

## Synthesis, Characterization, and Molecular Docking Against a Receptor Protein FimH of *Escherichia coli* (4XO8) of Thymidine Derivatives

Asraful Alam<sup>1</sup>, Mohammed Anowar Hosen<sup>1</sup>, Anowar Hosen<sup>2</sup>, Yuki Fujii<sup>3</sup>, Yasuhiro Ozeki<sup>4</sup>, Sarkar Mohammad Abe Kawsar<sup>1\*</sup>

<sup>1</sup>Laboratory of Carbohydrate and Nucleoside Chemistry (LCNC), Department of Chemistry, Faculty of Science, University of Chittagong, Chittagong-4331, Bangladesh.

<sup>2</sup>Centre for Advanced Research in Sciences, University of Dhaka, Dhaka-1000, Bangladesh.

<sup>3</sup>Laboratory of Functional Morphology, Graduate School of Pharmaceutical Sciences, Nagasaki International University, 2825-7 Huis Ten Bosch, Sasebo, Nagasaki 859-3298, Japan.

<sup>4</sup>Laboratory of Glycobiology and Marine Biochemistry, Department of Life and Environmental System Science, Graduate School of NanoBiosciences, Yokohama City University, 22-2 Seto, Kanazawa-ku, Yokohama 236-0027, Japan.

\*Corresponding author: Sarkar Mohammad Abe Kawsar, email: [akawsarabe@yahoo.com](mailto:akawsarabe@yahoo.com), Tel: +88 01762717081.

Received October 22<sup>nd</sup>, 2020; Accepted January 25<sup>th</sup>, 2021.

DOI: <http://dx.doi.org/10.29356/jmcs.v65i1.1464>

**Abstract.** Thymidine is known as a progenitor of nucleosides that have significant biological activity. The widening importance of nucleoside derivatives as unrivaled potential antimicrobial and therapeutic agents has attracted contemplation to the synthesis of thymidine derivatives. In the present study, thymidine was treated with various acyl halides to produce 5'-O-acyl thymidine derivatives by direct acylation method with an excellent yield. To obtain newer products for antimicrobial assessment studies, the 5'-O-thymidine derivatives were further modified into three series of 3'-O-acyl thymidine derivatives containing a wide variety of functionalities in a single molecular framework. The chemical structures of the newly synthesized compounds were elucidated by analyzing their physicochemical, elemental, and spectroscopic data. Additionally, the X-ray powder diffraction (XRD) of these acylated products was studied. For the computational investigation, we have selected eight synthesized thymidine derivatives, which have notable antibacterial activity, and performed molecular docking against bacterial lectin protein FimH of *Escherichia coli* (4XO8) to suggest a potent inhibitor against bacterial function. Molecular docking was performed using AutoDock Vina to calculate the binding affinities and interactions between the antibacterials and the FimH *E. coli* (4XO8). It was found that the selected thymidine derivatives have strongly interacted mainly with Tyr48, Tyr137, Asp140, Arg98, Gln133, Phe1, Asn23, Asn135, Lys76, Asp47, Ile13, and Ile52 residues. *In silico* pharmacokinetic properties were also predicted to search their absorption, metabolism, excretion, and toxicity. This computational examination showed that these thymidine derivatives might be used as potential inhibitors against the promising antibacterial activity for future studies.

**Keywords:** Thymidine; synthesis; *Escherichia coli*; molecular docking; DFT; inhibitor; lectin protein (FimH).

**Resumen.** Se prepararon varios derivados 5'-O-acil timidínicos por acilación directa con rendimientos excelentes que fueron transformados en tres series de derivados 3'-O-acil timidínicos con una amplia variedad de funcionalidades. Estos compuestos fueron la base de un estudio de docking dirigido a la lectina bacteriana FimH de *Escherichia coli* (4XO8) con la finalidad de proponer un inhibidor contra esta función bacteriana.

**Palabras clave:** Acilación de Timidina, Docking, lectina bacteriana (FimH), inhibidor, DFT.

## Abbreviations

DFT: Density Functional Theory

QM: Quantum Mechanical

LYP: Lee, Yang and Parr's

## Introduction

Nucleoside derivatives (NDs) have proven to be a fruitful field for drug discovery [1]. Due to the ubiquitous nature of NDs in biological settings, modifications in their structures can potentially lead to powerful biological effects. Much research has been conducted in this field producing many medicinally important compounds. NDs have proven particularly effective as anticancer agents [2,3]. These derivatives have also been successful in producing antiviral therapeutics [4,5].

Some of the most effective antiviral agents currently in use are purine or pyrimidine. Thymidine (Fig. 1), structurally known as deoxythymidine is a pyrimidine-based nucleoside that constitutes a major part of one of the four nucleotides in DNA and is listed as a chemical teratogen [6]. Modification of the hydroxyl (-OH) group at 3' and 5' position increases the activity of thymidine derivatives and bring about some lifesaving antiviral and antibacterial drug. Azidothymidine (AZT) is one of the most popular thymidine derivatives (antiviral drugs) in which 3' hydroxyl (-OH) of thymidine is modified by an azide group and now it's used worldwide for the treatment of HIV infection [7]. AZT suppresses the mode of reverse transcription, a ticklish phase in the life cycle of the virus. Edoxudine is another antiviral drug of thymidine, strongly working against herpes simplex virus [8-10]. Moreover, thymidine plays a key role in cell biology to synchronize cells. Thymidine analog Bromodeoxyuridine is globally used to identify the proliferating of cells in living tissues. Thymidine is also catabolized to find out TP-expressing tumor xenografts [11]. The modified compounds exhibited a broad-spectrum biological activity. For example, zidovudine, [12] with the azido group at the 3' position, is used for the treatment of HIV. Thymidine derivatives such as stavudine and zalcitabine are used in the treatment of HIV infections and tumors [13]. Hence, modifications in the sugar moiety of NDs have resulted in a variety of effective therapeutic applications.

Thus, we attempted to search for novel NDs with potent biological activity. Several efficient methods for selective modification were reported by chemists using numerous acylating agents and varying reaction conditions [14-18]. Various methods for acylation of carbohydrates and nucleosides have been developed and employed successfully [19-21]. Of these, the direct method is the most promising for the acylation of nucleosides [21].

FimH is a well-studied lectin, which is used by bacteria to adhere to glycosylated surfaces, such as cells [22]. Many natural, as well as synthetic mannosidic ligands of FimH, have been developed and examined for their inhibitory potencies [22]. Eosinophil cationic protein (ECP), also known as RNase 3, is an eosinophil secretion protein that plays a key role in innate immunity and displays antipathogen and proinflammatory activities. ECP has a high binding affinity for heterosaccharides, such as bacterial lipopolysaccharides and heparan sulfate found in the glycocalyx of eukaryotic cells [23]. ECP also belongs to the mammalian RNase A superfamily and its RNase activity is required for some of its biological properties [24]. Trp10 and Trp35 together with the adjacent stacking residue are critical for the damaging effect of ECP on mammalian cell lines. These residues are also crucial for the membrane disruption activity of ECP. Other exposed aromatic residues packed against arginine's (Arg75-Phe76 and Arg121-Tyr122) and specific cationic amino acids (Arg101 and Arg104) of ECP play a secondary role in the cell growth inhibition [23]. Again, residues Ile13, Ile52, Lys38, Lys76, Arg36, Arg98, and His15 are identified to be involved in the biological properties [23].

Computational chemistry is a popular tool to predict physicochemical, spectral, and biological properties of newly synthesized chemicals [25]. The prime aim of this study attempted to synthesize new biologically potent thymidine derivatives and investigate the binding affinities and interactions of antibacterial thymidine derivatives against the *E. coli* (4XO8) using computational tools to point out a potent inhibitor against bacterial function. Additionally, synthesized thymidine derivatives were optimized to minimize energy and some selected derivatives (having better antibacterial activity) were employed for molecular docking against bacterial protein FimH of *E. coli* (4XO8) to understand their nonbonding interactions, binding mode, and

binding affinity as the first time. Moreover, to ensure the drug-like ability standard computational pharmacokinetic parameters *in silico* (ADMET) has been predicted.

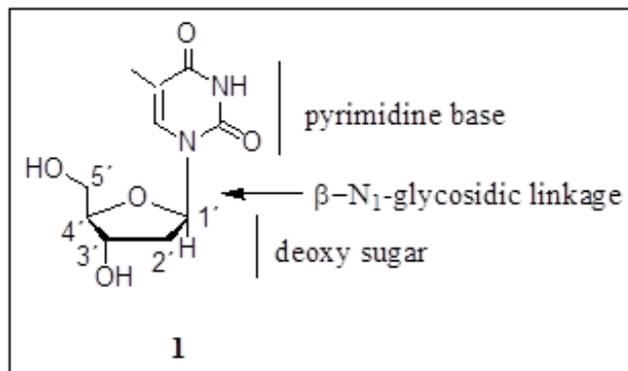


Fig. 1. Structure of the thymidine (1).

## Experimental

### Materials and methods

Thin-layer chromatography (TLC) was performed on Kieselgel GF<sub>254</sub> (Germany), and the chromatogram was visualized by spraying the plates with 1% H<sub>2</sub>SO<sub>4</sub>, followed by heating the plates at 150–200°C until coloration appeared. Melting points (mp) were determined on an electrothermal melting point apparatus. Evaporations were performed under diminished pressure on a Büchi rotary evaporator. The solvents used were of analytical grade and were purified using standard procedures. Infrared spectral analyses were recorded using a Fourier-transform infrared (FTIR) spectrophotometer (IR Prestige-21, Shimadzu, Japan) within 200–4000 cm<sup>-1</sup> at the Department of Chemistry, University of Chittagong, Bangladesh. Mass spectra of the synthesized compounds were obtained by liquid chromatography-electrospray ionization tandem mass spectrometry in positive ionization mode. A Bruker advance DPX 400 MHz using tetramethylsilane as an internal standard was employed for recording the proton nuclear magnetic resonance (<sup>1</sup>H-NMR) spectra in CDCl<sub>3</sub> ( $\delta$  in ppm) at WMSRC, JU, Bangladesh. XRD patterns were obtained using an XRD-53 analyzer diffractometer (Rigaku, Japan) with a back monochromator and a Cu target and K $\alpha$  ( $\lambda$  = 1.5406 nm) in  $2\theta$  = 2–70° at the CARS, Dhaka University, Bangladesh. Column chromatography was performed using silica gel G<sub>60</sub>. CHCl<sub>3</sub>/CH<sub>3</sub>OH in different proportions was employed as the solvent system for TLC analyses. All reagents used were commercially available Sigma-Aldrich (Germany) and were used as received unless otherwise specified.

### Synthesis

Over the past several years, our laboratory has been involved in the synthesis of nucleoside derivatives containing various acyl groups to explore their antimicrobial properties [26, 27].

#### Procedure for the synthesis of 5'-O-(pivaloyl)thymidine (2)

A solution of thymidine (1) (200 mg, 0.82 mmol) having hydroxyl group(s) was dissolved in dry pyridine (approximately 3 ml). The solution was cooled to 0°C, and pivaloyl chloride (0.11 ml, 1.1 molar eq.) was added dropwise with continuous stirring. The reaction mixture was allowed to attain room temperature, after which 4-dimethylaminopyridine (DMAP, catalytic amount) was added. The mixture was further stirred for 5–6 h, and TLC (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 4/1, v/v) indicated the presence of a faster moving product. Excess pivaloyl chloride was decomposed by adding approximately 3 ml of ice water. The organic product was extracted with chloroform (CHCl<sub>3</sub>, 3 × 10 ml), and this organic layer was washed successively with 5% HCl,

saturated aqueous NaHCO<sub>3</sub> solution, and brine. The organic layer was dried over MgSO<sub>4</sub> and concentrated in a vacuum to obtain a residue. Chromatographic purification of the residue (elution with CHCl<sub>3</sub>/CH<sub>3</sub>OH) yielded the pivaloyl derivative (2) (82.7 mg) in the pure form as a crystalline solid, which was used in the next stage. White crystalline solid in 78.4% yield. mp: 66–68 °C, R<sub>f</sub> = 0.52 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 4/1, v/v); FTIR (KBr) ν<sub>max</sub> 1690 (-CO), 3410~3535 (br) (-OH) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 9.01 (1H, s, -NH), 7.31 (1H, d, J = 2.3 Hz, H-6), 6.32 (1H, t, J = 6.6 Hz, H-1'), 4.41 (1H, dd, J = 12.0 and 4.5 Hz, H-5'a), 4.36 (1H, dd, J = 12.1 and 4.5 Hz, H-5'b), 4.00 (1H, m, H-3'), 3.91 (1H, ddd, J = 3.6, 4.6 and 4.2 Hz, H-4'), 3.45 (1H, s, 3'-OH), 2.48 (1H, ddd, J = 13.5, 6.5 and 4.0 Hz, H-2'a), 2.15 (1H, ddd, J = 13.5, 6.5 and 6.7 Hz, H-2'b), 1.92 (3H, d, J = 1.3 Hz, 5-CH<sub>3</sub>), 1.38 {9H, s, (CH<sub>3</sub>)<sub>3</sub>CCO-}. LC-MS [M+1]<sup>+</sup> 327.31. Analysis calcd for C<sub>14</sub>H<sub>22</sub>O<sub>5</sub>N<sub>2</sub>CO: C, 55.16, H, 6.74%, found: C, 55.18, H, 6.73%.

### General procedure for the direct 3'-O-acylation of 5'-O-(pivaloyl)thymidine derivatives (3-6)

A suspension of the 5'-O-pivaloylthymidine (2) (62.7 mg, 0.82 mmol) in dry pyridine (3 ml) was cooled to -5°C, after which octanoyl chloride (0.10 ml, 2.2 molar eq.) was added. The solution was stirred at this temperature for 6 h and then allowed to stand at room temperature for 5–6 hrs. The progress of the reaction was assessed by TLC (CHCl<sub>3</sub>/CH<sub>3</sub>OH) which indicated full conversion of the starting material into a single product. Excess reagent was decomposed by the addition of a few pieces of ice, and the reaction mixture was extracted with CHCl<sub>3</sub> (3×10 ml). The combined organic extracts were washed successively with dil HCl, aqueous saturated NaHCO<sub>3</sub>, and distill H<sub>2</sub>O. The organic layer was dried over (MgSO<sub>4</sub>) and filtered, and the filtrate was evaporated off. The resulting syrup was purified by column chromatography (with CHCl<sub>3</sub>/CH<sub>3</sub>OH as eluant) to yield the octanoyl derivative (3) (90 mg) as a needle.

### 3'-O-Octanoyl-5'-O-pivaloylthymidine (3)

Off white crystalline solid in 80.91% yield: mp 59-61 °C, R<sub>f</sub> = 0.54 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 8/1, v/v); FTIR (KBr) ν<sub>max</sub> 1704 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 9.01 (1H, s, -NH), 7.30 (1H, d, J = 1.8 Hz, H-6), 6.31 (1H, t, J = 6.6 Hz, H-1'), 4.49 (1H, m, H-3'), 4.41 (1H, dd, J = 12.0 and 4.5 Hz, H-5'a), 4.24 (1H, dd, J = 12.2 and 3.6 Hz, H-5'b), 4.20 (1H, m, H-4'), 2.51 (1H, ddd, J = 13.6, 6.6 and 4.6 Hz, H-2'a), 2.33 (1H, ddd, J = 12.5, 6.1 and 6.2 Hz, H-2'b), 2.32 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>5</sub>CH<sub>2</sub>CO-}, 1.93 (3H, d, J = 1.3 Hz, 5-CH<sub>3</sub>), 1.68 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>4</sub>CH<sub>2</sub>CH<sub>2</sub>CO-}, 1.29 {8H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>4</sub>(CH<sub>2</sub>)<sub>2</sub>CO-}, 1.37 {9H, s, (CH<sub>3</sub>)<sub>3</sub>CCO-}, 0.88 {3H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>6</sub>CO-}. LC-MS [M+1]<sup>+</sup> 453.31. Analysis calcd for C<sub>22</sub>H<sub>36</sub>O<sub>6</sub>N<sub>2</sub>CO: C, 61.02, H, 7.96%, found: C, 61.05, H, 7.98%.

Similar reaction and purification procedures were applied to prepare compound 4 (104 mg) compound 5 (113 mg), and compound 6 (176 mg).

### 3'-O-Myristoyl-5'-O-pivaloylthymidine (4)

White solid in 95.2% yield: mp 78–80 °C, R<sub>f</sub> = 0.50 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 8/1, v/v); FTIR (KBr) ν<sub>max</sub> 1705 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 9.01 (1H, s, -NH), 7.30 (1H, d, J = 1.3 Hz, H-6), 6.21 (1H, t, J = 6.5 Hz, H-1'), 4.48 (1H, m, H-3'), 4.41 (1H, dd, J = 12.0 and 4.5 Hz, H-5'a), 4.21 (1H, dd, J = 12.0 and 3.5 Hz, H-5'b), 4.05 (1H, ddd, J = 3.5, 4.5 and 3.9 Hz, H-4'), 2.38 (1H, ddd, J = 13.5, 6.5 and 4.0 Hz, H-2'a), 2.33 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>11</sub>CH<sub>2</sub>CO-}, 2.20 (1H, ddd, J = 13.5, 6.5 and 6.7 Hz, H-2'b), 1.91 (3H, d, J = 1.3 Hz, 5-CH<sub>3</sub>), 1.72 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>10</sub>CH<sub>2</sub>CH<sub>2</sub>CO-}, 1.64 {20H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>10</sub>CH<sub>2</sub>CH<sub>2</sub>CO-}, 1.36 {9H, s, (CH<sub>3</sub>)<sub>3</sub>CCO-}, 0.87 {3H, t, J = 6.8 Hz, CH<sub>3</sub>(CH<sub>2</sub>)<sub>12</sub>CO-}. LC-MS [M+1]<sup>+</sup> 537.21. Analysis calcd for C<sub>28</sub>H<sub>48</sub>O<sub>6</sub>N<sub>2</sub>CO: C, 64.89, H, 8.95%, found: C, 64.90, H, 8.97%.

### 3'-O-Cinnamoyl-5'-O-pivaloylthymidine (5)

White crystalline solid in 89.8% yield: mp 66–68 °C, R<sub>f</sub> = 0.51 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 7/1, v/v); FTIR (KBr) ν<sub>max</sub> 1735 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.88 (1H, s, -NH), 7.73 (1H, d, J = 12.0 Hz, PhCH=CHCO-), 7.50 (2H, m, Ar-H), 7.43 (1H, d, J = 1.8 Hz, H-6), 7.39 (3H, m, Ar-H), 6.49 (1H, d, J = 12.1 Hz, PhCH=CHCO-), 6.39 (1H, t, J = 6.4 Hz, H-1'), 4.51 (1H, m, H-3'), 4.38 (1H, dd, J = 12.1 and 4.5 Hz, H-5'a), 4.29 (1H, dd, J = 12.0 and 3.6 Hz, H-5'b), 3.98 (1H, ddd, J = 3.8, 4.8 and 3.8 Hz, H-4'), 2.60 (1H, ddd, J = 13.4, 6.4 and 4.4 Hz, H-2'a), 2.19 (1H, ddd, J = 13.5, 6.6 and 6.8 Hz, H-2'b), 1.89 (3H, d, J = 1.6 Hz, 5-CH<sub>3</sub>),

1.36 {9H, s, (CH<sub>3</sub>)<sub>3</sub>CCO-}. LC-MS [M+1]<sup>+</sup> 457.11. Analysis calcd for C<sub>23</sub>H<sub>28</sub>O<sub>6</sub>N<sub>2</sub>CO: C, 63.11, H, 6.14%, found: C, 63.14%, H, 6.18%.

### 5'-O-Pivaloyl-3'-O-tritylthymidine (6)

White crystalline solid in 94.35% yield: mp 142–144 °C, R<sub>f</sub> = 0.50 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 8/1, v/v); FTIR (KBr) ν<sub>max</sub> 1738 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.96 (1H, s, -NH), 7.74 (6H, m, Ar-H), 7.35 (9H, m, Ar-H), 7.14 (1H, d, J = 1.8 Hz, H-6), 6.34 (1H, t, J = 6.4 Hz, H-1'), 4.43 (1H, m, H-3'), 4.40 (1H, dd, J = 12.1 and 4.5 Hz, H-5'a), 4.36 (1H, dd, J = 12.0 and 3.6 Hz, H-5'b), 3.96 (1H, ddd, J = 3.8, 4.8 and 3.8 Hz, H-4'), 2.46 (1H, ddd, J = 13.4, 6.4 and 4.4 Hz, H-2'a), 2.12 (1H, ddd, J = 13.5, 6.6 and 6.8 Hz, H-2'b), 1.92 (3H, d, J = 1.6 Hz, 5-CH<sub>3</sub>), 1.35 {9H, s, (CH<sub>3</sub>)<sub>3</sub>CCO-}. LC-MS [M+1]<sup>+</sup> 569.08. Analysis calcd for C<sub>33</sub>H<sub>36</sub>O<sub>5</sub>N<sub>2</sub>CO: C, 71.79, H, 6.33%, found: C, 71.81, H, 6.36%.

### Procedure for the synthesis of 5'-O-(acetyl)thymidine (7)

A solution of thymidine (1) (200 mg, 0.82 mmol) in dry pyridine (3 ml) was cooled to -5°C, after which acetic anhydride (0.085 ml, 1.1 molar eq.) was added to it. The mixture was stirred at this temperature for 6–7 h and further at room temperature overnight. T.L.C. examination (CHCl<sub>3</sub>/CH<sub>3</sub>OH) indicated full conversion of the starting material into a single product. Work-up as usual and purification by silica gel column chromatography (CHCl<sub>3</sub>/CH<sub>3</sub>OH, as eluant) yielded the title compound (7) (92.4 mg) as a white crystalline solid. White crystalline solid in 90.96% yield: mp 58–60 °C, R<sub>f</sub> = 0.52 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 8/1, v/v); FTIR (KBr) ν<sub>max</sub> 1708 (-CO), 3418–3480 (br) (-OH) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 9.01 (1H, s, -NH), 7.16 (1H, d, J = 1.3 Hz, H-6), 6.80 (1H, m, H-1'), 4.21 (1H, dd, J = 12.0 and 4.7 Hz, H-5'a), 4.15 (1H, dd, J = 12.1 and 4.5 Hz, H-5'b), 4.08 (1H, m, H-3'), 4.05 (1H, ddd, J = 3.6, 4.6 and 4.2 Hz, H-4'), 3.37 (1H, br s, 3'-OH), 2.35 (1H, ddd, J = 13.7, 6.6 and 4.4 Hz, H-2'a), 2.25 (1H, ddd, J = 13.6, 6.6 and 6.8 Hz, H-2'b), 2.11 (3H, s, CH<sub>3</sub>CO-), 1.70 (3H, d, J = 1.3 Hz, 5-CH<sub>3</sub>). LC-MS [M+1]<sup>+</sup> 285.29. Analysis calcd for C<sub>11</sub>H<sub>16</sub>O<sub>5</sub>N<sub>2</sub>CO: C, 50.65, H, 5.63%, found: C, 50.67, H, 5.66%.

### General procedure for the direct 3'-O-acylation of 5'-O-(acetyl)thymidine derivatives (8–11)

A solution of the acetyl derivative (7) (60 mg, 0.21 mmol) in dry pyridine (3 ml) was treated with myristoyl chloride (0.110 gm, 2.2 molar eq.). The mixture was stirred at 0°C for 6–7 h and further kept overnight at room temperature. Excess reagent was decomposed by the addition of a few pieces of ice, and the contents were extracted with CHCl<sub>3</sub> (3×10 ml). The combined organic extract was washed successively with dilute HCl, saturated aqueous NaHCO<sub>3</sub>, and distilled H<sub>2</sub>O. The CHCl<sub>3</sub> layer was dried (Na<sub>2</sub>SO<sub>4</sub>), filtered, and evaporated to dryness. The residue was purified by passing it through a column with CHCl<sub>3</sub>/CH<sub>3</sub>OH as eluent, and the process yielded the 3'-O-myristoyl derivative (8, 101 mg) as a crystalline solid.

### 5'-O-Acetyl-3'-O-myristoylthymidine (8)

White crystalline solid in 97.11% yield: mp 69–71 °C, R<sub>f</sub> = 0.55 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 7/1, v/v); FTIR (KBr) ν<sub>max</sub> 1701 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.91 (1H, s, -NH), 7.31 (1H, d, J = 1.3 Hz, H-6), 6.28 (1H, t, J = 6.5 Hz, H-1'), 4.48 (1H, m, H-3'), 4.39 (1H, dd, J = 12.0 and 4.5 Hz, H-5'a), 4.20 (1H, dd, J = 12.0 and 3.5 Hz, H-5'b), 4.05 (1H, ddd, J = 3.5, 4.5 and 3.9 Hz, H-4'), 2.36 (1H, ddd, J = 13.5, 6.5 and 4.0 Hz, H-2'a), 2.35 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>11</sub>CH<sub>2</sub>CO-}, 2.24 (1H, ddd, J = 13.5, 6.5 and 6.7 Hz, H-2'b), 2.10 (3H, s, CH<sub>3</sub>CO-), 1.90 (3H, d, J = 1.3 Hz, 5-CH<sub>3</sub>), 1.68 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>10</sub>CH<sub>2</sub>CH<sub>2</sub>CO-}, 1.42 {20H, br m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>10</sub>CH<sub>2</sub>CH<sub>2</sub>CO-}, 0.89 {3H, t, J = 6.8 Hz, CH<sub>3</sub>(CH<sub>2</sub>)<sub>12</sub>CO-}. LC-MS [M+1]<sup>+</sup> 495.64. Analysis calcd for C<sub>25</sub>H<sub>42</sub>O<sub>6</sub>N<sub>2</sub>CO: C, 63.07, H, 8.49%, found: C, 63.10, H, 8.52%.

Similar reaction and purification methods were employed to synthesize a compound (9) (164 mg), compound (10) (152 mg), and compound (11) (99.5 mg).

### 5'-O-Acetyl-3'-O-palmitoylthymidine (9)

White crystalline solid in 77.49% yield: mp 88–90 °C, R<sub>f</sub> = 0.51 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 8/1, v/v); FTIR (KBr) ν<sub>max</sub> 1710 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 9.29 (1H, s, -NH), 7.31 (1H, d, J = 1.3 Hz, H-6), 6.31 (1H, t, J = 6.5 Hz, H-1'), 5.24 (1H, m, H-3'), 4.42 (1H, dd, J = 12.2 and 4.6 Hz, H-5'a), 4.38 (1H, dd, J = 12.1 and 3.6 Hz, H-5'b), 4.15 (1H, ddd, J = 3.5, 4.5 and 3.9 Hz, H-4'), 2.48 (1H, ddd, J = 13.5, 6.5 and 4.0 Hz,

H-2'a), 2.36 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>11</sub>CH<sub>2</sub>CO-}, 2.32 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>13</sub>CH<sub>2</sub>CO-}, 2.26 (1H, ddd, *J* = 13.5, 6.5 and 6.7 Hz, H-2'b), 2.12 (3H, s, CH<sub>3</sub>CO-), 1.96 (3H, d, *J* = 1.3 Hz, 5-CH<sub>3</sub>), 1.72 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>12</sub>CH<sub>2</sub>CH<sub>2</sub>CO-}, 1.28 {26H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>13</sub>CH<sub>2</sub>CO-}, 0.91 {3H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>14</sub>CO-}. LC-MS [M+1]<sup>+</sup> 523.32. Analysis calcd for C<sub>27</sub>H<sub>46</sub>O<sub>6</sub>N<sub>2</sub>CO: C, 64.33, H, 8.81%, found: C, 64.36, H, 8.84%.

### 5'-O-Acetyl-3'-O-stearoylthymidine (10)

White crystalline solid in 93.11% yield: mp 77–79 °C, *R<sub>f</sub>* = 0.54 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 8/1, v/v); FTIR (KBr)  $\nu_{\max}$  1711 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  9.00 (1H, s, -NH), 7.31 (1H, d, *J* = 2.2 Hz, H-6), 6.21 (1H, t, *J* = 6.5 Hz, H-1'), 4.68 (1H, m, H-3'), 4.50 (1H, dd, *J* = 12.0 and 4.5 Hz, H-5'a), 4.37 (1H, m, H-5'b), 4.24 (1H, m, H-4'), 2.38 (1H, ddd, *J* = 13.5, 6.5 and 4.0 Hz, H-2'a), 2.32 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>15</sub>CH<sub>2</sub>CO-}, 2.22 (1H, ddd, *J* = 13.5, 6.5 and 6.7 Hz, H-2'b), 2.11 (3H, s, CH<sub>3</sub>CO-), 1.92 (3H, d, *J* = 1.3 Hz, 5-CH<sub>3</sub>), 1.26 {30H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>15</sub>CH<sub>2</sub>CO-}, 1.23 {9H, s, (CH<sub>3</sub>)<sub>3</sub>CCO-}, 0.91 {3H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>16</sub>CO-}. LC-MS [M+1]<sup>+</sup> 551.82. Analysis calcd for C<sub>29</sub>H<sub>50</sub>O<sub>6</sub>N<sub>2</sub>CO: C, 65.36, H, 9.08%, found: C, 65.39, H, 9.10%.

### 5'-O-Acetyl-(3'-bromobenzoyl)thymidine (11)

White crystalline solid in 89.75% yield: mp 69–71 °C, *R<sub>f</sub>* = 0.52 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 16/1, v/v); FTIR (KBr)  $\nu_{\max}$  1706 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  9.02 (1H, s, -NH), 8.22 (1H, d, *J* = 7.7 Hz, Ar-H), 8.12 (1H, s, Ar-H), 8.01 (1H, d, *J* = 7.6 Hz, Ar-H), 7.31 (1H, t, *J* = 7.6 Hz, Ar-H), 6.66 (1H, d, *J* = 1.7 Hz, H-6), 6.62 (1H, t, *J* = 6.6 Hz, H-1'), 5.36 (1H, m, H-3'), 4.84 (1H, dd, *J* = 12.1 and 4.7 Hz, H-5'a), 4.43 (1H, dd, *J* = 11.9 and 3.8 Hz, H-5'b), 4.16 (1H, ddd, *J* = 3.7, 4.8 and 4.2 Hz, H-4'), 2.40 (1H, m, H-2'a), 2.35 (1H, m, H-2'b), 2.08 (3H, s, CH<sub>3</sub>CO-), 1.55 (3H, d, *J* = 1.6 Hz, 5-CH<sub>3</sub>). LC-MS [M+1]<sup>+</sup> 468.28. Analysis calcd for BrC<sub>18</sub>H<sub>19</sub>O<sub>6</sub>N<sub>2</sub>CO: C, 48.79, H, 4.07%, found: C, 48.80, H, 4.11%.

### Procedure for the synthesis of 5'-O-(4-*t*-butylbenzoyl)thymidine (12)

A solution of thymidine (1) (200 mg, 0.82 mmol) was dissolved in dry pyridine (3 ml) and cooled to 0°C, after which 4-*t*-butylbenzoyl chloride (173 mg, 1.1 molar eq.) was added. The mixture was stirred at 0°C for 6–8 h and allowed to stand at room temperature overnight. The conventional work-up procedure followed by chromatographic purification with CHCl<sub>3</sub>/CH<sub>3</sub>OH (4:1) as eluent yielded the title derivative (12, 90.4 mg) as a crystalline solid. White crystalline solid in 83.21% yield: mp 65–67 °C, *R<sub>f</sub>* = 0.52 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 9/1, v/v); FTIR (KBr)  $\nu_{\max}$  1701 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  9.00 (1H, s, -NH), 7.96 (2H, m, Ar-H), 7.34 (2H, m, Ar-H), 7.01 (1H, d, *J* = 1.8 Hz, H-6), 6.32 (1H, t, *J* = 6.4 Hz, H-1'), 4.68 (1H, dd, *J* = 12.2 and 4.6 Hz, H-5'a), 4.62 (1H, dd, *J* = 12.1 and 3.7 Hz, H-5'b), 4.10 (1H, m, H-3'), 4.06 (1H, ddd, *J* = 3.8, 4.8 and 3.8 Hz, H-4'), 3.13 (1H, br s, 3'-OH), 2.46 (1H, ddd, *J* = 13.6, 6.5 and 4.5 Hz, H-2'a), 2.19 (1H, ddd, *J* = 13.6, 6.7 and 6.9 Hz, H-2'b), 1.75 (3H, d, *J* = 1.6 Hz, 5-CH<sub>3</sub>), 1.26, 1.24, 1.21 {9H, 3s, (CH<sub>3</sub>)<sub>3</sub>C-}. LC-MS [M+1]<sup>+</sup> 403.75. Analysis calcd for C<sub>20</sub>H<sub>26</sub>O<sub>5</sub>N<sub>2</sub>CO: C, 62.57, H, 7.19%, found: C, 62.59, H, 7.22%.

### General procedure for the direct 3'-O-acylation of 5'-O-(4-*t*-butylbenzoyl)thymidine derivatives (13-16)

A solution of 5'-O-(4-*t*-butylbenzoyl)thymidine (12) (84 mg, 0.21 mmol) in dry pyridine (3 ml) was cooled to 0°C, after which acetic anhydride (0.05 ml, 0.5 mmol) was added. The mixture was stirred at 0°C for 6–8 h and further allowed to stand at room temperature overnight. Work-up as described earlier and chromatographic purification with CHCl<sub>3</sub>/CH<sub>3</sub>OH as eluant yielded the 4-*t*-butylbenzoyl derivative (13) (90 mg) as a crystalline solid.

### 5'-O-(4-*t*-Butylbenzoyl)-3'-O-hexanoylthymidine (13)

White crystalline solid in 88.46% yield: mp 63–65 °C, *R<sub>f</sub>* = 0.51 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 8/1, v/v); FTIR (KBr)  $\nu_{\max}$  1700 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  9.06 (1H, s, -NH), 8.05 (2H, m, Ar-H), 7.51 (2H, m, Ar-H), 7.28 (1H, d, *J* = 1.3 Hz, H-6), 6.37 (1H, t, *J* = 6.5 Hz, H-1'), 5.25 (1H, m, H-3'), 4.68 (1H, dd, *J* = 12.2 and 4.6 Hz, H-5'a), 4.54 (1H, dd, *J* = 12.1 and 3.6 Hz, H-5'b), 4.44 (1H, ddd, *J* = 3.5, 4.5 and 3.9 Hz, H-4'), 2.36 (1H, ddd, *J* = 13.5, 6.5 and 4.0 Hz, H-2'a), 2.33 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>3</sub>CH<sub>2</sub>CO-}, 2.28 (1H, ddd, *J* = 13.5, 6.5 and 6.7 Hz, H-2'b), 1.69 (3H, d, *J* = 1.3 Hz, 5-CH<sub>3</sub>), 1.64 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO-}, 1.28 {4H, m,

CH<sub>3</sub>(CH<sub>2</sub>)<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO-}, 0.90 {3H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>4</sub>CO-}, 1.21, 1.19, 1.16 {9H, 3s, (CH<sub>3</sub>)<sub>3</sub>C-}. LC-MS [M+1]<sup>+</sup> 501.85. Analysis calcd for C<sub>26</sub>H<sub>36</sub>O<sub>6</sub>N<sub>2</sub>CO: C, 64.69, H 7.19%, found: C, 64.72, H, 7.21%.

A similar procedure was used to yield compound (14, 120 mg.), compound (15, 123.9 mg), and compound (16, 84.54 mg, 79.2%) as needles (mp 61–63°C).

#### 5'-O-(4-*t*-Butylbenzoyl)-3'-O-lauroylthymidine (14)

White crystalline solid in 98.36% yield: mp 57–59 °C, R<sub>f</sub> = 0.54 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 8/1, v/v); FTIR (KBr) ν<sub>max</sub> 1707 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 9.0 (1H, s, -NH), 8.01 (2H, m, Ar-H), 7.48 (2H, m, Ar-H), 7.28 (1H, d, *J* = 2.3 Hz, H-6), 6.21 (1H, t, *J* = 6.4 Hz, H-1'), 4.78 (1H, m, H-3'), 4.75 (1H, dd, *J* = 11.8 and 4.6 Hz, H-5'a), 4.70 (1H, dd, *J* = 12.1 and 4.5 Hz, H-5'b), 4.51 (1H, ddd, *J* = 4.5, 5.5 and 4.9 Hz, H-4'), 2.38 (1H, ddd, *J* = 12.9, 6.7 and 4.4 Hz, H-2'a), 2.34 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>9</sub>CH<sub>2</sub>CO-}, 2.34 (1H, ddd, *J* = 13.6, 6.2 and 6.8 Hz, H-2'b), 1.95 (3H, d, *J* = 1.3 Hz, 5-CH<sub>3</sub>), 1.64 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>8</sub>CH<sub>2</sub>CH<sub>2</sub>CO-}, 1.26 {16H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>8</sub>CH<sub>2</sub>CH<sub>2</sub>CO-}, 1.24, 1.21, 1.14 {9H, 3s, (CH<sub>3</sub>)<sub>3</sub>C-}, 0.92 {3H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>10</sub>CO-}. LC-MS [M+1]<sup>+</sup> 586.01. Analysis calcd for C<sub>32</sub>H<sub>48</sub>O<sub>6</sub>N<sub>2</sub>CO: C, 67.69, H, 8.20%, found: C, 67.73, H, 8.22%.

#### 5'-O-(4-*t*-Butylbenzoyl)-3'-O-myristoylthymidine (15)

White crystalline solid in 97.3% yield: mp 74–76 °C, R<sub>f</sub> = 0.52 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 8/1, v/v); FTIR (KBr) ν<sub>max</sub> 1702 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.92 (1H, s, -NH), 8.0 (2H, m, Ar-H), 7.45 (2H, m, Ar-H), 7.27 (1H, d, *J* = 1.3 Hz, H-6), 6.24 (1H, t, *J* = 6.5 Hz, H-1'), 4.46 (1H, m, H-3'), 4.40 (1H, dd, *J* = 12.0 and 4.5 Hz, H-5'a), 4.26 (1H, dd, *J* = 12.0 and 3.5 Hz, H-5'b), 4.15 (1H, ddd, *J* = 3.5, 4.5 and 3.9 Hz, H-4'), 2.38 (1H, ddd, *J* = 13.5, 6.5 and 4.0 Hz, H-2'a), 2.34 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>11</sub>CH<sub>2</sub>CO-}, 2.25 (1H, ddd, *J* = 13.5, 6.5 and 6.7 Hz, H-2'b), 1.92 (3H, d, *J* = 1.3 Hz, 5-CH<sub>3</sub>), 1.67 {2H, m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>10</sub>CH<sub>2</sub>CH<sub>2</sub>CO-}, 1.26 {20H, br m, CH<sub>3</sub>(CH<sub>2</sub>)<sub>10</sub>CH<sub>2</sub>CH<sub>2</sub>CO-}, 1.25, 1.22, 1.15 {9H, 3s, (CH<sub>3</sub>)<sub>3</sub>C-}, 0.88 {3H, t, *J* = 6.8 Hz, CH<sub>3</sub>(CH<sub>2</sub>)<sub>12</sub>CO-}. LC-MS [M+1]<sup>+</sup> 614.07. Analysis calcd for C<sub>34</sub>H<sub>52</sub>O<sub>6</sub>N<sub>2</sub>CO: C, 68.51, H, 8.48%, found: C, 68.53, H, 8.51%.

#### 5'-O-(4-*t*-Butylbenzoyl)-3'-O-(4-nitrobenzoyl)thymidine (16)

White crystalline solid in 79.2% yield: mp 61–63 °C, R<sub>f</sub> = 0.50 (CHCl<sub>3</sub>/CH<sub>3</sub>OH = 8/1, v/v); FTIR (KBr) ν<sub>max</sub> 1700 (-CO) cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 9.09 (1H, s, -NH), 8.06, 7.52 (2 × 2H, 2 × d, *J* = 8.7 Hz, Ar-H), 8.0 (2H, m, Ar-H), 7.31 (2H, m, Ar-H), 7.27 (1H, d, *J* = 1.3 Hz, H-6), 6.24 (1H, t, *J* = 6.5 Hz, H-1'), 4.46 (1H, m, H-3'), 4.40 (1H, dd, *J* = 12.0 and 4.5 Hz, H-5'a), 4.26 (1H, dd, *J* = 12.0 and 3.5 Hz, H-5'b), 4.15 (1H, ddd, *J* = 3.5, 4.5 and 3.9 Hz, H-4'), 2.38 (1H, ddd, *J* = 13.5, 6.5 and 4.0 Hz, H-2'a), 2.25 (1H, ddd, *J* = 13.5, 6.5 and 6.7 Hz, H-2'b), 1.72 (3H, d, *J* = 1.3 Hz, 5-CH<sub>3</sub>), 1.36, 1.28, 1.21 {9H, 3s, (CH<sub>3</sub>)<sub>3</sub>C-}. LC-MS [M+1]<sup>+</sup> 552.81. Analysis calcd for C<sub>27</sub>H<sub>29</sub>O<sub>8</sub>N<sub>3</sub>CO: C, 60.89, H, 5.25%, found: C, 60.91, H, 5.28%.

### X-ray powder diffraction

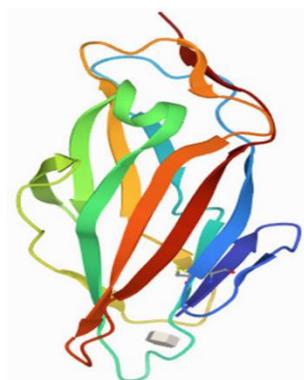
XRD is a rapid analytical technique primarily used for phase identification of crystalline material. It can provide information on unit cell dimensions [28]. The synthesized tested compounds to be analyzed were finely ground, homogenized, and average bulk composition was determined.

### Computational details

Geometrical optimization of the selected thymidine derivatives and calculation of vibrational frequency were performed using Gaussian 09 software package [29]. Density functional theory (DFT) with Beck's (B) [30] three-parameter hybrid model, Lee, Yang, and Parr's (LYP) [31] correlation functional under 3-21G basis set has been used to optimize the ligands.

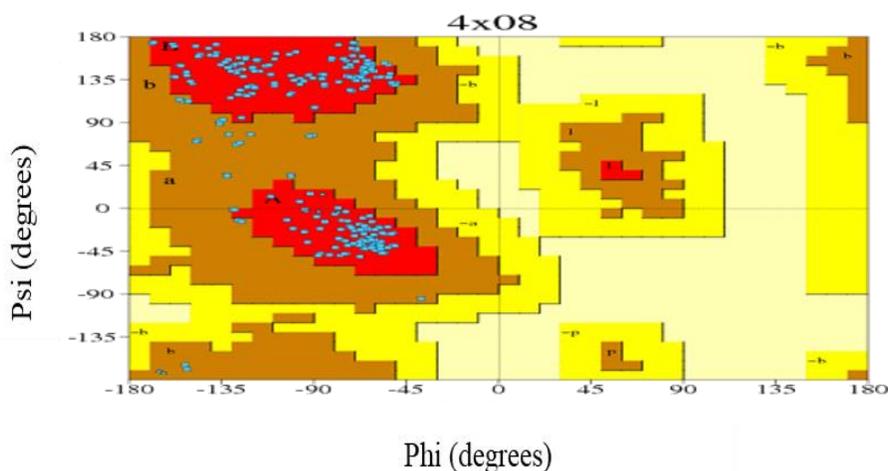
The three-dimensional crystal structure of a bacterial protein FimH of *E. coli* (PDB ID: 4XO8) (Fig. 2) was retrieved in pdb format from the protein data bank [32]. All hetero atoms and water molecules were removed using PyMol (version 1.3) software packages [33]. Swiss-Pdb viewer software (version 4.1.0) was employed for energy minimization of the protein [34]. Then optimized thymidine ligands were subjected for molecular docking study against *E. coli* (4XO8). In fine, molecular docking simulation was rendered by PyRx software (version 0.8) [35] considering the protein as a macromolecule and the drug as a ligand. AutodockVina was employed for docking analysis, and AutoDock Tools (ADT) of the MGL software package was used to convert pdb into a pdbqt format to input protein and ligands. The size of the grid box in AutoDockVina was

kept at 33.3452, 30.6390, and 60.5649 Å for X, Y, Z directions respectively. After the completion docking, both the macromolecule and ligand structures were saved in. pdbqt format needed by Accelrys Discovery Studio (version 4.1) to explore and visualize the docking result and search the nonbonding interactions between ligands and amino acid residues of receptor protein [36]. Lectin FimH is an eosinophil secretion protein with antipathogen activities involved in the host immune defense system. The bactericidal capacity of the receptor relies on its action on both the plasma membrane and the bacterial wall. It interacts with bacterial lipopolysaccharide (LPS) and lipoteichoic acid (LTA). *In vitro* interacts with and insert into lipid bilayers composed of dioleoyl phosphatidylcholine and dioleoyl phosphatidylglycerol. *In vitro*, tends to form amyloid-like aggregates at pH 3, but not at pH 5, nor 7. Since FimH of *E. coli* 4XO8 has its crystal structure in a state that represents the pharmacological target for the development of new drugs, it is selected for computational studies.



**Fig. 2.** Crystal structure of FimH *E. coli* (4XO8).

Protein validation was performed by the PROCHECK server (<https://servicesn.mbi.ucla.edu/PROCHECK/>) and fork out 96.9697 overall quality factors in ERRAT ([http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list\\_uids=8401235&dopt=Abstract](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=8401235&dopt=Abstract)), 80% score in VERIFY 3D (<https://www.ncbi.nlm.nih.gov/pubmed/1853201?dopt=Abstract>) PDBsum (<http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html>) server also used to test the validation of receptor with Ramachandran plot (Fig. 3) which exhibited 88.8% residues in the allowed region.



**Fig. 3.** Ramachandran plot of 4XO8.

### Docking validation protocol

The validation of docking results was done by extracting the co-crystallized ligand of the FimH protein (PDB ID: 4X08) and re-docking it into the same position. The lowest energy pose gained on re-docking and the co-crystallized ligands were superimposed, and its root means square deviation (RMSD) was calculated between these two superimposed ligands. To validate the docking process, The RMSD must be within a reliable range of 2 Å [37, 38].

### In silico pharmacokinetics ADMET prediction

To search potential drug candidates, the ADMET properties were developed for the preliminary prediction of the pharmacokinetic and physicochemical parameters in the discovery drug process. *In silico* study gives a manner the accession of pharmacokinetic parameters (adsorption, distribution, metabolism, excretion, and toxicity; ADMET) [39], its absorption in the human intestine, percolation of the blood-brain barrier and the central nervous system, the metabolism indicates the chemical biotransformation of a drug by the body, total clearance of drugs and the toxicity levels of the molecules.

## Results and discussion

### Chemistry

The main objective of this study was to perform selective pivaloylation, acetylation, and 4-*t*-butylbenzoylation (Scheme 1) of thymidine (**1**) using the direct acylation method. A series of derivatives of the resulting products were synthesized using several acylating agents. Selective acylation of a particular hydroxyl group is vital in nucleoside chemistry because the resulting acylation products might provide useful precursors for the synthesis of newer and biologically effective products. Additionally, the resulting acyl derivatives might have a great utility as versatile intermediates for the synthesis of various other nucleosides of fundamental biological importance.

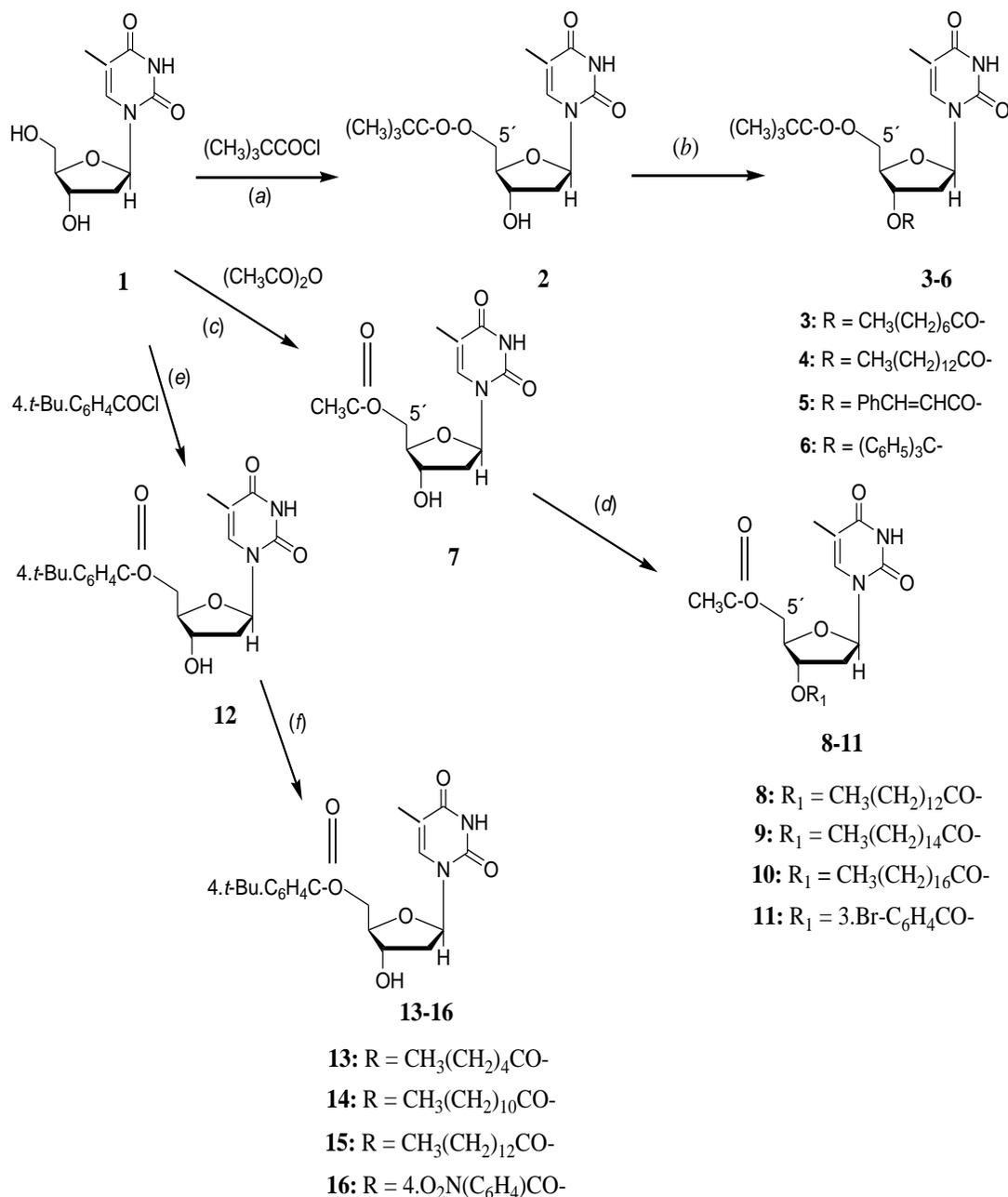
### Characterization and selective pivalylation of thymidine

Thymidine (**1**) was initially treated with unimolecular amount of pivaloyl chloride in dry pyridine at low temperature, followed by purification by silica gel chromatography, which yielded the pivaloyl derivative (**2**) as a crystalline solid. The infrared spectrum showed absorption bands at 1690 and 3410–3535 (broad)  $\text{cm}^{-1}$  corresponding to carbonyl and hydroxyl stretchings, respectively, and thereby indicating the introduction of an acyl group. The  $^1\text{H-NMR}$  spectrum showed a nine-proton singlet at  $\delta$  1.38 [ $(\text{CH}_3)_3\text{CCO-}$ ] was because of the presence of methyl protons of the pivaloyl group. Additionally, the  $^1\text{H-NMR}$  spectrum demonstrated a downfield shift of the H-5' proton to  $\delta$  4.41 (dd,  $J = 12.0$  and  $4.5$  Hz, H-5'a), and  $\delta$  4.36 (dd,  $J = 12.1$  and  $4.5$  Hz, H-5'b) from their precursor value [40], indicating the attachment of the pivaloyl group at position 5'. The mass spectrum of compound **2** contained a molecular ion peak at  $m/z$   $[\text{M}+1]^+$  327.31 that corresponded to the same molecular formula. The rest of the  $^1\text{H-NMR}$  spectrum was in accordance with the structure. The formation of compound (**2**) may be explained by assuming that pivaloyl chloride attaches with the most reactive and less sterically hindered primary -OH group of the ribfuranose moiety at 5' position, thereby forming the 5'-*O*-(pivaloyl)thymidine (**2**) as the sole product.

The pivaloyl derivative (**2**) was transformed into its octanoyl derivative (**3**) (80.91%) as needles (mp 59–61 °C) with octanoyl chloride in dry  $\text{C}_6\text{H}_5\text{N}$ . The  $^1\text{H-NMR}$  spectrum of the octanoate was in accordance with the structure assigned to it. The two two-proton multiplets at  $\delta$  2.32 [ $\text{CH}_3(\text{CH}_2)_5\text{CH}_2\text{CO-}$ ] and 1.68 [ $\text{CH}_3(\text{CH}_2)_4\text{CH}_2\text{CH}_2\text{CO-}$ ], an eight-proton multiplet at  $\delta$  1.29 [ $\text{CH}_3(\text{CH}_2)_4(\text{CH}_2)_2\text{CO-}$ ], and a three-proton multiplet at  $\delta$  0.88 [ $\text{CH}_3(\text{CH}_2)_6\text{CO-}$ ] were indicative of the presence of one octanoyl group. Compound (**2**) was also converted into the myristoate (**4**) with a yield of 95.2% as white needles (mp 78–80 °C). Mass spectrometry provided a molecular ion peak at  $m/z$   $[\text{M}+1]^+$  537.21, which corresponded to the aforementioned molecular formula. Complete analysis of the FTIR and  $^1\text{H-NMR}$  spectra of this compound conformed to the structure.

The formation of the octanoate (**3**) and the myristoate (**4**) further confirmed the structure of the compound (**2**). The structure of the pivaloyate (**2**) was finally confirmed by its conversion into the cinnamoyl

(5) and trityl (6) derivatives. In the  $^1\text{H-NMR}$  spectrum of compound (5), two one-proton doublets at  $\delta$  7.73 (as d,  $J = 12.0$  Hz,  $\text{PhCH=CHCO-}$ ) and  $\delta$  6.49 (as d,  $J = 12.1$  Hz,  $\text{PhCH=CHCO-}$ ) indicated the presence of cinnamoyl group in the molecule. The structure of the tritylate was assigned as 5'-*O*-pivaloyl-3'-*O*-tritylthymidine (6) by analyzing its  $^1\text{H-NMR}$  and FTIR spectra.



**Scheme 1.** Reagents and conditions: (a) dry  $\text{C}_6\text{H}_5\text{N}$ ,  $-5^\circ\text{C}$ , 5-6 h; (b) dry pyridine, several acyl halides (**3–6**); (c) dry  $\text{C}_6\text{H}_5\text{N}$ ,  $-5^\circ\text{C}$ , 6-7 h; (d) dry pyridine, acyl halides (**8–11**), (e) dry  $\text{C}_6\text{H}_5\text{N}$ ,  $-5-7^\circ\text{C}$ , 6-8 h; (f) dry pyridine, acyl halides (**13–16**),  $0^\circ\text{C}$  to rt, DMAP, stirrer for 6–8 h.

### Characterization and selective acetylation of thymidine

Thymidine (**1**) was further subjected to acetylation using the same method to observe the selectivity of the acetyl group. Thus, acetylation of compound (**1**) yielded acetyl derivative (**7**) (90.96%) as a white crystalline solid (mp 58–60°C). The <sup>1</sup>H-NMR spectrum, one three-proton singlet at δ 2.11 was because of the methyl protons of one acetyloxy group. The downfield shifts of H-5' to δ 4.21 (as dd, *J* = 12.0 and 4.7 Hz, H-5'a) and 4.15 (as, dd, *J* = 12.1 and 4.5 Hz, H-5'b) as compared with the reference value [41] indicated the attachment of the acetyl group at position 5'. Complete analysis of the rest of the spectra helped to deduce the structure of the acetyl derivative as 5'-*O*-(acetyl)thymidine (**7**). Its conversion to myristoyl derivative (**8**), palmitoate (**9**), stearoate (**10**), and 3-bromobenzoyl (**11**) and their identification further supported the structure of the acetyl derivative (**7**).

### Characterization and selective 4-*t*-butylbenzoylation of thymidine

Finally, the 4-*t*-butylbenzoylation of the thymidine (**1**) was performed, where it was allowed to react with an equimolar amount of 4-*t*-butylbenzoyl chloride in dry pyridine to yield the 4-*t*-butylbenzoylate (**12**) (83.21%) as a crystalline solid (mp 65–67°C). The FTIR spectrum demonstrated the absorption bands at 1701 cm<sup>-1</sup> because of -CO and 3429–3505 (br) cm<sup>-1</sup> for -OH stretching. In its <sup>1</sup>H-NMR spectrum, two two-proton multiplets at δ 7.96 (Ar-H) and δ 7.34 (Ar-H) and three singlets at δ 1.26, δ 1.24 and δ 1.21 [9H, 3s, (CH<sub>3</sub>)<sub>3</sub>C-] corresponded to the presence of one 4-*t*-butylbenzoyl group and downfield shift of C-5' to δ 4.68 (as dd, *J* = 12.2 and 4.6 Hz, H-5'a) and 4.62 (as dd, *J* = 12.1 and 3.7 Hz, H-5'b) from its usual δ value (approximately 4.0) [42]. The rest of the <sup>1</sup>H-NMR spectrum was in accordance with the structure assigned to the 4-*t*-butylbenzoylate as 5'-*O*-(4-*t*-butylbenzoyl)thymidine (**12**).

The structure of the 4-*t*-butylbenzoylate was further confirmed by synthesizing its fatty acids derivatives: hexanoyl, lauroyl, and myristoyl. Thus, the reaction of the 4-*t*-butylbenzoylate derivative (**12**) with hexanoyl chloride in pyridine, followed by conventional work-up and column chromatography yielded the hexanoylate (**13**) (88.46%) as a crystalline solid (mp 63–65°C). The <sup>1</sup>H-NMR spectrum revealed two two-proton multiplets at δ 2.33 [CH<sub>3</sub>(CH<sub>2</sub>)<sub>3</sub>CH<sub>2</sub>CO-] and δ 1.64 [CH<sub>3</sub>(CH<sub>2</sub>)<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO-], one four-proton multiplet at δ 1.28 [CH<sub>3</sub>(CH<sub>2</sub>)<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO-], and a three-proton multiplet at δ 0.90 [CH<sub>3</sub>(CH<sub>2</sub>)<sub>4</sub>CO-] indicating the presence of one hexanoyl group. Analysis of the FTIR, <sup>1</sup>H-NMR, and elemental data of this compound enabled us to assign the structure of the hexanoylate derivative as 5'-*O*-(4-*t*-butylbenzoyl)-3'-*O*-hexanoylthymidine (**13**). Similarly, lauroylation and myristoylation of compound (**12**) with the corresponding halides yielded the lauroylate (**14**) and myristoyl (**15**) in good yields. The structure of these derivatives (**14** and **15**) was confirmed by a complete analysis of their infrared and NMR spectra. The structure accorded to the 4-*t*-butylbenzoylate derivative (**12**) was finally confirmed by synthesis and identification of its 4-nitrobenzoyl derivative (**16**). In its <sup>1</sup>H-NMR spectrum, two low field two-proton doublets at δ 8.06 and δ 7.52 (2 × 2H, 2 × d, *J* = 8.7 Hz, Ar-H) corresponded to one 4-nitrobenzoyl group. The structure of this 4-nitrobenzoylate was assigned as 5'-*O*-(4-*t*-butylbenzoyl)-3'-*O*-(4-nitrobenzoyl)thymidine (**16**). Thus, selective pivaloylation, acetylation, and 4-*t*-butylbenzoylation of thymidine (**1**) using the direct acylation method were unique because all the reactions provided single-crystalline derivatives in good yields.

### X-ray powder diffraction

The XRD method is principally used for the quantitative identification of crystalline compounds, whereas the single-crystal X-ray diffraction method is principally used for structure determination. If *h*, *k*, and *l* represent the miller indices, the rules of the determination of crystal lattice type are as follows (Table 1).

**Table 1.** Rules for the determination of crystal lattice type.

Lattice type	Rules for reflection to be observed
Primitive, P	None
Body centered, I	<i>hkl</i> ; <i>h+k+l</i> = 2 <i>n</i>
Face centered, F	<i>hkl</i> ; <i>h,k,l</i> either all odd or all even
Side centered, C	<i>hkl</i> ; <i>h+k</i> = 2 <i>n</i>
Rhombohedral	<i>hkl</i> ; (- <i>h</i> )+ <i>k+l</i> = 3 <i>n</i> or ( <i>h-k</i> )+ <i>l</i> = 3 <i>n</i>

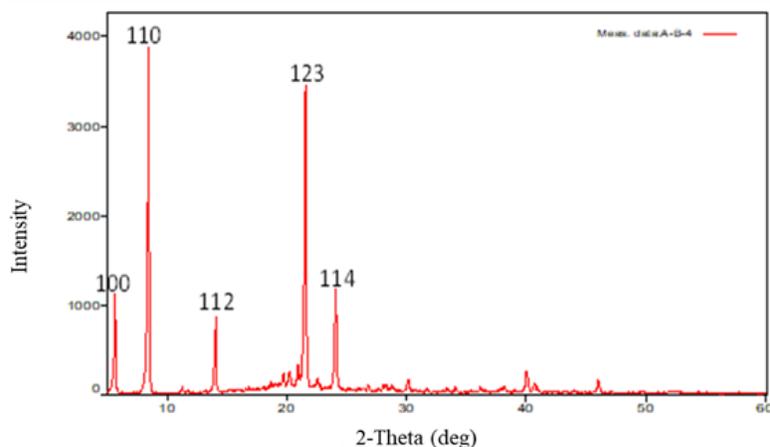
\**h*, *k*, and *l* are the Miller indices

The crystallographic structure of the compounds **4**, **8**, **9**, **10**, and **15** were evaluated by XRD at room temperature, and the pattern is presented in Table 2.

**Table 2.** The peak list of compounds **4**, **8**, **9**, **10**, and **15**.

Entry	Relative intensity	2 $\theta$ (deg.)	$\theta$ (deg.)	Sin $^2\theta$	Ratio	H $^2$ +k $^2$ +l $^2$	(h,k,l)	d (ang.)
<b>4</b>	20.87	5.618(4)	2.809	0.0024	1.00	1.00	100	15.718(11)
	100	8.37989	4.189	0.0053	2.00	2.00	110	10.5427
	18.86	14.035(3)	7.017	0.0149	6.00	6.00	112	6.3048(16)
	82.13	21.5457	10.772	0.0349	14.00	14.00	123	3.6951(8)
	31.22	24.064(6)	12.039	0.0434	18.00	18.00	114	3.6951(8)
<b>8</b>	24.26	5.562(3)	2.781	0.0023	1.00	1.00	100	15.875(7)
	100	8.364(6)	4.182	0.0053	2.00	2.00	110	10.563(8)
	21.33	13.953(3)	6.9765	0.0147	6.00	6.00	112	6.342(3)
	45.14	21.468(8)	10.772	0.0346	14.00	14.00	123	4.1358(16)
	23.87	23.990(17)	12.039	0.0435	18.00	18.00	114	3.706(3)
<b>9</b>	64.18	7.372(4)	3.686	0.0041	1.00	1.00	100	11.982(7)
	45.36	9.864(12)	4.932	0.0074	2.00	2.00	110	8.960(11)
	100	10.087(3)	5.0435	0.0077	2.00	2.00	110	8.762(2)
	24.19	11.159(3)	5.5795	0.0094	2.00	2.00	110	7.922(2)
	30.67	21.491(6)	10.7455	0.0347	9.00	9.00	122	4.1313(12)
<b>10</b>	100	6.564(5)	3.282	0.0033	1.00	1.00	100	13.455(9)
	21.47	10.993(15)	5.4965	0.0092	3.00	3.00	111	8.042(11)
	75.54	21.408(4)	10.704	0.0344	10.00	10.00	310	4.1473(8)
	25.50	23.993(3)	11.9965	0.0432	13.00	13.00	320	3.707(4)
<b>15</b>	22.60	5.562(6)	2.781	0.0024	1.00	1.00	100	15.877(16)
	100	8.353(5)	4.1765	0.0053	2.00	2.00	110	10.576
	24.83	13.974(6)	6.987	0.0147	6.00	6.00	112	6.3048(16)
	9.51	18.954(4)	9.477	0.0271	11.00	11.00	113	4.6781(10)
	6.40	45.975(7)	22.9875	0.1525	64.00	64.00	800	1.9724(3)

The compounds **4** (Fig. 4), **8**, **9**, **10**, and **15** showed several lines with high intensity in their XRD pattern, which indicates that all the compounds are crystalline. The XRD pattern of the pure compound synthesized under optimum conditions was displayed in the 2 $\theta$  range of 0 $^\circ$ –50 $^\circ$ . The peaks at 2 $\theta$  corresponded to 8.379 and 21.545 (h, k, l: 110 and 123), 8.364 and 21.468 (h, k, l: 110 and 123), 7.372 and 21.491 (h, k, l: 100 and 122), and 6.564 and 21.408 (h, k, l: 100 and 310) for compounds **4**, **8**, **9**, and **10**, respectively. These peaks indicated the formation of typical phases of compounds **4**, **8**, **9**, and **10**. According to the phase analysis, compounds synthesized by this method have high purity, and no impurities were detected in the XRD pattern.



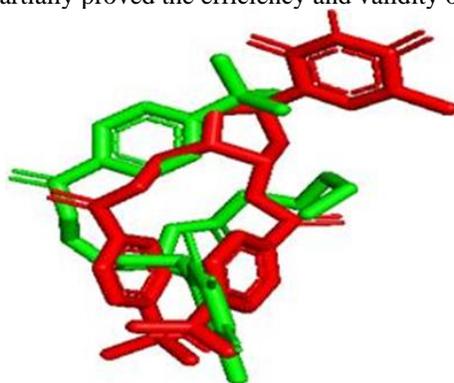
**Fig. 4.** X-ray diffraction peak of compound 4.

### Computational studies

The newly synthesized derivatives of thymidine were designated and used in this study were according to the reaction profile. Thymidine (**1**) and some of its derivatives were entitled and optimized in the quantum mechanical method.

### Docking validation study

To interpret the ability of docking algorithms to determine the conformation of the protein-bound ligand, re-docking of the co-crystallized ligand was employed to validate the accuracy of the docking procedure. Fig. 5 resolved the superimposed view between the docked ligand conformation and the co-crystallized ligand conformation and the RMSD value is 1.695 Å. The complex was then found to interact with the same amino acid residues compared to the ones reported in the present study. The bulky symmetric molecules can be exchanged in the binding site during docking, as the case in this investigation; the RMSD would be at a very high level. On the contrary, the small compounds can gain low RMSD easily even when placed randomly. Some reported studies [43-45] have suggested a new benchmark for the quality of docking poses based on visual inspection. For visual inspection, Fig. 6 shows the 2D visualization of the interactions between a generated docking pose and the experimental ligand conformation. The results of this visual inspection show the same interactions as in the experimental binding mode, as observed in Fig. 7. This result reveals that alone is not a reliable parameter for the quality of docking poses for docking validation and the use of visual inspection as a new reference is essential. This partially proved the efficiency and validity of the docking protocol.



**Fig. 5.** Superimpose view of re-docking, (blue) onto co-crystallized complex (red) in the active site using PyMOL (RMSD = 1.695 Å).

### Molecular docking studies

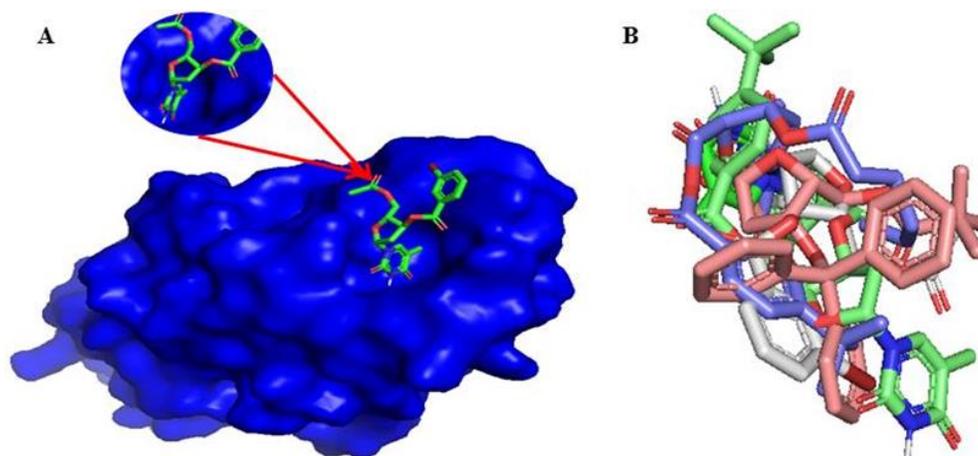
By using the Autodock Vina, molecular docking was performed to suggest the best candidates among the nine antibacterials (1, 3, 4, 5, 6, 11, 13, 14, and 16) based on their binding affinities. All selected candidates were docked into the same binding pocket of FimH of *E. coli* (4XO8) using similar optimized docking conditions to identify their binding mode. The results of the docking analysis showed that all ligands, along with the parent ligand, obtain binding affinities ranging from -4.8 to -6.8 kcal/mol. As shown in Table 3, derivatives (3, 4, 5, 13, and 14) showed comparatively lower binding affinities compared to the parent drug, thymidine (1), while others exhibited nearly to it. These results indicated that alteration of -OH group, along with an aromatic ring or a long carbon chain molecule increased the binding affinity, while the addition of hetero groups like Cl, Br, and NO<sub>2</sub> made some fickleness in binding affinities; however, modification with halogenated aromatic rings increased the binding affinity. The docked pose clearly showed that the drug molecules bind within the active site of the FimH of *E. coli* (4XO8) macromolecular structure (Fig. 6). Fig. 7 showed that derivatives (1, 6, 11, 13, and 16) (binding affinities -6.8, -6.1, -6.7, -5.4 and -6.5 kcal/mol) binds firmly through hydrophobic bonds with residues Tyr48, Ile52, and Ile13, where, these residues exhibited alkyl, pi-alkyl and pi-pi stacked interaction. Besides, compounds (6, 11, 13, and 16) showed a conventional hydrogen bond with residues Tyr137, Gln133, Asn135, Gln138, and Asp140. Again, electrostatic bonds (*Pi*-cation interaction) were observed for the compounds 13 and 16 with the residue phe1. Compounds 3 and 4 showed a similar binding site with residues Ala106 (shorter distance 3.85267 Å), but the compound 4 strongly interacted by a hydrogen bond with Gly79 in a closer distance of 1.96863 Å. It is clear from the structural contrast compounds (6, 11, and 16) have an additional aromatic (halogenated ring and nitrobenzene ring) substituent in the parent structure, indicating a high density of electron in the molecule leading to a comparatively higher binding affinity -6.1, -6.7, and -6.5 kcal/mol.

**Table 3.** Binding affinity (kcal/mol) and nonbonding interactions of thymidine and 5'-*O*- and 3'-*O*-acyl derivatives.

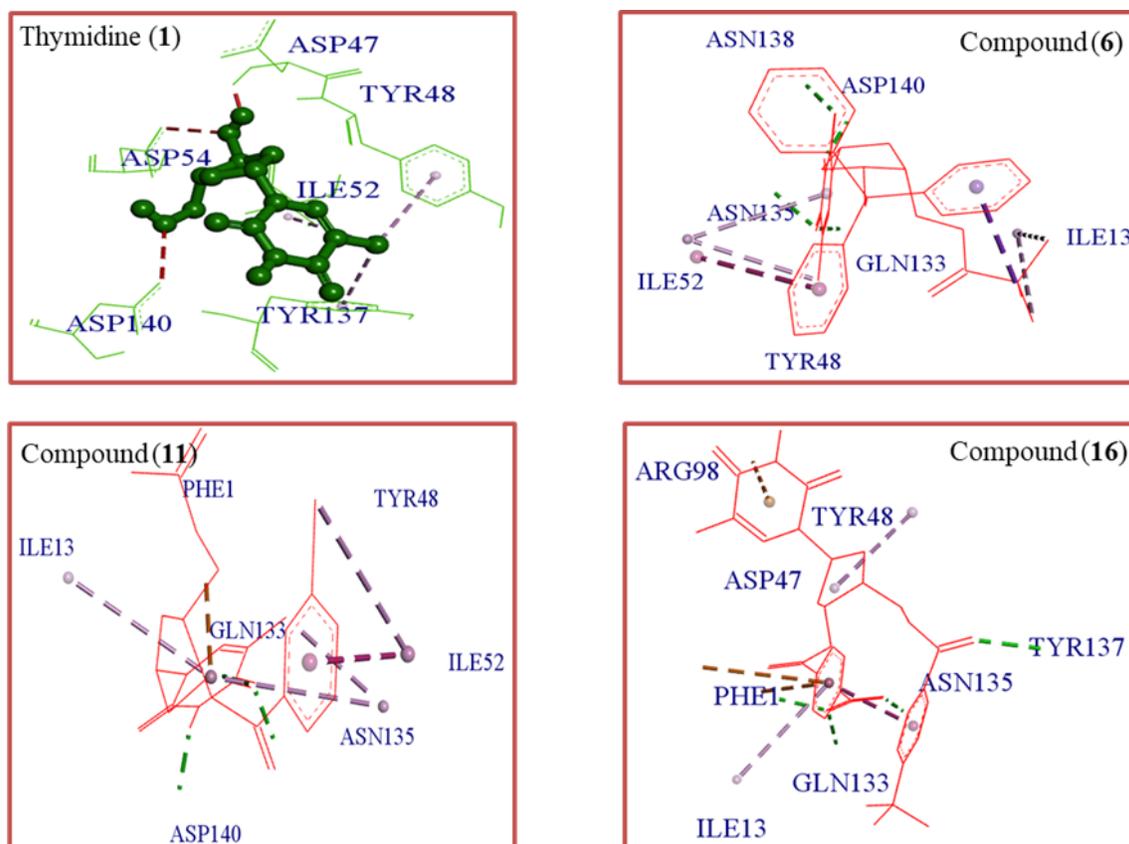
Entry	Protein	Binding affinity	Bond category	Residues in contact	Interaction types	Distance (Å)
1	4XO8	-6.8	Hydrophobic	ILE52	A	5.46042
			Hydrophobic	TYR48	PA	4.78341
			Hydrophobic	TYR137	PA	4.35259
3	4XO8	-5.4	Hydrophobic	PRO102	A	4.35065
			Hydrophobic	VAL35	A	5.05992
			Hydrophobic	ALA106	A	3.85267
			Hydrophobic	PRO102	PA	4.94622
			Hydrophobic	TYR108	PA	5.40235
4	4XO8	-4.8	Hydrogen	SER39	H	2.11714
			Hydrogen	SER39	H	2.43895
			Hydrogen	GLY79	H	1.96863
			Hydrophobic	PRO104	A	4.58817
			Hydrophobic	LYS76	PA	5.21545
			Hydrophobic	ALA106	PA	4.45017
5	4XO8	-5.9	Hydrogen	ASN23	H	2.05035
			Hydrogen	ALA25	H	2.4098
			Hydrogen	ASP37	H	2.23441
			Hydrogen	ASP37	C	3.44485
			Hydrophobic	VAL22	PA	5.15916
6	4XO8	-6.1	Hydrogen	ASP140	H	2.08949
			Hydrogen	GLN133	H	2.73867
			Hydrogen	ASN135	H	2.36646
			Hydrogen	ASN138	H	2.33626
			Hydrophobic	TYR48	PPS	4.18193

			Hydrophobic	ILE52	A	4.33738
			Hydrophobic	ILE52	PA	4.62052
11	4XO8	-6.7	Hydrogen	ASP140	H	2.0469
			Hydrogen	GLN133	H	2.58068
			Hydrogen	ASN135	H	2.04567
			Hydrogen	PHE1	PCa	3.26395
			Hydrophobic	TYR48	PPS	3.72993
			Hydrophobic	ILE52	A	3.9085
			Hydrophobic	ILE13	PA	5.2305
			Hydrophobic	ILE52	PA	5.49927
			Hydrophobic	TYR48	PA	4.71671
13	4XO8	-5.4	Hydrogen	GLN133	H	2.63951
			Hydrogen	TYR137	H	2.37081
			Electrostatic	PHE1	PCa	3.95948
			Hydrophobic	TYR48	PPS	3.93748
			Hydrophobic	ILE13	A	5.20327
			Hydrophobic	ILE13	A	4.51687
			Hydrophobic	ILE13	PA	5.14383
			Hydrophobic	TYR48	PA	4.30093
14	4XO8	-5.7	Hydrogen	ASN23	H	2.48688
			Hydrophobic	ALA25	PPS	3.92541
			Hydrophobic	VAL35	A	5.17729
			Hydrophobic	VAL22	A	3.89963
			Hydrophobic	VAL36	A	4.69495
16	4XO8	-6.5	Hydrophobic	VAL22	PA	5.4726
			Hydrogen	PHE1	H	2.31339
			Hydrogen	GLN133	H	2.46634
			Hydrogen	ASN135	H	2.13715
			Hydrogen	TYR137	H	2.67076
			Electrostatic	PHE1	PCa	3.88597
			Electrostatic	ARG98	PCa	4.17162
			Electrostatic	ASP47	PAn	4.91681
			Hydrophobic	ILE13	PA	4.88066
Hydrophobic	TYR48	PA	5.01157			

H = Conventional hydrogen bond; C = Carbon hydrogen bond; A= Alkyl; PA = Pi-alkyl; PAn = Pi-Anion; PCa = Pi-cation; PPS = Pi-Pi stacked.



**Fig. 6.** Docked conformation of compound (11) at inhibition bounding site of 4XO8 (A) and superimposed view of all compounds after rigid docking (B).



**Fig. 7.** Non-bonding interactions of compound (16, 11, and 16) with the amino acid residues of 4XO8 made by Discovery Studio.

On the other hand, compounds **5** and **14** exhibited similar three binding sites with Asn23, Ala25, and Val22 through both hydrogen and hydrophobic bonds. A new type of interaction i.e. pi-pi stacked observed in the case of compound (**14**) with Ala25. Again, these two derivatives maintained some uniformity with the ring-containing compounds (i.e. **6**, **11**, and **16**) by the active residues of Ser39, Lys76, Pro104, Val35, and Val36.

Along with Phe1, compounds (**11**, **13**, and **16**) displayed the maximum  $\pi$ -alkyl,  $\pi$ -cation, and  $\pi$ - $\pi$  interactions with the Tyr48 and Ile13 indicating the tight binding with the active site. It suggested that Tyr48 is considered as the principal component of the PPS, responsible for the accessibility of small molecules to the active site. Binding affinity and binding specialty were increased in the case of compounds (**4**, **5**, **6**, **11**, **13**, and **16**) due to their significant hydrogen bonding. It was observed that modifications of the -OH group of thymidine (**1**) increased the  $\pi$ - $\pi$  interactions with the residues of the active site while increasing their polarity resulted in the formation of hydrogen bonding interactions. The most prominent H-bonds were obtained in the compound (**11**), forming with Asp140, Gln133, Asn135, and Phe1 residues.

In contrast, compounds (**5**, **6**, and **16**) formed similar numbers of H-bonds with different residues. Again, compounds **11** and **16** showed the similar H-bond interactions with Phe1, Gln133, and Asn135, despite having different bonding distances. Hydrogen-bonds executed a vital function in shaping the specificity of ligand binding with the receptor, drug design in chemical and biological processes, molecular recognition, and biological activity [46]. The hydrogen bond surface and hydrophobic surface of the compound (**11**) consequently represented in Fig. 8. We realized that the analyzed bind within some of the catalytic active site such as isoleucines (Ile13 and Ile52), Arg98, Lys76, Tyr48, Tyr137, and Phe1 of the FimH of 4XO8, which is responsible for several activities of *E. coli* bacteria. The calculated binding affinities varied in the range of (-4.6 to -6.7 kcal/mol) suggesting the molecules can spontaneously interact within the binding site of FimH of *E. coli*. Among all the compounds, the inhibition activity of the compound (**11**) was found to be the highest (-6.7 kcal/mol).

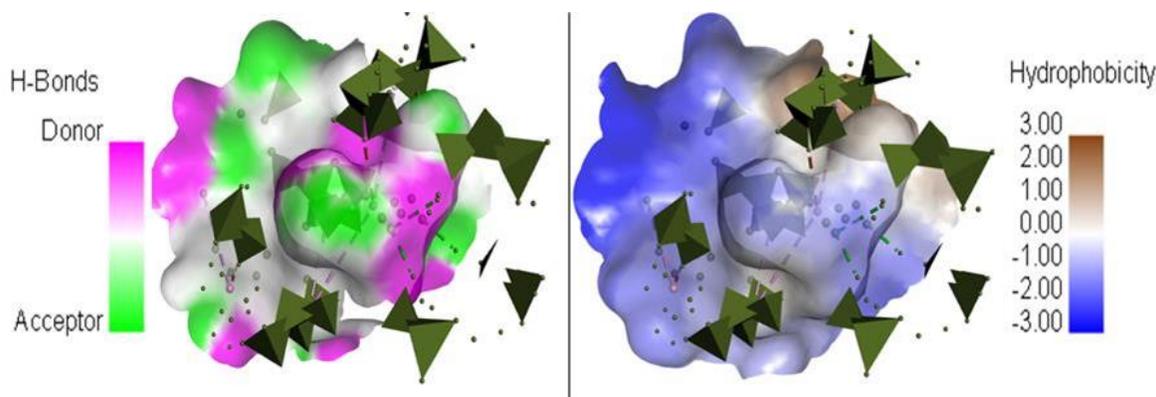


Fig. 8. Hydrogen bond and hydrophobic surface of FimH of *E. coli* of 4XO8 with compound (**11**).

### ADMET analysis

All newly modified thymidine derivatives have shown the potential activities. Therefore, to ensure that the modified compounds are viable drugs, we used the *in silico* pharmacokinetic parameters ADMET. The pkCSM online server [47] was employed to calculate *in silico* ADMET properties (Table 4). The absorbance value below 30% indicated poor absorbance but most of the compounds displayed a value above 70%, which revealed a good absorbance in the human intestine.

The volume of distribution (VDss) is thought high if the value is higher than 0.45. In addition, blood-brain barrier (BBB) and central nervous system (CNS) permeability standard values ( $>0.3$  to  $<-1$  Log BB and  $>-2$  to  $<-3$  Log PS), respectively. For a given compound a Log BB  $<-1$  is poorly distributed to the brain, while Log BB  $>0.3$  are the potential to cross BBB and LogPS  $>-2$  considered to penetrate the CNS, while Log PS  $<-3$  are difficult to move in the CNS [48]. It was observed that most of the compounds have the best significant potential to cross the barriers except compound (**7**). The enzymatic metabolism ensures the chemical

biotransformation of a designed drug in the body, which played a key role in the transformation of drug compounds. In the body, drugs produce several enzymatic metabolites, which contributed to catalyzing the reaction with various drug concentrations [49]. It is essential to consider their metabolism of drugs, which showed several physicochemical and pharmacological parameters. The cytochrome P450 (CYP450) depicted a major role in drug metabolism because of the major liver enzyme system involved in phase I metabolism. Some selective CYP genes CYP1, CYP2, CYP3, and CYP4 families were found to be involved in drug metabolism, with CYP (1A2, 2C19, 2D6, and 3A4) being causes the biotransformation of greater than 90% of drugs undergoing phase I metabolism. Therefore, among these members, CYP3A4 is the most important inhibition in this study [50]. All newly modified compounds were found to be the substrate and the inhibitor of CYP3A4. Clearance is a constant that indicated the relationship between drug concentration in the body and the rate of elimination of the drug. Therefore, all synthesized derivatives showed a somewhat high value, but still acceptable in the persistence of the drug in the body. Moreover, it is essential to test whether if the calculated compounds are non-toxic because this plays a critical role in the selection of drugs. Besides this, all the compounds we designed are non-toxic. Overall, compounds (3–6) and (8–16) have better *in silico* pharmacokinetic properties.

## Conclusion

In this study, we reported an easily applicable and efficient method for the synthesis of thymidine derivatives with good yields. These acylation's were found to be very promising since in all the reactions a single mono-substitution product was isolated in reasonably high yields. The structure and purity of all the compounds were indicated by their spectral and physical data. Molecular docking has been successfully employed to determine the best antibacterials against FimH of *E. coli* (4XO8). Among the studied eight thymidine derivatives, compounds (6, 11, 13, and 16) showed the highest binding affinity and strong interactions with at least one of the catalytic residues (Tyr48, Tyr137, Asp140, Gln133, Phe1, Asn135, Asn23, Arg98, Lys76, Ile13, and Ile52) of the FimH 4XO8. These compounds exhibited several non-covalent interactions, such as hydrogen bonding, hydrophobic, and electrostatic interactions. These blind molecular docking analyses may provide a potential approach for the application of antibacterials drugs as expected inhibitors *E. coli* protein FimH (4XO8). The docking validation process revealed that RMSD is in the standard range. Visual inspection demonstrated very convincing results in the molecular docking validation process. The ADMET prediction has been shown a promising result for *in silico* properties which indicate that all the modified compounds have an improved pharmacokinetic profile. Therefore, it concluded that most of the selected antibacterials showed promising features and might be used to design effective antibacterial drugs against *E. coli* protein FimH (4XO8).

## Acknowledgements

The authors are grateful to the Research and Publication Cell, University of Chittagong, Bangladesh for providing financial support to carry out this research project. The authors are grateful to the Director, Wazed Miah Science Research Centre, JU, Dhaka, Bangladesh for providing the spectra.

**Table 4.** In silico ADMET prediction of thymidine and 5'-O- and 3'-O-acyl derivatives.

Entry	Absorption	Distribution			Metabolism						Excretion	Toxicity
	Intestinal absorption (human)	VDss (human)	BBB permeability	CNS permeability	Substrate		Inhibitor				Total Clearance	AMES toxicity
					CYP							
					2D6	3A4	1A2	2C19	2D6	3A4		
Numeric (% Absorbed)	Numeric (Log L/kg)	Numeric (Log BB)	Numeric (Log PS)	Categorical (Yes/No)						Numeric (Log ml/min/kg)	Categorical (Yes/No)	
1	60.668	0.395	-0.982	-2.649	No	Yes	Yes	Yes	No	Yes	0.729	No
2	63.11	0.368	-0.103	-2.094	No	Yes	No	Yes	No	Yes	0.013	No
3	78.398	0.490	0.124	-2.515	No	Yes	Yes	Yes	No	Yes	0.935	No
4	75.942	0.481	0.438	-1.535	No	Yes	Yes	Yes	No	Yes	0.233	No
5	76.483	0.557	0.066	-2.194	No	Yes	Yes	Yes	No	Yes	0.716	No
6	89.479	0.464	0.941	-1.843	No	Yes	Yes	Yes	No	Yes	0.318	No
7	58.304	0.314	-0.918	-3.658	No	Yes	No	Yes	No	Yes	0.835	No
8	77.35	0.406	0.384	-2.902	No	Yes	Yes	Yes	No	Yes	0.324	No
9	78.801	0.296	0.106	-1.879	No	Yes	Yes	Yes	No	Yes	0.288	No
10	80.252	0.514	0.328	-1.855	No	Yes	Yes	Yes	No	Yes	0.317	No
11	73.153	0.493	0.136	-2.228	No	Yes	Yes	Yes	No	Yes	0.165	No
12	74.851	0.592	0.014	-2.377	No	Yes	Yes	Yes	No	Yes	0.993	No
13	84.691	0.470	0.120	-1.553	No	Yes	Yes	Yes	No	Yes	0.604	No
14	88.995	0.427	0.251	-1.475	No	Yes	Yes	Yes	No	Yes	0.690	No
15	90.445	0.476	0.294	-1.451	No	Yes	Yes	Yes	No	Yes	0.719	No
16	85.189	0.441	0.554	-1363	No	Yes	Yes	Yes	No	Yes	0.191	No

## References

1. Jordheim, L. P.; Durantel, D.; Zoulim, F.; Charles, D. *Nat. Rev. Drug Discov.* **2013**, *12*, 447–464.
2. Galmarini, C. M.; Mackey, J. R.; Dumontet, C. *Lancet Oncol.* **2002**, *3*, 415–424.
3. Damaraju, V. L.; Damaraju, S.; Young, J. D.; Stephen, A. B.; John, M.; Michael, B. S.; Carol, E. C. *Oncogene* **2003**, *22*, 7524–7536.
4. Eyer, L.; Nencka, R.; de Clercq, E.; Katherine, S. R.; Růžek D. *Antivir. Chem. Chemother.* **2018**, *26*, 1–28.
5. Vittori, S.; Ben, D. D.; Lambertucci, C.; Marucci, G.; Volpini, R.; Cristalli, G. *Curr. Med. Chem.* **2006**, *13*, 3529–3552.
6. Sax, N. I.; Lewis, R. J. *The Following List of Chemicals Known or Believed to be Teratogens is Drawn Primarily from Dangerous Properties of Industrial Materials*, 7th ed., John Wiley & Sons, Hoboken, N.J. **1988**.
7. Ohru, H. *Proc. Jpn. Acad. Ser.* **2011**, *87B*, 53–65.
8. Jordheim, L. P.; Durantel, D.; Zoulim, F.; Dumontet, C. *Nat. Rev. Drug Discov.* **2013**, *12*, 447–464.
9. Bacon, T. H.; Levin, M. J.; Leary, J. J.; Robert, Sarisky, T.; David, S. *Clin. Microbiol. Rev.* **2003**, *16*, 114–128.
10. Hamuy, R.; Berman, B. *Drugs Today (Barc.)* **1998**, *34*, 1013–1025.
11. Tabata, S.; Yamamoto, M.; Goto, H.; Akiyoshi, H.; Maki, O.; Takuya, K.; Atsushi, M.; Ryuji, I.; Misako, H.; Kohichi, K.; Yoshinari, S.; Kentaro, M.; Atsuro, S.; Masaki, H.; Yasuhiko, N.; Saburo, S.; Hiroyasu, E.; Masaru, T.; Tomoyoshi, S.; Tatsuhiko, F.; Shin-Ichi, A. *Cell Report* **2017**, *19*, 1313–1321.
12. Mitsuya, H.; Wetnholdt, K. J.; Furmant, P. A.; Clairt, M. H.; Nusinoff L. S.; Gallos, R. C.; Bologesit, D.; Arryt, D. W.; Broder, S. *Proc. Nalt. Acad. Sci.* **1985**, *82*, 7096–7100.
13. Hurst, M.; Noble, S. *Drugs* **1999**, *58*, 919–949.
14. Itoh, M.; Hagiwara, D.; Notani, J. *Synthesis* **1975**, *1975*, 456–457.
15. Tsuda, Y.; Haque, M. E. *Chem. Pharm. Bull. (Japan)* **1983**, *31*, 1437–1439.
16. Adinolfi, M.; Barone, G.; Iadonisi, A.; Schiattarella, M. *Tetrahedron Lett.* **2003**, *44*, 4661–4663.
17. Yan, Y. L.; Guo, J. R.; Liang, C. F. *Chem. Asian J.* **2017**, *12*, 2471–2479.
18. Ch, R.; Tyagi, M.; Patil, P. R.; Kartha, K. P. R. *Tetrahedron Lett.* **2011**, *52*, 5841–5846.
19. Andary, C.; Wylde, R.; Laffite, C.; Privat, G.; Winternitz, F. *Phytochemistry* **1982**, *21*, 1123–1127.
20. Ishji, H.; Nakamura, M.; Seo, S.; Tori, K.; Tozoy, T.; Yoshimura, Y. *Chem. Pharm. Bull. (Japan)* **1980**, *28*, 2367–2373.
21. Kabir, A. K. M. S.; Dutta, P.; Anwar, M. N. *Chittagong Univ. J. Sci.* **2005**, *29*, 1–8.
22. Mirja, H.; Thisbe K. L. *Eur. J. Org. Chem.* **2011**, *20-21*, 3583–3609.
23. Carreras, E.; Boix, E.; Navarro, S.; Rosenberg, H. F.; Cuchillo, C. M.; Nogues, M. V. *Mol. Cell. Biochem.* **2005**, *272*, 1–7.
24. Marc, T.; Nogues, M. V.; Boix, E. *J. Mol. Recognit.* **2011**, *24*, 90–100.
25. Chatfield, D.; Christopher, J. *Theor. Chem. Accounts Theor. Comput. Model (Theoretica Chim. Acta)* **2002**, *108*, 367–368.
26. Kawsar, S. M. A.; Hamida, A. A.; Sheikh, A. U.; Hossain, M. K.; Shagir, A. C.; Sanaulah, A. F. M.; Manchur, M. A.; Imtiaj, H.; Ogawa, Y.; Fujii, Y.; Koide, Y.; Ozeki, Y. *Int. J. Org. Chem.* **2015**, *5*, 232–245.
27. Shagir, A. C.; Bhuiyan, M. M. R.; Ozeki, Y.; Kawsar, S. M. A. *Curr. Chem. Lett.* **2016**, *5*, 83–92.
28. Said, S. A. J.; Anwar, U. H.; Abdul, R. I.; Mohammed, S. S. *Powder Technol.* **2007**, *175*, 115–121.

29. Gaussian09, R. A.; Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. Gaussian, Inc, Wallingford CT. **2009**.
30. Becke, A. D. *Phys. Rev.* **1988**, 38A, 3098–3100.
31. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev.* **1988**, 37B, 785–789.
32. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H. *Nucleic Acids Res.* **2000**, 28, 235–242.
33. Delano, W. L. *The PyMOL molecular graphics system*, San Carlos, CA, USA, **2002**.
34. Guex, N.; Peitsch, M. C. *Electrophoresis*, **1997**, 18, 2714–2723.
35. Dallakyan, S.; Olson, A. J. *Small-Molecule Library Screening by Docking with Py Rx*. In: Hempel, J. E.; Williams, C. H.; Hong, C. C. (eds.). *Chemical Biology: Methods and Protocols*, Springer, New York, USA, **2015**, 243-250.
36. *Version ADS 4.0, Accelrys*, San Diego, USA, **2017**.
37. Onodera, K.; Satou, K.; Hirota, H. *J. Chem. Inf. Model.* **2007**, 47, 1609–1618.
38. Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. *J. Med. Chem.* **2006**, 49, 5912–5931.
39. Ferreira, L. L. G.; Andricopulo, A. D. *Drug Discov. Today* **2019**, 24, 1157–1165.
40. Devi, S. R.; Jesmin, S.; Rahman, M.; Manchur, M. A.; Fujii, Y.; Kanaly, R. A.; Ozeki, Y.; Kawsar, S. M. A. *ACTA Pharm. Sci.* **2019**, 57, 47–68.
41. Jesmin, S.; Devi, S. R.; Rahman, M.; Islam, M.; Kanaly, R. A.; Fuji, Y.; Hayashi, N.; Ozeki, Y.; Kawsar, S. M. A. *J. Bang. Chem. Soc.* **2017**, 29, 12–20.
42. Arifuzzaman, M.; Mirajul, M. I.; Monjur, M. R.; Atiar, M. R.; Kawsar, S. M. A. *ACTA Pharm. Sci.* **2018**, 56, 7–22.
43. Kroemer, R. T.; Vulpetti, A.; McDonald, J. J.; Rohrer, D. C.; Trosset, J. Y.; Giordanetto, F.; Cotesta, S.; McMartin, C.; Kihlen, M.; Stouten, P. F. W. *J. Chem. Inf. Comput. Sci.* **2004**, 44, 871–881.
44. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, 267, 727–748.
45. Kontoyianni, M.; McClellan, L. M.; Sokol, G. S. *J. Med. Chem.* **2004**, 47, 558–565.
46. Perlstein, J. *J. Am. Chem. Soci.* **2001**, 123, 191–192.
47. Pires, D. E. V.; Blundell, T. L.; Ascher, D. B. *J. Med. Chem.* **2015**, 58, 4066–4072.
48. Clark, D. E. *Drug Discov. Today* **2003**, 8, 927–933.
49. Kok-Yong, S.; Lawrence, L. *Drug Distribution and Drug Elimination. Basic Pharmacokinetic Concepts and Some Clinical Applications*, InTechOpen, London, SW7 2QJ, UK, **2015**.
50. Thapar, M. M.; *Pharmacokinetics and Dynamics of Atovaquone and Proguanil*, Malarone®, Karolinska University Press, Stockholm, Sweden, **2004**.