UNIWERSYTET TECHNOLOGICZNO-PRZYRODNICZY IM. JANA I JĘDRZEJA ŚNIADECKICH W BYDGOSZCZY ZESZYTY NAUKOWE NR 265 TELEKOMUNIKACJA I ELEKTRONIKA 19 (2016) 5-18

COULD MATLAB HELP TO CURE HIV?

Anna Marciniak¹, Sylwester Kloska¹, Daniel Bujnowski³, Vinesh Badloe², Elio Abbondanzieri², Mahipal Ganji²

¹Nicolaus Copernicus University Ludwik Rydygier Collegium Medicum in Bydgoszcz, Department of Medicine, Faculty of Biotechnology, ul. Jagiellońska 13-15, 85-067 Bydgoszcz, Poland

> ²Delft University of Technology, Department of Bionanoscience, Postbus 5, 2600 AA Delft, the Netherlands

³UTP University of Science and Technology, Faculty of Telecommunications, Computer Science and Electrical Engineering, al. Prof. S. Kaliskiego 7, 85-796 Bydgoszcz, Poland

Summary: The human immunodeficiency virus (HIV) is a virus that causes HIV infection and can lead to acquired immunodeficiency syndrome (AIDS). HIV infects cells of the immune system - especially those, which are responsible for the activation of immune response. Every year a huge amount of people die due to diseases that would not be fatal if their immune system was working properly. Scientists from every country want to create an effective drug that helps to cure the infection and prevent development of AIDS. It is necessary to learn everything about HIV to create a drug that will help to save a lot of lives. There is a lot of information discovered by now but also there are some things that remain unknown and should be revealed. One of the most important enzymes for HIV is reverse transcriptase (RT). Thanks to this enzyme virus can re-write its genetic material from RNA (ribonucleic acid) to more stable cDNA (complementary DNA). Finding out the requirements for proper work of RT will help to block and stop the enzyme. A good way to study RT is to observe it under a laser microscope. Laser microscope allows observing single molecules. It is possible to see how RT works with different lengths of DNA (deoxyribonucleic acid) constructs and how does obstacles effect the activity of RT. Results from microscope observations can be analysed using MATLAB software. Special scripts are necessary to analyse binding events and how long they last.

Keywords: HIV, reverse transcriptase, FRET, MATLAB calculations

1. INTRODUCTION

The human immunodeficiency virus (also known as HIV) is responsible for HIV infection and development of acquired immunodeficiency syndrome (AIDS). HIV virus is a lentivirus (subgroup of retroviruses) and contains RNA as genetic material. HIV virus is divided to two types: HIV-1 and HIV-2. Those two types differ between virulence, infectivity and prevalence. HIV-1 type is more virulent and infective than

type HIV-2 and occurs globally, when HIV-2 type occurs mostly in West Africa. Nowadays, HIV is the best known virus, although it is characterized by high volatility, which can manifest itself even in one patient (in various stages of development of the infection). As a retrovirus, HIV virus has to integrate to host genome to replicate itself. To accomplish it, it is necessary to have integrase enzyme and reverse transcriptase (RT). Those two enzymes allow transcription from RNA to cDNA (complementary DNA) and integration to host genome (Fig. 1) [1].

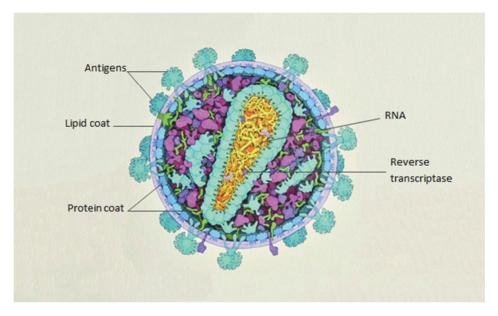


Fig. 1. HIV virus construction – genetic material (RNA) and enzyme (reverse transcriptase) is covered with both protein and lipid coats with antigens, which allows virus to penetrate the host cell [2]

Reverse transcriptase is an enzyme that allows copying single-stranded RNA, which is highly unstable, to more stable double-stranded cDNA. This enzyme is typical for the retrovirus family and some of hepadnaviruses (viruses that contains DNA as genetic material) [1]. RT has 3 functions: it synthesizes DNA on RNA template, synthesizes DNA on DNA template and hydrolysis RNA on DNA template [3]. RT differs between species, mostly in parameters such as molecular weight or number of subunits (Fig. 2). In human cells RT can be found as well. In that case RT is responsible for maintaining length of the telomeres of eukaryotic chromosomes [1, 4].

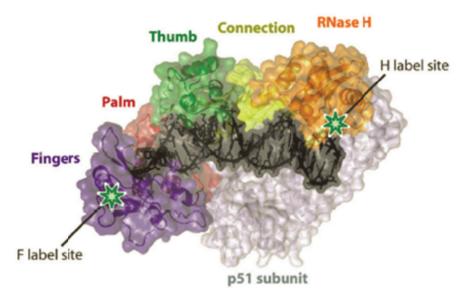


Fig. 2. The structure of HIV-1 reverse transcriptase. Labelling sites for Cy-3 on RT are highlighted by green stars [3]

When virus penetrates the host cell (e.g. human) then its lipid coat disappears and genetic material is released. Then there comes reverse transcription – a process in which the single-stranded RNA is transcribed on cDNA, which is double-stranded on the end of this process. cDNA is capable of integration with genome of the host. This integration plays a key role and is necessary for the next step of virus life cycle. Integrated genetic material uses DNA replication machinery of host to multiply itself (repopulate). When a proper number of copies is attained then the lipid coat is recreated and genetic material is packed inside. New-born particles of virus are being released by the disintegration of a host cell and then they infect other cells of host organism. Thanks to this system, the infection remains dormant for a long time. Cells cannot fight this kind of infection because virus mostly attacks the population of helper lymphocytes (Th), which are responsible for stimulation of immune response. Reverse transcriptase (RT) is an indispensable enzyme for a virus. Thanks to RT virus can replicate itself without any cost because everything that it needs is provided by the host cell [5].

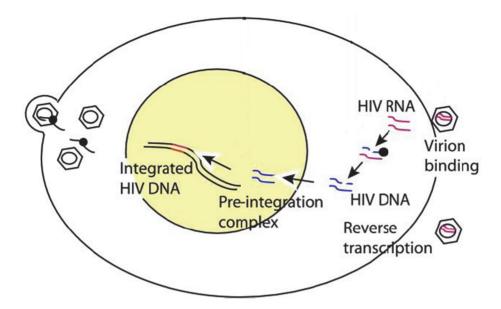


Fig. 3. HIV virus life cycle. HIV virus penetrates host cell. After that, released HIV RNA undergoes reverse transcription process, and then created HIV DNA (cDNA) integrates to host genome and replicates virus particles. In the end new virus particles are released by destroying host cell [5]

AIDS, which is caused by HIV virus was the reason of death in over than one billion people (worldwide) in 2014 [6]. Understanding the mechanism of action of RT may form the basis for the development of effective drugs (or even vaccines) against HIV. Most of the drugs that are used in treating of HIV infection work like inhibitors of reverse transcriptase [7, 8]. They should not allow RT to transcript the genetic material from RNA to cDNA. Unfortunately, so far there have not been invented a drug that would completely stop RT.

In this research we wanted to check total binding time and amount of binding events using new methods of analysis.

2. MATERIALS AND METHODS

The RT, which was used in this experiment, was in-home generated and labelled with fluorophore Cy-3 (cyanine), which maximum of absorption is 532-nm (nanometres). Maximum absorption of Cy-5 is 635-nm. The various length DNA strands (which were a substrate) were labelled with Cy-5, which maximum of absorption is the same as the donor's emission. The beam of 532 nm aroused the labelled enzyme but it did not arouse Cy-5 on the substrate. Cy-5 was aroused only if the donor emitted beam with a proper wavelength. To understand all of this we should take a look on the Jabłoński diagram, which illustrates intramolecular processes of redistribution and excitation energy dissipation of chemical molecule following the absorption of a photon and lead to emissions, i.e. fluorescence or phosphorescence. It allowed the detection of Förster Resonance Energy Transfer (FRET) (Fig. 4) [9, 10].

FRET is a mechanism in which chromophores are capable of transmitting energy, but only if they are close enough to each other (<10 nm). If 2 proteins share a distance of less than 10 nm, it means that the reaction is going on.

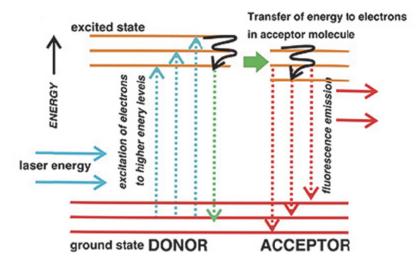


Fig. 4. The scheme of FRET. Laser energy excites donor, then energy is transferred to electrons in acceptor molecule [11]

To observe the FRET we need a laser microscope, which consists of a source of light, mirrors, the beam divider that allows distributing the light on red and green, a prism and CCD camera (charge coupled device). To work a CCD camera needs to be cooled to temperature about -80°C. This temperature is needed because of the speed of taking photos, which can lead to overheating. To watch the results we need a program (e.g. LabView). It is necessary to write proper scripts, which allow alternate switching of green and red laser beams and watching sample on a microscope (Fig. 5).

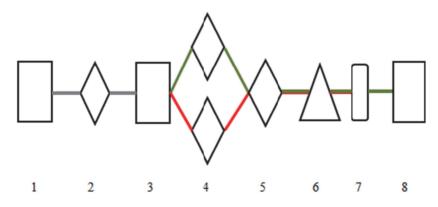


Fig. 5. Scheme of laser microscope. 1 – source of light, 2 – mirror, 3 – separator, 4 – mirrors, 5 – lens, 6 – prism, 7 – glass slide, 8 – CCD camera

To be able to run this experiment it is necessary to prepare a sample in a proper way. Glass slide, which is usually used in microscope observations, was modified this time. To create the flow cells 5 holes was drilled on the longer edges of a slide. A glass slide prepared this way was then covered with polyethylene glycol (PEG) and stored in a temperature of -86°C. Before being used in an experiment the slide was covered with PEG one more time for about an hour. Then the slide was washed with distilled water and then dried. Double-sided tape was stuck between the drilled holes and after that a cover glass was put on top of it. Last part was to seal the edges with glue to prevent leaking. After the glue has dried, the glass slide was ready for further stages of analysis.

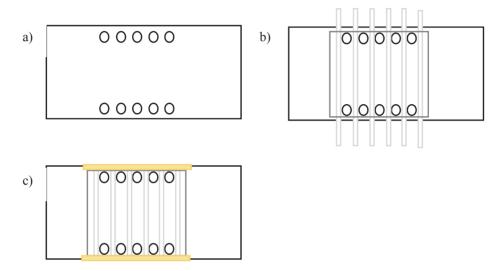


Fig. 6. Glass slide preparation steps: a) drilling holes in glass slide, b) sticking tape and putting on cover glass, c) sealing the edges with glue

In the experiment there was used a linkage between biotin and streptavidin. It is the strongest binding naturally occurring. The strength of this bond is influenced by a few factors, like high complementarity of shapes between so called pocket of streptavidin and biotin. There is also a very extensively network of hydrogen bonds, which stabilizes biotin, when it is in an appropriate position to bind with streptavidin. The "pocket" for biotin has a hydrophobic character (it does not like water). Hydrophobic interactions and the van der Waals bonds, which are also present in there, contribute to the high affinity of biotin and streptavidin. The last thing that should be mentioned in here is an elastic loop, which closes bonded biotin in the "pocket" and thereby contributes to slowing the dissociation. Because of the strength of binding of biotin to streptavidin or avidin, it is used in various fields of molecular biology (during Roche454 sequencing technology), in microbiology and immunology (enzyme-linked immunosorbent assay (ELISA)) to increase the sensitivity of detection [12, 13].

In the experiment the DNA sample was used because of its less demanding requirements of storage and usage. Template was 63 bp (base pairs) long. Primer was 40 nt (nucleotides) long (38 nt complementary and 2 nt non-complementary flap). We tested the activity of RT in various constructs – to described basis and primer we attached constructs in which 15 nucleotides was paired (complementary) and 3, 6, 9, 14

and 23 was not complementary. We also tested the construct without any noncomplementary flap. In total there have been 7 constructs tested. We tested the attitude of RT in each of them: the total binding time and the number of binding events.

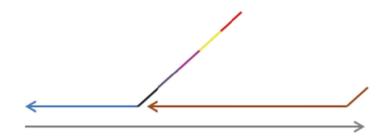


Fig. 7. Scheme of DNA construct. With different colours are marked different flap lengths. The grey colour is 63 bp template, the brown colour is primer (38+2nt). The blue construct is 15 nt long. The additional flaps are 3 nt (black), 6 nt (violet), 9 nt (pink), 14 nt (yellow) and 23 nt (red)

In a first stage of microscopic assay streptavidin was applied to the flow cell and incubated for one minute. After that time the flow cell was washed out with TE buffer plus sodium chloride. Afterwards DNA sample was applied and incubated for one minute. When the incubation was over the last part was to add imaging buffer, consisting of water, 50 mM Tris-HCl buffer (pH = 8.0), 10 nM reverse transcriptase (RT), 1x glucose oxidase (GOx), 0.2 mg/ml bovine serum albumin (BSA), 5% glucose, 0,2% Triton-X100, 100 mM sodium chloride (NaCl), 6 mM magnesium chloride (MgCl₂) and 2 mM Trolox. Imaging buffer was applied to add the reverse transcriptase to the flow cell and the components that prevent rapid ordination of fluorophores. Thus prepared, the preparation can be seen under a laser microscope.

Component	Function
Water	Dissolvent for other components
Tris-HCl buffer	Maintain proper pH
Various length of DNA strand	Substrate for reverse transcriptase
Reverse transcriptase	Enzyme that carries out process
Glucose oxidase	Used to eliminate oxygen from the reaction
Bovine serum albumin	Used to stabilize enzyme and prevent adhesion
Glucose	Substrate for glucose oxidase
Triton x100	Detergent, reduce surface tension
Sodium chloride (NaCl)	Different concentration of this component allows modulating reaction speed
Magnesium chloride (MgCl ₂)	Supplies Mg ²⁺ ions for proper working of RT
Trolox	Antioxidant, protects DNA sample from damage

Table 1. Components necessary to run microscopic assay. Each of them is added to the solution with specific purpose

The preparation was placed under the laser microscope. Properly written scripts in LabView software allowed alternately irradiation with the green and red laser and recording the movie. A highly sensitive, monochrome CCD digital camera was used to save images. It used the active cooling system. This type of cameras is dedicated to work in the field of cell biology that requires short exposure times (fluorescence). Movie consisted of a series of photos made with CCD camera each 0.1 sec for around 210 sec. To correlate the molecules that were irradiated with various laser colours, a mapping function was used. Mapping allowed to match and merger molecules from two screens (Fig. 9). Then there was selected area, from which molecules were chosen for the next stages of analysis. For this purpose was used ImageJ software (Image Processing and Analysis in Java) [14]. Molecules that were on edges have been removed from further analysis to avoid false results. ImageJ software chooses the local maxima depending on determined noise tolerance. Position of these molecules was presented in Cartesian – each molecule had X and Y coordinate (Fig. 10). Then those coordinates was saved as a list in "*.txt" file. The next stage was extraction of the data in MATLAB software.

```
% Import data from file `txt'
PeaksFileName = dir('*.txt');
Peak=importdata(PeaksFileName(1,1).name);
Cy5Xcoordinate=Peak.data(:,2)+1;
Cy5ycoordinate=Peak.data(:,3)+1;
```

Listing 1 Load values for the coordinates of the txt file

Data are processed in three MATLAB files. First of them is responsible for preprocessing data. Listing 1 shows only the most important lines of code, how to import data for appropriate channels.

Afterwards it is necessary to obtain coordinates in green channel. Coordinates are calculated from mapping and peaks of the red channel. Code showed on Linsting 2 is responsible for create three matrixes needed to obtain coordinates in green channel. In two of them $(p_r_x \& p_r_y)$ are stored data corresponding to coordinates in red channel. In 3rd matrix are calculated data to 4th order polynomial.

```
% Obtain the coordinates in green channel from mapping and
peaks from red channel
for i=1:length(Cy5Xcoordinate)
    if Cy5Xcoordinate(i)>0
        p_r_x(k)=Cy5Xcoordinate(i)-256;
        %4th order polynomial
        args(k,:)=[1 p_r_x(k) p_r_y(k) p_r_x(k)*p_r_y(k) p_r_x(k)^2
p_r_y(k)^2 p_r_y(k)*p_r_x(k)^2 p_r_x(k)*p_r_y(k)^2 p_r_x(k)^3
p_r_y(k)^3 p_r_x(k)^3*p_r_y(k) p_r_x(k)^2*p_r_y(k)^2
p_r_x(k)*p_r_y(k)^3 p_r_x(k)^4 p_r_y(k)^4];
        k=k+1;
    end
end
```

Listing 2 Matrixes storing data for the green channel

Then in the decision process are calculated coordinates for the green channel, depending on the level of the red channel (listing 3).

```
for i=1:length(args(:,1))
% the transformation type is polynamial
if p_r_x(1,i) <=140
        p_g_x(i,:) = args(i,:) * mytform1.tdata(:,1);
        p_g_y(i,:) = args(i,:) * mytform1.tdata(:,2);
elseif p_r_x(1,i) >140 & p_r_x(1,i) <=370
        p_g_x(i,:) = args(i,:) * mytform2.tdata(:,1);
        p_g_y(i,:) = args(i,:) * mytform2.tdata(:,2);
elseif p_r_x(1,i) >370
        p_g_x(i,:) = args(i,:) * mytform3.tdata(:,1);
        p_g_y(i,:) = args(i,:) * mytform3.tdata(:,2);
end
end
```

Listing 3 Calculating data for the green channel

Next the data are aggregated and regions of interest are saved in to file (listing 4)

```
FnamaGreen=[fname 'GreenRegionsOfInterest' num2str(l) '.mat'];
save(FnamaGreen,'GreenRegionOfInterest','-mat') % saves all
the extracted green regions of interests
FnamaRed=[fname 'RedRegionsOfInterest' num2str(l) '.mat'];
save(FnamaRed,'RedRegionOfInterest','-mat')% saves all the
extracted red regions of interests
```

Listing 4 Save the selected regions to files

Next file contain the code which is reduce level of noise in an image. Algorithm import pre-prepared data and process them to reduce level of noise. Code shown on listing 5 is responsible for calculate level of threshold which is used to decide is whether the data or noise. The process of reduce noise in green channel is performed similarly.

Listing 5 Process to reduce noise in red channel

Last step is to display all processed data on plot to compare and analyse them. Responsible for this is code shown on listing 6. Presented code is used to set limit of axis, colours of bar, method display data etc.

```
figure()
h=bar(q,prob,1,'b','EdgeColor','k');
set(gca, 'FontSize', 18, 'LineWidth', 3)
set(h, 'FaceColor', [0.75 0.75 0])
xlim([-0.20 1.25]) %set limit on X axis
hold on
[fo,gof]=fit(g',prob','gauss2','startpoint',[0.025, 0.25, 0.2]
0.021, 1, 0.4])
hold on
plot(-0.2:0.01:1.3, fo(-0.2:0.01:1.3)', 'color', [0 0 0],
'linewidth',2.5)
xlabel('FRET','fontweight','b','fontsize',22)
ylabel('Fraction','fontweight','b','fontsize',22)
title('63/38+2/15nt DNA', 'fontweight', 'b', 'fontsize', 16)
RatioOfFretAreas=sum(fo(-0.2:0.005:0.68))/
sum(fo(0.68:0.005:1.2))
```

Listing 6 Chart of analysed data

During this stage the results of a single molecule was saved. We observed if there occurred the binding of RT with DNA construct (Fig. 8). The selection of molecules was made manually. The DNA construct was labelled with Cy-5. The red colour was showing the signal emitted by excited with the laser dye Cy-5. The green fluorescence comes from the reverse transcriptase which was labelled with Cv-3. There is a binding event if on the analysed image is a significant (distinctive from the background) increase of the red signal or increase of both red and green signals simultaneously. If there is increase of the green fluorescence without red signal, there is no binding event – these signals should be removed with the help of proper scripts in MATLAB.[3] On this stage it is important to remove the background "noise" and false results, which were generated by RT that was not tied with the DNA. After removing of all noises, there are only binding events on the screen. Then the results were subjected to statistical analysis (Gaussian distribution). Thanks to the results it is possible to learn e.g.: how many binding events is in a single molecule during the checked time, how long they last and it is possible to compare the activity of RT in different conditions by changing the reaction mix (for example the concentration of sodium chloride (NaCl)).

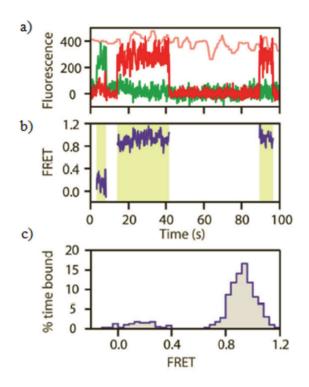


Fig. 8. The results of analysis. a) Increase of red signal shows a binding of RT and DNA. b) Noises were removed from the background, binding time and FRET values are shown. c) FRET distribution histogram [3]

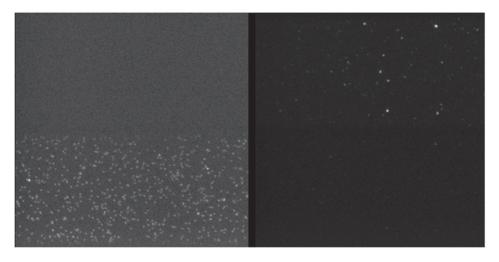


Fig. 9. Image obtained from the laser microscope. On the left side is shown result of red laser irradiation (DNA construct). On the right side is shown result of green laser irradiation (reverse transcriptase particles)

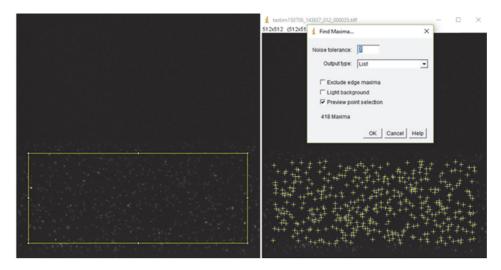


Fig. 10. ImageJ analysis. On the left side is shown clipping area for further analysis. On the right side is shown creating list of molecule coordinates

3. RESULTS

In this experiment we obtained information about number and time of binding RT to DNA construct. This data can be used to calculate chemical parameter value, such as dissociation constant (K_d) or statistic values such as standard deviation (e.g. from K_d) and Gaussian distribution (e.g. of number of binding).

4. CONCLUSIONS

Understanding of reverse transcriptase is a crucial way to discover a new way to treat HIV infections. The mortality of HIV infection is very high and the ways to treat that kind of infection are not good enough. Described method of analysis works for testing the activity of reverse transcriptase in different DNA constructs. It can be helpful in finding new ways to stop RT activity. To carry out the analysis in a way described in this paper, it is necessary to be familiar with LabView and MATLAB software. Properly written scripts allow performing correct analysis, obtaining trustful results and process automation. Reproducible results are necessary to draw appropriate conclusions. One of the ways it can be accomplished is the common work of biologists and IT specialists. To find new ways for successful analysis it is necessary to combine those two science disciplines. Merger of biology and computer sciences, so called bioinformatics, is the future of nature sciences and medicine.

BIBLIOGRAPHY

- Hu W.S., Hughes S.H., 2012. HIV-1 reverse transcription, Cold Spring Harb Perspect Med. 2, 1-22.
- [2] Johnson G.T., Goodsell D.S., Autin L., Forli S., Sanner M.F., Olson A.J., 2014. 3D molecular models of whole HIV-1 virions generated with cellPACK., Faraday Discuss. 169, 23-44.
- [3] Abbondanzieri E.A., Bokinsky G., Rausch J.W., Zhang J.X., Le Grice F.J., Zhuang X., 2009. Dynamic binding orientations direct activity of HIV reverse transcriptase, Cancer. 453, 184-189.
- [4] Stone D.M., Mihalusova M., O'Connor M.C., Prathapam R., Collins K., Zhuang X., 2007. Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein, Nature. 446, 458-461.
- [5] Murray J.M., Kelleher A.D., Cooper D.A., 2011. Timing of the components of the HIV life cycle in productively infected CD4+ T cells in a population of HIVinfected individuals., J. Virol. 85, 10798-10805.
- [6] WHO HIV Department, Global summary of the AIDS epidemic, (2014). http://www.who.int/hiv/data/epi_core_july2015.png?ua=1.
- [7] Das K., Arnold E., 2013. HIV-1 reverse transcriptase and antiviral drug resistance. Part 1., Curr. Opin. Virol. 3 (2013) 111-118.
- [8] Das K., Arnold E. HIV-1 reverse transcriptase and antiviral drug resistance. Part 2., Curr. Opin. Virol. 3, 119-128.
- [9] Myong S., Bruno M.M., Pyle M.A., Ha T., 2007. pring-Loaded Mechanism of DNA Unwinding by Heptitis C Virus NS3 Helicase, Science. 317, 513-516.
- [10] Kupfer S.S., Torres J.B., Hooker S., Anderson J.R., Skol A.D., Ellis N.A., et al., 2009. Novel single nucleotide polymorphism associations with colorectal cancer on chromosome 8q24 in African and European Americans., Carcinogenesis. 30, 1353-1357.
- [11] Bio-Imaging Unit, FRET, (n.d.) http://www.ncl.ac.uk/bioimaging/techniques/fret/.
- [12] Diamandis E.P., Christopoulos T.K., 1991. The biotin-(strept)avidin system: Principles and applications in biotechnology, Clin. Chem. 37, 625-636.
- [13] Frampas E., Rousseau C., Bodet-Milin C., Barbet J., Chatal J.-F., Kraeber-Bodéré F., 2013. Improvement of radioimmunotherapy using pretargeting., Front. Oncol. 3, 159.
- [14] ImageJ. http://imagej.nih.gov/ij/.

CZY MATLAB MOŻE POMÓC WYLECZYĆ HIV?

Streszczenie

Infekcja wywołana ludzkim wirusem niedoboru odporności (HIV) może prowadzić do zespołu nabytego niedoboru odporności (AIDS). Wirus HIV infekuje komórki układu odpornościowego – zwłaszcza te, które odpowiedzialne są za aktywację odpowiedzi immunologicznej. Każdego roku wiele osób umiera z powodu chorób, które w wypadku prawidłowego działania układu odpornościowego nie byłyby śmiertelne. Aktualnie naukowcy próbują opracować skuteczny lek, który pomoże leczyć infekcję wirusem HIV i będzie zapobiegać rozwojowi AIDS. Aby to osiągnąć konieczne jest jak najlepsze poznanie cząsteczki wirusa HIV i sposobu jego działania. Do dnia dzisiejszego odkryto wiele informacji o wirusie HIV, jednak wiele jego właściwości pozostaje nieznane. Jednym z niezbędnych enzymów wirusa HIV jest odwrotna transkryptaza (RT). Dzięki temu enzymowi wirus może przepisać swój materiał genetyczny z RNA na bardziej stabilne cDNA (ang. complementary DNA). Poznanie warunków, w których działa odwrotna transkryptaza pomoże zablokować jej aktywność. Dobrym sposobem na poznanie tego enzymu jest jego obserwacja pod mikroskopem laserowym. Mikroskop laserowy umożliwia obserwację pojedynczych cząstek. Możliwa staje się obserwacja reakcji RT z konstruktami DNA o różnej długości. Wyniki otrzymane z obserwacji pod mikroskopem mogą być analizowane za pomocą programu MATLAB. W tym celu konieczne jest napisanie odpowiednich skryptów, które pozwolą na dokładną analizę aktywności odwrotnej transkryptazy.

Słowa kluczowe: HIV, odwrotna transkryptaza, FRET, obliczenia w MATLABie