Applying the Bioinformatics Methods to Design and Evaluate the SapM-M13 pIX Fusion Protein and Its Theoretical Role in the Phage ELISA System

Mozafar Mohammadi[1, 2], Peyman Bemani[2*], Neda Zarei[3]

Abstract

Phage ELISA is a common method used to confirm binding of obtained phages from phage display technique to related antigens. Enzyme-conjugated antibody directed against the major capsid protein (pVIII) or enzyme-conjugated secondary antibody against the primary antibody is used as a detection system in phage ELISA. We suggested expression of the secreted acid phosphatase (SapM) enzyme on M13 pIX minor coat protein directly, and evaluated this hypothesis using *In Silico* techniques. 3D structure model of the fusion protein (SapM+M13 pIX) was generated and evaluated by related software. MD simulation and TMHMM program results showed a stable fusion protein which is anchored to the inner membrane of *E. coli* by membrane spanning region suggesting a proper assembling on M13 phage. In theory, SapM enzyme on the phage surface can catalyze the p-nitrophenyl phosphate as substrate and creates yellow color which can be measured at OD=405 nm by microtiter plate reader. We believe that decreasing the antibody layers in phage ELISA will significantly increase the reliability and reproducibility of the test and reduce its time.

1. *Applied Biotechnology Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran*
2. *Recombinant antibody laboratory, Department of Immunology, Shiraz University of Medical Sciences, Shiraz, Iran*
3. *Department of Biotechnology, School of Veterinary Medicine, Shiraz University, Shiraz, Iran*

*Corresponding Author
Peyman Bemani
Recombinant antibody laboratory, Department of immunology, Shiraz University of Medical Sciences, Shiraz, Iran
E-mail: peymanbemani@gmail.com

Introduction

Phage display technology is a powerful method which allows selection of exogenous (poly) peptides. This technology has been used for different aspects such as selection of monoclonal antibodies against a given antigen, affinity maturation of monoclonal antibodies and identification of tumor-associated antigens [1, 2]. *Escherichia coli* filamentous bacteriophages (f1, fd, M13) are commonly used phages for phage display [3]. Each M13 virion contains 2,700 copies of pVIII protein (as major coat protein), plus 3 to 5 copies of minor proteins at the filament ends, including pVI and pIII at one end and pVII and pIX at the other end of filamentous phage [4-6]. Screening of the library for specific antibodies is performed through a process called panning during which phages expressing the desired antibody on their surfaces are selected against a coated antigen [7].

To evaluate the efficacy of the panning selection process and also determining binding ability of a single clone binder phage derived from each round of panning, the resultant (polyclonal) phage pools are tested by phage ELISA. Briefly, during this process the polyclonal phage is incubated in a target molecule-coated plates [7, 8], then primary unconjugated antibody against pVIII and enzyme-conjugated secondary antibody against primary antibody or enzyme-conjugated antibody directed against the major capsid protein (pVIII) are added [9-11]. Phage-ELISA is a lengthy, time-consuming and laborious process which requires several steps of incubations and washing [12, 13].

The necessity of using one or two antibody as detecting reagent (HRP-conjugated antibody or primary unconjugated antibody and secondary HRP-conjugated antibody) results in increasing the number of layers (coated target molecule, phage displaying binder, and antibodies, respectively). In order to simplify and accelerate the rate of screening process and reduce number of reagent layers in phage ELISA test, we have suggested expressing an acid phosphatase (SapM) enzyme on M13 phage pIX minor coat protein. All steps of our hypothesis were performed by bioinformatics techniques.

Materials and Methods

Retrieval of nucleotide and amino acid sequences

Nucleotide sequences of secreted acid phosphatase of *Mycobacterium tuberculosis H37Rv* (SapM) (Gene ID: 887988) and M13KO7 pIX (Gene ID:14964) genes and related amino acids were retrieved from NCBI [14] protein reference sequences (Protein RefSeq) in FASTA format.

Construction of fusion protein sequence

Based on the Pfam server [15] information, among 299 amino acids (aa), 213 aa related to the functional domain of SapM enzyme (amino acids from 70 to 282) which had phosphoesterase activity were kept. A flexible linker containing 15 amino acids (GGGGS)3 [16] was added between the truncated SapM and M13 pIX subunit.

Prediction of fusion protein 3D structure

Tertiary structure of truncated SapM enzyme was generated by pGenTHREADER program (Profile Based Fold

Recognition) [17] on PSIPRED server [18]. pGen-THREADER generates models by transferring coordinates from aligned regions only with no loop modeling [17]. Afterward, assembling the final fusion protein construct including truncated SapM, linker and M13 pIX segments was performed by Modeller 9.11 software [19]. Modeller applies an automated strategy for comparative or homology protein structure modeling by satisfaction of spatial restraints. Also, Chimera 1.10.2 [20] was implemented for visualizing the tertiary structures of proteins.

Tertiary structure validation
VADAR (Volume Area Dihedral Angle Reporter) is a comprehensive web server for quantitative protein structure evaluation [21]. It was used to evaluate the protein 3D models. Ramachandran plot was derived from VADAR program for structure validation.

Molecular dynamics simulation
Entire modeled protein was subjected to MD simulation studies. For this purpose, modeled PDB structure was embedded in a box with dimensions equal to 1 nm from the edges of the molecule. Subsequently, the system was solvated with spc216 water model. The structure was relaxed through the energy minimization (EM) process. The main run of MD was finally performed after removing all restrains from the system. GROMOS force field implemented in Gromacs 4.5.3. Calculation of RMSD and RMSF plots were performed using the commands implemented in Gromacs.

Cellular subsorting of fusion protein
To investigate the bacterial subcellular localization of fusion protein (SapM + linker + pIX) the TMHMM program was utilized [22]. Also, SACS server [23] was applied to predict the transmembrane segments of protein and generate a TOPO2 image of the prediction.

Results

Construction of fusion protein sequence
Pfam is a database for protein family features including their annotations, domains and multiple sequence alignments generated using hidden Markov models [24]. According to its data, residues from 70 to 282 were related to phosphoesterase family. These enzyme family catalyze the hydrolysis of phosphoric monoesters and release inorganic phosphate [25]. Based on this information, 69 amino acids from N-terminal and 16 amino acids from C-terminal were deleted. The rest of the protein (213 amino acids), which contained phosphoesterase activity, was retained.

For decreasing the unfavorable interactions between two protein domains, a 15-amino acid flexible linker (GGGGS)3 [16] was added between truncated SapM (C-terminal) and M13 pIX subunit (N-terminal) (Table 1). Flexible linkers are usually applied when joined domains require a certain degree of mobility or interaction. They mainly contain small, non-polar (e.g. Gly) or polar (e.g. Ser or Thr) amino acids [26].

Prediction of fusion protein 3D structure
The tertiary structure model of the fusion protein was generated by pGenTHREADER program (Profile Based Fold Recognition) [17] and Modeller 9.11 software [19].

Modeller is stand-alone software used for homology or comparative modeling of protein three-dimensional structures [27].

Table 1. Amino acid sequence of different domains of fusion protein.

SapM Sequence
KSAPFINSLAANGAMMAQAFAETHPSEPNYLALFAGN TFGLTKNTCPVNGGALPNLGSELLSAGYTFMGFAEDL PAVGSTVCSAGKYARKHVPWVNFSNVPTTLSVPFSAF PKPQNYPGLPTVSFVIPNADNDMHDGSIAQGDAWLNR HLSAYANWAKTNNSLLVVTWDEDDGSSRNQIPTVFY GAHVRPGTYNETISHYNVLSTLEQIYGLP

Linker	GGGGSGGGGSGGGGS
M13 pIX	SVLVYSFASFVLGWCLRSGITYFTRLMETSS

Modeling was performed in the intensive mode of homology modeling. *Francisella tularensis* acid phosphatase A (PDB code: 2D1G) and the structure of fd filamentous bacteriophage coat protein (PDB code: 1NH4) were used as template to create the 3D model of the fusion protein (Fig. 1).

Evaluation of predicted structure
Obtained data from the Ramachandran plot on VADAR program indicated that 94% of amino acids are located in the low energy areas and they had proper dihedral angles. According to their φ and ψ torsion angles 82% of residues were in the most favored (phi psi core) regions and 12% of residues were in additional allowed (phi psi allowed) regions. Moreover, free folding energy of tertiary structure of this fusion protein was -228.03 kcal/mol. Since expected free energy deduced for well-structured proteins is -240.39 kcal/mol, this calculated value confirmed the accuracy of this structure (Fig. 2).
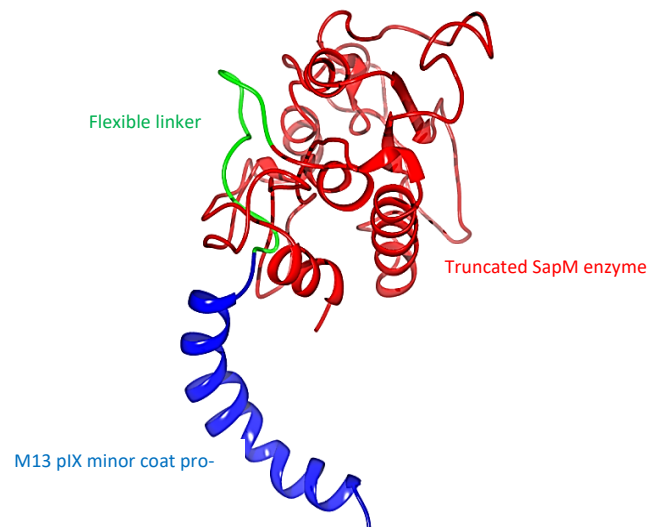

Figure 1. Flatted ribbon style representation of predicted fusion protein's 3D structure. Red, green and blue flatted ribbons relate to the truncated SapM enzyme, flexible linker and M13 PIX minor coat protein, respectively.
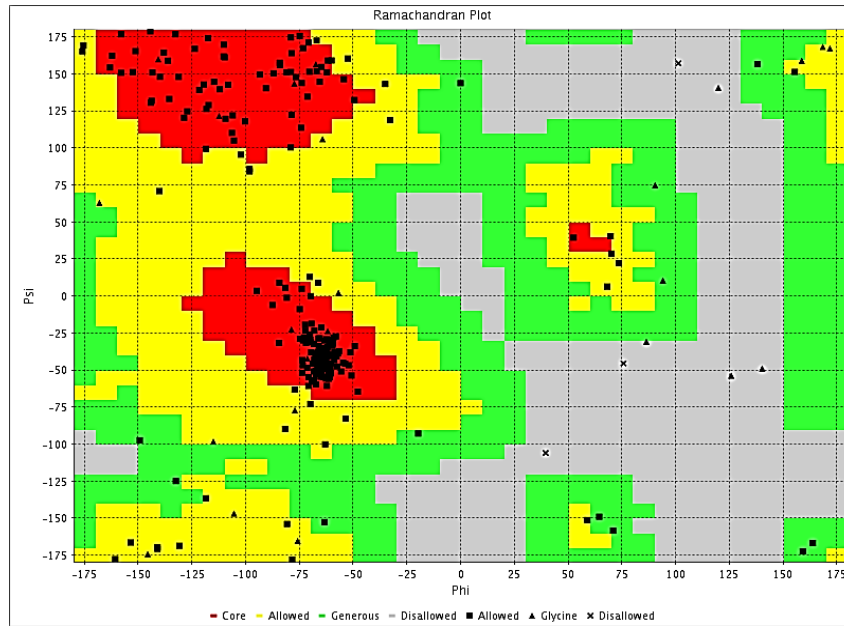
Figure 2. Qualification of fusion protein three dimensional model using Ramachandran plot derived from VADAR program. The Ramachandran plots of the model indicate that 82% of residues are located in phi psi core regions, 12% in phi psi allowed regions, 3% in generously allowed regions and 1% in disallowed regions of the plot. Core and the allowed regions are demonstrated in red and yellow, respectively. The generously allowed regions are indicated in green, and the disallowed regions are in gray.

Molecular dynamics simulation

Molecular Dynamics (MD) simulation, as a function of time can provide the ultimate detail concerning individual atom motion [28]. Root mean square deviation (RMSD) was measured for finding the deviations and changes of the protein structure over the simulation period.

As observed in Fig. 3A, there was a minimum change (RMSD<1Å) from the 2th nanosecond until end of simulation course. These deviation values indicated that the interest protein has had a suitable behavior during the simulation. In addition, to investigate the more local changes, it is measured the root mean square fluctuation (RMSF). According to the RMSF plot, the C-terminal region including 12 amino acids from 248 to 259 of pIX subunit fluctuated much more than any other part of the recombinant protein. As it is seen in Fig. 3B, this region (248 to 259) located in the cytoplasmic space of the bacterial cell and it doesn't have any roles during the process of phage assembly in the bacterial host.

Cellular subsorting of fusion protein

During the process of phage assembly, the pVII and pIX are located at the same end of the phage particle that emerges first from the membrane of Gram-negative bacteria [29]. In the Gram-negative organisms (such as *E. coli*) possible subcellular localization sites of proteins are cytoplasm, cytoplasmic membrane, periplasm, outer membrane and extracellular space. TMHMM program was used to investigate the bacterial subcellular localization of the fusion protein. It is a suitable method for predicting transmembrane helices based on hidden Markov model and developed by Anders Krogh and Erik Sonnhammer [22].

As shown in Fig. 4, the TMHMM program predicts the transmembrane helices of single sequences with a high level of reliability and accuracy.

As observed in Fig. 5, amino acids from 225 to 244 participate in the membrane spanning region. Four residues belonging to the flexible linker including Gly225, Gly226, Gly227 and Ser228, and sixteen residues belonging to pIX protein including Ser229, Val230, Lue231, Val232, Thy233, Ser234, Phe235, Ala236, Ser237, Phe238, Val239, Lue240, Gly241, Trp242, Cys243 and Lue244 are present in transmembrane region. Presence of more hydrophobic amino acids than polar and non-polar residues has an important role in creating the helical transmembrane region. Accordingly, the fusion protein is anchored to the inner membrane by membrane spanning region and sapM enzyme is located in periplasmic area while the end sixteen residues of pIX are placed in cytoplasmic space.

Suggested ELISA procedure

SapM enzyme has shown 100% activity for p-nitrophenyl phosphate (p-NPP), a-Naphthyl phosphate and phosphoenolpyruvate [30]. Typically, p-NPP is used as substrate for acid phosphatase assays which absorbs light at 405 nm[31]. During acid phosphatase assessment in our improved phage ELISA, SapM-conjugated phages bind to corresponding antigen in the 96-well plate, and then in the presence of p-NPP, SapM enzyme produces a yellow color due to the releasing of p-NP molecules; since p-nitrophenol product absorbs light at 405 nm and absorbance at this wave-length is monitored as a measure of specific bound phages (Fig.6)
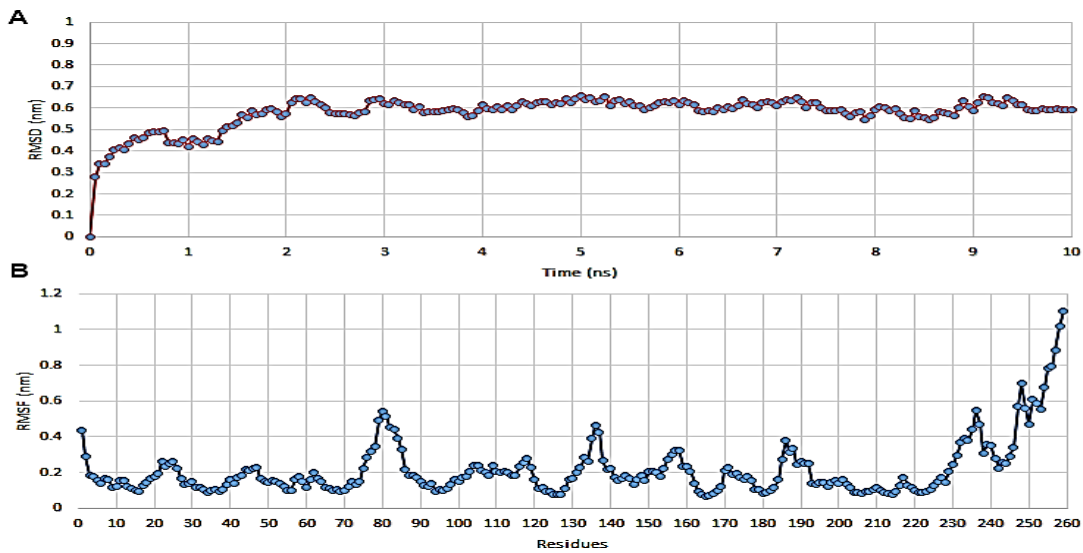
Figure 3. Trajectory plots obtained from MD Simulation, A: RMSD vs time, B: RMSF vs time.
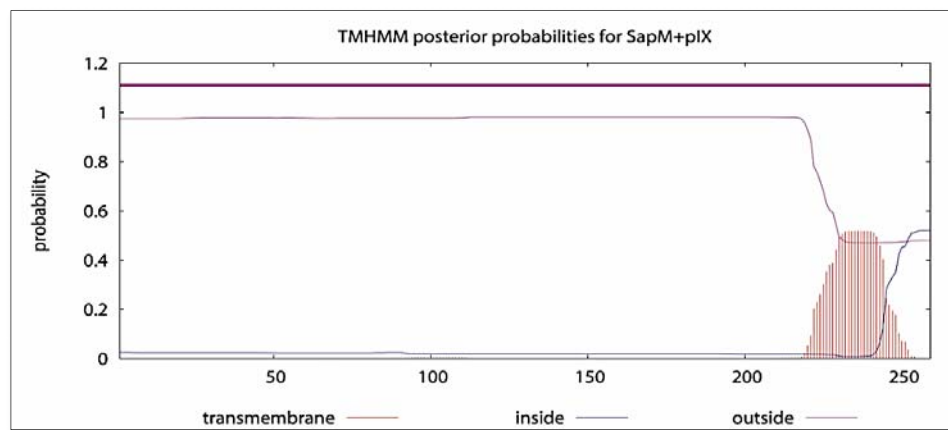

Figure 4. Probability prediction of potential transmembrane regions in the fusion protein (SapM+linker+ pIX). Vertical red lines represent transmembrane region.

Discussion

Phage-ELISA is a lengthy, time-consuming and laborious process which requires several steps of incubations and washing. The necessity of using antibodies as detecting reagents (HRP-conjugated antibody or primary unconjugated antibody and secondary HRP-conjugated antibody) result in increasing the number of layers, thus reducing test reliability. We proposed a modified method of phage ELISA system without using the primary and secondary antibody. Thereby, it is suggested to directly express an acid phosphatase (SapM) enzyme on the M13 pIX minor coat protein.

Theoretically, all phage coat proteins pIII, pVI, pVII, pVIII and pIX, are able to carry foreign proteins on their N-terminus. But with increasing the copies of coat proteins such aspVIII subunit, the length of displayed proteins would be decreased up to 12 residues [11]. Minor coat protein PIII with 3-5 copy, is traditionally and successfully used for displaying large proteins such as Fab and scFv antibody fragments. On the other hand, several studies showed that by accommodating a proper signal sequence between the N-terminal of pIX and the desired protein, it becomes able to develop a stable protein display system [32-37]. Gao *et al*., constructed optimized vector pCGMT-1b in which pIX was used as anchor protein [29]. Also Nilssen and co-workers developed a modified helper phage termed Delta Phage that allows for high-valence protein display on M13 pIX subunit [38]. Although, previous studies showed the ability of pIX minor coat protein for displaying of foreign protein, but none of them proposed this ability to display a proper enzyme to employ in the phage ELISA system.
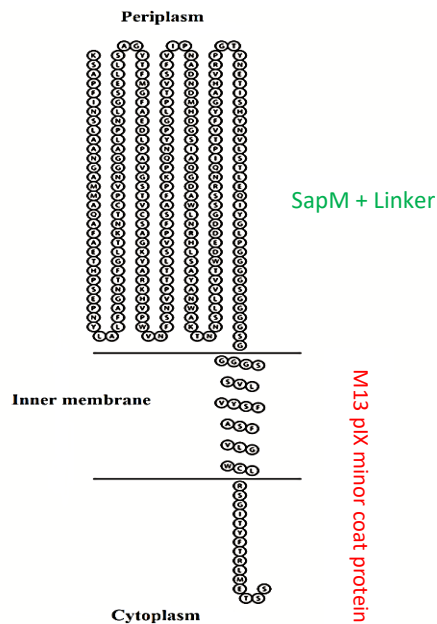
Figure 5. TOPO2 image of predicted transmembrane segments in fusion protein. Amino acid are depicted as circles.
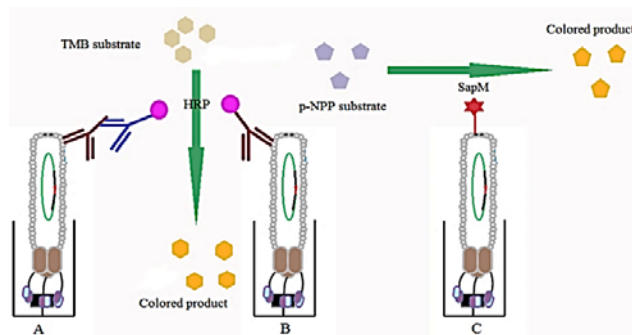

Figure 6. Schematic representation of conventional phage ELISA (A&B) and SapM enzyme-conjugated phage ELISA (C). Three (A) or two (B) layers use in HRP conjugated antibody phage ELISA. While, there is one (C) layer in SapM enzyme conjugated phage ELISA.

In this study a proper tertiary structure of the fusion protein (SapM+pIX) was provided and evaluated using related programs. Also, structural stability of the fusion protein verified using Molecular Dynamic simulation. TMHMM program results showed a stable fusion protein which is anchored to the inner membrane of *E. coli* bacteria by membrane spanning region very well. Theoretically, expressed enzyme on the phage surface, could catalyze the p-nitrophenyl phosphate as substrate and yellow-color product could be measured at OD=405 nm by microtiter plate reader.

Conclusion
In conclusion, we believe that applying this suggested approach can decrease the macromolecular layers in the phage ELISA, and will significantly increase the reliability and reproducibility of the test and reduce the required time.

References
1. Pavoni, E., Vaccaro, P., Pucci, A., Monteriu, G., Beghetto, E., Barca, S., Dupuis, M.L., De Pasquale Ceratti, A., Lugini, A., Cianfriglia, M., Cortesi, E., Felici, F., Minenkova, O., Identification of a panel of tumor-associated antigens from breast carcinoma cell lines, solid tumors and testis cDNA libraries displayed on lambda phage. *BMC Cancer*, 2004, Vol. 4, pp. 78-85.
2. Minenkova, O., Pucci, A., Pavoni, E., De Tomassi, A., Fortugno, P., Gargano, N., Cianfriglia, M., Barca, S., De Placido, S., Martignetti, A., Felici, F., Cortese, R., Monaci, P., Identification of tumor-associated antigens by screening phage-displayed human cDNA libraries with sera from tumor patients. *Int J Cancer*, 2003, Vol. 106, pp. 534-44.
3. Bazan, J., Całkosiński, I., Gamian, A., Phage display—A powerful technique for immunotherapy: 2. Vaccine delivery. *Hum vaccin immunother*, 2012, Vol. 8, pp. 1829-1835.
4. Dantas-Barbosa, C., Brigido, M.D., Maranhao, A.Q., Antibody phage display libraries: contributions to oncology. *Int J Mol Sci*, 2012, vol. 13, pp. 5420-5440.
5. Huang, J.X., Bishop-Hurley, S.L., Cooper, M.A., Development of Anti-Infectives Using Phage Display: Biological agents against Bacteria, Viruses, and Parasites. *Antimicrob Agent Chemother*, 2012, Vol. 56, pp. 4569-4582.
6. Arap, M.A., Phage display technology applications and innovations. *Genet Mol Biol*, 2005, Vol. 28, pp. 1-9.
7. Hammers, C.M., Stanley, J.R., Antibody phage display: technique and applications. *J Invest Dermatol*, 2014, Vol. 134, pp. 1-13.
8. Lee, C.M.Y., Iorno, N., Sierro, F., Christ, D., Selection of human antibody fragments by phage display. *Nat Protoc*, 2007, Vol. 2, pp. 3001-3008.
9. Wu, D., Li, G., Qin, C., Ren, X., Phage displayed peptides to avian H5N1 virus distinguished the virus from other viruses. *PLoS One*, 2011, Vol. 6, pp. e23058.
10. Hu, D., Hu, S., Wan, W., Xu, M., Du, R., Zhao, W., Gao, X., Liu, J., Liu, H., Hong, J., Effective optimization of antibody Affinity by phage display integrated with high-throughput DNA Synthesis and Sequencing Technologies. *PLoS One*, 2015, Vol. 10, pp. e0129125.
11. Qi, H., Lu, H., Qiu, H.J., Petrenko, V., Liu, A., Phagemid vectors for phage display: properties, characteristics and construction. *J Mol Biol*, 2012, Vol. 417, pp. 129-43.
12. Rajaram, K., Losada-Pérez, P., Vermeeren, V., Hosseinkhani, B., Wagner, P., Somers, V., Michiels, L., Real-time analysis of dual-display phage immobilization and autoantibody screening using quartz crystal microbalance with dissipation monitoring. *Int J Nanomed*, 2015, Vol. 10, pp. 5237-5242.
13. Shukla, G.S., Krag, D.N., Phage-displayed combinatorial peptide libraries in fusion to β-lactamase as reporter for an accelerated clone screening: Potential uses of selected enzyme-linked affinity reagents in downstream applications. *Comb Chem High Throughput Screen*, 2010, Vol. 13, pp. 75-81.
14. http://www.ncbi.nlm.nih.gov/gene.
15. http://pfam.xfam.org.
16. Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M.S., Novotny, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E., Crea, R., Protein engineering of antibody binding sites:

recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. *Proc Natl Acad Sci USA*, 1988, Vol. 85, pp. 5879-5883.

17. Lobley, A., Sadowski, M.I., Jones, D.T., pGenTHREADER and pDomTHREADER: new methods for improved protein fold recognition and superfamily discrimination. *Bioinformatics*, 2009, Vol. 25, pp. 1761-1767.

18. Buchan, D.W., Minneci, F., Nugent, T.C., Bryson, K., Jones, D.T., Scalable web services for the PSIPRED protein analysis work bench. *Nucleic Acids Res*, 2013, Vol. 41, pp. 349-357.

19. Sali, A., Blundell, T.L., Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*, 1993, Vol. 234, pp. 779-815.

20. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, 2004, Vol. 25, pp. 1605-1612.

21. Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R.F., Sykes, B.D., Wishart, D.S., VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res*, 2003, Vol. 31, pp. 3309-3316.

22. Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*, 2001, Vol. 305, pp. 567-580.

23. http://www.sacs.ucsf.edu/cgi-bin/hmmtop.py.

24. Finn, R.D., Tate, J., Mistry, J., Coggill, P.C., Sammut, S.J., Hotz, H.R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L., Bateman, A., The Pfam protein families database. *Nucleic Acids Res*, 2008, Vol. 36, pp. 281-288.

25. Henneberry, M.O., Engel, G., Grayhack, J.T., Acid phosphatase. *Urol Clin North Am*, 1979, Vol. 6, pp. 629-641.

26. Argos, P., An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *J Mol Biol*, 1990, Vol. 211, pp. 943-58.

27. Webb, B., Sali, A., Comparative protein structure modeling using Modeller. *Curr Protoc Protein Sci*, 2014, Chapter 2: Unit 2.9.

28. Caflisch, A., Paci, E., Molecular Dynamics Simulations to Study Protein Folding and Unfolding, in Protein Fold ing Handbook, J. Buchner and T. Kiefhaber, Editors. 2005, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

29. Gao, C., Mao, S., Kaufmann, G., Wirsching, P., Lerner, R.A., Janda, K.D., A method for the generation of combinatorial antibody libraries using pIX phage display. *Proc Natl Acad Sci USA*, 2002, Vol. 99, pp. 12612-6.

30. Saleh, M.T., Belisle, J.T., Secretion of an acid phosphatase (SapM) by Mycobacterium tuberculosis that is similar to eukaryotic acid phosphatases. J Bacteriol, 2000, vol. 182, pp. 6850-3.

31. Yang, T.T., Sinai, P., Kain, S.R., An acid phosphatase assay for quantifying the growth of adherent and nonadherent cells. *Anal Biochem*, 1996, Vol. 241, pp. 103-108.

32. Kaushik Rajaram, Veronique Vermeeren, Klaartje Somers, Veerle Somers, Michiels, L., Construction of helper plasmid-mediated dual-display phage for autoantibody screening in serum. Appl Microbiol Biotechnol, 2014, Vol. 98, pp. 6365–6373.

33. Georgieva. Y, Konthur. Z, Design and screening of M13 phage display cDNA libraries. *Molecules*, 2012, Vol. 16, pp. 1667–1681.

34. Loset GA, Sandlie I, Next generation phage display by use of pVII and pIX as display scaffolds. *Methods*, 2012, Vol. 58, pp. 40–46.

35. Ploss, M., Kuhn, A., Membrane insertion and assembly of epitope-tagged gp9 at the tip of the M13 phage. *BMC Microbiol*, 2011, Vol. 11, pp. 211.

36. Shi, L., Wheeler, J.C., Sweet, R.W., Lu, J., Luo, J., Tornetta, M., Whitaker, B., Reddy, R., Brittingham, R., Borozdina, L., Chen, Q., Amegadzie, B., Knight, D.M., Almagro, J.C., Tsui, P., De novo selection of high-affinity antibodies from synthetic fab libraries displayed on phage as pIX fusion proteins. *J Mol Biol*, 2010, Vol. 397, pp. 385-96.

37. Tornetta, M., Baker, S., Whitaker, B., Lu, J., Chen, Q., Pisors, E., Shi, L., Luo, J., Sweet, R., Tsui, P., Antibody Fab display and selection through fusion to the pIX coat protein of filamentous phage. *J Immunol Methods*, 2010, Vol. 360, pp. 39-46.

38. Nicolay R. Nilssen, Terje Frigstad, Sylvie Pollmann, Norbert Roos, Bjarne Bogen, Inger Sandlie, Løset, G.A., Delta Phage a novel helper phage for high-valence pIX phagemid display. *Nucleic Acids Res*, 2012, Vol. 40, pp. e120.