Asian Biomedicine Vol. 9 No. 3 June 2015; 299 - 311 DOI: 10.5372/1905-7415.0903.398

Original article

Sexually transmitted diseases in symptomatic and asymptomatic Thai women and girls: a study from Bangkok and nearby

Naraporn Somboonna[a],[b] Ilada Choopara[a], Kanchapan Sukhonpan[c], Jarun Sayasathid[d]
[a]*Department of Microbiology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand*
[b]*Omics Sciences and Bioinformatics Center, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand*
[c]*Department of Obstetrics and Gynecology, Buddhachinaraj Phitsanulok Hospital, Phitsanulok 65000, Thailand*
[d]*Cardiac Center, Naresuan University Hospital, Phitsanulok 65000, Thailand*

Background: *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, and human papillomavirus (HPV) infections among women and girls may be symptomatic or asymptomatic.
Objectives: To survey and identify *C. trachomatis*, *N. gonorrhoeae*, and HPV infections in symptomatic and asymptomatic women and girls in Bangkok and surrounding neighborhoods using molecular techniques, evaluate the use of doxycycline treatment for clinically symptomatic patients infected with *C. trachomatis*, and identify possible genetic mutations associated with persistence and/or drug resistance.
Method: We enrolled 150 women and girls with inflamed urogenital tracts and 134 asymptomatic controls, both without a history of HIV infection or cervical cancer in this prospective study. Genomic DNAs were extracted, and *C. trachomatis*, *N. gonorrhoeae*, and HPV infections were detected using established PCR primers and protocols. PCR controls included no confirmatory template controls or human β-globin. Patients infected with *C. trachomatis* were treated with doxycycline and re-examined after treatment. *C. trachomatis* major outer membrane (*ompA*) and tryptophan synthesis A (*trpA*) genes were sequenced to identify possible genetic mutations associated with persistence and/or drug resistance.
Results: *C. trachomatis*, *N. gonorrhoeae*, and HPV were detected in 22%, 6%, and 48% of symptomatic, and in 3%, 16%, and 10% of asymptomatic women and girls, respectively. Coinfection with *C. trachomatis* and HPV were frequent in the 15-34 year age group, and associated with upper urogenital tract symptoms. Doxycycline was not considered effective for *C. trachomatis* infection. Several nonconserved amino acid changes were detected in *C. trachomatis* ompA and trpA.
Conclusion: We found different distributions of these pathogens among symptomatic and asymptomatic patients. We also found doxycycline treatment failures, and mutated trpA supported persistent *C. trachomatis* infections.

Keywords: *Chlamydia trachomatis*, drug-resistant *C. trachomatis*, human papillomavirus, *Neisseria gonorrhoeae*, sexually transmitted disease, tryptophan

While human papillomavirus (HPV) is the most common sexually-transmitted infection (STI) worldwide. *Chlamydia trachomatis* and *Neisseria gonorrhea* are major bacterial causes of STIs [1, 2]. Rugpao et al. [3] cultivated bacteria from cervices of women with clinical symptoms. Among these, 7.8% were *C. trachomatis* and 4.8% *N. gonorrhea*. Celentano et al. [4] reported *C. trachomatis* to be even more prevalent (22%) than *N. gonorrhea* (6%) by culture- and DNA-based techniques. They reported that ages 20–25 years and subjects reporting more than 2 heterosexual partners were associated with a higher risk of the STIs [4]. The prevalence was probably underestimated. These infections were also found in apparently healthy men and women. For example, *C. trachomatis* infections in humans are initially often asymptomatic until they become severe, causing pelvic inflammatory disease, ectopic pregnancy, and infertility [2, 5-8]. We therefore focused this study on the presence or absence of *C. trachomatis*, *N. gonorrhoeae*, and HPV infections in women and girls with symptomatic

Correspondence to: Naraporn Somboonna, Department of Microbiology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand. E-mail: Naraporn.S@chula.ac.th

inflamed urogenital tracts and also on asymptomatic apparently healthy controls.

C. trachomatis is known for long-term, persistent infections that could be aided by low-levels of human interferon-γ immune responses, degrading tryptophan, which is an essential amino acid required for bacterial growth [8, 9]. This amino acid-starved condition leads to persistent *C. trachomatis* infection and enhances bacteria to remain viable in a nonculturable (dormant) state. Such persistent *C. trachomatis* infection can be resistant to the bacteriostatic action of commonly prescribed *C. trachomatis* antibiotics. In brief, persistent *C. trachomatis* expresses minimal gene transcription and translation, which is a target for doxycycline and tetracycline [8, 10, 11]. Importantly, persistent *C. trachomatis* infection can last for several years in the female urogenital tract [6-8].

An outer membrane protein A gene (*ompA*) encodes the major outer membrane and *ompA* sequencing represents an established method to categorize *C. trachomatis* into 19 serovars. These are A to C, including Ba, which cause trachoma, D to K including Da, which cause noninvasive urogenital infections and trachoma in newborns, and L_1 to L_3, which cause an invasive urogenital infection called lymphogranuloma venereum [2, 12]. Globally, *C. trachomatis* serovars D, E, and F are the most common sexually transmitted diseases (STDs) [2].

C. trachomatis tryptophan synthesis, α-subunit (*trpA*) sequence, was included in this study because of the association between *C. trachomatis* persistence and possible *trpA* mutation. Tryptophan is essential for chlamydiae to replicate. Limited tryptophan, through the effect of human interferon-γ, can induce chlamydial persistence [8, 10]. Interferon-γ stimulates a tryptophan-degrading enzyme named indoleamine-2,3-dioxygenase. Polymorphisms in *trpA* in various *C. trachomatis* serovars can affect the structure and function of TrpA and, for instance, result in truncated TrpA in serovars A–C. This mutation also disrupts chlamydial tryptophan synthesis from indole substrates. TrpA in most urogenital serovars is functional [8, 13, 14].

This study included the molecular detection of HPV and *Neisseria gonorrhoeae*, because these pathogens share transmission routes and are often found associated. Importantly, long-term *C. trachomatis* infection may increase the risk for acquisition and retention of the high-risk HPV variants that cause cervical cancer. Likewise, long-term HPV infection may increase the risk of persistence, and for more severe clinical outcomes from *C. trachomatis* infection [8, 15]. This study aimed to describe STD prevalence in women and girls seeking treatment in the Bangkok region and to determine the efficiency of doxycycline treatment for *C. trachomatis* along with the genetic characteristics of *C. trachomatis* *ompA* and *trpA*.

Materials and methods

Participants

Participants included 284 Thai women and girls aged 15–54 years attending the Buddhachinaraj Phitsanulok Hospital, Bangrak STD clinic, which is part of the Office of Disease Prevention and Control 1 Bangkok, and Office of Disease Prevention and Control 3 of Chonburi province. All participants were HIV-negative, and had neither cervical cancer nor previous cervical cancer treatment. Participants were grouped as (i) upper genital tract (UGT), (ii) lower genital tract (LGT), and (iii) undefined UGT or LGT infection; based on U.S. Centers for Disease Control and Prevention definitions [2]. Samples were obtained by clinicians, following a standard cervical swab sample collection method and a medium using M4RT (Illinois, USA). For the symptomatic group, 100 mg of doxycycline, twice daily for 14 days, was prescribed, and refrain from sexual contact during treatment until follow-up sample collection 14 days after treatment was completed [2, 7, 11, 12]. All samples were collected during September 2011–2012 for Buddhachinaraj Phitsanulok Hospital, and March 2013–2014 for the other sites.

The study was approved by the Institutional Review Board of Buddhachinaraj Phitsanulok Hospital (101/54), and the Ethics Committee for Research in Human Subjects of the Department of Disease Control, Bangkok (FWA00013622). Written informed consent was obtained from all subjects for their participation in the study. Clinical samples were identified by a unique identification number with no link to patient names.

Molecular detections for C. trachomatis, N. gonorrhoeae, HPV, and human β-globin

DNA was extracted from a 100 µL aliquot of each clinical sample was according to High Pure Template Preparation Kit instructions (Roche Diagnostics Corporation, Indianapolis, IN, USA). Concentration and quality of the extracted DNA were measured

by $A_{260\ nm}$ and $A_{260}/A_{280\ nm}$ spectrophotometry. The extracted DNA had an average A_{260}/A_{280} between 1.80 and 2. *C. trachomatis*, *N. gonorrhoeae*, HPV, and human β-globin were detected by PCRs following established protocols [12, 16-18]. Primers and PCR annealing temperatures were listed in **Table 1**: Ct.ompA.60UF and Ct.ompA.VB3 were primers for *C. trachomatis ompA* first-half; Ct.ompA.MVF3 and Ct.ompA.220DR for *ompA* second-half; Ng.opa.F and Ng.opa.R to Ng.opa.R2 for primary *N. gonorrhoeae*; Ng.porA.F and Ng.porA.R for confirmed positive *N. gonorrhoeae*; PGMY11-A to PGMY11-D, and PGMY09-F to HMB01 as forward and reverse primer cocktails for HPV L1 gene; and globinF and globinR for human β-globin. Sizes and semi-quantitation of the PCR products were determined from agarose gel electrophoresis. All negative PCR results were repeated to confirm the negative results.

Table 1. PCR primers and annealing temperatures

Primer name	Primer sequence (5′ → 3′)	Tm (°C)
Ct.ompA.60UF	GTCCCGCCAGAAAAAGATAG	45
Ct.ompA.VB3	CATCGTAGTCAATAGAGGCAT	
Ct.ompA.MVF3	TGTAAAACGACGGCCAGTGCCCGTGCAGCTTTGTGGGAATGT	45
Ct.ompA.220DR	GCGCTCAAGTAGACCGATATAGTA	
F.Ct.trpA	ATTAGCCACCGATGAAGAG	50
R.Ct.trpA	ATGTTGAATTAGGAGAGTTGTTAT	
Ng.opa.F	TTGAAACACCGCCCGGAA	60
Ng.opa.R	TTTCGGCTCCTTATTCGGTTTAA	
Ng.opa.R1	TTTCGGCTCCTTATTCGGTTTGA	
Ng.opa.R2	TTTCGGCTCCTTATTCGGTTTGA	
Ng.porA.F	CAGCATTCAATTTGTTCCGAGTC	60
Ng.porA.R	GAACTGGTTTCATCTGATTACTTTCCA	
PGMY11-A	GCACAGGGACATAACAATGG	50
PGMY11-B	GCGCAGGGCCACAATAATGG	
PGMY11-C	GCACAGGGACATAATAATGG	
PGMY11-D	GCCCAGGGCCACAACAATGG	
PGMY11-E	GCTCAGGGTTTAAACAATGG	
PGMY09-F	CGTCCCAAAGGAAACTGATC	
PGMY09-G	CGACCTAAAGGAAACTGATC	
PGMY09-H	CGTCCAAAAGGAAACTGATC	
PGMY09-I	GCCAAGGGGAAACTGATC	
PGMY09-J	CGTCCCAAAGGATACTGATC	
PGMY09-K	CGTCCAAGGGGATACTGATC	
PGMY09-L	CGACCTAAAGGGAATTGATC	
PGMY09-M	CGACCTAGTGGAAATTGATC	
PGMY09-N	CGACCAAGGGGATATTGATC	
PGMY09-P	GCCCAACGGAAACTGATC	
PGMY09-Q	CGACCCAAGGGAAACTGGTC	
PGMY09-R	CGTCCTAAAGGAAACTGGTC	
HMB01	GCGACCCAATGCAAATTGGT	
globinF	GAAGAGCCAAGGACAGGTAC	58
globinR	CAACTTCATCCACGTTCACC	

C. trachomatis ompA and *trpA* sequencing and bioinformatics analyses

trpA PCR were performed using F.Ct.trpA and R.Ct.trpA primers (**Table 1**). The *ompA* and *trpA* PCR products were DNA purified, and sequenced at Macrogen Inc (Seoul, Korea). All sequences were validated using an electropherogram (BioEdit version 7.0.0, Carlsbad, CA, USA). The sequences were deposited in GenBank under accession numbers KM369928 to KM369961. The *ompA* sequences were identified as *C. trachomatis* serovars by BLASTN against the GenBank nucleotide nonredundant database [19]. Multiple sequences were aligned using MEGA 5 software (Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, PA, USA), and compared with 19 reference *C. trachomatis* serovars A-L_3: A/Har-13 (NC007429), B/TW-5 (DQ064281 *ompA*), Ba/Apache-2 (DQ064282 *ompA*, AY096806 *trpA*), C/TW-3 (NC023060), D/UW-3 (NC000117), Da/TW-448 (X62921 *ompA*, KM369947 *trpA*), E/Bour (NC020971), F/IC-Cal3 (DQ064287 *ompA*, AY096810 *trpA*), G/UW-57 for *ompA* (AF304326), G/UW-524 *trpA* (AY096811), H/UW-4 (AE304328 *ompA*, AY096812 *trpA*), I/UW-12 (DQ064290 *ompA*, AY096813 *trpA*), Ia/IU-4168 (AF063201 *ompA*, KM369958 *trpA*), J/UW-36 (DQ064292 *ompA*, AY096814 *trpA*), Ja/IU-37538 *ompA* (AF202458), Ja/UW-92 *trpA* (KM369959), K/UW-31 (DQ064293 *ompA*, AY096815 *trpA*), L_1/440 (DQ064294 *ompA*, AY096816 *trpA*), L_2/434 (DQ064295 *ompA*, AY096817 *trpA*), L_2a/TW-396 (AF304858 *ompA*, KM369961 *trpA*), and L_3/404 (DQ064296 *ompA*, AY096818 *trpA*). A variant was defined as having ≥ 1 nucleotide (nt) difference from the sequence of the reference strain. The p-distance and neighbor-joining tree with 1,000 bootstrap replicates were constructed using MEGA 5.

Comparative structure modeling of C. trachomatis TrpA and active TrpA-TrpB complex

C. trachomatis TrpA three-dimensional structure was constructed using consensus TrpA sequences from reference serovars D to K, by MODWEB (http://salilab.org/modweb) [20]. MODWEB uses PSI-BLAST [19] and IMPALA [21] to obtain representative three-dimensional structures from the Protein Data Bank (PDB) (http://www.rcsb.org), and then uses MODELLER to align our query sequence with the representative PDB structure [22]. Every modeled structure was given a GA341 model score. The GA341 of >0.7 means reliable structure prediction, with >95% probability of the correct fold [20-22], and all of our modeled structures had GA341 of 1.0. The structures were visualized using the UCSF Chimera tool (http://www.cgl.ucsf.edu/chimera) [23]. Critical sites, including active sites, binding sites and TrpB communication (COMM) domain, were annotated based on the structural alignment to the representative PDB structures (1qopA for TrpA and 1qopB for TrpB, from *Salmonella typhimurium*) [24], and literature reviews [25, 26]. The structural alignments were accomplished using the MultAlign Viewer program [23].

Results

Prevalence of C. trachomatis, N. gonorrhoeae, and HPV

Of 284 total participants, HPV (86 individuals: 30.3%) was detected most frequently, followed by *C. trachomatis* (37 individuals: 13.0%) and *N. gonorrhoeae* (30 individuals: 10.6%), respectively. After categorizing based on symptomatic versus healthy or asymptomatic groups, the ranking remained valid for symptomatic urogenital group (HPV 48%, *C. trachomatis* 22%, and *N. gonorrhoeae* 6%), but asymptomatic group (*N. gonorrhoeae* 16%, HPV 10%, and *C. trachomatis* 3%) (**Table 2**). For the symptomatic group, coinfections with *C. trachomatis* and HPV were found commonly, compared with *C. trachomatis* and *N. gonorrhoeae* coinfections (**Tables 2 and 3**). Samples with no *C. trachomatis*, *N. gonorrhoeae*, or HPV were positive for the human -globin gene (data not shown). This was to confirm the quality of the DNA, because all samples must be at least positive for this human housekeeping gene.

Characterization of symptoms based on the progression of the diseases (UGT or LGT) showed that the pathogens were more frequently detected in the UGT and LGT than in the undefined, and the asymptomatic groups; except for *N. gonorrhoeae* (**Table 2**). HPV and *C. trachomatis* were found even more frequently when symptoms reached the UGT than LGT (23% HPV, and 12% *C. trachomatis*). Although samples were collected at the cervix and the number of participants with UGT was smaller than the number of participants with LGT. Additionally, categorization of symptomatic women and girls by age showed that those in the

15–34 age range had more than 70% of infections with *C. trachomatis* (76%), *N. gonorrhoeae* (78%), and HPV (72%) (**Table 2**).

For symptomatic women and girls, those not engaged in commercial sex (non-CSWs) had a lower prevalence of *C. trachomatis* (82% CSW compared to 18% in non-CSW), *N. gonorrhoeae* (67% CSW, 33% non-CSW), and HPV (68% CSW, 32% non-CSW) infections. In those women and girls who appeared healthy, *C. trachomatis* was found relatively more commonly in CSWs (**Table 2**).

C. trachomatis ompA serovar distributions and persistence

Nineteen random symptomatic patients with *C. trachomatis* infections were *ompA* sequenced to identify serovars: 6 belonging to D, 4 E, 4 F, 2 G, 2 H, and 1 K. These patients underwent a complete course of doxycycline treatment. While the majority of patients did not come back for *C. trachomatis* follow-up examination after treatment, all patients who came back had persistence of *C. trachomatis* of the same serovar found before treatment (**Table 3**).

Table 2. Prevalence of STIs in symptomatic and asymptomatic urogenital tract females

Symptom location* (n)	*Ct*□ (%)	*Ng*□ (%)	HPV (%)	Characterized by ages (%)	*Ct*□ (%)	*Ng*□ (%)	HPV (%)	Characterized by occupation (%)	*Ct* (%)	*Ng* (%)	HPV (%)
Symptomatic											
UGT (57)	18 (12)	4 (3)	35 (23)	15–24	12 (36)	2 (22)	18 (25)	Commercial sex worker	1 (3)	2 (22)	9 (13)
				25–34	5 (15)	1 (11)	8 (11)				
				35–44	1 (3)	1 (11)	6 (8)				
								Non-commercial sex worker	17 (52)	2 (22)	26 (36)
				45–54	0	0	1 (1)				
				N/A	0	0	2 (2)				
LGT (75)	14 (9)	5 (3)	29 (19)	15–24	4 (12)	0	11 (15)	Commercial sex worker	4 (12)	1 (11)	10 (14)
				25–34	4 (12)	4 (44)	11 (15)				
				35–44	5 (15)	0	7 (10)				
								Non-commercial sex worker	10 (30)	4 (44)	19 (26)
				45–54	1 (3)	1 (11)	0				
				N/A	0	0	0				
Undefined (18)	1 (1)	0	8 (5)	15–24	0	0	1 (1)	Commercial sex worker	1 (3)	0	4 (6)
				25–34	0	0	3 (4)				
				35–44	0	0	0				
				45–54	0	0	0	Non-commercial sex worker			(6)
				N/A	1 (3)	0	4 (6)				
Total (150)	33 (22)	9 (6)	72 (48)	15–24	16 (48)	2 (22)	30 (42)	Commercial sex worker	6 (18)	3 (33)	23 (32)
				25–34	9 (27)	5 (56)	22 (31)				
				35–44	6 (18)	1 (22)	13 (18)				
								Non-commercial sex worker	27 (82)	6 (67)	49 (68)
				45–54	1 (3)	1 (11)	1 (1)				
				N/A	1 (3)	0	6 (8)				
Asymptomatic											
None (134)	4 (3)	21 (16)	14 (10)	15–24	0	7 (33)	6 (43)	Commercial sex worker	3 (75)	10 (47.62)	7 (50)
				25–34	4 (100)	6 (29)	1 (7)				
				35–44	0	4 (19)	4 (29)				
								Non-commercial sex worker	1 (25)	11 (52.38)	7 (50)
				45–54	0	3 (14)	3 (21)				
				N/A	0	1 (5)	0				

*Symptom location includes UGT (upper urogenital tract), LGT (lower urogenital tract), Undefined (undefined UGT or LGT), and none for asymptomatic

□*Ct* represents *C. trachomatis*

□*Ng* represents *N. gonorrhoeae*

N/A means age information is not available

Table 3. *C. trachomatis* serovar distribution and infection pattern after doxycycline treatment

Symptomatic	No. of patients (%)	Doxycycline treatment					
		C. trachomatis		*N. gonorrhoeae*		HPV	
		Before	After	Before	After	Before	After
Serovar D	6 (32)	6	2/4*	0	0	3	1/2
Serovar E	4 (21)	4	1/3	0	0	4	1/3
Serovar F	4 (21)	4	1/3	1	0/1	4	1/3
Serovar G	2 (11)	2	0/2	0	0	1	0/1
Serovar H	2 (11)	2	0/2	0	0	1	0/1
Serovar K	1 (5)	1	0/1	0	0	0	0

*The number after the solidus (/) represents the number of patients who did not come back after the treatment for *C. trachomatis*, *N. gonorrhoeae*, or HPV

Polymorphisms of ompA among 14 clinical isolates from Thai women and girls

Full-length *ompA* sequences of 14 clinical isolates from Thai women and girls, compared with respective reference serovars, demonstrated that all clinical Ds had 2–4 *ompA* nt polymorphisms compared to reference D, and D/U14bf and D/L11bf had 0.003–0.006 ± 0.002–0.003 p-distance from D (**Table 4**). Note that D/U1bf and D/U9bf harbored the same *ompA* polymorphisms that caused 0.000 p-distance against D (**Table 4**). *ompA* variants were also observed in 1 clinical F and 2 Hs. Most polymorphisms were of different nt positions, and many encoded for amino acid changes and thus high dN/dS. Bidirectional sequencing indicated H/U20bf and H/L23bf containing a substitution that caused a stop codon (W295). However, these *ompA* variants remained phylogenetically clustered with their respective reference strains (**Figure 1**).

Polymorphisms of trpA among 14 Thai clinical isolates

Clinical D, Gs, and H demonstrated *trpA* polymorphisms; many of which encoded for nonsynonymous amino acid changes, causing high dN/dS. Moreover, the *trpA* polymorphisms were at the positions that affected the phylogenetic relatedness among serovars. These clinical *trpA* variants became grouped with the different reference serovars: D/L11bf instead clustering with H, Ia, J, and Ja; and G/U18bf, G/U31bf, and H/U20bf clustered with F and I (**Table 4 and Figure 2**).

Analysis of C. trachomatis trpA polymorphisms on TrpA and TrpA–TrpB modeling structures

To investigate the possible effects of clinical *trpA* variants, three-dimensional structures of TrpA and active TrpA-TrpB complex were computed. The computed structures yielded the GA341 model scores of 1.0, which represented perfect modeling scores. This indicated that greater than 95% of the modeling structures were correct [20]. **Figure 3A** demonstrates the clinical *trpA* polymorphisms at around TrpA critical sites. For instance, C177Y was one of the active TrpA residues, and Q37R/R37Q was located close to another active TrpA residue. Further, for tryptophan synthesis to proceed, TrpA and TrpB subunits must form an active, planar TrpA–TrpB–TrpB–TrpA complex (Fig. 3B). During the process, the COMM domain of TrpB joined and lifted up TrpA to broaden the tunnel lining TrpA and TrpB active residues, allow indole substrate to be converted to tryptophan. In **Figure 3B**, C177Y is also positioned close to the interacting COMM domain. By contrast, when tryptophan is available, TrpA and TrpB do not interact and the COMM domain of TrpB shifts downward [24-26].

Discussion

Our findings that HPV is the most common STI, and *that C. trachomatis* is more common than *N. gonorrhoeae* (Table 2) are consistent with previous reports worldwide; including those from Thailand [1-3]. There was a greater prevalence of HPV and *C. trachomatis* in symptomatic subjects with UGT and LGT than in asymptomatic women and girls. Table 2 highlights that these pathogens cause a clinical burden more often than *N. gonorrhoeae.* In addition, coinfection between *C. trachomatis* and HPV was common (**Tables 2 and 3**). This is consistent with previous studies that reported *C. trachomatis* infection likely promoted acquisition and retention of high-risk HPV types that also cause cervical cancer [8, 15].

Table 4. Polymorphisms of *C. trachomatis ompA* and *trpA* among 14 Thai clinical isolates compared with the respective reference serovars

		ompA				*trpA*			
Reference *ompA* genotype	**Patient name**	**p-distance ± S.E.**	**nt differences (type and position of changes)**	**dN/dS* (type and position of changes)**	**Serovars that the clinical sequence is most similar to †**	**p-distance ± S.E.**	**nt differences (type and position of changes)**	**dN/dS (type and position of dN changes)**	**Serovars that the clinical sequence is most similar to**
D	D/U1bf	.000 ± .000	3 (G903A, A927G, A937G) Y313C)	3/0 (D301N, S309G,	**D**	.000 ± .000	0	–	**D**, K
	D/U9bf	.000 ± .000	3 (G903A, A927G, A937G)	3/0 (D301N, S309G, Y313C)	**D**	.000 ± .000	0	–	**D**, K
	D/U14bf	.006 ± .003	4 (C258G, C594T, T601C, G603C)	0/4	**D**	.000 ± .000	0	–	**D**, K
	D/L11bf	.003 ± .002	2 (C258G, G428C)	1/1 (S143T)	**D**	.005 ± .003	3 (A110G, C344T, G530A)	3/0 (Q37R, A115V, C177Y)	H, Ia, J, Ja
E	E/U2bf	.000 ± .000	0	–	**E**	.000 ± .000	0	–	**E**
	E/U4bf	.000 ± .000	0	–	**E**	.000 ± .000	0	–	**E**
	E/U28bf	.000 ± .000	0	–	**E**	.000 ± .000	0	–	**E**
	E/L8bf	.000 ± .000	0	–	**E**	.000 ± .000	0	–	**E**
F	F/U29bf	.000 ± .000	0	–	**F**	.000 ± .000	0	–	**F**, I
	F/L1bf	.000 ± .000	2 (G769C, G775C)	2/0 (R257T, S259T)	**F**	.000 ± .000	0	–	**F**, I
G	G/U18bf	.000 ± .000	0	–	**G**	.002 ± .002	1 (C10T)	0/1	F, I
	G/U31bf	.000 ± .000	0	–	**G**	.002 ± .002	1 (C10T)	0/1	F, I
H	H/U20bf	.000 ± .000	2 (G795A, G885A)	2/0 (G265R, W295.)	**H**	.004 ± .002	2 (G110A, T344C)	2/0 (R37Q, V115A)	F, I
	H/L23bf	.000 ± .000	4 (C270A, A271C, C850T, G885A)	4/0 (N90H, T284I, W295.)	**H**	.000 ± .000	0	–	**H**, Ia, J, Ja
K	K/U15bf	.000 ± .000	0	–	**K**	.000 ± .000	0	–	D, **K**

*dN/dS represents the ratio of non-synonymous (dN) to synonymous (dS) amino acid changes

†If >1 serovars are most similar to the *ompA* or *trpA* sequence of the clinical isolate, the serovars are listed in alphabetical order (A, Ba, C, D, Da, E, F, G, H, I, Ia, J, Ja, K, L1, L2, L2a, L3) and bold for the respective reference serovar. For *trpA* analysis, serovar B is excluded because of its lack of *trpA*

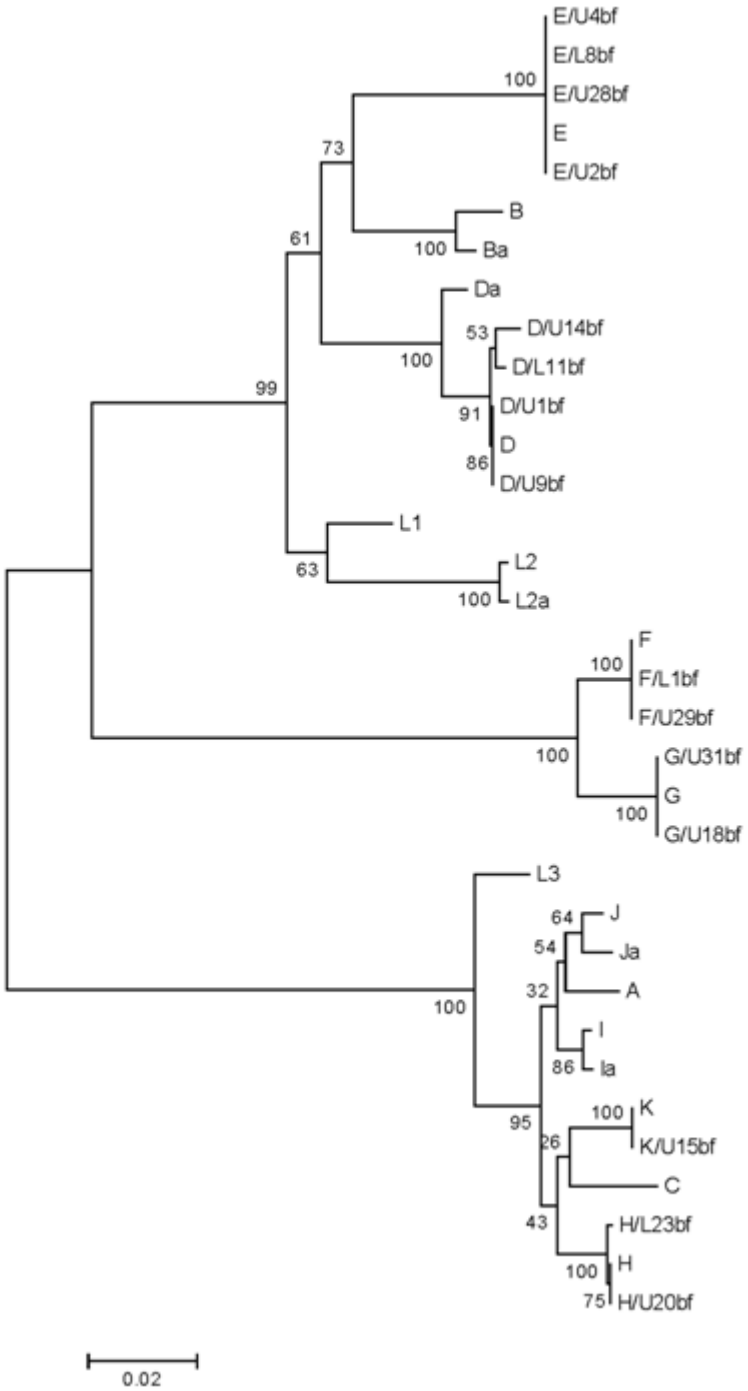

Figure 1. Neighbor-joining tree representing evolutionary relatedness of full-length *ompA* sequences from 19 *C. trachomatis* reference strains and 14 clinical isolates of serovars D, E, F, G, H, and K from Thai women and girls. All clinicals represent the isolates before antibiotic treatment. The length of branch is proportional to the distance between sequences, and number at node is the percent bootstrap confidence for clustering of strains.

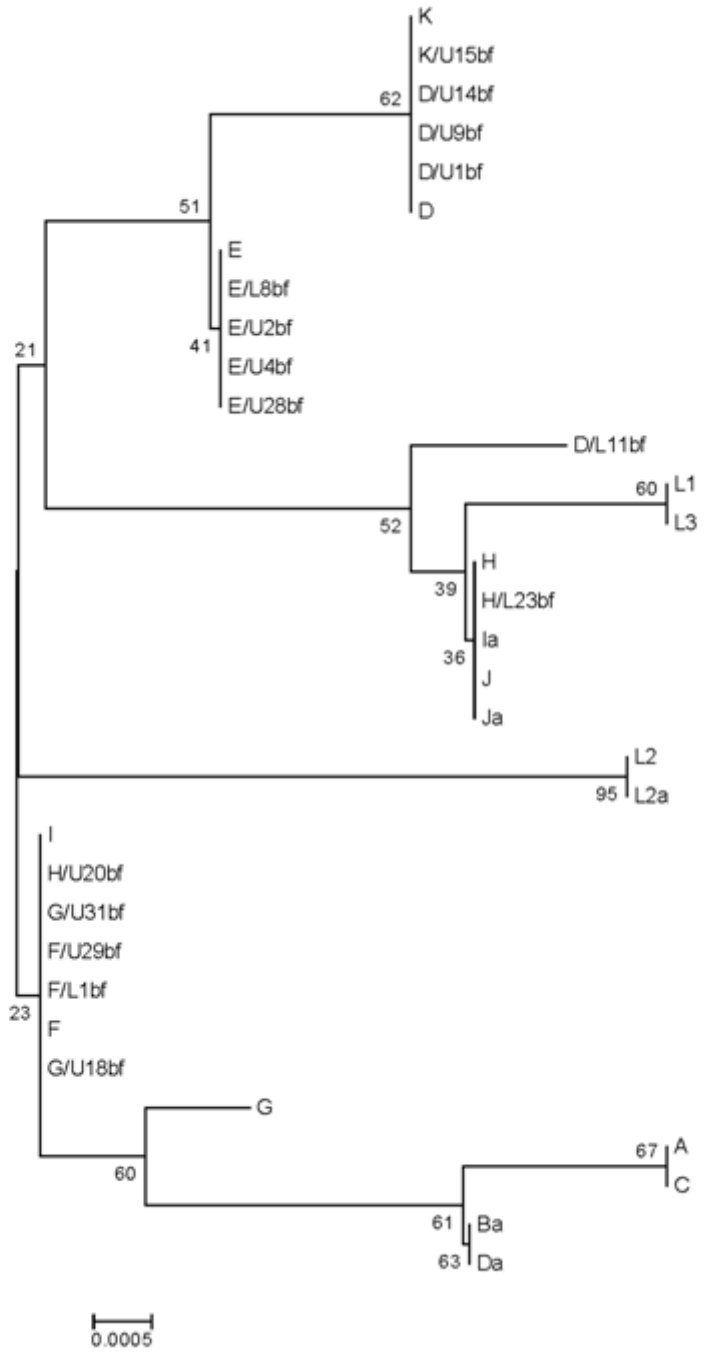

Figure 2. Neighbor-joining tree representing evolutionary relatedness of full-length *trpA* sequences from 19 *C. trachomatis* reference strains and 14 clinical isolates from Thai women and girls (serovars D, E, F, G, H and K). All clinical isolates represent the first collected isolates prior to antibiotic treatment. Length of branch is proportional to distance between sequences, and number at node is the percent bootstrap confidence for clustering of strains.

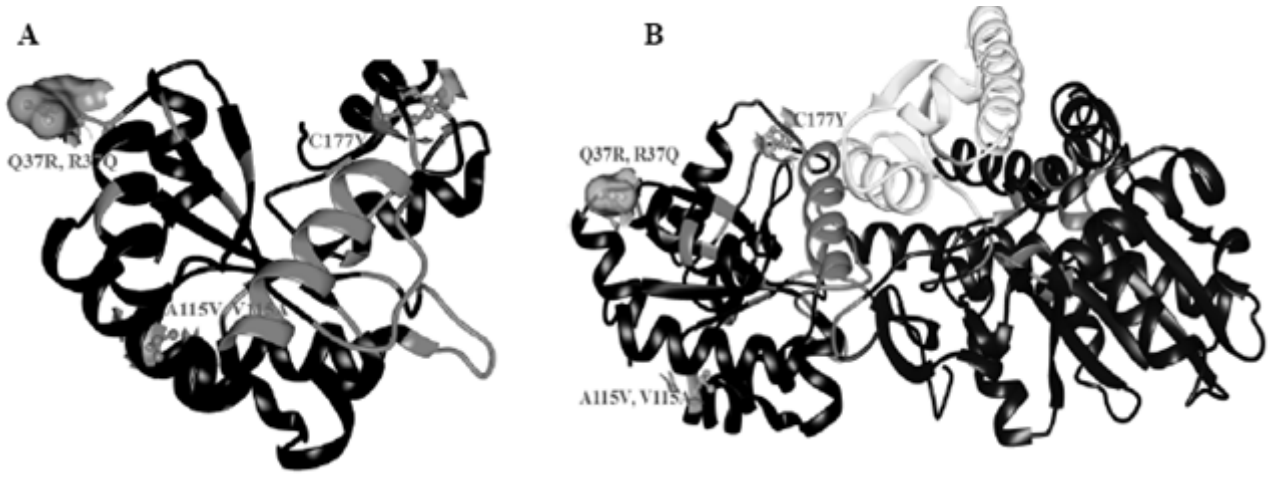

Figure 3. Three-dimensional structures of non-active TrpA (**A**) and active TrpA–TrpB complex (**B**), with the nonsynonymous amino acid polymorphisms in respected to the clinical D/L11bf and H/U20bf (Q37R/R37Q, A115V/V115A, and C177Y). Clinical polymorphisms were represented in gray with atomic bonds, critical TrpA and TrpB residues (including C177Y) in gray, and TrpB COMM domain in white.

Younger symptomatic patients (15–34 years), and non-CSWs were more often found with the STIs (**Table 2**: symptomatic group). Younger age might represent higher risks because of unprotected sex, multiple sex partners, and shared needle use [4]. *C. trachomatis* and HPV infections were also more common in the UGT than the LGT (**Table 2**), supporting the fact that *C. trachomatis* could reach and induce excess inflammation at the UGT including ovary and spermatic tubes. This can result in sterility [2, 5, 8, 10]. The greater finding of *C. trachomatis* in women and girls with UGT symptoms is associated with multiple disease complications [5, 8].

Asymptomatic women and girls, who reported no clinical symptom and had normal urogenital tracts on examination, were nevertheless infected by STIs (38 patients or 28%). Such infected individuals can transmit the diseases unwittingly, and might have clinical symptoms later in their lives.

Table 3 supports the inconsistent success of common *C. trachomatis* approved antibiotics and the persistent state of *C. trachomatis* infection and/or resistance where drugs had been used [10]. These included tetracycline, azithromycin, and doxycycline, which target bacterial protein synthesis. A state of persistent *C. trachomatis* infection with no or only minimal protein synthesis, offers no target for the action of these antibiotics [7, 8, 10-12]. Previous studies examined whether Chlamydiae that were antibiotic-resistant in vivo were because of antibiotic-resistant strains or not. They found that all Chlamydiae, once cultivated in vitro in a normal life cycle, remained sensitive to these antibiotics [7]. In vitro, interferon-γ induces persistence of *C. trachomatis*, which is able to resume a normal life cycle upon tryptophan supplementation. Unfortunately, only few of our patients returned for follow-up after antibiotic treatment. This suggests that there may have been undetected treatment failures and therefore further risk for transmission [2, 8, 15].

Our study of *C. trachomatis ompA* serovar distribution adds knowledge of the serovar dynamic in Thailand and is consistent with previous reports [3, 4]. Serovars D, E, and F remain the most common, followed by G and H, and K in this order (**Table 3**). This might also highlight the successful phenotypic characteristics of these serovars. Some *ompA* mutations were identified in Ds, F, and Hs. G885A in 2 and Hs caused an early stop codon at W295 compared with the reference H. A similar scenario was previously reported for clinical Da/TW-448 (GenBank no. X62921), but at a different nt position. Therefore, the 2 clinical Hs from Thai patients with symptomatic cases and the clinical Da might harbor some similar evolutionary pathway. A similar mutation pattern was diagnosed; yet this still requires validation. Nonetheless, Thai *ompA* variants in this study remained clustered with their respective references (**Table 4 and Figure 1**). Therefore, these *ompA* polymorphisms might cause no effect on the OmpA phenotype between the Thai and the reference strains.

To date, there are limited studies on *trpA* variants in clinical samples from Asia. Meanwhile, *trpA* was reported to contain several hotspots that could affect phylogenetic clustering and the tryptophan synthesis ability of Chlamydiae [13, 14]. This study included *trpA* sequencing and TrpA three-dimensional structure analysis to understand better the clinical *trpA* representing recent clinical strains in Thailand. Multiple sequence alignment and phylogenetic trees could not cluster some Thai clinical *trpA* with the respective references (**Table 4 and Figure 2**). It is further signified that some Thai clinical *trpA* mutations might be of importance, possibly correlate with altered TrpA phenotypes, and might help understand chlamydial pathogenesis.

To help evaluate clinical effects of *trpA* polymorphisms, three-dimensional structures of TrpA and active TrpA-TrpB complex were modeled with confident score (GA341 = 1.0), and the nonsynonymous amino acid mutations were analyzed. While diminished tryptophan could lead to chlamydial eradication, a restricted tryptophan level could induce *C. trachomatis* persistence, as TrpA allows *Chlamydia* to resume a normal life cycle by synthesizing tryptophan when indole substrate is available [14]. TrpA and TrpB form an active complex where the COMM domain of TrpB shifts up to broaden the TrpA–TrpB tunnel, and allows indole to catalyze to tryptophan [25-28]. Interestingly, the C177Y mutation in D/L11bf was at an active TrpA site (**Figure 3A**), so the mutation might affect TrpA activity. C177Y, which is also located close to the TrpA–TrpB interaction (**Figure 3B**), further affecting the connection between TrpA and TrpB [24, 25]. The change from cysteine (polar with thiol side chain (R–SH)) to tyrosine (aromatic R group) was critical, because the thiol side chain is often involved in enzymatic reaction [24, 25, 29]. Moreover, cysteine often forms a disulfide bond (S–S) with another cysteine to tighten and stabilize the structural fold, so its mutation from cysteine to tyrosine might affect the structure of the protein. This could be a way to favor chlamydial persistence and strain persistence in spite of antibiotic treatment (**Table 3**). For instance, the disrupted fold by C177Y might impair the tryptophan synthesis ability, causing the chlamydiae to remain persistent even when indole substrate is available. The other mutations, Q37R or R37Q and A115V or V115A (**Table 4**), served hotspots that distinguish between reference serovars [13]. For Q37R/R37Q, the substitution between glutamine and arginine could affect the charge type of protein, because glutamine is a negatively charged amino acid with acidic side chain, while arginine is a positively charged amino acid with basic side chain [29]. For A115V/V115A, valine is a big-size amino acid that is often found associated with β-sheet formation [29]. This substitution could affect TrpA conformation, because residue 115 was part of the helix (**Figure 3**).

Conclusions

Our study revealed the frequency of STIs in symptomatic and apparently healthy asymptomatic Thai women and girls, and evidence for *C. trachomatis* persistent infection in patients attending STD clinics in Bangkok and nearby. The high rate of HPV and *C. trachomatis* among symptomatic individuals suggests increased virulence of these infections compared with *N. gonorrhoeae*, and the common coinfection between *C. trachomatis* and HPV. Persistent *C. trachomatis* infection is serious, as it promotes severe clinical outcomes (i.e. pelvic inflammatory disease, ectopic pregnancy, and sterility) and predisposes to high-risk HPV types and antibiotic treatment failures. These findings are consistent with previous reports. Persistence of *C. trachomatis,* despite treatment with doxycycline, is of growing concern. While clinical *ompA* sequences were still relatively close to reference sequences, clinical *trpA* sequences contained mutations that might affect the TrpA structure and function. Knowledge of these *trpA* variants may help to understand chlamydial evolution and pathogenesis in the Thai population.

Acknowledgements

The authors thank D. Dean for research advice. This work was supported by the National Science and Technology Development Agency (NSTDA), and Chulalongkorn University through the Omics Sciences and Bioinformatics Center. Choopara acknowledges the Research Professional Development Project under Science Achievement Scholarship of Thailand.

Conflict of interest statement

The authors have no conflicts of interest to declare.

References

1. World Health Organization. Sexually transmitted diseases. [online] 2007. [cited 2007 Nov 29]; Available from: *http://www.who.int/vaccine_research/diseases/soa_std/en/print.html*

2. Centers for Disease Control and Prevention (CDC). Sexually Transmitted Diseases. U.S. Department of Health and Human Services, Atlanta, GA. [online] 2014. [cited 2014 Sep 3]; Available from *http://www.cdc.gov/std/default.htm*
3. Rugpao S, Sirirungsi W, Vannareumol P, Leechanachai P, Wongtrangarn S, Niyomka P, et al. Isolation of *Chlamydia trachomatis* among women with symptoms of lower genital tract infection. J Med Assoc Thai. 1993; 76:475-81.
4. Celentano DD, Sirirojn B, Sutcliffe CG, Quan VM, Thomson N, Keawvichit R, et al. Sexually transmitted infections and sexual and substance use correlates among young adults in Chiang Mai, Thailand. Sex Transm Dis. 2008; 35:400-5.
5. den Hartog JE, Morre SA, Land JA. *Chlamydia trachomatis*-associated tubal factor subfertility: immunogenetic aspects and serological screening. Hum Reprod Update. 2006; 12:719-30.
6. Bragina EY, Gomberg MA, Dmitriev GA. Electron microscopic evidence of persistent chlamydial infection following treatment. J Eur Acad Dermatol Venereol. 2001; 15:405-9.
7. Dean D, Suchland RJ, Stamm WE. Evidence for long-term cervical persistence of *Chlamydia trachomatis* by *omp1* genotyping. J Infect Dis. 2000; 182:909-16.
8. Hogan RJ, Mathews SA, Mukhopadhyay S, Summersgill JT, Timms P. Chlamydial persistence: beyond the biphasic paradigm. Infect Immun. 2004; 72:1843-55.
9. Leonhardt RM, Lee S, Kavathas PB, Cresswell P. Severe tryptophan-starvation blocks onset of conventional persistence and reduces reactivation of *Chlamydia trachomatis*. Infect Immun. 2007; 75: 5105-17.
10. Singla M. Role of tryptophan supplementation in the treatment of *Chlamydia*. Medical Hypo. 2007; 68: 278-80.
11. Wang SA, Papp JR, Stamm WE, Peeling RW, Martin DH, Holmes KK. Evaluation of antimicrobial resistance and treatment failures for *Chlamydia trachomatis*: a meeting report. J Infect Dis. 2005; 191:917-23.
12. Somboonna N, Mead S, Liu J, Dean D. Discovering and differentiating new and emerging clonal populations of *Chlamydia trachomatis* with a novel shotgun cell culture harvest assay. Emerg Infect Dis. 2008; 14:445-53.
13. Caldwell HD, Wood H, Crane D, Bailey R, Jones RB, Mabey D, et al. Polymorphisms in *Chlamydia trachomatis* tryptophan synthase genes differentiate between genital and ocular isolates. J Clin Invest. 2003; 111:1757-69.
14. Wood H, Fehlner-Gardiner C, Berry J, Fischer E, Graham B, Hackstadt T, et al. Regulation of tryptophan synthase gene expression in *Chlamydia trachomatis*. Mol Microbiol. 2003; 49:1347-59.
15. Shew ML, Ermel AC, Weaver BA, Tong Y, Tu W, Kester LM, et al. Association of *Chlamydia trachomatis* infection with redetection of human papillomavirus after apparent clearance. J Infect Dis. 2013; 208:1416-21.
16. Gravitt PE, Peyton CL, Alessi TQ, Wheeler CM, Coutlee F, Hildesheim A, et al. Improved amplification of genital human papillomaviruses. J Clin Microbiol. 2000; 38:357-61.
17. Goire N, Nissen MD, LeCornee GM, Sloots TP, Whiley DM. A duplex *Neisseria gonorrhoeae* real-time polymerase chain reaction assay targeting the gonococcal *porA* pseudogene and multicopy *opa* genes. Diagn Microbiol Infect Dis. 2008; 61:6-12.
18. Dictor M, Warenholt J. Single-tube multiplex PCR using type-specific *E6/E7* primers and capillary electrophoresis genotypes 21 human papillomaviruses in neoplasia. Infect Agent Cancer. 2011; 6. doi:10.1186/1750-9378-6-1
19. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997; 25:3389-402.
20. Eswar N, John B, Mirkovic N, Fiser A, Ilyin VA, Pieper U, et al. Tools for comparative protein structure modeling and analysis. Nucleic Acids Res. 2003; 31: 3375-80.
21. Schaffer AA, Wolf YI, Ponting CP, Koonin EV, Aravind L, Altschul SF. IMPALA: matching a protein sequence against a collection of PSI-BLAST-constructed position-specific score matrices. Bioinformatics. 1999; 15:1000-11.
22. Pieper U, Eswar N, Braberg H, Schneidman-Duhovny D, Fan H, Kim SJ, et al. MODBASE, a database of annotated comparative protein structure models, and associated resources. Nucleic Acids Res. 2004; 32: D217-22.
23. Meng EC, Pettersen EF, Couch GS, Huang CC, Ferrin TE. Tools for integrated sequence-structure analysis with UCSF Chimera. BMC Bioinformatics. 2006; 7: 339.
24. Weyand M, Schlichting I. Crystal structure of wild-type tryptophan synthase complexed with the natural substrate indole-3-glycerol phosphate. Biochem. 1999;

38:16469-80.
25. Lee SJ, Ogasahara K, Ma J, Nishio K, Ishida M, Yamagata Y, et al. Conformational changes in the tryptophan synthase from a hyperthermophile upon $\alpha_2\beta_2$ complex formation: crystal structure of the complex. Biochem. 2005; 44:11417-27.
26. Schneider TR, Gerhardt E, Lee M, Liang PH, Anderson KS, Schlichting I. Loop closure and intersubunit communication in tryptophan synthase. Biochem. 1998; 37:53394-406.
27. Akers JC, Tan M. Molecular mechanism of tryptophan-dependent transcriptional regulation in *Chlamydia trachomatis*. J Bacteriol. 2006; 188: 4236-43.
28. Dunn MF, Aguilar V, Brzovic P, Drewe WF Jr, Houben KF, Leja CA, et al. The tryptophan synthase bienzyme complex transfers indole between the α- and β- sites via a 25–30 □ long tunnel. Biochem. 1990; 29:8598-607.
29. Betts MJ, Russell RB. Amino acid properties and consequences of substitutions. In: Barnes MR, Gray IC, editors. Bioinformatics for geneticists. West Sussex, England: John Wiley & Sons; 2003. p. 289-316.