection of bacteria with the progression of different types of cancer has confirmed in various studies (Arthur *et al.*, 2012; Khan, 2015; Magat *et al.*, 2020). The Gram-negative microbe *Chlamydia pneumoniae* is an intracellular obligate bacterium. *C. pneumoniae* is transmitted due to acute respiratory infection and potentially associated with lung disease including bronchitis, asthma, atherosclerosis, and high risk of lung cancer however, the causal pathogenetic mechanisms are not well understood. The infection of *C. pneumoniae* has been reported to alter several pathways of infected host cell including calcium-mediated nuclear factor- $\kappa$ B (NF- $\kappa$ B) pathway, apoptosis, necrosis etc (Chaturvedi *et al.*, 2010; Littman *et al.*, 2005; Wahl *et al.*, 2003). NF- $\kappa$ B has crucial role in proliferation, inflammation, and regulation of apoptosis. Alteration in program cell death due to NF- $\kappa$ B pathway enigmatically associated with cancer

\*Address correspondence to: Shahanavaj Khan, sdkhan@ksu.edu.sa; Ashwag Shami, Ayshami@pnu.edu.sa

Received: 13 March 2021; Accepted: 17 May 2021



(Wahl *et al.*, 2003). Inflammation may be connected with carcinogenic process due to chronic infection of bacteria in human (Littman *et al.*, 2005). The lungs of the smoker may be more susceptible for the localization of *C. pneumoniae* (Von Hertzen, 1998). The radicals of superoxide oxygen, tumor necrosis factor, IL-8 and IL-1 $\beta$  are generated by induced monocytes. The inflammatory factors may trigger the process of carcinogenesis in lung cells (Koyi *et al.*, 1999).

*C. pneumoniae* contained the capability to interfering with the host cell's apoptotic apparatus. Alteration in the apoptosis has been related to survival scheme of intracellular *C. pneumoniae*. Many reports of *C. pneumoniae* infection have been demonstrated to inhibit host cell apoptosis (Airenne *et al.*, 2002; Fischer *et al.*, 2001; Wahl *et al.*, 2003). However, some contradictory study has shown the pro-apoptotic activities of CP during the infection (Sessa *et al.*, 2009). Members of the inhibitor of apoptosis protein family are important factors that regulate apoptotic cell death. The infection of *C. pneumoniae* showed the induction of the expression of mRNA and protein of cellular inhibitor of apoptosis 2 (c-IAP2) in human cell line Mono Mac 6 that may help to intracellular survival of the bacterium (Wahl *et al.*, 2003). The carcinogenic process of lung cancer due to the infection of *C. pneumoniae* is still infancy stage.

Cancer is a serious problem around the world and arises due to various factors including infection, mutations, smoking and disturbance in homeostasis of various biological process. Endoplasmic reticulum (ER) is the important subcellular organelle of eukaryotes which transfers various proteins to their subcellular location or on surface of the cell (Braakman and Bulleid, 2011). ER has various molecular chaperones which involved in protein synthesis, protein folding and maturation. Several biological processes are controlled by the normal functioning of ER. Interestingly, there are several connections to activation of ER stress response constitutes a cellular process that can be triggered by a great variety of pathological human conditions such as Alzheimer's disease and cancer. Different conditions including inflammatory bowel disease, diabetes and various types of cancers are related to the activation of the ER stress response pathway. ER stress is induced in numerous pathological and physiological conditions through accretion of misfolded proteins in the lumen of ER.

In the current work, we predicted the *C. pneumoniae* protein targeting in the ER of host cell and their possible involvement in the growth of lung cancer. We utilized the bioinformatics approach to decipher the possible localization of whole protein sequences of *C. pneumoniae* in ER of host cells. Furthermore, we predicted the possible role of molecular weight and isoelectric point of ER targeted proteins in protein targeting and growth of lung cancer. We expect that our results might be very helpful, valuable to provide right direction to experimental research in wet lab experiments.

## Materials and Methods

### Selection of database

Various databases were served in order to obtain maximum proteins sequences of *C. pneumoniae*. In the current study we were served the different database such as NCBI,

Uniprot and EMBL to retrieve the complete proteome of *C. pneumoniae*.

### *C. pneumoniae* proteome retrieval

The bacterium *C. pneumoniae* was infecting the host as obligate intracellular bacteria and act as possible factor in the growth of lung cancer (Hua-Feng *et al.*, 2015; Xiong *et al.*, 2019). Six isolates (CWL029, AR39, J138, B21, TW-183, LPCoLN) of bacteria were available in Uniprot database with huge information of protein sequences (Kalman *et al.*, 1999; Myers *et al.*, 2009; Read *et al.*, 2000; Shirai *et al.*, 2000). The whole protein sequences of bacteria TW-183 isolate were retrieved from Uniprot database. These protein sequences were utilized to decipher the ER targeted proteins using bioinformatics approach.

### Deciphering the endoplasmic reticulum targeting proteins of *C. pneumoniae* and their role in lung cancer

Online available bioinformatics predictor Hum-mPLoc 2.0 was used in current study to decipher the *C. pneumoniae* proteins targeting in ER of human cells. This predictor has predicted subcellular localization of proteins in various organisms. (Shen and Chou, 2009). This predictor has predicted subcellular localization of proteins in various organisms. All proteins sequences of TW-183 isolate of *C. pneumoniae* were used to predict the protein targeting in host Endoplasmic Reticulum. Hum-mPLoc 2.0 was deciphered the protein targeting in various sub cellular organelles such as endoplasmic reticulum, cytoplasm, mitochondrion, nucleus, centriole, cytoskeleton, endosome, extracell, golgi apparatus, lysosome etc. The predictor has used sequential evolution information and functional domain information using ensemble classifier. The protocol of prediction showed in general scheme (Fig. 1).

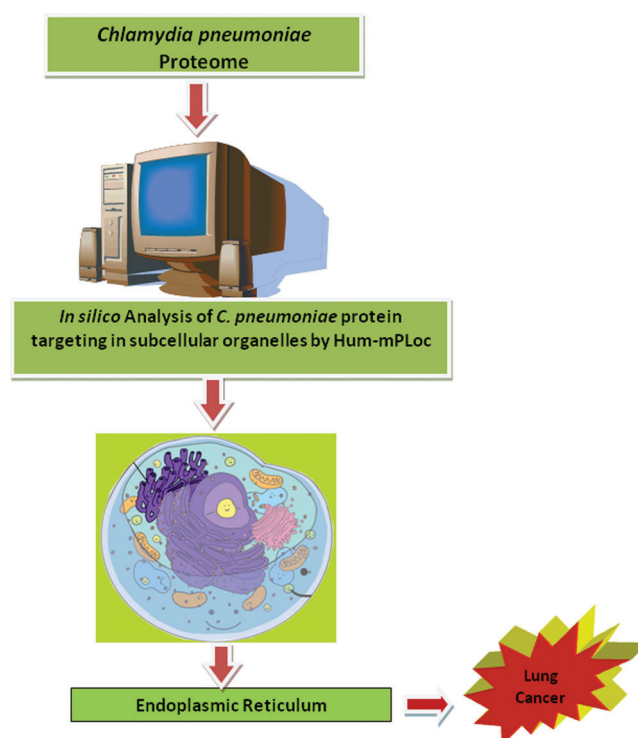


FIGURE 1. Illustration showed the plan of prediction strategies.

*Decipher the possible relation of particular molecular weight of C. pneumoniae proteins in endoplasmic reticulum targeting and their role in lungs cancer*

ExPASy online bioinformatics predictor compute pI/Mw tool was used to decipher the molecular weight in whole protein sequences of TW-183 bacteria. This tool was used to analyse the various characters of proteins in proteomics research. The tool was worked on experimentally available information of particular amino acid and provided accurate results.

*Decipher the possible relation of particular isoelectric point of C. pneumoniae proteins in endoplasmic reticulum targeting and their role in lung cancer*

Similarly, ExPASy online bioinformatics predictor compute pI/Mw tool was utilized to decipher the isoelectric point (pI) in whole protein sequences of TW-183 bacteria.

## Results

*Selection of database and retrieve of proteome*

In the present work, we were selected the Uniprot database to retrieving the protein sequences of *C. pneumoniae*. Uniprot database was contained information of most of the bacterium and considered a reliable source for bioinformatic study.

*To retrieve the proteome of C. pneumonia*

TW-183 isolate of *C. pneumoniae* was selected for the retrieving of whole protein sequences of bacterium due to their large size proteome. The proteome was contained maximum number of protein sequences.

*Decipher the endoplasmic reticulum targeting proteins of C. pneumoniae and their role in lungs cancer*

The results of Hum-mPLoc 2.0 were showed that only 48 proteins of TW-183 isolate of *C. Pneumonia* targeted in ER of host cells. The descriptions of proteins such as Accession number, possible functions, and number of amino acids are shown in [Tab. 1](#).

The normal functions of host cells may alter through different strategies due to the protein-protein interaction in the lumen of ER as existences of *C. Pneumoniae* proteins. ER plays very important role in protein folding and formation of active proteins/enzymes.

*Decipher the possible relation of particular molecular weight of C. pneumoniae proteins in endoplasmic reticulum targeting and their role in lungs cancer*

Whole protein sequences of *C. pneumoniae* were used to calculate the theoretical molecular weight (MW) using compute pI/Mw tool ([https://web.expasy.org/compute\\_pi/](https://web.expasy.org/compute_pi/)). This tool is highly useful to decipher the MW and theoretical pI value in unknown protein sequence. The relations of specific range of molecular weight of ER target proteins and whole proteins of TW-183 isolate of *C. pneumoniae* were demonstrated in [Figs. 2A](#) and [2B](#).

The results of our current study were clearly demonstrated that the increases in molecular weight consistently decreased the targeting of proteins in ER of host cells. The proteins contain lowest range of molecular

weight (0–20 kDa were showed high potential to target in ER of host cell while the protein having highest molecular (>80 kDa) weight were showed very little potential to target in ER of host cells.

*Decipher the possible relation of particular isoelectric point of C. pneumoniae proteins in endoplasmic reticulum targeting and their role in lung cancer*

Similarly, the values of isoelectric point (pI) in whole proteins sequences of TW-183 isolate of *C. pneumoniae* were deciphered using online tool compute pI/Mw. The relations of specific range of pI values of ER target proteins and whole proteins of TW-183 isolate of *C. pneumoniae* were demonstrated in [Figs. 3A](#) and [3B](#).

The previous report demonstrated that pathogen protein translocation is supported through a neutral isoelectric point (pI) between 6.0 to 8.4. Nevertheless, for proteins >50 kDa in size, pI alone does not clarify their localisation ([Smigielski et al., 2019](#)). The results were showed that the value of isoelectric point (pI) was not showing any constant pattern for protein targeting in ER of host cells ([Fig. 3A](#) and [3B](#)). It was observed that the pI value 9–10 have the high potential (24 Proteins) to target in ER of host cells.

## Discussion

Although wet lab research and experiments for analysing protein targeting signals can provide comprehensible evidence and distinguish between maintenance and return signals but performing wet lab research and experiments are generally more time-taking and costly. Therefore, *in-silico* computational research is recognized as an alternative approach that presents informative, valuable, and useful direction to experimental research. Cell is considered the small unit of life, which holds numerous molecules including proteins and enzymes. Various subcellular locations generally called cell organelles contain different proteins which have the important functions for the cell's survival. ER is a main subcellular organelle in the cells of eukaryotes. Approximate one-third the proteins are transported to various subcellular locations to the cells through the ER ([Hetz et al., 2015](#)). It is involved in regulation of various essential biological processes including protein folding, post-translational modification, cell metabolism and protein synthesis ([Schwarz and Blower, 2016](#)).

In our current study 48 different proteins were predicted to be targeted to the ER of human cells from the entire proteome of *C. pneumoniae*. These targeted proteins of *C. pneumoniae* may alter the different functions of human cell during the course of infection. This study has showed that ER of infected epithelial cells mostly contained chlamydial major outer membrane protein (MOMP), lipopolysaccharide (LPS) and the inclusion membrane protein A (IncA) ([Giles and Wyrick, 2008](#)). The *in-silico* to decipher the protein targeting into the human cell is very essential to discover the growth of cancer, particularly when the progression of cancer is related to the bacterial intracellular infection. Various studies have showed that many targeted proteins of bacteria in different subcellular organelles of host cell (including mitochondria, Golgi complex, cytoplasm, and nucleus) exert adverse effect and act as the etiological factors

TABLE 1

The data in the table showed the details information of *C. pneumoniae* proteins, which targeted in the endoplasmic reticulum of host cell

S. No.	Accession No.	Proteins/Enzymes name	No. of Amino acid	Function in bacteria	pI Values	Mol. Wt.	Localization by Hum-mPLoc 2.0
1.	Q9Z7Y4	1-acyl-sn-glycerol-3-phosphate acyltransferase, putative (Glycerol-3-P Acyltransferase)	212	Transferase activity, transferring acyl groups	9.68	24,449	Endoplasmic reticulum
2.	Q9Z943	Uncharacterized protein	146	nuclease activity	4.88	16,716	Endoplasmic reticulum
3.	Q9Z8D4	Uncharacterized protein	161	Unknown	4.96	18,609	Endoplasmic reticulum
4.	Q9Z6V8	Glycogen synthase (EC 2.4.1.21) (Starch [bacterial glycogen] synthase)	476	Glycosyltransferase, Transferase, Glycogen biosynthesis	4.97	53,146	Endoplasmic reticulum
5.	Q9Z778	Secretion protein (Type III secretion protein SctI) (Yop proteins translocation lipoprotein J)	335	protein secretion	5.44	36,245	Endoplasmic reticulum
6.	Q9Z8P4	Cyclic nucleotide-binding protein (cAMP-Dependent Protein Kinase Regulatory Subunit)	137	kinase activity	5.52	15,555	Endoplasmic reticulum
7.	Q9Z7G4	Transcription elongation factor GreA (Transcript cleavage factor GreA)	722	Transcription, Transcription regulation	5.52	82,642	Endoplasmic reticulum
8.	Q7VPT1	Anti-sigma factor antagonist	115	regulation of transcription	5.77	12,860	Endoplasmic reticulum
9.	Q9Z8A8	Phospholipase D family protein (Phospholipase D superfamily) (Putative phospholipase D)	476	catalytic activity	5.99	54,470	Endoplasmic reticulum
10.	Q9Z8U6	Glucose-6-phosphate 1-dehydrogenase (G6PD)	512	Oxidoreductase, glucose-6-phosphate dehydrogenase activity	6.25	58,740	Endoplasmic reticulum
11.	Q9Z6U9	Axial Filament Protein (Cytosolic axial filament protein cafA and ribonuclease E) (Ribonuclease G)	515	ribonuclease activity	6.32	59,551	Endoplasmic reticulum
12.	Q9Z779	Putative uncharacterized protein CPn0827 Homolog (Uncharacterized protein)	277	Unknown	8.18	31,821	Endoplasmic reticulum
13.	Q9Z7C1	ExbB (Macromolecule transporter) (MotA/TolQ/ExbB proton channel family protein)	232	protein transporter activity	8.48	25,575	Endoplasmic reticulum
14.	Q9Z809	Probable metal transport system membrane protein	293	Transport, ATPase activity	8.72	32,243	Endoplasmic reticulum
15.	Q7VPR2	ABC transporter	591	transport	8.82	66,061	Endoplasmic reticulum
16.	Q9Z978	Cytochrome D ubiquinol oxidase, subunit II	344	oxidation reduction process	8.85	38,174	Endoplasmic reticulum
17.	Q9Z8W7	Uncharacterized protein	228	Unknown	8.87	25,657	Endoplasmic reticulum
18.	Q9Z6P9	UPF0056 membrane protein	202	Unknown	8.95	22,652	Endoplasmic reticulum
19.	Q9Z8J7	Probable metal transport system membrane protein	324	Transport, ATPase activity, coupled to transmembrane movement of substances Source: InterPro	9.07	35,249	Endoplasmic reticulum
20.	Q9Z816	Alanine/glycine transporter protein (D-Ala/Gly Permease) (D-Ala/Gly permease)	449	alanine: sodium symporter activity	9.12	49,481	Endoplasmic reticulum

(Continued)

**Table 1 (continued).**

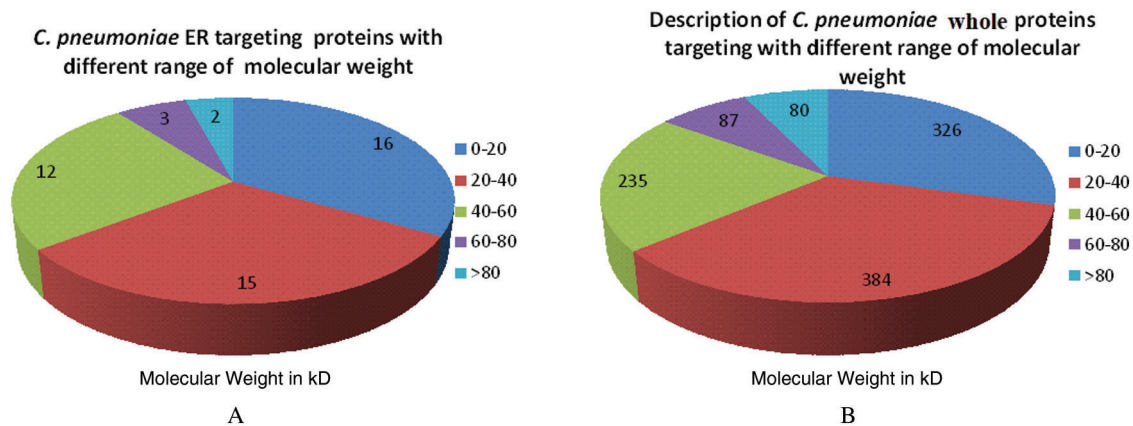
S. No.	Accession No.	Proteins/Enzymes name	No. of Amino acid	Function in bacteria	pI Values	Mol. Wt.	Localization by Hum-mPLoc 2.0
21.	Q9Z7U0	ADP,ATP carrier protein 2 (ADP/ATP translocase 2)	540	Transport, ATP:ADP antiporter activity	9.15	60,050	Endoplasmic reticulum
22.	Q9Z6V2	Uncharacterized protein	450	Unknown	9.17	51,759	Endoplasmic reticulum
23.	Q7VPV0	Succinate dehydrogenase subunit C	297	Unknown	9.25	33,328	Endoplasmic reticulum
24.	Q9Z739	Rod Shape Protein (Rod shape protein) (Stage V sporulation protein E)	415	cell cycle	9.32	46,663	Endoplasmic reticulum
25.	Q7VPQ7	Uncharacterized protein	61	Unknown	9.43	7,085	Endoplasmic reticulum
26.	Q7VPV8	Uncharacterized protein	831	Unknown	9.43	7,085	Endoplasmic reticulum
27.	Q9K1X8	Uncharacterized protein	232	Unknown	9.48	26,313	Endoplasmic reticulum
28.	Q7VQ03	Uncharacterized protein	53	Unknown	9.52	6,188	Endoplasmic reticulum
29.	Q9Z8J2	ADP,ATP carrier protein 1 (ADP/ATP translocase 1)	515	Transport, ATP:ADP antiporter activity	9.52	57,197	Endoplasmic reticulum
30.	Q9Z7D9	Uncharacterized protein	98	Unknown	9.57	11,577	Endoplasmic reticulum
31.	Q9Z767	Phosphatidylserine decarboxylase proenzyme	301	Decarboxylase, Lyase, Lipid biosynthesis	9.59	34,827	Endoplasmic reticulum
32.	Q9Z782	Translocation protein S (Type III secretion inner membrane protein SctS) (Yop proteins translocation protein S)	95	protein secretion	9.6	10,664	Endoplasmic reticulum
33.	Q9Z6Z3	Uncharacterized protein	62	Unknown	9.66	7,465	Endoplasmic reticulum
34.	Q9Z6Y5	2-acylglycerophospho-ethanolamine acyl transferase/acyl carrier	253	transferase activity	9.67	29,519	Endoplasmic reticulum
35.	Q9Z9D4	Uncharacterized protein	66	Unknown	9.73	7,880	Endoplasmic reticulum
36.	Q7VPQ0	Uncharacterized protein	60	Unknown	9.74	7,306	Endoplasmic reticulum
37.	Q9Z817	Lipoprotein signal peptidase (Prolipoprotein signal peptidase)	168	Aspartyl protease, Hydrolase, Protease	9.74	19,365	Endoplasmic reticulum
38.	Q9Z7S5	Protein translocase subunit SecY	457	Protein transport, Translocation, Transport	9.95	50,286	Endoplasmic reticulum
39.	Q9Z8M6	Uncharacterized protein	151	Unknown	10.36	17,315	Endoplasmic reticulum
40.	Q7VQ15	Uncharacterized protein	56	Unknown	10.51	6,658	Endoplasmic reticulum
41.	Q9Z8N1	Glycogen phosphorylase	824	Glycosyltransferase, Transferase, glycogen phosphorylase activity	5.83	94,517	Endoplasmic reticulum
42.	Q7VPU0	Uncharacterized protein	65	Unknown	8.8	7,382	Endoplasmic reticulum
43.	Q9Z979	Cytochrome D ubiquinol oxidase, subunit I	448	Unknown	8.93	50,674	Endoplasmic reticulum
44.	Q9Z6M5	Putative outer membrane protein	262	Unknown	5.39	30,001	Endoplasmic reticulum
45.	Q9K1U7	Uncharacterized protein	366	Unknown	9.4	40,828	Endoplasmic reticulum

(Continued)

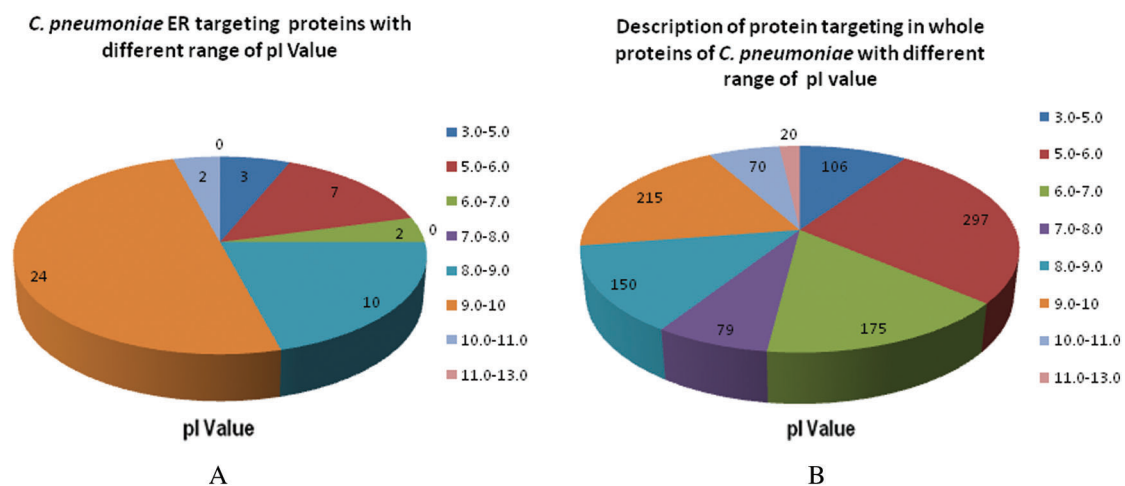


Table 1 (continued).

S. No.	Accession No.	Proteins/Enzymes name	No. of Amino acid	Function in bacteria	pI Values	Mol. Wt.	Localization by Hum-mPLoc 2.0
46.	Q9Z8B6	Probable Na(+)-translocating NADH-quinone reductase subunit B	503	Oxidoreductase, Ion transport, Sodium transport, Transport	9.02	55,280	Endoplasmic reticulum
47.	Q9Z6U3	Lipid-A-disaccharide synthase	604	Glycosyltransferase, Transferase, Lipid A biosynthesis	9.29	69,014	Endoplasmic reticulum
48.	Q9Z7D6	Uncharacterized protein	264	Unknown	8.4	31,052	Endoplasmic reticulum



**FIGURE 2.** Protein targeting blueprint. (A) Graph showed the protein targeting of *C. pneumoniae* in endoplasmic reticulum of host cells with respect to various ranges of molecular weight. (B) Graph showed the protein targeting of *C. pneumoniae* whole protein in endoplasmic reticulum of host cells with respect to various ranges of molecular weight.



**FIGURE 3.** Protein targeting blueprint. (A) Graph showed the protein targeting of *C. pneumoniae* in endoplasmic reticulum of host cells with respect to various ranges of pI value. (B) Graph showed the protein targeting of *C. pneumoniae* whole protein in endoplasmic reticulum of host cells with respect to various ranges of pI value.

in the progression of various types of cancer (Khan et al., 2016; Khan et al., 2020; Khan et al., 2017). The proteins and enzymes that exist in the ER are known as ER-resident proteins. ER-resident proteins are an important topic in ER-related studies. Several of the ER-resident proteins and enzymes have specific protein sorting signals including KDEL or KXXX, while some others proteins do not having such signals (Stornaiuolo et al., 2003). Information of proteins and enzymes localization in different subcellular location may be helpful to know the

clues of different functions. It is prerequisite to reveal the information and functions of proteins/enzymes located in intricate pathways of different cell organelles.

Infections are considered as an important part of the natural path of cancer. Pneumonia is very common disease which influencing approximate 450 million persons every year and finding all over in the world (Biscevic-Tokic et al., 2013). Approximate 50%–70% patients has been showed lung cancer with pneumonia (Akinosoglou et al., 2013). It is

a major reason of deaths with about 7% of total deaths worldwide every year (Ruuskanen *et al.*, 2011). The cases of lung cancer and pneumonia enhanced noticeably (Shen *et al.*, 2016). The molecular mechanism of association of *C. pneumoniae* with lung cancer is poorly investigated till date. We were predicted targeting of various enzymes and proteins in ER of human cell which includes Glycogen synthase (Accession No. Q9Z6V8), Glucose-6-phosphate 1-dehydrogenase (G6PD) (Accession No. Q9Z8U6), ABC transporter (Accession No. Q7VPR2), Phosphatidylserine decarboxylase proenzyme (Accession No. Q9Z767), Lipoprotein signal peptidase (Accession No. Q9Z817), Glycogen phosphorylase (Accession No. Q9Z8N1), Lipid-A-disaccharide synthase (Accession No. Q9Z6U3). Although in current study we have deciphered the 48 proteins to target in ER but the functions of various proteins still unknown. The targeted enzymes and proteins may change the normal functioning of protein folding in ER of during infection and help in the growth of lung cancer. Accumulation of *C. pneumoniae* unfolded enzymes and proteins in the human ER possible act as a factor of ER stress and promote the cleavage of ATF6, which consequently facilitates the activation of unfolded protein response (UPR). The UPR activation has associated with the progression of cancer and lung diseases (Barabutis, 2020; Piton *et al.*, 2016; Wang *et al.*, 2014). It have reported that the expression of G6PD is enhanced in various tumors such as colon cancers, leukemia, gastrointestinal cancers, breast cancers, endometrial carcinomas, liver cancer and lung cancer (Batetta *et al.*, 1999; Jiang *et al.*, 2013; Rao *et al.*, 1997; van Driel *et al.*, 1999).

The functions and other informations of the ER targeted proteins of *C. pneumoniae* were showed in Tab. 1. Hence, an important basic purpose in proteomics and molecular cell biology is to decipher the subcellular targeting of proteins and enzymes in the whole cell. It is also very essential for prioritizing and selecting the accurate goals for drug development. Unfortunately, both processes are time consuming, laborious with and high expenses. To confirm the subcellular location of proteins and enzymes in various organelles of cell is completely based on laboratory works.

The cells constantly translate the great variety of proteins and enzymes. The translated proteins need to be folded correctly, post-translationally modification, accumulated into complexes and located to their ultimate destinations of subcellular organelle. If some mistakes arise in the maturation process the normal functioning of the cell disturbed. In the current study we are focused to decipher the protein targeting of *C. pneumoniae* in endoplasmic reticulum of host cell.

With the availability of next generation sequences data of proteins sequences, we needed highly desired *in silico* approaches for rapidly and efficiently identifying the subcellular locations of uncharacterized proteins based on their sequence information alone. ER is the main site of protein modification and protein folding. Misfolded proteins are responsible to stimulates the UPR in the ER (Schmidt *et al.*, 2019), which increases the phenomena of protein folding to restore homeostasis (Choi and Song, 2020). UPR dysregulation signalling contributes to the pathogenesis during infection. The ER stress is related to pathogenesis of

different diseases, including diabetes, neurodegenerative disorders, cancer, etc. Nevertheless, the pathogen proteins which have targeted in ER may disturbed the homeostasis and responsible for alteration in normal functioning of host cells. These changes may connect with the abnormal growth of infected cells and act as a prominent factor of cancer growth.

## Conclusion

The ER is responsible for various housekeeping functions in the cell and act as a key organelle for the final maturation of proteins that controls its homeostasis state. Change in homeostasis has been associated with the progression of various situations including cancer, which is the highly challenging problem in the present time. The *C. pneumoniae* bacterial protein targeting in ER of human lung cells may disturbed the normal functioning of ER and act as a potential factor for the growth of lung cancer. Alterations in the UPR connected with the growth and development of cancer due to accumulation of pathogen proteins. In the present study we deciphered the *C. pneumoniae* protein targeting in ER and their association with lung cancer. We observed that the intercellular infection of *C. pneumoniae* in human cells acts as an etiological factor in lung cancer. Therefore, our novel finding provides the new direction to researchers, scientists and clinicians for the management and treatment of *C. pneumoniae* associated lung cancer.

**Acknowledgement:** The authors extend their appreciation to the College of Applied Medical Sciences Research Center and the Deanship of Scientific Research, King Saud University, Riyadh, Saudi Arabia for their kind support.

**Availability of Data and Materials:** All data generated or analyzed during this study are included in this published article.

**Authors' Contributions:** YL and AAC separately assessed electronic search, eligibility of the studies, inclusion criteria; HAS and AM design the idea, help in manuscript drafting; SK and AS supervised the whole study including formulation of research idea, literature review, data collection, manuscript drafting with critical revision in final draft of manuscript.

**Funding Statement:** This research was funded by the Deanship of Scientific Research at Princess Nourah bint Abdulrahman University through the Fast-track Research Funding Program.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- Airenne S, Surcel HM, Tuukkanen J, Leinonen M, Saikku P (2002). Chlamydia pneumoniae inhibits apoptosis in human epithelial and monocyte cell lines. *Scandinavian Journal of Immunology* 55: 390–398.
- Akinosoglou KS, Karkoulas K, Marangos M (2013). Infectious complications in patients with lung cancer. *European Review for Medical and Pharmacological Sciences* 17: 8–18.
- Arthur JC, Perez-Chanona E, Muhlbauer M, Tomkovich S, Uronis JM *et al.* (2012). Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* 338: 120–123.

- Barabutis N (2020). Unfolded protein response in lung health and disease. *Frontiers in Medicine* **7**: 344.
- Batetta B, Pulisci D, Bonatesta RR, Sanna F, Piras S et al. (1999). G6PD activity and gene expression in leukemic cells from G6PD-deficient subjects. *Cancer Letters* **140**: 53–58.
- Biscevic-Tokic J, Tokic N, Musanovic A (2013). Pneumonia as the most common lower respiratory tract infection. *Medical Archives* **67**: 442–445.
- Braakman I, Bulleid NJ (2011). Protein folding and modification in the mammalian endoplasmic reticulum. *Annual Review of Biochemistry* **80**: 71–99.
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* **68**: 394–424.
- Chaturvedi AK, Gaydos CA, Agreda P, Holden JP, Chatterjee N et al. (2010). Chlamydia pneumoniae infection and risk for lung cancer. *Cancer Epidemiology and Prevention Biomarkers* **19**: 1498–1505.
- Choi JA, Song CH (2020). Insights into the role of endoplasmic reticulum stress in infectious diseases. *Frontiers in Immunology* **10**: 3147.
- Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F (2015). Cancer incidence and mortality worldwide: Sources methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer* **136**: E359–E386.
- Fischer SF, Schwarz C, Vier J, Hacker G (2001). Characterization of antiapoptotic activities of *Chlamydia pneumoniae* in human cells. *Infection & Immunity* **69**: 7121–7129.
- Giles DK, Wyrick PB (2008). Trafficking of chlamydial antigens to the endoplasmic reticulum of infected epithelial cells. *Microbes and Infection* **10**: 1494–1503.
- Hetz C, Chevet E, Oakes SA (2015). Proteostasis control by the unfolded protein response. *Nature Cell Biology* **17**: 829–838.
- Hua-Feng X, Yue-Ming W, Hong L, Junyi D (2015). A meta-analysis of the association between *Chlamydia pneumoniae* infection and lung cancer risk. *Indian Journal of Cancer* **52**: e112–e115.
- Jiang P, Du W, Yang X (2013). A critical role of glucose-6-phosphate dehydrogenase in TAp73-mediated cell proliferation. *Cell Cycle* **12**: 3720–3726.
- Kalman S, Mitchell W, Marathe R, Lammel C, Fan J et al. (1999). Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*. *Nature Genetics* **21**: 385–389.
- Khan S (2015). Potential role of *Escherichia coli* DNA mismatch repair proteins in colon cancer. *Critical Reviews in Oncology/Hematology* **96**: 475–482.
- Khan S, Imran A, Khan AA, Abul Kalam M, Alshamsan A (2016). Systems biology approaches for the prediction of possible role of Chlamydia pneumoniae proteins in the etiology of lung cancer. *PLoS One* **11**: e0148530.
- Khan S, Zaidi S, Alouffi AS, Hassan I, Imran A, Khan RA (2020). Computational proteome-wide study for the prediction of *Escherichia coli* protein targeting in host cell organelles and their implication in development of colon cancer. *ACS Omega* **5**: 7254–7261.
- Khan S, Zakariah M, Rolfo C, Robrecht L, Palaniappan S (2017). Prediction of mycoplasma hominis proteins targeting in mitochondria and cytoplasm of host cells and their implication in prostate cancer etiology. *Oncotarget* **8**: 30830–30843.
- Koyi H, Branden E, Gnarpe J, Gnarpe H, Arnholm B, Hillerdal G (1999). *Chlamydia pneumoniae* may be associated with lung cancer. *Preliminary Report on a Seroepidemiological Study. Apmis* **107**: 828–832.
- Littman AJ, Jackson LA, Vaughan TL (2005). Chlamydia pneumoniae and lung cancer: Epidemiologic evidence. *Cancer Epidemiology and Prevention Biomarkers* **14**: 773–778.
- Magat EM, Balanag GA, Carino AM, Fellizar A, Ortin TS et al. (2020). Clostridioides difficile antibody response of colorectal cancer patients versus clinically healthy individuals. *Bioscience of Microbiota, Food and Health* **39**: 123–127.
- Myers GS, Mathews SA, Eppinger M, Mitchell C, O'Brien KK et al. (2009). Evidence that human *Chlamydia pneumoniae* was zoonotically acquired. *Journal of Bacteriology* **191**: 7225–7233.
- Piton N, Wason J, Colasse E, Cornic M, Lemoine F et al. (2016). Endoplasmic reticulum stress, unfolded protein response and development of colon adenocarcinoma. *Virchows Archiv* **469**: 145–154.
- Rao KN, Elm MS, Kelly RH, Chandar N, Brady EP et al. (1997). Hepatic hyperplasia and cancer in rats: Metabolic alterations associated with cell growth. *Gastroenterology* **113**: 238–248.
- Read TD, Brunham RC, Shen C, Gill SR, Heidelberg JF et al. (2000). Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. *Nucleic Acids Research* **28**: 1397–1406.
- Ruuskanen O, Lahti E, Jennings LC, Murdoch DR (2011). Viral pneumonia. *Lancet* **377**: 1264–1275.
- Schmidt RM, Schessner JP, Borner GH, Schuck S (2019). The proteasome biogenesis regulator Rpn4 cooperates with the unfolded protein response to promote ER stress resistance. *eLife* **8**: e43244.
- Schwarz DS, Blower MD (2016). The endoplasmic reticulum: Structure, function and response to cellular signaling. *Cellular and Molecular Life Sciences* **73**: 79–94.
- Sessa R, Di Pietro M, Schiavoni G, Maccone A, Maras B et al. (2009). Chlamydia pneumoniae induces T cell apoptosis through glutathione redox imbalance and secretion of TNF-alpha. *International Journal of Immunopathology and Pharmacology* **22**: 659–668.
- Shen HB, Chou KC (2009). A top-down approach to enhance the power of predicting human protein subcellular localization: Hum-mPLoc 2.0. *Analytical Biochemistry* **394**: 269–274.
- Shen Y, Tian Z, Lu D, Huang J, Zhang Z et al. (2016). Impact of pneumonia and lung cancer on mortality of women with hypertension. *Scientific Reports* **6**: 20.
- Shirai M, Hirakawa H, Kimoto M, Tabuchi M, Kishi F et al. (2000). Comparison of whole genome sequences of *Chlamydia pneumoniae* J138 from Japan and CWL029 from USA. *Nucleic Acids Research* **28**: 2311–2314.
- Smigielski L, Aguilar GB, Kwaaitaal M, Zhang WJ, Thordal-Christensen H (2019). The isoelectric point of proteins influences their translocation to the extrahaustorial matrix of the barley powdery mildew fungus. *Cellular Microbiology* **21**: e13091.
- Stornaiuolo M, Lotti LV, Borgese N, Torrisi MR, Mottola G et al. (2003). KDEL and KKXX retrieval signals appended to the same reporter protein determine different trafficking between endoplasmic reticulum, intermediate compartment, and Golgi complex. *Molecular Biology of the Cell* **14**: 889–902.
- van Driel BE, Valet GK, Lyon H, Hansen U, Song JY, Van Noorden CJ (1999). Prognostic estimation of survival of colorectal cancer patients with the quantitative histochemical assay of G6PDH activity and the multiparameter classification program CLASSIF1. *Cytometry* **38**: 176–183.



- von Hertzen LC (1998). *Chlamydia pneumoniae* and its role in chronic obstructive pulmonary disease. *Annals of Medicine* **30**: 27–37.
- Wahl C, Maier S, Marre R, Essig A (2003). *Chlamydia pneumoniae* induces the expression of inhibitor of apoptosis 2 (c-IAP2) in a human monocytic cell line by an NF-kappaB-dependent pathway. *International Journal of Medical Microbiology* **293**: 377–381.
- Wang WA, Groenendyk J, Michalak M (2014). Endoplasmic reticulum stress associated responses in cancer. *Biochimica et Biophysica Acta* **1843**: 2143–2149.
- Xiong WM, Xu QP, Xiao RD, Hu ZJ, Cai L, He F (2019). Genome-wide DNA methylation and RNA expression profiles identified RIPK3 as a differentially methylated gene in *Chlamydia pneumoniae* infection lung carcinoma patients in China. *Cancer Management Research* **11**: 5785–5797.