Budapest University of Technology and Economics
Faculty of Chemical Technology and Biotechnology
George A. Oláh Doctoral School

Key structural details regulate oligomerization, cellular localization and inhibition of trimeric dUTPases

Thesis book

Author: András Benedek
Supervisor: Dr. Beáta G. Vértessy

Laboratory of Genome Metabolism and Repair

Department of Applied Biotechnology and Food Science
Budapest University of Technology and Economics

Institute of Enzymology
Research Centre for Natural Sciences
Hungarian Academy of Sciences Centre of Excellence

2019

1. Introduction

1.1. The physiological role of dUTPase

The dUTPase (2'-deoxyuridine 5'-triphosphate nucleotidohydrolase) enzyme family plays a key preventive role in maintenance of DNA integrity. This enzyme is supposed to prevent uracil incorporation into DNA via hydrolysis of 2'-deoxyuridine 5'-triphosphate (dUTP) into its monophosphate (dUTP), thereby dissipating the amount of energy being stored in the triphosphate group for incorporation of the uracil base [1].

Exclusion of uracil from DNA is the consequence of the relative instability of cytosines. Cytosine is one out of the four canonic bases of the genetic code, but unfortunately it is able to lose its amino group in a spontaneous process and it transforms into uracil. Uracil carries the same genetic information as thymine (5-methyluracil), therefore this change would lead to a stable point mutation in the genetic code. To avoid this, uracils in DNA are found and excised during the base excision repair process mediated by uracil-DNA glycosylase enzymes. The excised base is then going to be replaced by the correct complementary base of the opposite strand [1].

The base excision repair process would have been a perfectly operating process if conventional DNA polymerases were strictly specific for thymine against uracil, but they are not. Most of these enzymes are able to incorporate uracil instead of thymine depending on the intracellular dUTP/dTTP ratio. The resulting U:A base pairs carry the correct genetic information, but most uracil-DNA glycosylases recognize and cleave these uracils from DNA, as well. Therefore the base excision repair process carries a risk factor. If two uracils are situated in close proximity of each other on the opposite DNA strands and are excised at the same time, a double strand break is formed during the repair process. Simul-

taneously formed double strand breaks due to increased uracil incorporation may carry the risk of overloading the capacity of the repair process and can easily lead to cell death, referred to as „thymine-less" cell death. The enzyme dUTPase is responsible for prevention of this process [1].

2. Research background

2.1 Our basic motivation is to find a way for species-specific dUTPase inhibition

As dUTPase plays a key role in prevention of base excision repair hyperactivation, inhibition of the enzyme may be a useful tool for fighting against those group of diseases which are associated with extensive DNA synthesis bypassing self-regulatory mechanisms of the body. Accordingly, we may aim to inhibit the human dUTPase in a combination therapy to fight against cancer cells [2], but we may also aim at inhibition of dUTPase homologues encoded by certain parasites [3]. To use dUTPase inhibition in antimicrobial therapy, development of species-specific dUTPase inhibitors is needed which do not block the human dUTPase in their effective dose. To reach this goal, we are dedicated to characterize those structural differences which are present among dUTPase homologues encoded by different species.

2.2 Characterization of the covalent pseudo-heterotrimeric dUTPase structure

The vast majority of dUTPases adopt a homotrimeric, β-pleated structure. Five conserved sequence motifs being responsible for active site formation are found in the polypeptide sequence. Three active sites are formed via synchronized interaction of all three subunits throughout their conserved sequence motifs. The active site

1. Vertessy, B. G. & Toth, J. (2009) Accounts of chemical research. 42, 97-106.
2. Ladner, R. D. et al. (2000) Cancer research. 60, 3493-503.
3. Recio, E., et al. (2011) European journal of medicinal chemistry. 46, 3309-14.

architecture of different trimeric dUTPase homologues is almost identical, but those sequence sections which are not involved in active site formation may be slightly or even widely different depending on their evolutionary distance [1], rising hope for design of species-specific inhibitors.

Evolution of the so-called covalent pseudo-heterotrimeric dUTPase structure has left space for species-specific structural variations. In these cases trimer-like protein assembly is additionally ensured on the level of the genetic code, as the gene encoding dUTPase (*dut* gene) is triplicated and these copies are located in a single open reading frame. This structural assembly was first discovered in the genome of *Caenorhabditis elegans* and it was assumed that its dUTPase is translated as one single polypeptide in which the subunits are linked covalently by short peptide sequences [1]. As all three subunits have their own *dut* gene copy, point mutations accumulated during evolutionary development of the enzyme, making this structural arrangement even more unique.

Based on gene and domain data base searches we supposed that another nematode named *Caenorhabditis briggsae* and a fruitfly named *Drosophila virilis* also possess its dUTPase in a covalently linked form. The starting objective of my doctoral studies was to verify that the triplicated *dut* gene is indeed translated into one single trimer-like polypeptide. Thereinafter I wished to characterize this rare structural arrangement on the example of the *D. virilis* dUTPase. As the triplicated genetic code is probably emerged via two independent gene duplications, I aimed to reconstruct the ancient duplicated form of the *dut* genes and to express a covalent homodimeric *D. virilis* dUTPase from this form of genetic code. I was dedicated to verify experimentally if the assumed evolutionary intermediate form of dUTPase could indeed have been able encode a catalytically competent covalent pseudo-dimeric dUTPase.

I was also intended to reconstruct and express the putative ancient homotrimeric form of the *D. virilis* dUTPase encoded by a single

dut gene copy in order to compare its catalytic competence and stability with the currently existing form of the enzyme.

Finally, I wished to examine how point mutations being present among the three subunits of the covalent pseudo-heterotrimeric dUTPase affect intersubunit interactions.

2.3 Protein Stl as a species-specific dUTPase inhibitor

Almost all presently known dUTPase inhibitors are substrate analogue molecules binding to the active site of the enzyme. These molecules are generally not species-specific and they are poorly penetrating across biological membranes due to their high negative surface charge. However, according to a study of 2010, in certain strains of *Staphylococcus aureus* a protein interaction partner of dUTPase named $Stl_{SaPibov1}$ is present. Stl is a repressor protein of *S. aureus* pathogenicity islands (SaPi-s), a kind of mobile genetic elements acting as molecular parasites of certain bacteriophages. Namely, if the *S. aureus* cell is infected by either of the θ11 or 80α bacteriophages, the dUTPase of the bacteriophage interacts with protein Stl which abolishes repression of the SaPi-s. Consequently, the pathogenicity islands are excised from the genome, amplified and then packed into the newly formed phage capsids instead of the phage DNA itself, creating opportunity for horizontal gene transfer of SaPi-s. As the interaction between Stl and the phage dUTPase was reported to suspend the repressor function of Stl [4], my colleagues started to investigate if the activity of dUTPase is also inhibited upon complex formation. They demonstrated that the activity of Φ11 phage dUTPase may be completely inhibited *in vitro* [5]. In addition, they found that the hydrolysis reaction rate catalyzed by *Mycobacterium tuberculosis* and human dUTPases are inhibited by 84% and 70% upon Stl addition, respectively [6]. Based on these results on the one hand I wished to examine if protein Stl

4. Tormo-Mas et al. (2010) Nature. 465, 779-82.
5. Szabo, JE. et al. (2014) Nucleic Acids Res. 42, 11912-20.
6. Hirmondo, R. et al. (2015) DNA Repair (Amst). 30, 21-7.

may function as a universal dUTPase inhibitor. On the other hand I was dedicated to discover which amino acid side-chains or structural motifs of dUTPase homologues may be decisive in effective inhibition of dUTPase activity upon complex formation.

3. Experimental

Proteins were expressed recombinantly in *Escherichia coli* BL21-Rosetta strains. Covalent pseudo-trimeric translation of the triplicated *dut* gene copies was verified by western blot analysis using a primary antibody developed against *D. melanogaster* dUTPase. The most probably ancient *D. virilis* dUTPase sequence was identified by phylogenetic analysis. Putative ancient dUTPase gene sequences were reconstructed by site-directed mutagenesis and gene synthesis.

Catalytic efficiency of dUTPases was assessed spectrophotometrically. The reaction is accompanied by proton formation which is detected by addition of phenol-red indicator to the reaction buffer. This method was used for Stl inhibition measurements, as well, but an initial 5-minute incubation period was inserted to the protocol to promote complex formation. Heat stability was followed by differential scanning fluorimetry (thermofluorimetry), heat inactivation measurements and in case of *D. virilis* dUTPases by circular dichroism spectrometry. Oligomeric state of proteins was assessed by mass spectrometry and dynamic light scattering.

Interaction between dUTPases and protein Stl was detected by size-exclusion chromatography, thermofluorimetry and native gel electrophoresis. Isothermal titration calorimetric measurements provided information on complex formation and the strength of interaction at the same time. Position of the point mutation which has made *E. coli* dUTPase inhibitable by protein Stl was found by multiple sequence alignment of *E. coli* dUTPase with four inhibitable dUTPase homologues. Effect of the point mutation on dUTPase

structure was analysed by X-ray diffraction measurement performed on the *E. coli* Q93H point mutant dUTPase crystal.

4. Results

4.1. The triplicated dut gene indeed encodes for covalent pseudo-heterotrimeric dUTPases

Using the fruitfly *Drosophila melanogaster* encoding its dUTPase in a homotrimeric form as a reference, we separated the total protein content of *D. melanogaster, D. virilis, C. elegans* és *Trichinella spiralis* tissue samples by gel electrophoresis, then visualized the dUTPase protein bands in their corresponding molecular weight position using western blot technique. Sodium-dodecyl-sulphate content of the gel electrophoresis buffer disrupted non-covalent intersubunit interactions. Consequently, the homotrimeric reference dUTPase appeared in its monomeric form upon antibody treatment, while all other dUTPase bands appeared at the trimeric position, arguing for possession of a covalently linked architecture.

4.2. A covalent dimeric D. virilis dUTPase representing a putative evolutionary intermediate enzyme architecture is a stable and effective catalyst of dUTP hydrolysis

Using phylogenetic analysis I have made clear that the ancient homotrimeric and putative covalent pseudo-dimeric forms of *D. virilis* dUTPase are sequentially really close to subunit “B” of the extant “ABC” covalent pseudo-heterotrimeric dUTPase form. Therefore they are labelled as B^*_3 and B*B* dUTPases, respectively. Both reconstructed putative ancient enzyme forms are potent catalysts of dUTP hydrolysis and their heat stability is almost equal to that of the extant “ABC” dUTPase (Table 1.).

dUTPase	k_{cat} (s^{-1})	Melting point (°C)
ABC	4.7 ± 0.1	65.0 ± 0.5
B^*_3	4.6 ± 0.3	65.9 ± 0.3
B^*B^*	3.4 ± 0.3	66.2 ± 0.4

Table 1. *The extant covalent pseudo-heterotrimeric (ABC), ancient homotrimeric (B^*_3) and putative covalent pseudo-homodimeric (B^*B^*) forms of D. virilis dUTPase compared by their reaction rate constants (k_{cat}) and melting temperatures.*

According to my measurements the B*B* dUTPase was able to accomplish its functional role if has indeed existed in this form as an evolutionary intermediate.

4.3. Point mutations among the subunits of D. virilis dUTPase are not independent from each other

To examine the impact of point mutations on intersubunit interactins, I have placed the triplicated *D. virilis dut* gene copies into separate open reading frames. I assumed that the monomeric polypeptides expressed from the separated gene copies will be able to form homotrimers in a spontaneous process, teherefore I named these proteins A_3, B_3 and C_3 dUTPases. My hypothesis was directly verified by mass spectrometric analysis and indirectly strengthened by isothermal titration calorimetry for the A_3 dUTPase. In the latter case importin-α binding ability of the wild-type "ABC" and the artificial A_3 dUTPase was compared. The change in enthropy was approximately three times higher for the A_3 dUTPase than for the wild-type "ABC" dUTPase, arguing for the presence of three nuclear localization signals on the A_3 dUTPase sequence.

While the existing point mutations among the three subunits do not prevent homotrimer formation from the artificially separated subunits, catalytic efficiency of these non-natural homotrimers is one order of magnitude lower compared to the wild-type dUTPase.

Accordigly, their heat stability is 20-30% lower than that of the "ABC" or the ancient B^*_3 enzyme. These results provide a strong evidence for the interdependence of point mutations among the three pseudo-subunits and argue for their complementary relationship.

4.4. Protein Stl only partially inhibits the Drosophila melanogaster dUTPase (40% decrease in reaction rate) and does not inhibit the Escherichia coli dUTPase

As I wished to investigate if protein Stl may be considered as a universal dUTPase inhibitor, I decided to test its inhibitory potential on dUTPase homologues of two evolutionary distinct laboratory model organisms. Namely, the *D. melanogaster* and *E. coli* dUTPases were used as test enzymes. I have concluded that the *D. melanogaster* dUTPase may only be partially inhibited by protein Stl, while the *E. coli* dUTPase is not inhibited at all (Fig. 1).

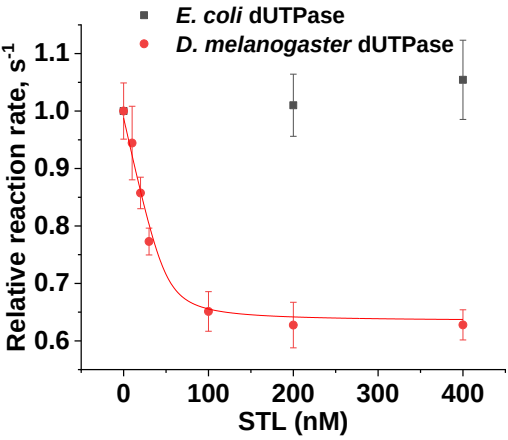

Fig. 1. *Protein Stl does not inhibit the E. coli dUTPase, and exerts only 40% inhibition on D. melanogaster dUTPase.*

4.5 E. coli dUTPase and protein Stl form complex despite the lack of inhibition

I was intended to verify complex formation between *D. melanogaster* dUTPase and protein Stl and I wished to perform the same experiments on the *E. coli* dUTPase – Stl mixture as control measurements. Strikingly a complex was forrned between the latter two proteins, as well (Fig. 2). Consequently, the *E. coli* dUTPase is also an interaction partner of protein Stl, despite of the lack of inhibition upon complex formation.

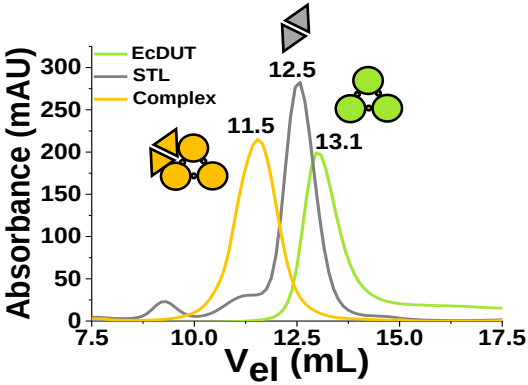

Fig. 2 *Detection of complex formation between the E. coli dUTPase and protein Stl using size-exclusion chromatography.* EcDUT – *E. coli dUTPase.*

4.6. Mutation of the 93rd glutamine side-chain of E. coli dUTPase either to histidine or arginine results in Stl-inhibitable point mutants

I assumed that complex formation but the lack of inhibition at the same time might be explained by minor sequential alterations among the *E. coli* dUTPase and the "Stl-sensitive" homologues. Even one amino acid difference may result in inhibition-less binding if the conserved dUTPase sequence motifs are affected by complex formation. Towards this end I was searching for such amino acids in the *E. coli* dUTPase which are markedly different in their

side-chain characteristics from the inhibitable dUTPase homologues at a given position along the protein sequence (Fig. 3).

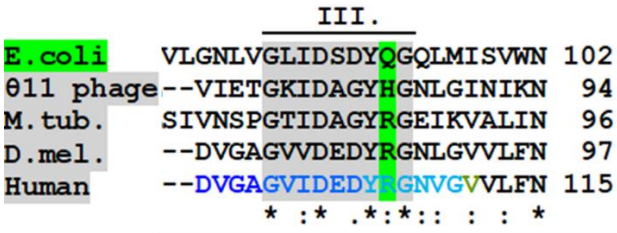

Fig. 3 Multiple sequence alignment of four dUTPases being inhibitable by protein Stl (names on grey background) and the non-inhibitable *E. coli* dUTPase (name on green background). Within the 3rd conserved sequence motif (grey background) a remarkable difference in side-chain characteristics of the *E. coli* enzyme is found. Namely, the 93rd glutamine side-chain possesses an amide group, contrary to the positively charged histidine or arginine side-chain of Stl-inhibitable dUTPases.

After identification of some promising positions by multiple sequence alignment I started to generate point mutant versions of the *E. coli* dUTPase. Changing the Q93 side-chain either to histidine (Q93H) or arginine (Q93R) resulted in "Stl-sensitive" point mutants (Fig. 4).

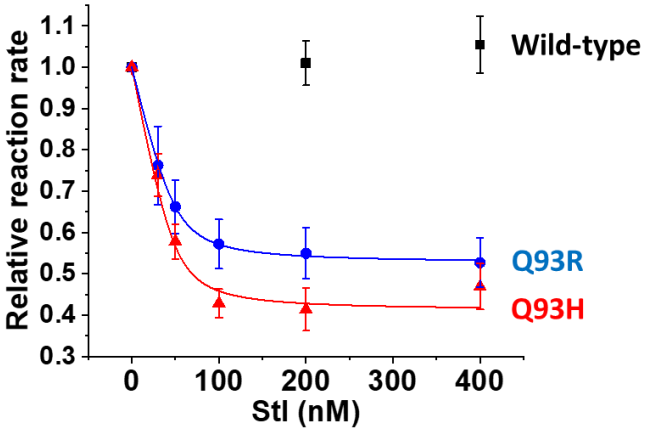

Fig. 4*. The Q93H and Q93R point-mutants became inhibitable by protein Stl, contrary to the wild-type E. coli dUTPase.*

4.7 The C-terminal arms became visible in crystal structure of the E. coli Q93H dUTPase

In order to explain the structural reasons of inhibitability, I successfully crystallized the *E.coli* Q93H dUTPase and we resolved its structure based on the crystal's X-ray diffraction pattern at 1.82 Å resolution (PDB-ID: 6HDE). The novelty of the structure is that the C-terminal arms became visible in the model, contrary to any previously published *E. coli* dUTPase crystal structure. These sequence segments play a key role in closure of the active site upon substrate binding. I analysed the structure and pointed out that the Phe135 residue of the C-terminal arm establishes an aromatic overlapping interaction with the His93 side-chain, and at the same time the Glu142 side-chain on its other end is fixed by a water molecule mediated hydrogen-bond network to the His93 side-chain. Consequently, the Q93H mutation decreased conformational flexibility of the C-terminal arm, which must be in connection with the point mutant's increased sensitivity towards protein Stl.

4.8 Mobility of the C-terminal arm is decreased upon substrate analogue binding in the E. coli Q93H dUTPase, contrary to the wild-type enzyme

We wished to verify decreased conformational flexibility of the C-terminal arm in a solution-phase experiment, as well. Therefore we introduced a tryptophan residue into the C-terminal arms of the wild-type and the Q93H and Q93R point-mutant enzymes. This modification enabled us to follow arm movements via tryptophan fluorimetry. Our results indicated that the C-terminal arms of the *E. coli* Q93H dUTPase are significantly less exposed to solution upon substrate analogue binding than in their apo state, contrary to the C-terminal arms of the quasi-wild type dUTPase bearing only the arm tryptophan mutation.

5. New scientific results

1. We have shown that the triplicated genetic code of *Drosophila virilis* dUTPase is indeed translated as one single polypeptide, but artificial *D. virilis* dUTPase homotrimers may be created if the triplicated *dut* gene copies are engineered into separate open reading frames [P1, P4].

2. I have proven that the putative evolutionary intermediate of *D. virilis* dUTPase is a functional enzyme in the form of a covalent pseudo-dimer [P1].

3. I have shown that the point-like sequence alterations among the subunits of *D. virilis* dUTPase are complementary to each other and hence these subunits became slightly different in terms of flexibility and heat stability [P1].

4. I have shown that a remarkable difference may be present in the maximal degree of Stl mediated inhibition among dUTPase homologues. Consequently, trimeric dUTPase protein sequences involved in binding to Stl are only partially conserved through evolution from bacteria to eukaryotes [P2, P3].

5. I have made clear that protein Stl may be a universal interaction partner but not a universally effective inhibitor of dUTPases. The explanation for this lies behind the existence of minor structural differences among dUTPase homologues. These minor differences may serve as a basis for species-specific dUTPase inhibitory peptide design [P3].

6. I have pointed out that most probably the high flexibility of the C-terminal arm is responsible for the lack of Stl-mediated inhibition in the wild-type *E. coli* dUTPase [P3].

6. Possible applications

Analysis of species-specific structural differences among dUTPase homologues provides an opportunity to develop species-specific dUTPase inhibitors based on this knowledge. Species-specificity has utmost importance if dUTPase inhibition is to be used in anti-microbial therapies, as the human cells also encode a trimeric dUTPase homologue.

Development of a dUTPase inhibitor harnessing the covalent pseudo-heterotrimeric enzyme organization may provide a new, species-specific antibiotic against the human and mammalian pathogen nematode, *Trichinella spiralis*.

Fine-tuning the interaction between protein Stl and dUTPase homologues is a promising way of species-specific dUTPase inhibitor development.

7. Publications

7.1. Papers used for the preparation of the thesis

(IF: impact factor, IC: independent citation, AC: author contribution among those persons who has not obtained their PhD degree).

P1. **Benedek A**, Horváth A, Hirmondó R, Ozohanics O, Békési A, Módos K, Révész Á, Vékey K, Nagy GN, Vértessy BG. Potential steps in the evolution of a fused trimeric all-β dUTPase involve a catalytically competent fused dimeric intermediate. (2016) FEBS J. 2016 Sep;283(18):3268-86. Editor's choice. doi:10.1111/febs.13800. IF (2016): 4.237. IC: 1. AC: 100%.

P2. **Benedek A**, Pölöskei I, Ozohanics O, Vékey K, Vértessy BG. The Stl repressor from *Staphylococcus aureus* is an efficient inhibitor of the eukaryotic fruitfly dUTPase. (2017) FEBS Open Bio. 2017 Dec 27;8(2):158-167. doi:10.1002/2211-5463.12302.

IF (2017/2018): 1.782. AC: 95%. Fig. 1A chosen for cover illustration of FEBS Open Bio.

P3. **Benedek A**, Temesváry-Kis F, Khatanbaatar T, Leveles I, Surányi ÉV, Szabó JE, Wunderlich L, Vértessy BG. (2019) The Role of a Key Amino Acid Position in Species-Specific Proteinaceous dUTPase Inhibition. Biomolecules. 2019 Jun 6;9(6). doi:10.3390/biom9060221. IF (2019): 4.694. AC: 70%.

P4. Róna G, Pálinkás HL, Borsos M, Horváth A, Scheer I, **Benedek A**, Nagy GN, Zagyva I, Vértessy BG. (2014) NLS copy-number variation governs efficiency of nuclear import--case study on dUTPases. FEBS J. 2014 Dec;281(24):5463-78. doi:10.1111/febs.13086. IF (2014): 4.001. IC: 1. AC: 15%.

7.2. Verbal lectures on conferences

1. Benedek A, Temesváry-Kis F, Khatanbaatar T, Leveles I, Surányi ÉV, Szabó JE, Wunderlich L, Vértessy BG. Insights into the structural background and mechanism of inhibition of dUTPases by a proteinaceous inhibitor. Hungarian Molecular Life Sciences 2019. Eger, Hungary.

2. Benedek A, Pölöskei I, Vértessy BG. A comparative study of dUTPase inhibition by Stl, a staphylococcal repressor protein. PhD Scientific Days 2016. Budapest, Hungary.

3. Benedek A, Pölöskei I, Vértessy BG. Az Stl bakteriális represszor dUTPázokra gyakorolt hatásának összehasonlító elemzése. BME ABÉT Tanszéki Doktoráns Konferencia, 2017. január 31. Budapest, Hungary

7.3. Poster presentations on conferences

1. Benedek A, Temesváry-Kis F, Khatanbaatar T, Leveles I, Surányi ÉV, Szabó JE, Wunderlich L, Vértessy BG. (2019) The Role of a Key Amino Acid Position in Species-Specific dUTPase Inhibition. International School on Biological Crystallization. Granada, Spanyolország, 2019. május 26-31.

2. Benedek A, Temesváry-Kis F, Khatanbaatar T, Leveles I, Surányi ÉV, Vértessy BG. Insights into the structural background and mechanism of inhibition of dUTPases by a proteinaceous inhibitor. 31st European Crystallographic Meeting, Oviedo, Spanyolország, 2018. augusztus 22-27.

3. Benedek A, Ozohanics O, Vékey K, Pölöskei I, Temesváry-Kis F, Khatanbaatar, T, Vértessy BG. A Comparative Study of Protein Stl's Inhibitory Effect on a Eukaryotic and a Prokaryotic dUTPase. Hungarian Molecular Life Sciences 2017, Eger, Hungary.

4. Benedek A, Pölöskei I, Vértessy BG. A comparative study of protein Stl's inhibitory effect on a eukaryotic and a prokaryotic dUTPase. EMBO Young Sientists' Forum 2016. Lisbon, Portugal, 2016. szeptember 1-2.

5. Benedek A, Pölöskei I, Vértessy BG. Comparative analysis of dUTPase enzyme - Stl inhibitor protein interaction effect on a prokaryotic and a eukaryotic enzyme model. FEBS Advanced Methods in Marcomolecular Crystallization VII. Nové Hrady, Czech Republic, 2016. június 27. – július 2.

6. Benedek A, Pölöskei I, Vértessy BG. Protein Stl, a *Staphylococcus aureus* related transcription factor inhibits a eukaryotic dUTPase. Hungarian Molecular Life Sciences 2015. Eger, Hungary.

7. Benedek A, Horváth A, Nagy GN, Vértessy BG. Role of subunit interactions in structural competence of *Drosophila virilis*

dUTPase. Chemistry of Metals in Biological Systems, Louvain-la-neuve, Belgium, 2014. szeptember 7-14.