

Regular Article

Molecular characterization of *TatD* DNase gene from *Staphylococcus pasteurii* RA3^T

Amin Alborzian Deh Sheikh and Ramin Hosseini*

Department of Agricultural Biotechnology, Imam Khomeini International University, Qazvin, Iran

Corresponding author Email: raminh_2001@yahoo.com

Staphylococcus pasteurii strain RA3^T a gram positive, non-spore forming, non-motile, catalase positive was isolated from surface soil. The identification of the bacterium was confirmed by 16S rDNA phylogeny and biochemical tests. TatD is a cytoplasmic protein and exhibits a magnesium-dependent DNase activity. The TatD family has no involvement in protein export by the Tat system. A TatD DNase gene was isolated and cloned from *Staphylococcus pasteurii* strain RA3^T genome. Nucleotide sequence analysis revealed 774 nucleotides in length encoding a protein of 257 amino acid residues. The TatD gene showed a high similarity to homologs belonging to *Staphylococcus warneri* SG1. The deduced polypeptide sequence harbors a typical catalytic site, HHRPEDHRHFSSGEDPLN and its calculated molecular mass and its predicted isoelectric point are 29656.7 Da and 5.05, respectively. The deduced amino acid sequence showed a high degree of similarity to TatD DNase isoforms from *Staphylococcus* genus and other sources. Three dimensional predictions of TatD confirmed the active site and its theoretical functions as DNase.

Keywords: Soil bacteria, TatD DNase, *Staphylococcus*.

Abbreviations: TatD: Twin arginine translocation type D; DNase: Deoxyribonuclease

Staphylococcus aureus is a human pathogen that can cause a variety of disease states, including minor skin and soft tissue infections and life threatening systemic and pulmonary infections. *Staphylococcus aureus* is a phylogenetically well defined species. Orthologous pairs of housekeeping coding sequences exhibit in general, 2% nucleotide diversity within *S. aureus*, which is at least 10-fold lower than the diversity between *S. aureus* and the most closely related *Staphylococcus* species (Enright et al., 2000; Poyart et al., 2001; Drancourt and Raoult,

2002; Ghebremedhin et al., 2008). In 1993, *S. pasteurii* was introduced as an environment and clinical contaminant (Chesneau et al., 1993). The translocation of proteins across biological membranes usually requires the existence of N-terminal signal sequences (Blobel, G. and Dobberstein, 1975). Bacterial signal sequences are generally constituted by a positively charged N-terminal (n-) region, followed by an uncharged hydrophobic (h-) region and in most cases by a C-terminal cleavage site (c-) region. Although signal sequences of secreted proteins are similar in

their overall structure, they have features which direct proteins to either the general secretory (Sec), or to the twin-arginine translocation (Tat) systems (Pugsley, 1993; Bruser and Sanders, 2003). The twin-arginine translocation (Tat) pathway was originally identified in chloroplasts and has recently been found in bacteria and archaea. It is distinct from the Sec pathway in that; Tat substrates are secreted in a folded conformation, Tat signal peptides contain a highly conserved twin-arginine motif, the energy driving translocation is provided solely by the proton motive force and the Tat pathway is not a universally conserved secretion mechanism (Kieran et al, 2003). Signal sequences of Tat substrates contain usually a 'twin-arginine motif' in their n-region (Stanley et al, 2000). In addition, they are relatively long, less hydrophobic in their h-region and often charged in their c-region (Bruser et al, 1998; Cristobal et al, 1999). Tat substrates can be translocated in a folded state (Rodrigue et al, 1999; Thomas et al, 2001). As incorrectly folded proteins are not efficiently translocated, it has been proposed that a further determinant of Tat substrates is their folded state (Halbig et al, 1999; Sanders et al, 2001). The twin arginine signal peptides, target the precursor protein to the recently discovered Tat protein export system that is mechanistically and structurally related to the Δ pH-dependent thylakoid import pathway of chloroplasts (Margaret et al, 2000). The Sec system is a major secretion system in various bacterial species and the role of the Tat system is unknown in a number of bacterial species (Keiko et al, 2007).

In this project, *Staphylococcus pasteurii* was isolated from soil. The presence of TatD DNase was studied. Finally, TatD gene from *S. pasteurii* was cloned, sequenced and characterized bioinformatically. This is the first report about *Staphylococcus pasteurii*.

Materials and Methods

Bacteria isolation

Fifty surface soil samples were collected from experimental farm of Imam Khomeini International University-Qazvin, Iran (36°16'09"N 50°00'10"E, 36.2693°N 50.0029°E). The soils were diluted serially to 10⁻⁶ concentration in autoclaved water (w/v). One ml of dilution was streaked on LB broth agar. After 3-7 days, colonies were examined by some physiological and biochemical tests.

DNA extraction and PCR

Bacterial genomic DNA was isolated using the method described by Zhou et al (1996). Quality of the DNA was monitored by 0.8% agarose gel electrophoresis and the concentration was determined by they spectrophotometric measurement. Amplification of 16S rDNA gene was performed by using two sets of constructed oligonucleotide universal primers 27F 5'-AGA GTT TGA TCM TGG CTC AG-3' and 1525R 5'-AAG GAG GTG WTC CAR CC-3'. The reaction mixture in a final volume of 50 μ l consisted of 20 mM Tris-HCl (pH 8.8), 10 mM (NH₄)₂SO₄, 10 mM KCl, 2 mM MgSO₄, 0.1% Triton X-100, 0.1 mg/ml BSA, 500 μ M of each dNTP, 50 pM of each primer (forward and reverse primers), 100 ng template DNA and 1.25 units *Pfu* DNA polymerase (Fermentas). The PCR reaction was carried out using a thermal cycler (Techne, U.K.) programmed under the following conditions: 3 min at 94 °C as initial denaturation, 30 s at 94 °C, 1 min at 55° C, 1 min at 72°C for 35 cycles and the final extension of 10 min at 72°C. The PCR product was separated on a 0.8% agarose gel and then the amplified fragments with the expected size were excised from the gel and purified using GF-1 PCR Clean Up Kit (Vivantis). The nucleotide sequence of the purified fragments was determined in both directions by dideoxynucleotide sequencing method (Bioneer Sequencing Company, South Korea).

Bacteria phylogeny

The 16S rDNA sequence was extracted from EMBL/NCBI databases. After identification of the most similar organism, multiple alignment and construction of phylogenetic trees were performed using ClustalW (<http://clustalw.genome.ad.jp>) and MEGA4 software with the most similar species. A bootstrap test with 1000 replicates was performed for the phylogenetic studies (Tamura et al, 2007; Felsenstein, 1985).

PCR and cloning of gene encoding TatD DNase

Amplification of the *S. pasteurii* TatD gene was performed by direct PCR using two sets of constructed degenerated oligonucleotides as forward 5'-ATG YTA ATM GAT ACA CAT GTW CAT TTA AA-3' and reverse 5'-TTA TKY TTT TAA WTT AAA CAA WCG TTC TGC-3' primers based on the available three sequences from *Staphylococcus aureus* Mu50, *Staphylococcus carnosus* TM300 and *Staphylococcus saprophyticus* ATCC 15305 (NC_002758.2, NC_012121.1 and NC_007350.1, respectively), identified by the BLAST program (<http://www.ncbi.nlm.nih.gov>). The upstream oligonucleotides were synthesized homologous to the coding strand and the downstream oligonucleotides were complementary and reverse to the coding strand. The reaction mixture in a final volume of 50 µl consisted of 20 mM Tris-HCl (pH 8.8), 10 mM (NH₄)₂SO₄, 10 mM KCl, 2 mM MgSO₄, 0.1% Triton X-100, 0.1 mg/ml BSA, 500 µM of each dNTP, 50 pM of each primer (forward and reverse primers), 100 ng template DNA and 1.25 units *Pfu* DNA polymerase (Fermentas). The PCR reaction was carried out using a thermal cycler (Techne, U.K.) programmed under the following conditions: 3 min at 94 °C as the initial denaturation, 30 s at 94 °C, 1 min 60 °C, 1 min 72 °C for 35 cycles and the final extension of 10 min at 72°C. The PCR product was separated on a 0.8% agarose gel then the amplified fragments with the expected size

were excised from the gel and purified using GF-1 PCR Clean Up Kit (Vivantis). The purified fragments were cloned into the pTG19-T plasmid vector (Vivantis) to generate the pTatD DNase plasmid. The ligation sample was directly used to transform the competent *Escherichia coli* strain DH5α cells as described by Sambrook and Russell (2001). After screening, the recombinant plasmids were isolated and purified. Then the nucleotide sequence of the insert was determined in both directions by dideoxynucleotide sequencing method (Bioneer Sequencing Company, South Korea).

Nucleotide and amino acid sequence analysis

The properties of deduced amino acid sequence were estimated using ProtScale, TMHMM, ScanProsite, CSS-PALM (links available at <http://www.expasy.ch>), ProtParam (Gasteiger et al, 2005) and Conseq (Ashkenazy et al, 2010) programs. The subcellular localization prediction of *S. pasteurii* TatD was performed by using a combination of three programs, TargetP (Emanuelsson et al, 2007), iPSORT (Bannai et al, 2002) and YLOC (Briesemeister et al, 2010), and secondary structure was determined by SOPMA (Geourjon et al, 1995). The deduced protein sequence was searched for homologous proteins in different databases using BLAST network services at the National Center for Biotechnology Information (NCBI) and several TatD DNases sequences were selected with the highest score from different organisms. Multiple alignment and construction of phylogenetic trees were performed using ClustalW) and MEGA4 software.

Prediction of 3D structure

The three-dimensional structure of *S. pasteurii* TatD was predicted using I-TASSER (Zhang, 2008) with the crystal structure of *Thermotoga maritima* (PDB ID code 1j6o) (Peterson et al,

2005) as a template. Cofactors of TatD enzyme were predicted using I-TASSER CoFACTOR program. Superimposition analysis of the 3D models of TatD DNase and its templates, such as; human (*Homo sapiens*) TatD (PDB ID code 2y1h), *E. coli* TatD (PDB ID code 1yix) and *Staphylococcus aureus* TatD (PDB ID code 2gzx) was carried out using 3-Dimensional Structural Superposition (3d-SS) service (Russell et al, 1992). Conserved amino acids at the protein surface were determined using ConSurf (Landau et al, 2005) and functionally important regions were also identified in protein by PatchFinder (Nimrod et al, 2008).

Accession numbers

The NCBI, EMBL and SwissProt accession numbers for the sequences described and mentioned in this study are as follows: 16S rDNA from *Staphylococcus pasteurii* HA3 (KC170006.1); *S. simiae* (DQ127902.1); *S. lutrae* (AB233333.1); *S. sciuri* (AB233331.1); *S. capitis* (AB233325.1); *S. schleiferi* (AB233334.1); *S. fleurettii* (AB233330.1); *S. cohnii* (AB233328.1); *S. hominis* (AB233326.1); *S. aureus* (L37597.1); *S. pasteurii* (JN102556.1) and *S. warneri* (NR_102499.1). TatD genes from *S. pasteurii* HA3 (KC608821); *S. warneri* (AGC91492.1); *S. caprae* (EFS18295.1); *S. epidermidis* (EKS26995.1); *S. capitis* (EGS39692.1); *S. hominis* (EHR89663.1); *S. lugdunensis* (EHS05591.1); *S. aureus* (pdb 2gzx); *S. haemolyticus* (BAE05829.1); *S. saprophyticus* (BAE19410.1); *S. carnosus* (CAL27054.1); *S. massiliensis* (EKU46147.1); *S. simiae* (EHJ07788.1); *S. simulans* (EKS23227.1); *S. pettenkoferi* (EHM65380.1); *S. arlettae* (EJY95897.1); *S. pseudintermedius* (ADX77539.1); *Homo sapiens* (pdb 2y1h) and *E. coli* (pdb 1yix).

Results and Discussion

Bacterium Isolation and Identification

A gram positive, motile, resistant to salt (up to 15g/L) and catalase positive bacterium was obtained and genomic DNA was

extracted. 16S rDNA gene was amplified and a unique band on the 0.8% agarose gel detected (Figure 1). Sequencing of the band showed the expected 16S rDNA gene. The 16S rDNA sequence was compared with the EMBL/NCBI databases and the bacteria genus was recognized as *Staphylococcus*. The 16S rDNA sequence shares a high degree of homology to *S. pasteurii*; 98% identity and 98% similarity and 98% identity and 98% similarity to *S. warneri*; (Accession numbers mentioned above). Phylogenetic studies of the 16S rDNA with species of *staphylococcus*, resulted *pasteurii* as the most similar species (Figure 2). Biochemical tests profile also ratified *pasteurii* as the species (Table 1). After phylogenetic studies, the 16S rDNA sequence was submitted to the NCBI GenBank under accession number KC170006 as *Staphylococcus pasteurii* RA3^T.

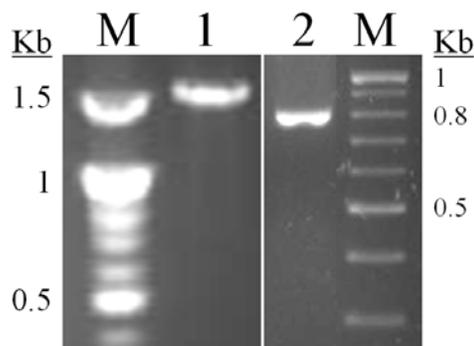


Figure 1: *Staphylococcus pasteurii* RA3^T 16S rDNA gene (1) and *TatD* gene (2).

Isolation, cloning and sequence analysis of gene encoding TatD DNase

Using the BLAST program, three sequences were identified from *Staphylococcus aureus* Mu50, *Staphylococcus carnosus* TM300 and *Staphylococcus saprophyticus* ATCC 15305 (NCBI GenBank accession numbers are NC_002758.2, NC_012121.1 and NC_007350.1, respectively) as putative TatD DNase. In order to characterize the full-length cDNA, the full-length open reading frame (ORF) was amplified by PCR using degenerate oligonucleotide primers

designated based on three putative TatD DNase gene sequences. The PCR product was a single PCR fragment of the expected size that was purified (Figure 1), cloned into pTG19-T vector (Vivantis) and sequenced by the dideoxyribonucleotide sequencing method. The *S. pasteurii* TatD DNase sequence (submitted to NCBI GenBank under

accession number KC608821) was 774 nucleotides long. This ORF with 36% G+C content encodes a protein of 257 amino acid residues, beginning at the initiation codon ATG and ending at the stop codon TAA of the cDNA. The predicted secondary structure was obtained using SOPMA (Figure 3).

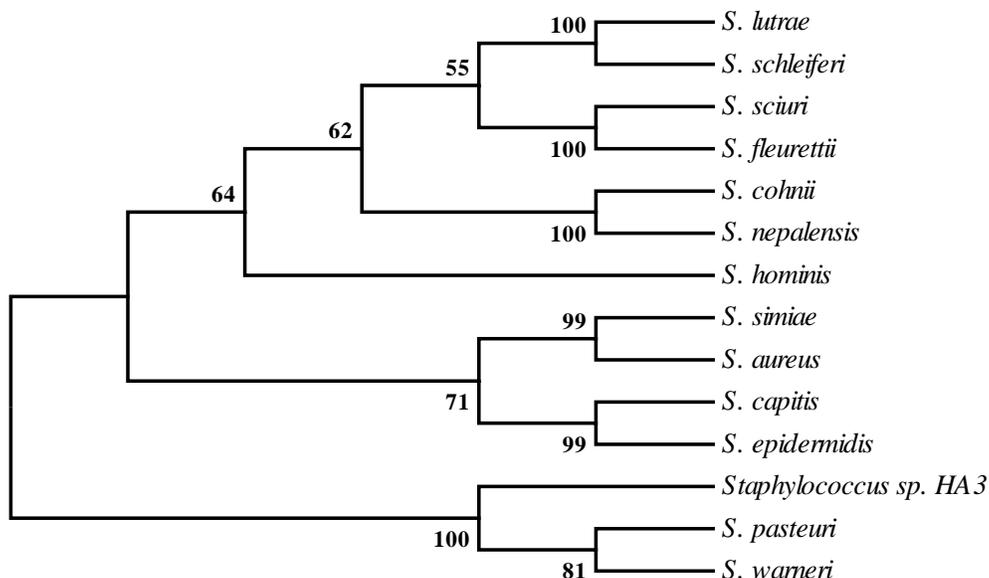


Figure 2: Phylogenetic tree of some *Staphylococcus* species 16S rDNA sequences using MEGA4. The program was performed using Neighbor Joining method with Bootstrap test with 1000 replicates.

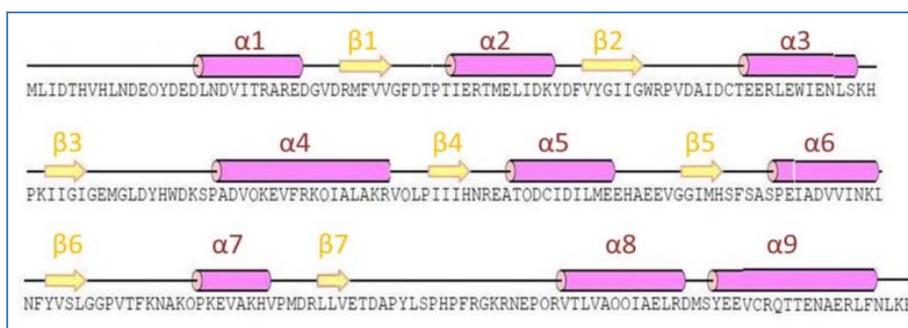


Figure 3: Predicted secondary structure of *S. pasteurii* TatD deduced amino acid sequence using the SOPMA program. The helix, sheet and coil are indicated by red, yellow and black boxes, respectively.

Table 1: Some differential phenotypic characteristics of *S. pasteurii* and similar bacteria

Characteristic	<i>S. pasteurii</i>					<i>S. warneri</i>			<i>S. hominis</i>	
	KC170006 RA3	BM10507	BM9363	BM10426	BM10427	BM9372	BM9408	BM9414	SM67	75611
Colony diam of >5 mm	+	-	+	-	-	-	+	+	-	-
Pigmentation (carotenoid)	-	+	+	-	+	+	-	-	+	+
Anaerobic growth	+	+	+	+	+	+	+	+	w	w
Urease activity	+	+	+	+	+	-	+	-	+	+
Acid produced (aerobically) from:										
D-Trehalose	+	+	+	+	+	+	+	+	+	+
D-Mannose	-	-	-	-	-	-	-	-	-	-
D-Ribose	-	-	-	-	-	+	+	+	-	-
a-Lactose	-	+	-	-	-	-	+	-	+	+
Maltose	+	+	+	+	+	+	+	+	+	+
D-Mannitol	+	+	-	+	+	+	-	-	-	-
D-Melezitose	-	+	+	-	w	-	-	-	w	w
D-Turanose	-	w	-	-	w	-	-	-	+	+
Resistance to: Lysostaphin (100 U)	+	+	+	+	+	+	-	-	-	+
Chloramphenicol (30 µg)	+	+	-	-	-	ND	ND	ND	ND	ND
Penicillin (6 µg)	-	-	-	-	+	ND	ND	ND	ND	ND

**S. pasteurii* isolated in this experiment. +, Character is present in all strains; -, character is absent in all strains; w, weakly positive reaction; not determined (Chesneau et al, 1993).

Table 2: Molecular characteristics of TatD molecules from different organisms

Characteristic	<i>S. p</i> TatD	2y1h	2gzx	1yix
MW (Da)	29616.3	30121.7	30346.4	29710.6
IP	5.34	6.39	5.40	5.19
Half time (h)	10	10	10	10
A.I	90.48	103.90	92.72	96.36
Trp	3	1	3	0
Phe	14	9	9	10
Met	5	3	6	4
Tyr	4	4	8	10
A.A	270	272	265	264

PDB codes prepared from RCSB Protein data bank. *S. p*, *Staphylococcus pasteurii* isolated in this experiment. A.I, Aliphatic Index, A.A, Amino Acid number.

Analysis of the deduced amino acid sequence

The deduced primary structure of *S. pasteurii* TatD DNase contains HHRPEDHRHF SSGEDPLN active site that is common to the

TatD molecules. The theoretical molecular mass of the deduced polypeptide is 29656.7 Da and the predicted pI is 5.05. The aliphatic index, regarded as a positive factor for increased thermostability (Kim et al, 2008),

was calculated as 90.48 by using ProtParam (Table 2). By calculating the hydrophobicity value by the method of Kyte and Doolittle (2008), it was determined that TatD was hydrophilic (Figure 4). Based upon studies of *T. maritima* TatD protein, it was demonstrated that Thr203 and Asp204 can play a particularly crucial role. Positions of other amino acids that are important for maintaining the tertiary structure and function in TatD based upon ConSeq, I-TASSER and PatchFinder servers are 6 His, 8 His, 63 Arg, 64 Pro, 92 Glu, 96 Asp, 128 His, 130 Arg, 153 His, 155 Phe, 156 Ser, 174 Ser,

176 Gly, 202 Glu, 204 Asp, 206 Pro, 208 Leu, 218 Asn. Using ConSurf program revealed that 155 Phe is highly variable, Ser 156, Gly 176 and Leu 208 are less variable and the other residues are completely conserved. TargetP program showed that the TatD protein is a signal peptide with a high confidence score about 0.978. However, by using a combination of three programs, TMHMM, YLOC (confidence score about 0.8510) and iPSORT, no signal peptide or transit peptide is present in the *S. pasteurii* TatD protein sequence, suggesting a cytoplasmic localization for this protein.

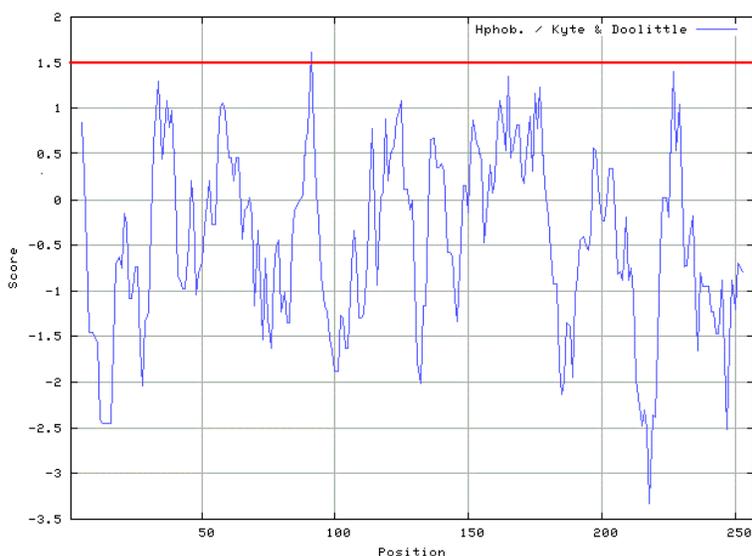


Figure 4: Hydrophathic index analysis of *S. pasteurii* TatD deduced amino acid sequence. Hydrophobic domains are indicated by positive numbers; hydrophobic domains are above the line, and hydrophilic domains are below.

Homology analysis of *S. pasteurii* TatD

The deduced polypeptide sequence for TatD DNase was compared with those of previously reported TatD proteins (Figure 5). The putative protein encoded by TatD shares a high degree of homology with *Staphylococcus warneri*; 99% identity and 99% similarity, *S. caprae*; 94% identity and 98% similarity, *S. epidermidis*; 94% identity and 98% similarity, *S. aureua*; 93% identity and 98% similarity. TatD of *S. pasteurii* is also 37.5% similar with a TatD from *Thermotoga*

maritima (pdb 1j60), used as a template for comparative modeling of the predicted 3D structure for TatD by applying I-TASSER simulation (Figure 6A). Superimposition analysis of the 3D models of TatD DNase and its templates are shown in Figure 6C. In contrast, *S. pasteurii* TatD shares lower degrees of identity to TatD of human (32% identity and 56% similarity), *E. coli* (36% identity and 58% similarity) and *Pseudomonas putida* (PDB 3rcm; 28% identity and 48% similarity). The overall sequence identity

between TatD with other *Staphylococcus* TatD, ranges from 74% to 99%. A phylogenetic tree was constructed by using different TatD DNases belonging to *Staphylococcus* species, human and *E. coli* (Figure 5). Prediction of subcellular localization of *S. pasteurii* TatD protein was also performed by using TargetP,

iPSORT and YLOC programs. The cell sorting prediction programs revealed that these sequences do not harbor a signal peptide or a transit peptide, suggesting a cytoplasmic localization. The previous reports confirmed this point (Wexler et al, 2000).

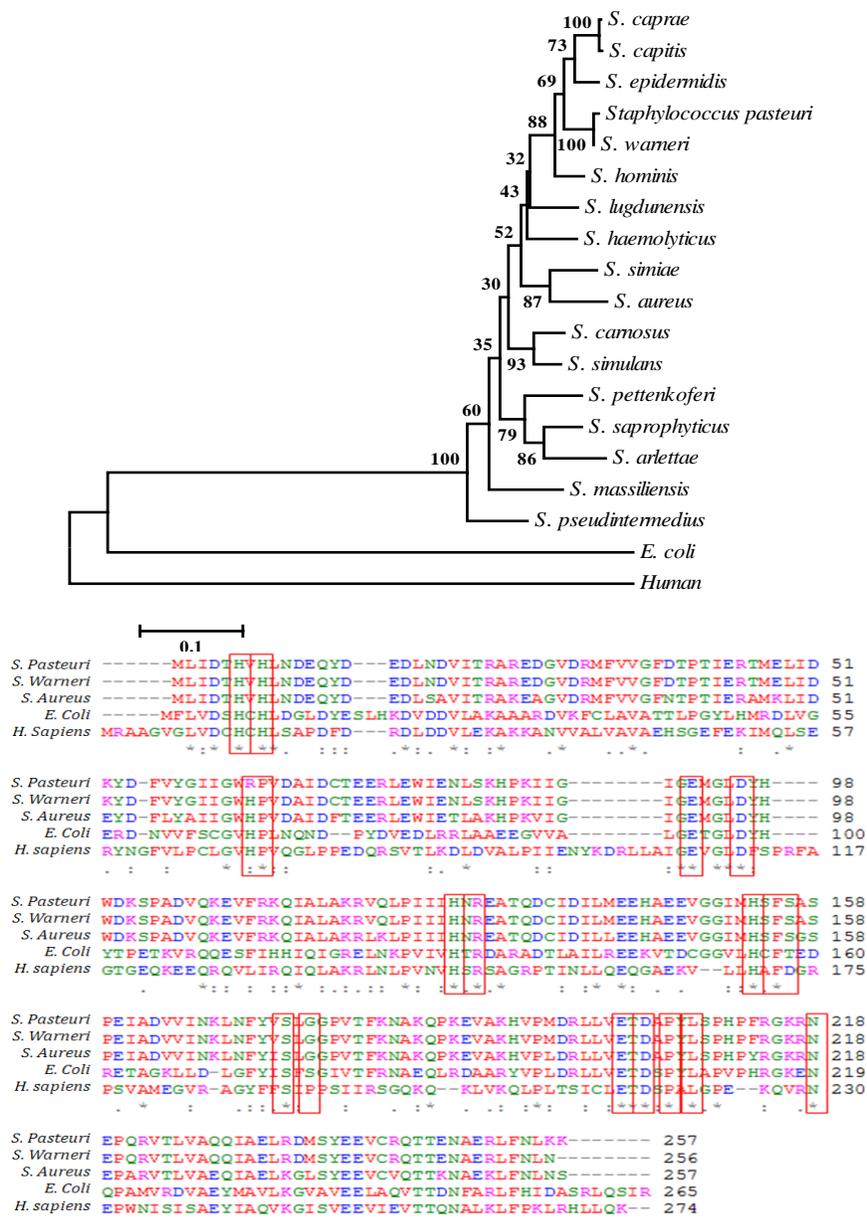


Figure 5: Multiple sequence alignment and phylogenetic tree of *S. pasteurii* TatD protein sequences using ClustalW and MEGA4 programs, respectively. The red rectangles show the putative active site of the enzyme predicted by Patchfider program. The green one shows the putative active site of TatD DNase predicted by I-TASSER program. The black rectangles show the difference between *S. pasteurii* and *R. pickettii* TatD amino acids. Accession numbers are

given in Materials and Methods. MEGA4 program was performed using Bootstrap test with 1000 replicates and Neighbor Joining method.

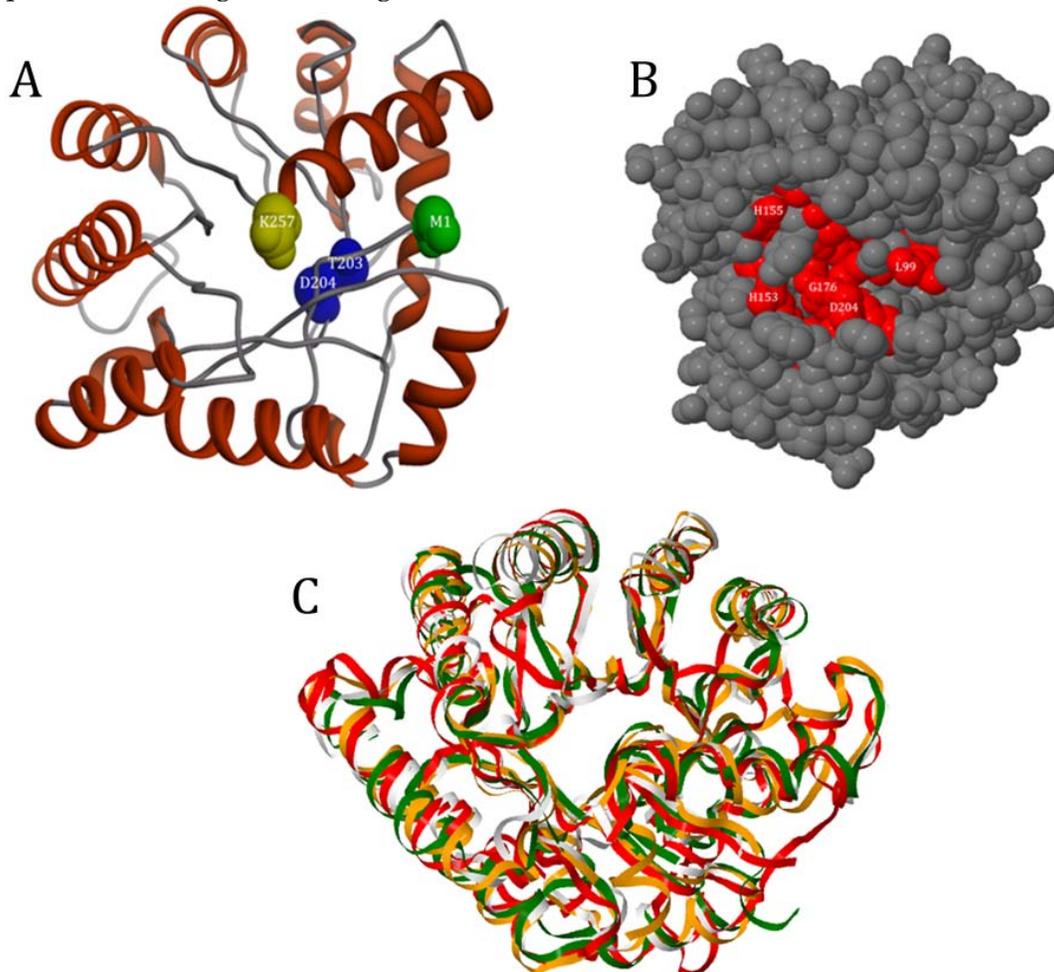


Figure 6: Three-dimensional model predictions for *S. pasteurii* TatD. (A), Cartoon display of the three-dimensional structure of *S. pasteurii* TatD. The Thr and Asp residues in the active site are indicated by ball-stick in blue color and α -helices, β -sheets, and coiled coil regions are colored in red, yellow and gray respectively. The N-terminal Met showed in green and C-terminal Lys showed in yellow (B), Active residue analysis of *S. pasteurii* TatD using Patchfinder program. Active residues showed a high conservation using ConSurf program that is shown in red. (C), Superimposition of 3D model of *S. pasteurii* TatD (green) and the top three templates of: *E. coli* TatD (white; PDB ID code 1yixA); TatD protein from *S. aureus* (orange; PDB ID code 2gzxA) and human TatD protein (red; PDB ID code 2y1hA), using 3d-SS (3 Dimensional Structural Superposition) service.

The amino acid sequence alignment revealed that the TatD DNase is a highly conserved gene in all organisms. Predicted active sites of molecules are conserved in human, *E. coli* and other bacteria (Figure 5). However, there are also differences in other positions, such as positions 63 Arg, 156 Ser and 176 Gly

(numbering based on *S. pasteurii* TatD). Using *T. maritima* TatD as a template for comparative modeling, a predicted 3D structure was determined for *S. pasteurii* TatD by applying I-TASSER simulation (Figure 6A). *S. pasteurii* TatD has the typical folding pattern consisting of 12 α -helices and the

active site HHRPEDHRHFSSGEDPLN is present (Figure 6B). In all organisms indicated in figure 5, position 63 (numbering based on *S. pasteurii* TatD) there is a His residue while in *S. pasteurii* this position is occupied by an Arg residue. This difference may change the efficiency or the function of the enzyme at all because the His residue is completely conserved in all organisms. In addition in position 156 in *Staphylococcus* genus always a Ser residue is observed while in *E. coli* and in human this position is occupied by a Thr and Asp residues, respectively. Another difference was observed in position 176, in all *Staphylococcus* species a Gly residue is present in this site, but in *E. coli* and human in the same position Ser and Pro residues are observed, respectively. These differences probably show a different function or efficiency in different organisms. However the predicted secondary structure using SOPMA showed 7 β -sheets and 9 α -helices in a $\alpha 1\beta 1\alpha 2\beta 2\alpha 3\beta 3\alpha 4\beta 4\alpha 5\beta 5\alpha 6\beta 6\alpha 7\beta 7\alpha 8\alpha 9$ topology (Figure 3). The secondary structure of *E. coli* TatD (PDB 1yix) showed eight β -sheets and 14 α -helices. The secondary structure topology is $\beta 1\alpha 1\alpha 2\beta 2\alpha 3\beta 3\alpha 4\beta 4\alpha 5\beta 5\alpha 6\alpha 7\beta 6\alpha 8\beta 7\alpha 9\alpha 10\alpha 11\beta 8\alpha 12\alpha 13\alpha 14$. The secondary structure of *S. aureus* TatD (PDB 2gzx) showed also 8 β -sheets and 14 α -helices in a $\beta 1\alpha 1\alpha 2\beta 2\alpha 3\beta 3\alpha 4\alpha 5\beta 4\alpha 6\beta 5\alpha 7\alpha 8\alpha 9\beta 6\alpha 9\beta 7\alpha 10\alpha 11\alpha 12\beta 8\alpha 13\alpha 14$ topology. *S. pasteurii* TatD protein 3D structure matches nearly perfectly with crystal structures from human (*Homo sapiens*) TatD (PDB 2y1h), *S. aureus* TatD (PDB 2gzx) and *E. coli* TatD proteins (1yix) (Figure 6C). Analysis of the evolutionary conservation of its surface amino acids was performed using ConSurf program and several residues were identified as functional and structural important regions in the protein by using PatchFinder and ConSeq servers. *E. coli* TatD amino acid sequence analysis showed that it possesses 260 residues and is a cytoplasmic protein. Pro151, Arg152 and Ser201 residues

are as the binding site that bind to Zn^{2+} , Mg^{2+} (Wexler et al, 2000) and its gene showed constitutive expression (Jack, 2001). By using CoFACTOR program and 1yix, 2y1h, 3rcm, 2gzx and 8 other sequences as templates it was revealed that Na^+ , Ni^{2+} , citric acid, acetate and Zn^{2+} are putative *S. pasteurii* TatD DNase cofactors that bind to Thr203 and Asp204 residues. In Gram-negative bacteria like *E. coli*, the Tat system is composed of TatABC subunits. TatE also participates in *E. coli*, but this seems to be a TatA paralogue of minor importance (Sargent et al, 1999). In *E. coli*, the tatABC genes are present in an operon together with a fourth gene, tatD, but TatD has been shown to be a cytoplasmic protein that is not involved in translocation (Wexler et al, 2000). It has been reported that bacterial Tat systems export folded proteins, including FeS proteins such as NrfC and NapG, which acquire their cofactors before translocation. NrfC and NapG are proofread by the Tat pathway, and misfolded examples are degraded after interaction with the translocon. They also identified TatD as a crucial component of a quality control system in *Escherichia coli* (Cristina et al, 2009). The *tatD* gene is variously predicted to encode a soluble cytoplasmically located protein (Sargent et al, 1998) or an integral membrane protein with a large cytoplasmic domain (Weiner et al, 1999). TatD is a soluble protein and is unlikely to be a permanent component of the Tat translocon. There is no information on its mode of action; the Δ tatD strain has no clear phenotype (Wexler et al, 2000), and although studies on the purified protein showed that it has DNase activity, it seems unlikely that this DNase activity is involved in this particular quality control pathway (Cristina et al, 2009). TatD binds to immobilized Ni^{2+} or Zn^{2+} affinity columns and exhibits magnesium-dependent DNase activity. TatD family proteins show significant sequence similarity to proteins assigned to the polymerase and histidinol phosphatase superfamily of metal-dependent

phosphoesterases (Park and Michel, 1996). For example, *E. coli* TatD exhibits 33% amino acid identity to the YabD protein of *Chlamydomonas reinhardtii*. The functional requirement for the different metal centers is still unclear. Interestingly, *Bacillus subtilis* N-acetylglucosamine-6-phosphate deacetylase (NAGA) contains two iron ions and thus belong to the β subset, whereas the *Thermotoga maritima* enzyme has only one iron bound at the β site and may belong to the $\alpha\beta$ subset (Vincent et al, 2004). The putative magnesium-dependent TatD DNase from *T. maritima* also displays a similar β -mononuclear metal center. The bacterial TatDs substitute the HXH with GXN. Structural comparison demonstrates that both $\alpha\beta$ and β subsets share similar binuclear active sites. TatD DNases have amino acid triplets for the conserved HXH. Sequence analysis of the HXH and GXN groups of TatD DNases shows that the former are solely present in the Eubacteria, while the latter are present in the Eubacteria, Archaea, and Eukaryotes. This would suggest that the latter is a more ancient form of the enzyme, which does not require the involvement of metal regulation. TatD has other activities and it will be important to determine how, and at which point, it is involved in the degradation of NrfC/NapG/FhuD (Cristina et al, 2009). Overexpression of TatD does not restore the turnover of these Tat substrates in the Δ tatD strain, and has the dual effects of rendering these Tat substrates more stable in wild-type cells while blocking the export of both NrfC and FhuD in Δ tatD cells. Overexpression of TatD might cause an imbalance between the two pathways (Cristina et al, 2009).

Conclusion

Putative TatD DNase gene isolated from the soil bacterium *Staphylococcus pasteurii* RA3^T was most similar to other TatD DNases from *Staphylococcus* genus. Amino acid sequence alignment and 3D structural studies

confirmed the active site and function as DNase. The similarity of TatD protein of *S. pasteurii*, *E. coli* and human showed that TatD has an essential role in approximately all cells. On the other hand, TatD-like proteins are coded by the genomes of the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, neither of which possesses proteins with twin arginine signal peptides or homologues of the essential Tat proteins. Thus no exact correlation exists between the ability of an organism to synthesize a TatD like protein and whether that organism has a Tat protein transport pathway (Klenk et al, 1997). This research was a beginning for investigation on the DNase activity of the soil bacteria such as *Staphylococcus pasteurii* RA3^T TatD proteins.

Reference

- Ashkenazy H., Erez E., Martz E., Pupko T., Ben-Tal N., ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids, *Nucl. Acid Res.*, 2010, 38, 529-533.
- Bannai H., Tamada Y., Maruyama O., Nakai K., Miyano S., Extensive feature detection of N-terminal protein sorting signals, *Bioinformatics*, 2002, 18, 298-305.
- Blobel, G. and Dobberstein, B. (1975) Transfer of proteins across membranes. *J. Cell Biol.* 67, 852-862.
- Briesemeister S., Rahnenführer J., Kohlbacher O. (2010) Going from where to why-interpretable prediction of protein subcellular localization, *Bioinformatics*, 26, 1232-1238.
- Brüser T, Deutzmann R and Dahl C (1998) Evidence against the double-arginine motif as the only determinant for protein translocation by a novel Sec-independent pathway in *Escherichia coli*. *FEMS Microbiol Lett* 164: 329-336.
- Bruser, T. and Sanders, C. (2003) An alternative model of the twin arginine translocation system. *Microbiol. Res.* 158, 7-17.

- Chesneau O., MORVAN A., GRIMONT F., LABISCHINSKI H. AND EL SOLH N. (1993) *Staphylococcus pasteun* sp. nov., Isolated from Human, Animal, and Food Specimens. *Int J of Systematic Biology*, 237-244.
- Cristina F.R.O. Matos, Alessandra Di Cola, Colin Robinson. (2009) TatD is a central component of a Tat translocon initiated quality control system for exported FeS proteins in *Escherichia coli*. *EMBO reports* 10, 474-479.
- Cristobal, S., de Gier, J.-W., Nielsen, H. and von Heijne, G. (1999) Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. *EMBO J.* 18, 2982-2990.
- Drancourt M, Raoult D. 2002. rpoB gene sequence-based identification of *Staphylococcus* species. *J Clin Microbiol.* 40:1333-1338.
- Emanuelsson O., Brunak S., Von Heijne G., Nielsen H., Locating proteins in the cell using TargetP, SignalP, and related tools, *Nat. Protoc.*, 2007, 2, 953-971.
- Enright MC, et al. 2000. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *J Clin Microbiol.* 38:1008-1015.
- Felsenstein J (1985) Confidence limit on phylogenies: an approach using the bootstrap. *Evolution* 39:783-791.
- Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., et al., Protein identification and analysis tools on the ExPASy server-the proteomics protocols handbook, John M, Walker, Totowa, Humana Press, 2005.
- Geourjon C., Deleage G. (1995) SOPMA: significant improvement in protein secondary structure prediction by consensus prediction from multiple alignments, *Cabios*, 11, 681-684.
- Ghebremedhin B, Layer F, Konig W, Konig B. 2008. Genetic classification and distinguishing of *Staphylococcus* species based on different partial gap, 16S rRNA, hsp60, rpoB, sodA, and tuf gene sequences. *J Clin Microbiol.* 46:1019-1025.
- Halbig, D., Wiegert, T., Blaudeck, N., Freudl, R. and Sprenger, G.A. (1999) The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of *Zymomonas mobilis* depends both on an intact twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding. *Eur. J. Biochem.* 263, 543-551.
- Jack R.L., Sargent F., Berks B.C., Sawers G., Palmer T. (2001). Constitutive expression of *Escherichia coli* tat genes indicates an important role for the twin-arginine translocase during aerobic and anaerobic growth. *J. Bacteriol.* 183:1801-1804.
- Keiko Yamada, Ikuyo Sanzen, Teruko Ohkura, Akira Okamoto, Keizo Torii, Tadao Hasegawa, Michio Ohta. (2007) Analysis of Twin-Arginine Translocation Pathway Homologue in *Staphylococcus aureus*. *Current Microbiology*, 55, 14-19.
- Kieran Dilks, R. Wesley Rose, Enno Hartmann and Mechthild Pohlschroder. (2003) Prokaryotic Utilization of the Twin-Arginine Translocation Pathway: a Genomic Survey. *J Bacteriol.* 185(4): 1478-1483.
- Kim Y.J., Shim J.S., Krishna P.R., Kim S.Y., In J.G., Kim M.K. (2008) Isolation and characterization of a glutaredoxin gene from *Panax ginseng* C. A. Meyer, *Plant Mol. Biol. Rep.*, 26, 335-349
- Klenk HP, Clayton RA, Tomb JF, White O, et al., (1997) The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* 390, 364-370.
- Kyte J., Doolittle R.F. (1982) A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol.*, 157, 105-132
- Landau M., Mayrose I., Rosenberg Y., Glaser F., Martz E., Pupko T. (2005) ConSurf: the projection of evolutionary conservation

- scores of residues on protein structures, *Nucl. Aci. Res.*, 33, 299-302.
- Nimrod G., Schushan M., Steinberg D.M., Ben-Tal N. (2008) Detection of functionally important regions in hypothetical proteins of known structure, *Structure*, 16, 1755-1763.
- Park, I. S., and Michel, L. O. (1996) Characterization of the mononickel metallocenter in H134A mutant RT urease. *J. Biol. Chem.* 271, 18632-18637.
- Peterson F.C., Lytle B.L., Sampath S., Vinarov D., Tyler E., Shahan M., et al. (2005) Solution structure of thioredoxin h1 from *Arabidopsis thaliana*, *Prot. Struct. Rep.* 14, 2195-2200.
- Poyart C, Quesne G, Boumaila C, Trieu-Cuot P. 2001. Rapid and Accurate species-level identification of coagulase-negative Staphylococci by using the sodA gene as a target. *J Clin Microbiol.* 39:4296-4301.
- Pugsley, A.P. (1993) The complete general secretory pathway in gram-negative bacteria. *Microbiol. Rev.* 57, 50-108.
- Rodrigue, A., Chanal, A., Beck, K., Muller, M. and Wu, L.-F. (1999) Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway. *J. Biol. Chem.* 274, 13223-13228.
- Russell R.B., Breed J., Barton G.J. (1992) Conservation analysis and secondary structure prediction of the SH2 family of phosphotyrosine binding domains, *FEBS Lett.*, 304, 15-20.
- Sambrook, J., Russell, D.W., editors. (2001). *Molecular Cloning, A Laboratory Manual*. 3rd ed. Cold Spring Harbor Laboratory Press, New York.
- Sanders, C., Wethkamp, N. and Lill, H. (2001) Transport of cytochrome c derivatives by the bacterial Tat protein translocation system. *Mol. Microbiol.* 41, 241-246.
- Sargent, F., Stanley, N. R., Berks, B. C., and Palmer, T. (1999). Sec-independent protein translocation in *Escherichia coli*. A distinct and pivotal role for the TatB protein. *J. Biol. Chem.* 274, 36073-36082
- Stanley, N.R., Palmer, T. and Berks, B.C. (2000) The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. *J. Biol. Chem.* 275, 11591-11596.
- Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596-1599.
- Thomas, J.D., Daniel, R.A., Errington, J. and Robinson, C. (2001) Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in *Escherichia coli*. *Mol. Microbiol.* 39, 47-53.
- Vincent, F., Yates, D., Garman, E., Davies, G. J., and Brannigan, J. A. (2004) The 3-D structure of the N-acetylglucosamine-6-phosphate deacetylase, NagA, from *Bacillus subtilis* : a member of the urease superfamily. *J. Biol. Chem.* 279, 2809-2816
- Weiner, J. H., Bilous, P. T., Shaw, G. M., Lubitz, S. P., Frost, L., Thomas, G. H., Cole, J. A., and Turner, R. J. (1998). A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* 93, 93-101.
- Wexler M , Sargent F , Jack RL , Stanley NR , Bogsch EG , Robinson C , Berks BC , Palmer T. (2000). TatD is a cytoplasmic protein with DNase activity. No requirement for TatD family proteins in sec-independent protein export. *J. Biol. Chem.* 275: 16717-22.
- Zhang Y. (2008) I-TASSER server for protein 3D structure prediction, *BMC Bioinformatics*, 9, 26.
- Zhou J, M A Bruns and J M Tiedje. (1996). DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol.* 62(2):316.