Reports of Biochemistry & Molecular Biology
Vol. 4, No. 1, Oct 2015

Original article

www.RBMB.net

In silico Analysis and Molecular Modeling of RNA Polymerase, Sigma S (RpoS) Protein in *Pseudomonas aeruginosa* PAO1

Mansour Sedighi[1], Mohsen Moghoofei[2], Ebrahim Kouhsari[1], Abazar Pournajaf[1], Behzad Emadi[1], Masoud Tohidfar[3], Mehrdad Gholami*[1]

Abstract

Background: Sigma factors are proteins that regulate transcription in bacteria. Sigma factors can be activated in response to different environmental conditions. The rpoS (RNA polymerase, sigma S) gene encodes sigma-38 (σ38, or RpoS), a 37.8 kDa protein in *Pseudomonas aeruginosa* (*P. aeruginosa*) strains. RpoS is a central regulator of the general stress response and operates in both retroactive and proactive manners; not only does it allow the cell to survive environmental challenges; it also prepares the cell for subsequent stresses (cross-protection).
Methods: The significance of RpoS for stress resistance and protein expression in stationary-phase *P. aeruginosa* cells was assessed. The goal of the current study was to characterize RpoS of *P. aeruginosa* PAO1 using bioinformatics tools.
Results: The results showed that RpoS is an unstable protein that belongs to the sigma-70 factor family. Secondary structure analysis predicted that random coil is the predominant structure followed by extended alpha helix. The three-dimensional (3D) structure was modeled using SWISS-MODEL Workspace.
Conclusion: Determination of sequence, function, structure, and predicted epitopes of RpoS is important for modeling of inhibitors that will help in the design of new drugs to combat multi-drug-resistant (MDR) strains. Such information may aid in the development of new diagnostic tools, drugs, and vaccines for treatment in endemic regions.

Keywords: Bioinformatics, *In silico*, *Pseudomonas aeruginosa*, RpoS, Therapy

Introduction

Pseudomonas aeruginosa (*P. aeruginosa*) is of increasing concern in hospital settings because of frequently-emerging antibiotic resistant strains implicated in nosocomial infections. *P. aeruginosa* is a gram-negative opportunistic pathogen and a major cause of nosocomial infections in immunocompromised individuals and patients with severe burns and cystic fibrosis (1). *P. aeruginosa* can persist during feast, famine, and stress conditions in many different environments including soil, water, plants, animals, and humans. The alternative sigma factor, RpoS, positively regulates many genes in stationary phase and is considered to be a master stress-response regulator (2, 3). The rpoS gene encodes sigma factor RpoS (also called σs and σ38), which was identified in several gram-negative bacteria (4). As in *Escherichia coli* (*E. coli*), the RpoS level in *P. aeruginosa* increases upon entry of the cells into the stationary phase (2, 5). RpoS affects and regulates the production of several virulence factors

1: Department of Microbiology, Faculty of Medicine, Iran University of Medical Sciences, Tehran, Iran
2: Department of Virology, Faculty of Medicine, Iran University of Medical Sciences, Tehran, Iran
3: Agricultural Biotechnology Research Institute of Iran, Tehran, Iran
***Corresponding author: Mehrdad Gholami; Tel: +98 936-638-3466; Fax: +98 2188058649, E-mail: mehrdad_gholami90@yahoo.com**
Received: Jan 7, 2015; Accepted: Feb 20, 2015

including extracellular alginate, exotoxin A, and biofilm formation (6, 7). Furthermore, RpoS influences the expression of more than 40% of individual genes that are controlled by quorum sensing, the cell-cell communication device in gram-negative bacteria, which was identified by transcriptome analysis (8, 9). Quorum sensing has been demonstrated to be involved in the control of rpoS transcription in *P. aeruginosa* (5). In addition, *P. aeruginosa* RpoS mutants form biofilms of increased biomass and enhanced resistance to the antibiotic tobramycin (7). RpoS alters RNA polymerase core specificity and switching gene expression in the stationary phase (5, 10). The study of protein structures provides valuable information that aids in the design and development of specific inhibitors (11-13). No experimental structural information is available for most protein sequences; thus, protein structure prediction has received considerable interest in recent years. Improved understanding of protein structures will aid in the design of drugs and inhibitors for the treatment of pathogens (14, 15). The purpose of this study was to characterize RpoS of *P. aeruginosa* using on- and off-line computational tools to aid in the development of new diagnostic tools, drugs, and vaccines for treatments in endemic regions.

Materials and Methods

Sequence analysis of RpoS

P. aeruginosa was selected as the candidate organism for the present study. Its complete genome sequence (GI: 110227054) is available at http://www.ncbi.nlm.nih.gov. The protein sequence of RpoS (NP_252312) was downloaded from NCBI (http://www.ncbi.nlm.nih.gov). The RpoS sequence was analyzed for similar sequences using BlastP at http://www.ncbi.nlm.nih.gov. Proteins that were most similar to RpoS were selected and multiple alignments were performed using the ClustalW web tool (http://www.genome.jp/tools/clustalw/). A phylogenic tree was obtained using Mega4 Software.

Functional analysis of RpoS

Functional characterizations of the RpoS sequence obtained from the Uniprot site (http://www.expasy.org/) are described next.

RpoS primary, secondary, and three-dimensional structure predictions

The primary structure was predicted using the Expasy ProtParam server (http://expasy.org/cgi-bin/protparam). Post-translational modifications were predicted by the http://www.cbs.dtu.dk/researchgroups/PTM.php web site.

Physicochemical parameters such as theoretical isoelectric point, molecular weight, number of atoms, and number of positively- (Lysine, Arginine, Histidine) and negatively- (Aspartic acid, Glutamic acid) charged amino acids were determined. Hydrophobic, hydrophilic, and aromatic amino acids were identified. The instability index, estimated charge at pH 7.00, isoelectric point (pI), aliphatic index, grand average of hydropathy (GRAVY), and half-life were predicted. All physicochemical properties were predicted using the PBIL-IBCP Lyon-Gerland server (http://pbil.ibcp.fr/htm/index.php) and the ExPASy prediction tools (http://web.expasy.org/). Conserved domains were predicted by the NCBI's Conserved Domain Database (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Locations of transmembrane, intracellular, and extracellular regions were predicted by the TMHMM data bank (http://www.cbs.dtu.dk/services/TMHMM/).
The secondary structure was predicted using GOR IV and SOPMA tools and results acquired from these two banks were compared. The 3D structure of RpoS was obtained from the http://www.ncbi.nlm.nih.gov/Structure/VAST/.

Epitope prediction for RpoS

To predict potential RpoS immune epitopes we used the immune epitope database analysis resource (IEDB-AR) web-based tools (http://tools.immuneepitope.org).

Results

In silico analysis is a suitable option for finding drug targets and choosing strategies to treat and control diseases. In first step, we used BlastP and selected for further analyses 10 sequences with the greatest identities and similarities with RpoS (Table 1). Multiple alignment performed on the 10 sequences identified nine groups. Using the ClustalW alignment

program the total alignment score was 90605 (Fig. 1 and Table 2). Conserved regions of the protein were found using multiple alignments that used 323 of 334 amino acids. These conserved regions were located at amino acids 4 through 334. After that, a phylogenic tree was drawn using MEGA4 software based on the 10 sequences with the most similarity with RpoS (Fig. 2). The phylogenetic tree was based upon similarities and differences in physical and genetic characteristics. The taxa joined together in the tree are implied to have descended from a common ancestor and consist of nine groups, 10 clades, 10 taxons and 10 branches that were constructed based on bootstrap values. The tree in Fig. 1 shows bootstrap values in the inner nodes; for example, 22 means that the amino acid sequences gi|489182889 | and gi|489204806| were siblings in 22% of the bootstrap replications. The seven means that the sequences gi|553756714|, gi|489182889 |, and gi|489204806| were grouped together in what is called a monophyletic clade in 7% of the bootstrap replications, and continues consecutively. The method of bootstrap is the multinomial non-parametric bootstrap as applied in the binomial setting for each bootstrap simulation step.

Table 1. Summary of BlastP (http://www.ncbi.nlm.nih.gov/BLAST/) of the RpoS sequence

Seq.	Subject ID	Length	Max score	Total score	Query cover	E value	Identities number (%)	Positive number (%)	Gap number (%)
1	WP_003092333.1	334	676	676	100%	0.0	334/3(100%)	334/3 (100%)	0/334(0%)
2	WP_003113871.1	334	675	675	100%	0.0	334/3(100%)	334/3 (100%)	0/334(0%)
3	WP_023089396.1	334	674	674	100%	0.0	333/3(99%)	334/3 (100%)	0/334(0%)
4	WP_024928436.1	334	674	674	100%	0.0	333/3(99%)	333/3 (99%)	0/334(0%)
5	WP_003119353.1	334	674	674	100%	0.0	333/3(99%)	333/3 (99%)	0/334(0%)
6	WP_023085982.1	334	674	674	100%	0.0	333/3(99%)	334/3 (100%)	0/334(0%)
7	WP_003109327.1	613	673	673	100%	0.0	333/3(99%)	333/3 (99%)	0/334(0%)
8	WP_003143264.1	613	672	672	100%	0.0	333/3(99%)	333/3 (99%)	0/334(0%)
9	WP_024916106.1	628	671	671	100%	0.0	332/3(99%)	333/3 (99%)	0/334(0%)
10	AID 86162.1	241	670	670	99%	0.0	331/3(100%)	331/3 (100%)	0/331(0%)

```
gi|489182889|ref|WP_003092333.      MSRTHARDNDMALKKEGPEFDHDDEVLLLEPGIMLDESSADEQPSPRATP
gi|489204806|ref|WP_003113871.      ----------MALKKEGPEFDHDDEVLLLEPGIMLDESSADEQPSPRATP
gi|553756714|ref|WP_023089396.      ----------MALKKEGPEFDHDDEVLLLEPGIMLDESSADEQPSPRATP
gi|640489202|ref|WP_024928436.      ----------MALKKEGPEFDHDDEVLLLEPGIMLDESSADEQPSPRATP
gi|489210523|ref|WP_003119353.      ----------MALKKEGPEFDHDDEVLLLEPGIMLDESSADEQPSPRATP
gi|553753188|ref|WP_023085982.      ----------MALKKEGPEFDHDDEVLLLEPGIMLDESSADEQPSPRATP
gi|489200093|ref|WP_003109327.      ----------MALKKEGPEFDHDDEVLLLEPGIMLDESSADEQPSPRTTP
gi|489234972|ref|WP_003143264.      ----------MALKKEGPEFDRDDEVLLLEPGIMLDESSADEQPSPRATP
gi|640475633|ref|WP_024916106.      ----------MALKKEGPEFDHDDEVLLLEPGIMLDELSADEQPSPRATP
gi|660536842|gb|AID86162.1|         ----MQGITTWHSKKEGPEFDHDDEVLLLEPGIMLDESSADEQPSPRATP
                                              ********:*************** *********:**

gi|489182889|ref|WP_003092333.        STLEEVGQEIGLTRERVRQIQVEALKRLREILEKNGLSSDALFQ
gi|489204806|ref|WP_003113871.        STLEEVGQEIGLTRERVRQIQVEALKRLREILEKNGLSSDALFQ
gi|553756714|ref|WP_023089396.        STLEEVGQEIGLTRERVRQIQVEALKRLRDILEKNGLSSDALFQ
gi|640489202|ref|WP_024928436.        STLEEVGQEIGLTRERVRQIQVEALKRLREILEKNGLSSDALFQ
gi|489210523|ref|WP_003119353.        STLEEVGQEIGLARERVRQIQVEALKRLREILEKNGLSSDALFQ
gi|553753188|ref|WP_023085982.        STLEEVGQEIGLTRERVRQIQVEALKRLREILEKNGLSSDALFQ
gi|489200093|ref|WP_003109327.        STLEEVGQEIGLTRERVRQIQVEALKRLREILEKNGLSSDALFQ
gi|489234972|ref|WP_003143264.        STLEEVGQEIGLTRERVRQIQVEALKRLREILEKNGLSSDALFQ
gi|640475633|ref|WP_024916106.        STLEEVGQEIGLTRERVRQIQVEALKRLREILEKNGLSSDALFQ
gi|660536842|gb|AID86162.1|           STLEEVGQEIGLTRERVRQIQVEALKRLREILEKNGLSSDALFQ
                                      ************:***************:**************
```

Fig. 1. Results of multiple alignments by ClustalW

Table 2. The nine alignment groups determined by ClustalW

Group	Sequences	Score
1	2	5457
2	3	5451
3	4	5447
4	5	5445
5	6	5447
6	7	5445
7	8	5440
8	9	5435
9	10	5421
Alignment Score = 90605		

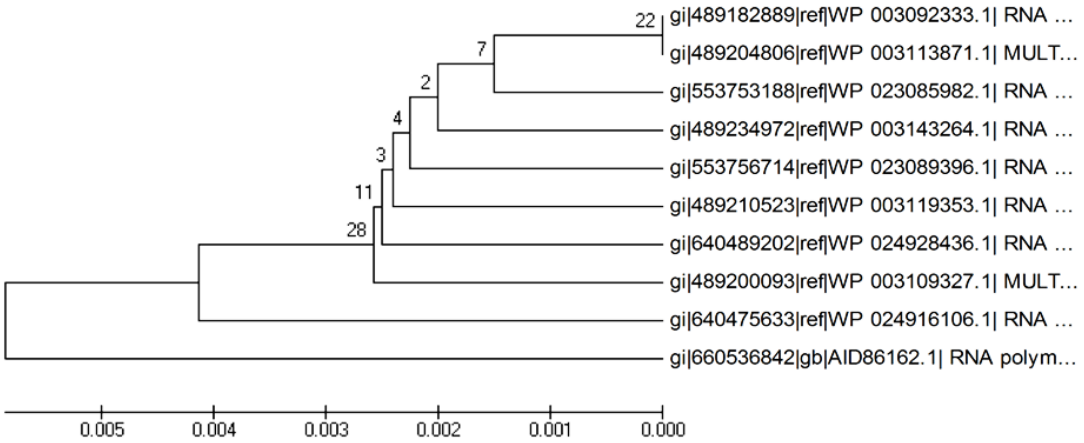

Fig. 2. RpoS phylogenic tree based on bootstrap values, by MEGA4 software

Functionality analysis of RpoS identified characteristic domains (Table 3). Two motifs were identified; these are the 20-amino acid HTH motif and the four-amino acid motif that interacts with the RpoC polymerase core subunit. RpoS contains four domains of 34, 71, 76, and 54 amino acids for the Sigma-70 factor. Sigma factors are initiation factors that promote the attachment of RNA polymerase to specific transcription initiation sites. This sigma factor is the master transcriptional regulator of the stationary phase and the general stress response.

Table 3. Domains, repeats, motifs, and features of RpoS

Name	Start	End	length	E value
H-T-H motif	293	312	20	-
Pfam: Sigma-70 factor domain-1	61	94	34	4.8e-16
Pfam: Sigma-70 factor domain-2	99	169	71	3e-25
Pfam: Sigma-70 factor domain-3	179	254	76	4.9e-22
Pfam: Sigma-70 factor domain-4	267	320	54	8.4e-18
Interaction with polymerase core subunit RpoC motif	123	126	4	-

Primary structure prediction of RpoS

Analysis of the primary structure of RpoS showed that theprotein contained 334 amino acids with an estimated molecular weight of 38235.3 (Table 4). Forty-seven positively-charged (Arg+Lys) and 60 negatively-charged (Asp+Glu) residues were identified.

Phosphorylation sites were predicted at serines 28, 29, 35, 47, 120, 148, 199, 228, 230, 235, and 291, and threonines 39, 44, 77, 158, 169, 191, 225, 238, 272 and 292. Tyrosines were not predicted to be phosphorylated. O-Glycosylation sites were found on serines 28, 35, 45, 47, and 226, and on other amino acids at residues 43, 44, and 225. N-glycosylation sites were identified on aasparagine 165 and lysines 4, 5, 49, 97, 109, 138, 193, 316, and 324. C-mannosylation sites were identified on tryptophans 153, 154, and 267. No acetylation sites were identified.

The protein contains 5415 atoms and its chemical formula is C1680H2724N482O523S6. The isoelectric point (pI) is 5.26. The aliphatic index of a protein is defined as the relative volume occupied by aliphatic side chains, which include alanine, valine, isoleucine, and leucine, and contributes to protein thermostability. The aliphatic index of RpoS is 95.75, indicating this protein is unstable for a wide range of temperatures. The instability index of RpoS is 49.78, indicting a relatively short half-life.

The GRAVY value is -0.547, indicating that RpoS is relatively non-polar The GRAVY value is determined by the sum of hydropathy values of all amino acids in a protein divided by the protein length. The calculation is based on the Kyte-Doolittle scale. RpoS contains 109 hydrophobic (ALIVMW), 170 hydrophilic (DEKNQRST), 17 aromatic (FYW), and 46 hydroxyl (STY) residues. The estimated half-life is 30 hours in mammalian reticulocytes *in vitro*, >20 hours in yeast *in vivo*, and >10 hours in *E. coli*, *in vivo*.

The half-life is a prediction of the time required for half of a protein in a cell to degrade after its synthesis. ProtParam relies on the "N-end rule", which relates the half-life of a protein to the identity of its N-terminal residue; the prediction is given for three model organisms; human, yeast, and *E. coli*. Identity of the N-terminal residue of a protein plays an important role in determining its stability *in vivo*. Proteins have strikingly different half-lives *in vivo*, from seconds to hours, depending on the nature of the amino acid at the N-terminus and the different models.

Table 4. Physicochemical properties of RpoS

Amino acid	No. of residues	% of residues
Ala (A)	19	5.7%
Arg (R)	30	9.0%
Asn (N)	9	2.7%
Asp (D)	24	7.2%
Cys (C)	1	0.3%
Gln (Q)	13	3.9%
Glu (E)	36	10.8%
Gly (G)	16	4.8%
His (H)	8	2.4%
Ile (I)	18	5.4%
Leu (L)	45	13.5%
Lys (K)	17	5.1%
Met (M)	5	1.5%
Phe (F)	9	2.7%
Pro (P)	16	4.8%
Ser (S)	20	6.0%
Thr (T)	21	6.3%
Trp (W)	3	0.9%
Tyr (Y)	5	1.5%
Val (V)	19	5.7%
Pyll (O)	0	0.0%
Sec (U)	0	0.0%

Five conserved domains were predicted in RpoS using PRS-Blast (Table 5). The first domains refer to the C-termini of four conserved domains. The second domain is the most conserved region of the protein. The third domain forms a discrete compact three helical domain within the sigma factor. The forth domain is region 1-2. The fifth domain is RNA polymerase sigma factor RpoS; Validated. The TMHMM program predicts the locations of transmembrane, intracellular, and extracellular domains. As seen in Fig. 3, RpoS is predicted to be entirely extracellular (Fig. 3).

Table 5. Predicted conserved domains of RpoS by PRS-Blast

Name	Accession	Description	Interval	E-value
Sigma70_r4	cd06171	Sigma70, region (SR) 4 refers to the most C-terminal of four conserved domains found. ...	262-319	2.11e-10
Sigma70_r2	pfam04542	Sigma-70 region 2; Region 2 of sigma-70 is the most conserved region of the entire protein. ...	99-169	6.95e-24
Sigma70_r1_2	pfam00140	Sigma-70 factor, region 1.2	61-96	8.25e-10
PRK05657	PRK05657	RNA polymerase sigma factor RpoS; Validated	1-333	0e+00

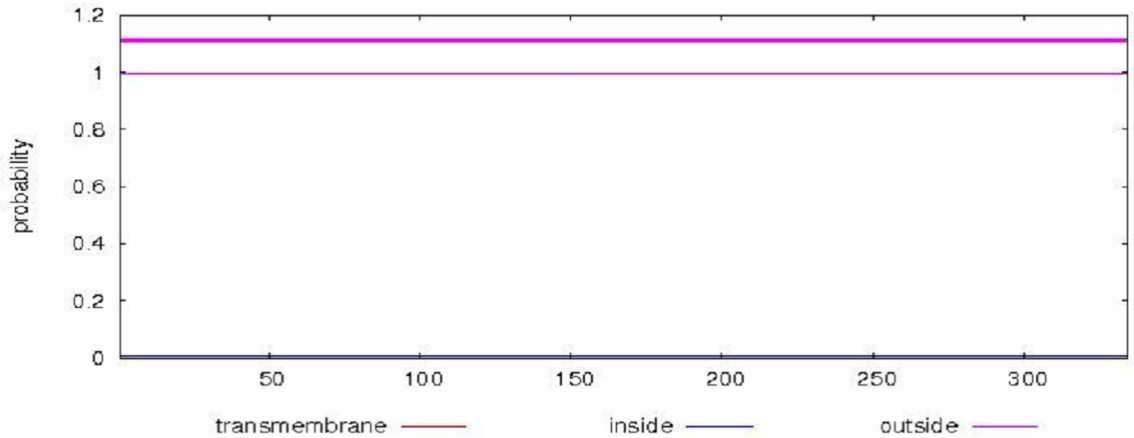

Fig. 3. Graphical representation of the location of RpoS protein by TMHMM: RpoS is predicted to be entirely extracellular

Secondary structure prediction of RpoS

The secondary structure of RpoS was predicted using GOR IV and SOPMA web-site tools. Table 6 shows the results obtained from the two programs. Both programs predicted that alpha helices (50.60 and 57.78%, respectively), random coils (40.42 and 28.14%, respectively), and extended strands (8.98 and 9.28%, respectively) were the most abundant structures. This is graphically represented in Fig. 4.

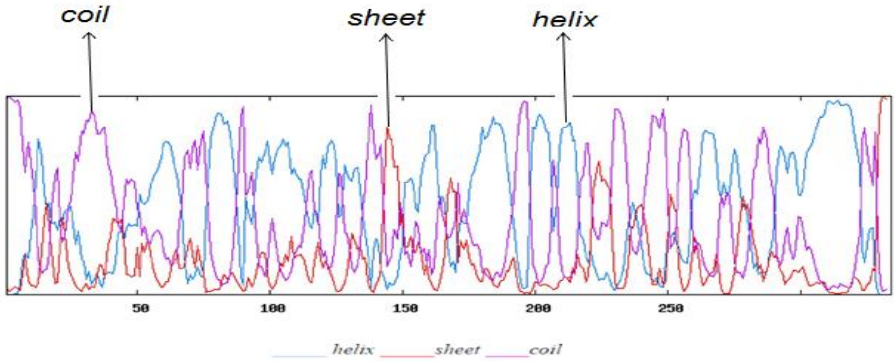

Fig. 4. Secondary structure prediction of RpoS by Gor IV

Table 6. Prediction of secondary structure of RpoS by SOPMA and GOR IV

Secondary structure	GOR IV		SOPMA	
	Length	Percent	Length	Percent
Alpha helix (Hh)	169	50.60%	193	57.78%
310 helix (Gg)	0	0.00%	0	0.00%
Pi helix (Ii)	0	0.00%	0	0.00%
Beta bridge (Bb)	0	0.00%	0	0.00%
Extended strand (Ee)	30	8.98%	31	9.28%
Beta turn (Tt)	0	0.00%	16	4.79%
Bend region (Ss)	0	0.00%	0	0.00%
Random coil (Cc)	135	40.42%	94	28.14%
Ambiguous states	0	0.00%	0	0.00%
Other states	0	0.00%	0	0.00%
Sequence length	334		334	

Three-dimensional structure prediction of RpoS

Because only primary sequence information about RpoS was available from NCBI, and no 3D structure information in the form of X-ray crystallographic data was available from the Protein Data Bank (PDB), modeling of RpoS was necessary to determine the protein's 3D structure. The 3D structure was predicted using the first protein obtained from the BlastP search against the PDB (Table 7). The first protein identified via this process was associated with locus 1L9U_H, encoding a protein from *Thermus aquaticus* (*T. aquaticus*) bacteria. This sequence has the greatest similarity with RpoS and its 3D structure was accessible in the PDB. 1L9U is an RNA polymerase holoenzyme containing core RNA polymerase (alpha2betabeta'omega) and the promoter specificity sigma subunit, including A, B, C, D, E, and H. The H chain (1L9U_H) is sigma factor sigA in *T. aquaticus* with the greatest similarity and identity with RpoS in the PDB BlastP analysis of RpoS, and IL9U_H resulted in a maximum score of 204 and 46% identity.

The modeling and 3D structure of RpoS was based on 1L9U_H. Here homology modeling was predicted using SWISS PDB Viewer. The structure of the protein was refined further to improve the model (Figs. 5-7). Fig. 5 is a spatial model of RpoS using yellow ribbons and cylinders that represent α-helices and β-sheets, respectively. Fig. 6 is related to charge distribution on the RpoS Connolly surface in the form of colored spheres. In Fig. 6 the blue, red, and gray spheres represent positively, negatively, and neutrally charged amino acids, respectively. Fig. 7 displays the hydrophobicity distribution on the RpoS Connolly surface in the form of colored spheres representing hydrophobic, hydrophilic, and neutral amino acids.

Table 7. Summary of the BlastP analysis (http://www.ncbi.nlm.nih.gov/BLAST/) of RpoS

Seq.	Subject ID	Length	Max score	Total score	Query cover (%)	E value	Identities number (%)	Positive number (%)	Gap number (%)
1	1L9U_H	332	204	204	72%	6e-62	112.2 (46%)	163.2 (66%)	5.2 (2%)
2	1L9Z_H	438	204	204	72%	7e-61	112.2 (46%)	163.2 (66%)	5.2 (2%)
3	1IW7_F	423	202	202	72%	2e-60	110.2 (45%)	163.2 (66%)	5.2 (2%)
4	3DXJ_F	423	202	202	72%	2e-60	110.2 (45%)	163.2 (66%)	5.2 (2%)
5	4G7H_F	443	202	202	72%	5e-60	110.2 (45%)	163.2 (66%)	5.2 (2%)
6	4LJZ_F	522	201	201	78%	6e-59	107.3(40%)	166.3(62%)	7.3(2%)
7	3IYD_F	613	201	201	78%	2e-58	107.3(40%)	166.3(62%)	7.3(2%)
8	4IGC_X	613	201	201	78%	2e-58	107.3(40%)	166.3(62%)	7.3(2%)
9	4JKR_F	628	201	201	78%	2e-58	107.3(40%)	166.3(62%)	7.3(2%)
10	1KU2_A	241	155	155	41%	2e-44	78.1 (56%)	104.1 (74%)	2.1(1%)

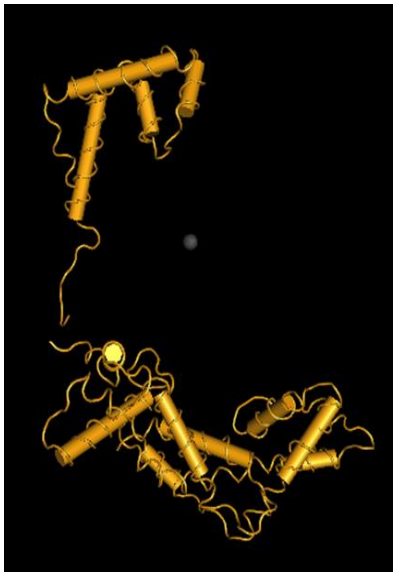

Fig. 5. Modeled spatial configuration of RpoS representing α-helices and β-sheets.

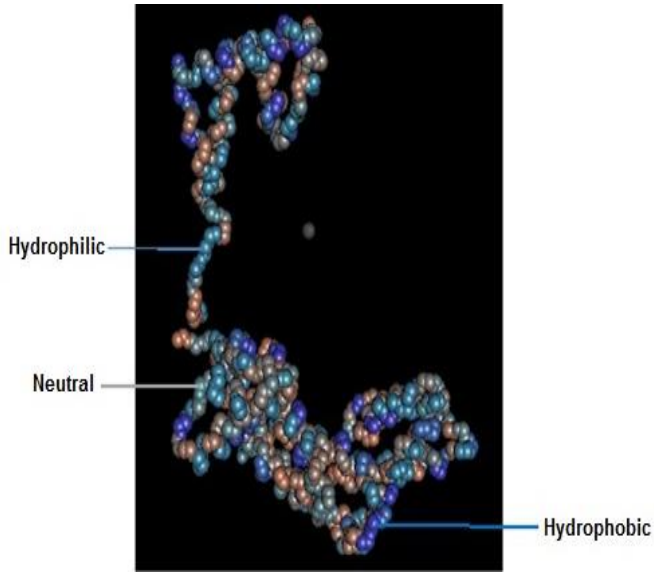

Fig. 6. Diagram of charge distribution on RpoS Connolly surface (the blue, red and grey spheres represent positively-charged, negatively-charged, and neutral amino acids, respectively).

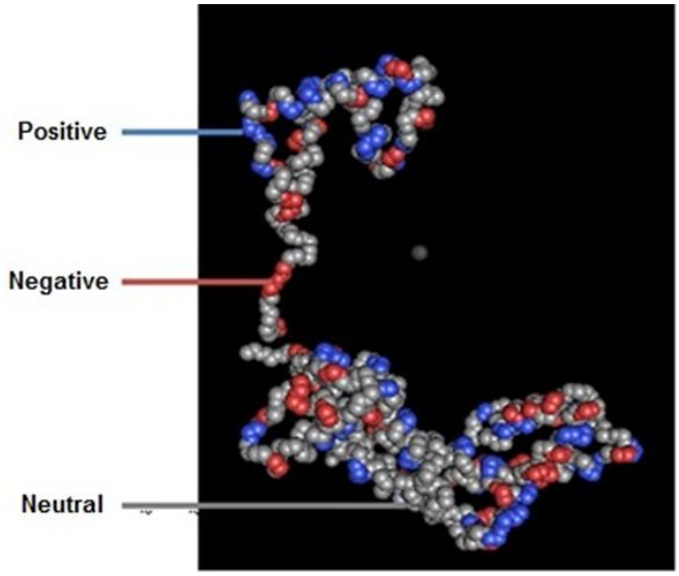

Fig. 7. Diagram of hydrophobicity distribution on RpoS Connolly surfaceThe dark blue, light blue and gray spheres represent hydrophobic, hydrophilic, and neutral amino acids, respectively.

An epitope, also known as antigenic determinant, is the part of an antigen that is recognized by the immune system, specifically by antibodies, B cells, or T cells. Predicted B-cell epitopes of RpoS are presented in Table 8. RpoS contains 14 epitopes with lengths from one to 24 amino acids. The longest epitope contains 24 residues, with six (25%) serines. This epitope probability can be considered as a potential binding site for an inhibitor or drug target. The shorter epitopes of one, two, and three residues are less likely to be regarded as potential inhibitor targets due to their small sizes.

Table 8. Predicted B-cell epitopes in RpoS

No.	Start Position	End Position	Peptide	Peptide Length
1	4	13	KKEGPEFDHD	10
2	26	49	DESSADEQPSPRATPKATTSFSSK	24
3	51	51	H	1
4	76	82	PLLTPEE	7
5	90	98	AQKGDPAGR	9
6	139	147	EKFD PER	9
7	196	205	HKLDHEPSPE	10
8	214	215	P	1
9	232	241	SVDVSLGPDS	10
10	247	258	DTLTDDRPTDPC	12
11	262	266	QDDDL	5
12	275	277	TEL	3
13	292	301	GLRGHESSTL	10
14	330	332	LEK	3

Discussion

P. aeruginosa is a leading cause of nosocomial infections in immune compromised individuals (16). Such infections are especially difficult to treat because of emerging antibiotic resistant strains, a trend that could be attributed to selective pressures from antibiotic treatment, as well as the organism's intrinsic ability to adapt to drug-stressed environments (17, 18). *P. aeruginosa* can persist during feast, famine, and stress conditions such as nutrient limitation, antibiotics, heat, osmotic, and high pH in many different environments (2). In particular, RpoS, a sigma factor, is a general stress response regulator that affects the transcription of genes that confer increased tolerance to stress conditions. Little information is available regarding the characterization and structure of RpoS in *P. aeruginosa* (19, 20). The focus of this study was to characterize RpoS of *P. aeruginosa* PAO1 using bioinformatics tools to apply the knowledge gained to this protein as a drug or inhibitor target. The aliphatic index of 95.75 indicates that the protein is unstable at a wide temperature range. In this work the primary, secondary, and 3D structures of RpoS were analyzed *in silico*. The primary structure determines the 3D shape known as the tertiary structure. The tertiary structure of a protein is related to protein's geometric shape and describes the folding of its secondary structural elements and specifies the position of each atom in the protein, including those of its side chains (21). The alpha-helixes and beta-pleated sheets fold into a compact globular structure (22). In addition to, globular proteins contains cores of hydrophobic amino acid residues. Also have surface regions of water-exposed, charged, hydrophilic residues. The common features of protein tertiary structures reveal much about the biological functions of the proteins and their evolutionary origins (23).

The IEDB was used to predict likely RpoS epitopes. These epitopes can be considered for the design of drugs or inhibitors. Currently, a wide range of antibiotics against the membrane, wall, enzymes, nucleic acids, ribosomes, and metabolic pathways of *P. aeruginosa* are used to treat infections caused by these bacteria. The emergence of MDR and XDR strains has resulted in resistance to many of these drugs (24); therefore, new and effective treatments are needed. Numerous *In silico* studies that examined the effects of various drugs on a variety of pathogenic bacteria indicated that the drug targets are mostly proteins and enzymes. For example, Munikumar et al., using pathway analysis, reported that the putative drug targets of the *Streptococcus pneumonia* that causes bacterial meningitis included 26 enzymes, 8 non-enzyme proteins, and 3 conserved hypothetical proteins (25). A similar *In silico* analysis by Neema et al. identified and characterized 3986 proteins, including seven novel membrane proteins, in *Edwardsiella tarda*. One hundred and seventy-one of these proteins were proposed as potential drug targets (26). Another

In silico study by Perumal et al., also using pathway analysis, identified eight putative drug targets in *P. aeruginosa* (27). Knowledge of the 3D structures of proteins will aid in the development of reasonable strategies for specific inhibitor design (28, 29). RpoS positively regulates many genes in stationary phase and is considered to be a master stress-response regulator (30). Understanding of this biosynthetic pathway and bioinformatics information are necessary for treatment of *P. aeruginosa* infections and will aid in the search for new anti-*Pseudomonas* agents. Therefore, the goal of this study was to examine the structure of *P. aeruginosa* RpoS at the molecular level using bioinformatics tools. Using *In silic*o analysis, the long-term goal is to design drugs and inhibitors against RpoS.

RpoS inhibitor modeling will aid in the design of new drugs to combat MDR strains of *P. aeruginosa*. This approach enables rapid potential drug target identification, thereby greatly facilitating the search for new inhibitors.

Acknowledgements

We would like to appreciate Dr. Gholamreza Irajian and Dr. Shabnam Razavi for their technical assistance.

References

1. Firme M, Kular H, Lee C, Song D. RpoS Contributes to Variations in the Survival Pattern of *Pseudomonas aeruginosain* Response to Ciprofloxacin. Journal of Experimentall Microbiology and Immunology. 2010; 14:21-27.

2. Jørgensen F, Bally M, Chapon-Herve V, Michel G, Lazdunski A, Williams P, Stewart GS. RpoS-dependent stress tolerance in *Pseudomonas aeruginosa*. Microbiology. 1999; 145:835-44.

3. Ishihama A. Functional modulation of *Escherichia coli* RNA polymerase. Annu. Rev. Microbiol. 2000; 54:499-518.

4. Suh SJ, Silo-Suh L, Woods DE, Hassett DJ, West SE, Ohman DE. Effect of rpoS mutation on the stress response and expression of virulence factors in *Pseudomonas aeruginosa*. J. Bacteriol. 1999; 181:3890-97.

5. Kojic M, Venturi V. Regulation of rpoS Gene Expression in Pseudomonas: Involvement of a TetR Family Regulator. Journal of bacteriology. 2001; 183(12):3712-20.

6. Balasubramanian D, Schneper L, Kumari H, Mathee K. A dynamic and intricate regulatory network determines *Pseudomonas aeruginosa* virulence. Nucleic Acids Research. 2012; 1-20.

7. Schuster M., Hawkins AC, Harwood CS, Greenberg EP. The *Pseudomonas aeruginosa* RpoS regulon and its relationship to quorum sensing. Mol. Microbiol. 2004; 51:973-985.

8. Whiteley M, Parsek MR, Greenberg EP. Regulation of Quorum Sensing by RpoS in *Pseudomonas aeruginosa*. Journal of Bacteriology. 2000; 182(15):4356-60.

9. Coggan KA, Wolfgang MC. Global Regulatory Pathways and Cross-talk Control *Pseudomonas aeruginosa* Environmental Lifestyle and Virulence Phenotype. Curr. Issues Mol. Biol. 2012; 14(2):47-70.

10. Farewell A, Kvint K, Nystrom T. Negative regulation by RpoS: a case of sigma factor competition. Molecular Microbiology. 1998; 29(4):1039-51.

11. Kopp J, Schwede T. The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models. Nucleic Acids Research. 2004; 32:230-34.

12. Singh Sh, Sharma DK. *In silico* modeling in conjunction with natural products: Paving the way for rational drug-design. Biotechnology and Molecular Biology Review. 2011; 6(4):88-93.

13. Daniel JM, Tanja K. Computer-aided design of functional protein interactions. Nature Chemical Biology. 2009; 5:797-807.

14. Jónsdóttir SO, Jørgensen FS, Brunak S. Prediction methods and databases within chemoinformatics: emphasis on drugs and drug candidates. Bioinformatics. 2005; 21(10):2145-60.

15. Lahti JL, Tang GW, Capriotti E, Liu T, Altman RB.Bioinformatics and variability in drug response: a protein structural perspective. J R Soc Interface. 2012; 9(72):1409-37.

16. Faghri J, Pourentezari M, Esmaeily M, Pirouzi S, Sedighi M. Prevalence of metallo-beta-lactamase genes blaVIM -1 and blaSPM-1 in *Pseudomonas aeruginosa* Clinical Isolates in Isfahan, Iran. Global Journal of Medicine Researches and Studies. 2014; 1(1):20-27.

17. Sedighi M, Peyvandi M, Dehghan M, Raissi V, Pourentezari M, Faghri J. Antimicrobial Resistant Pattern of *Pseudomonas aeruginosa* Strains Isolated from Patients Referring to Hospitals of Isfahan, Iran. Global Journal of Medicine Researches and Studies. 2014; 1(1): 6-10.
18. Firme M, Kular H, Lee C, Song D. RpoS Contributes to Variations in the Survival Pattern of *Pseudomonas aeruginosa* in Response to Ciprofloxacin. Journal of Experimentall Microbiology and Immunology. 2010; 14:21-27.
19. Chung HJ, Bang W, Drake MA. Stress Response of *Escherichia coli*. Comprehensive reviews in food science and food safety. 2006; 5:52-64.
20. Loewen PC, Hengge-Aronis R. The role of the sigma factor δs (katf) bacterial global regulation. annu. rev. Microbiol. 1994; 48:53-80.
21. Branden C, Tooze J. Introduction to Protein Structure. Garland Publishing. New York. 1991.
22. Govindarajan S, Recabarren R, Goldstein RA. Estimating the total number of protein folds. Proteins. 1999; 35 (4):408-414.
23. Hochreiter S, Chelbat N. Bioinformatics III, Structural Bioinformatics and Genome Analysis. Institute of Bioinformatics, Johannes Kepler University Linz. 2007: 1-126.
24. Gill MM, Usman J, Kaleem F, Hassan A, Khalid A, Anjum R, Fahim Q. Frequency and Antibiogram of Multi-drug Resistant *Pseudomonas aeruginosa*. Journal of the College of Physicians and Surgeons Pakistan. 2011; 21(9):531-534.
25. Munikumar M, Priyadarshini V, Pradhan D, Sandeep S, Umamaheswari A, Vengamma B. *In Silico* Identification of Common Putative Drug Targets among the Pathogens of Bacterial Meningitis. Biochem Anal Biochem. 2012; 1(8):1-7.
26. Neema M, Karunasagar I, Karunasagar I. *In silico* identification and characterization of novel drug targets and outer membrane proteins in the fish pathogen *Edwardsiella tarda*. Open Access Bioinformatics. 2011; 3:37-42.
27. Perumal D, Sing Lim C, Sakharkar MK. *In silico* Identification of Putative Drug Targets in *Pseudomonas aeruginosa* Through Metabolic Pathway Analysis. Pattern Recognition in Bioinformatics. 2007; 4774:323-36.
28. Kubinyi H. Structure-based design of enzyme inhibitors and receptor ligands. Current Opinion in Drug Discovery and Development. 1998; 1(1):4-15.
29. Rahuel J, Rasetti V, Maibaum J, Rüeger H, Göschke R, Cohen NC, Stutz S, Cumin F, Fuhrer W, Wood JM, Grütter MG. Structure-based drug design: the discovery of novel nonpeptide orally active inhibitors of human renin. Chemistry & Biology. 2000; 7:493-504.
30. Murakami K, Ono T, Viducic D, Kayama Sh, Mori M, Hirota K, et al. Role for rpoS gene of *Pseudomonas aeruginosain* antibiotic tolerance. FEMS Microbiology Letters. 2005; 242:161-67.