Current Synthetic and Systems Biology
ISSN: 2332-0737

Joshi et al., Curr Synthetic Sys Biol 2013, 1:1
DOI: 10.4172/2332-0737.1000103

Research Article | Open Access

Discovering Novel Anti-Malarial Peptides from the Not-coding Genome - A Working Hypothesis

Mukta Joshi[1], Shankar V Kundapura[1], Thirtha Poovaiah[1], Kundan Ingle[2] and Pawan K Dhar[3]*

[1]*Symbiosis School of Biomedical Sciences, Symbiosis International University, Pune, India*
[2]*VLife Sciences Technologies Pvt. Ltd., Pune, India*
[3]*Centre for Systems and Synthetic Biology, University of Kerala, Kerala, India*

Abstract

The not-coding regions of the genome describe sequences that do not have a history of transcription. They are also called 'dark matter' of the genome. Here we present a working hypothesis for finding novel anti-malarial peptides from such regions of the yeast genome that encode neither RNA nor protein. This is based on our previous experimental work where not-coding DNA sequences were artificially expressed leading to protein expression and phenotypic outcome. In this study, we explored the vast not-coding DNA space of Saccharomyces cerevisiae in search of novel antimalarial peptides. Given the lack of effective therapeutic solutions against malaria, there is an urgent unmet requirement to find novel antimalarial drugs and targets. Our initial efforts to find novel anti-malarial peptides have led to unexpected and interesting results. However, our work is preliminary and is based on computational studies only. In future, more computational and experimental work is needed to establish therapeutic potential of synthetic peptides that have origins in the not-coding genome space.

Keywords: Not-coding DNA; Synthetic peptides; Malaria; AMA 1; EBA-175; MSP-1(19)

Abbreviations: AMA-1: Apical Membrane Antigen-1; EBA-175: Erythrocyte Binding Antigen-175; MSP-1(19): Merozoite Surface Antigen 1(19)

Introduction

Malaria is one of the most life-threatening diseases that results in a global loss of over 2.7 million lives every year, with African continent representing bulk of the reported malarial cases [1]. Some of the key reasons for the difficulty in controlling malaria are: (a) rapid emergence of Plasmodium strains resistant to existing antimalarial therapeutics, (b) failure in effectively implementing vector control programs and (c) lack of an effective malaria vaccine. Recent studies indicate that number of malaria cases would double in 20 years if novel methods of control are not developed [2].

For the past 30 years, there has been a steady escalation in antimalarial drug resistance indicating ineffectiveness of the existing antimalarial therapeutic solutions [3]. Furthermore, the pipeline for discovering novel and affordable antimalarial drugs has almost dried up and some of the highly effective antimalarial drugs like artemisinin are prohibitively expensive. Chloroquine, one of the first successful antimalarial drugs was launched in 1940. However, its drug resistance has been observed in every region where Plasmodium falciparum infection occurs [4]. Many countries have adopted sulfadoxine/pyrimethamine therapy to fight the wide spread resistance of malarial parasite to chloroquine. However, there have been reports of resistance to sulfadoxine / pyrimethamine therapy from Southeast Asia, South America and Africa [4]. The situation is increasingly bleak with the resistance emerging to atovaquone within a year of its launch [4]. Thus, there is an urgent need for novel drugs for effectively treating malaria [5].

Currently, artemisinin is the considered as the drug of choice to combat malaria. The World Health Organisation (WHO) has stressed that artemisinin must be combined with other drugs that have different mechanisms of action and longer half-lives. Decreased sensitivity to artemisinin monotherapy, coupled with the rise of resistance of parasites to all partner drugs, threaten to place millions of patients at risk of inadequate treatment of malaria [6]. With very few antimalarial drugs in the pipeline, there is an urgent unmet need of a novel, efficacious and cost effective antimalarial therapy.

The key question remains: where will novel drug molecules come from? Given that the traditional drug discovery process has not led to significant breakthroughs in the recent times, it is important to develop new strategies in search of novel antimalarial molecules.

In a recent study [7] we showed that not-coding DNA is an untapped goldmine of functional peptides and proteins. As against the non-coding DNA, which describes RNA coding sequences, the 'not-coding' DNA refers to sequences that encode neither protein nor RNA i.e., they do not have any evidence of expression in their natural settings (Figure 1). This region is also sometimes referred to as the "dark matter of the genome" as it is functionally uncharacterised, vast and largely unexplored.

Methods and Preliminary Results

In 2009, we experimentally demonstrated for the first time, an ability to make user defined genes by using an artificial gene expression system [7]. Using not-coding genomic template of Escherichia coli, six unique intergenic regions were randomly selected, cloned using pBAD vector and expressed in the same host. The protein expression was verified using western blot. Of six proteins artificially expressed from not-coding regions of Escherichia coli, one showed significant growth inhibitory effects. By switching off the expression of this synthetic

***Corresponding author:** Pawan K Dhar, Director, Centre for Systems and Synthetic Biology, University of Kerala, Trivandrum, Kerala, India 695 581, E-mail: pawan@cssb.res.in

Received June 18, 2013; **Accepted** August 17, 2013; **Published** August 21, 2013

Citation: Joshi M, Shankar VK, Poovaiah T, Ingle K, Dhar PK (2013) Discovering Novel Anti-Malarial Peptides from the Not-coding Genome - A Working Hypothesis. Curr Synthetic Sys Biol 1: 103. doi: 10.4172/2332-0737.1000103

Copyright: © 2013 Joshi M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

gene, the cell growth was restored thereby showing a linear correlation between artificial gene expression and the observed phenotype.

After successfully providing the proof-of-the-concept, the next challenge was to find if the artificial conversion of not-coding DNA into coding DNA is scalable? Further, what are the best case scenarios and boundary conditions?

To address these questions, in this study we use malaria as an example. Our hypothesis is that given a sufficiently large space of not-coding DNA sequences, one can find novel peptides and proteins. We call such synthetic sequences, which originate from naturally non-expressed DNA and have potential to encode functional peptides as synpeps (synthetic peptides).

To explore an application of synpep approach, we used malaria as a test case and considered several key interaction points which, when blocked, could result in the failure of parasitic infection in humans. One such intervention point is the invasion of host erythrocytes by Plasmodium falciparum merozoites [8]. Merozoites specifically target and invade erythrocytes, as the erythrocytic haemoglobin is a rich nutritional source for these parasites. If merozoites are denied entry into erythrocytes, they perish due to starvation. Thus, delaying entry of merozoites into the erythrocytes retains them within the bloodstream for a longer period of time. This makes merozoites vulnerable to immune response from the host.

The entry of merozoite into host erythrocytes is facilitated by binding of a specific Plasmodium falciparum surface antigen to its corresponding receptor present on host erythrocytes. The parasite possesses different types of surface antigens that mediate erythrocyte invasion. The sialic acid dependent pathway, sialic acid independent pathway and tight junction formation interaction are some of the well-known invasion pathways. Malarial parasites employ any one of these invasion mechanisms to infect the host cell. Hence, it seems that a novel synthetic peptide targeted at disrupting erythrocyte invasion cycle would be a reasonable approach to block routes of parasite invasion.

Based on the published literature, we selected three Plasmodium falciparum membrane targets to design novel peptides.

1. Erythrocyte Binding Antigen-175 (EBA-175): The invasion of RBCs through sialic acid dependent pathway involves the parasitic EBA-175 interacting with human host receptor Glycophorin A present on the erythrocytes [9]. During this pathway, the dimerization of EBA-175 is considered to be an crucial step for the completion of the invasion [10]. The strategy is to hinder the dimerization of EBA-175, with the help of synpeps.

2. Merozoite Surface Protein 1 (MSP-1): The sialic acid independent pathway involves the interaction between the parasitic MSP-1 and band 3 anion transporter present on erythrocytes. As a requirement for merozoite entry into an erythrocyte, MSP-1 is synthesized as a 180–225-kDa polypeptide which undergoes two processing steps, the first at merozoite release from an infected cell and the second during invasion of an erythrocyte. At the end of the proteolytic cleavage steps, the fragment MSP-1(19) remains anchored to the merozoite membrane. MSP-1(19) interacts with its natural ligand, band 3 anion transporter and ensures erythrocyte invasion [11].

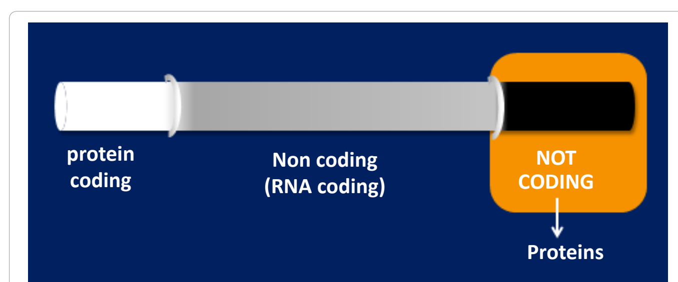

Figure 1: Representation of the genome based on functionality (not drawn to scale).

3. Apical Membrane Antigen 1 (AMA-1): AMA-1 is a surface protein present on Plasmodium falciparum. It is expressed at two stages in the life cycle of Plasmodium, sporozoite stage (invasion of hepatocyte) and merozoite stage (invasion of erythrocytes) [12]. The parasite injects a protein Plasmodium falciparum rhoptry neck protein-2 (PfRON-2) into membrane of erythrocyte. AMA-1 interacts with PfRON-2 to form the moving junction which aids in the process of invasion [13]. Mutation of the residues in the hydrophobic pocket of AMA1 results in inhibition of formation of AMA1-PfRON2 junction [14]. Some of the recent reports suggest that on blocking the moving junction formation, the parasite is not able to invade RBC, thus preventing the infection [15,16].

In this study, a library of synpeps was constructed upon translation of 1000 randomly selected not-coding DNA sequences of the yeast genome. The relevant synpeps were screened on the basis of sequence similarity with known ligands that bind to target proteins in their natural setting. This significantly reduced the number of candidate peptides to top nine candidate molecules i.e., three against each target. This number was further reduced to one lead peptide against each target on the basis of structural similarity with the naturally target binding ligands. Subsequently, the three best synpeps were docked with their respective targets to find preferred orientation of binding - important for stable complex formation. The not-coding parts of the yeast genome were identified from Saccharomyces Genome Database [17] via Yeastmine [18].

The not-coding sequences were computationally extracted, translated [19] and sequence matched with the regions of natural ligands that bind to the three targets selected [20]. A global sequence similarity of > 30% and gap less than 12% were considered for further studies (Table 1). The selected peptides were submitted to 3-D structure prediction softwares [21-26] which employ threading and ab-initio modelling methods. The predicted synpep structures were validated [27] for their structural correctness. The validated structures of the selected synpeps and natural ligands were superimposed the Root Mean Square Deviation (RMSD) was calculated [28]. The synpeps whose RMSDs with the natural ligands were less than 1 were chosen (Table 2) with an aim of finding peptides that structurally mimic the binding of natural ligand with the respective target and prevent further Plasmodium infection after entering the blood stream. The structures of the selected not-coding peptides after due validation and structural similarity with the natural ligands of the targets are illustrated in Figure 2. Finally, synpeps were docked against their targets to assess the correctness of fit. Docking jobs were performed using Cluspro [29-32] and HADDOCK [33,34] (Figure 3).

	Similarity%	Gap%
AMA1-SSBS-1	42.3%	11.5%
EBA 175-SSBS-2	36.0%	0%
MSP1 19-SSBS-3	44.4%	7.4%

Table 1: Comparison of sequences of binding region of natural ligands of targets and prospective synpeps.

Peptide		I-TASSER	QUARK	PEPFOLD	RMSD
SSBS-1	Q-Mean score	0.04	0.47	0.16	0.63
	Z score	-3.14	-0.78	-2.48	
SSBS-2	Q-Mean score	0.45	0.576	0.503	0.66
	Z score	-0.903	0.242	-0.64	
SSBS-3	Q-Mean score	0.56	0.65	0.54	0.80
	Z score	0.28	0.16	-0.41	

Table 2: Structure assessment and structural similarity assessment of not-coding peptides.

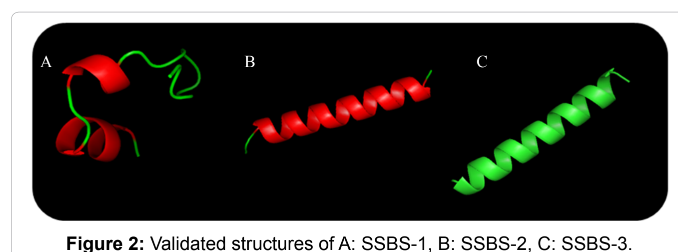

Figure 2: Validated structures of A: SSBS-1, B: SSBS-2, C: SSBS-3.

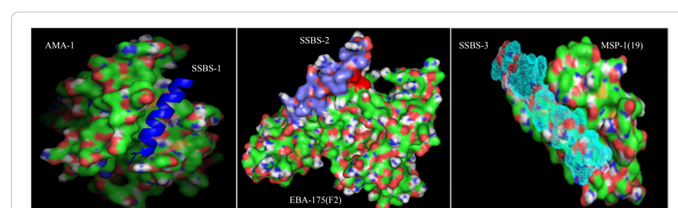

Figure 3: Docking poses of *Plasmodium* surface antigens with their respective *synpeps*.

The docking of AMA-1 with R1 peptide resulted in a docking score of -1735 as against docking score of -1616 of AMA-1 and SSBS-1 using Cluspro. The ClusPro results for peptide SSBS-3 and MSP-1(19) was -8821 and for docking of band3 and MSP-1 (19), it was -5661. The docking of EBA-175 dimer performed by HADDOCK showed a barely strong interaction at a score of -19.6 ± -14.7. However, the docking of EBA-175 and SSBS-2 showed a stronger interaction at -23.7 ± -3.7 indicating that EBA-175 might bind with greater affinity to SSBS-2 than its own monomer. Although these findings are encouraging, they are preliminary observations. More computational studies of not-coding DNA derived synpeps and their design would be required for future experimental studies.

A Critical Overview

The human genome project revealed a surprising finding that large chunk of genome is unexpressed. This was strongly supported by the recent ENCODE studies [35]. Why nature decided not to express a significant region of the genome remains a mystery?

Having successfully demonstrated that one can make artificial genes from non-expressed genome of Escherichia coli, the big question is: what kind of molecules can possibly be made from the un-expressed genome? Given an enormous combinatorial possibility, what are the best case scenarios?

A non-obvious advantage of using the not-coding DNA space is that synpeps are novel and not exposed to the organisms in the natural setting. Hence it may take greater effort and time for a pathogen to develop resistance towards synpeps in comparison to the naturally made therapeutic molecules by organisms.

Although synpeps have a potential therapeutic use, we recognize certain constraints too. Synpeps present themselves with a non-trivial possibility of not passing the cell membrane, unless their length is about 5 amino acids long [36]. This is because the entry of a molecule into a cell is regulated by channels, aquaporins and such; they do not allow macromolecules like proteins to pass through. As synpeps are possibly non-natural molecules, it is unclear whether the existing membrane transport mechanism will help. To address this issue, it would help to design novel cell penetrating peptides that can deliver synpeps into the cell [37]. Further, although synpeps can be designed to bind to a particular target one also needs to look at all the downstream and indirect effects. As synpeps are not exposed to natural immune system before, there are slim chances of body initiating an immune response like that of an antibody when bound to a prospective antigen.

The present paper presents preliminary evidence-based hypothesis and offers a novel approach of making novel antimalarial peptides from not-expressed regions of the genome. In future, significant experimental work needs to be performed to establish the efficacy and safety of synpeps as an alternative to traditional routes of drug discovery against malaria.

Acknowledgement

We would like to acknowledge and thank Prof. Ravindra Ghooi for his support and help. We would also like to thank Dr. Sudhir Kulkarni for his help.

Author Contribution

Pawan K Dhar conceptualized the project, provided the guidance and reviewed the manuscript. Mukta Joshi, Shankar V Kundapura, Thirtha Poovaiah carried the work and wrote the manuscript. Kundan Ingle provided technical support and mentored the structural modelling part of the work. All authors declare no conflict of interest.

References

1. World Health Origanization (2013) Media centre Malaria: WHO.
2. Doolan DL, Dobaño C, Baird JK (2009) Acquired immunity to malaria. Clin Microbiol Rev 22: 13-36, Table of Contents.
3. White NJ (2002) The assessment of antimalarial drug efficacy. Trends Parasitol 18: 458-464.
4. Biagini GA, O'Neill PM, Bray PG, Ward SA (2005) Current drug development portfolio for antimalarial therapies. Curr Opin Pharmacol 5: 473-478.
5. Bloland PB, Ettling M, Meek S (2000) Combination therapy for malaria in Africa: hype or hope? Bull World Health Organ 78: 1378-1388.
6. Fairhurst RM, Nayyar GM, Breman JG, Hallett R, Vennerstrom JL, et al. (2012) Artemisinin-resistant malaria: research challenges, opportunities, and public health implications. Am J Trop Med Hyg 87: 231-241.
7. Dhar PK, Thwin CS, Tun K, Tsumoto Y, Maurer-Stroh S, et al. (2009) Synthesizing non-natural parts from natural genomic template. J Biol Eng 3: 2.
8. Gaur D, Mayer DC, Miller LH (2004) Parasite ligand-host receptor interactions during invasion of erythrocytes by Plasmodium merozoites. Int J Parasitol 34: 1413-1429.
9. Duraisingh MT, Maier AG, Triglia T, Cowman AF (2003) Erythrocyte-binding antigen 175 mediates invasion in Plasmodium falciparum utilizing sialic acid-dependent and -independent pathways. Proc Natl Acad Sci U S A 100: 4796-4801.
10. Tolia NH, Enemark EJ, Sim BK, Joshua-Tor L (2005) Structural basis for the EBA-175 erythrocyte invasion pathway of the malaria parasite Plasmodium falciparum. Cell 122: 183-193.
11. Pizarro JC, Chitarra V, Verger D, Holm I, Pêtres S, et al. (2003) Crystal structure of a Fab complex formed with PfMSP1-19, the C-terminal fragment of merozoite surface protein 1 from Plasmodium falciparum: a malaria vaccine candidate. J Mol Biol 328: 1091-1103.

12. Srinivasan P, Beatty WL, Diouf A, Herrera R, Ambroggio X, et al. (2011) Binding of Plasmodium merozoite proteins RON2 and AMA1 triggers commitment to invasion. Proc Natl Acad Sci U S A 108: 13275-13280.

13. Lamarque M, Besteiro S, Papoin J, Roques M, Vulliez-Le Normand B, et al. (2011) The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites. PLoS Pathog 7: e1001276.

14. Bai T, Becker M, Gupta A, Strike P, Murphy VJ, et al. (2005) Structure of AMA1 from Plasmodium falciparum reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. Proc Natl Acad Sci U S A 102: 12736-12741.

15. Vulliez-Le Normand B, Tonkin ML, Lamarque MH, Langer S, Hoos S, et al. (2012) Structural and functional insights into the malaria parasite moving junction complex. PLoS Pathog 8: e1002755.

16. Richard D, MacRaild C, Riglar D, Chan J, Foley M, et al. (2010) Interaction between Plasmodium falciparum apical membrane antigen 1 and the rhoptry neck protein complex defines a key step in the erythrocyte invasion process of malaria parasites. J Biol Chem 285: 14815-14822.

17. Skrzypek MS, Hirschman J (2011) Using the Saccharomyces Genome Database (SGD) for analysis of genomic information. Curr Protoc Bioinformatics Chapter 1: Unit 1.

18. Balakrishnan R, Park J, Karra K, Hitz BC, Binkley G, et al. (2012) YeastMine--an integrated data warehouse for Saccharomyces cerevisiae data as a multipurpose tool-kit. Database (Oxford) 2012: bar062.

19. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, et al. (2012) ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 40: W597-603.

20. Rice P, Longden I, Bleasby A (2000) EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 16: 276-277.

21. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9: 40.

22. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5: 725-738.

23. Xu D, Zhang Y (2012) Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. Proteins 80: 1715-1735.

24. Maupetit J, Derreumaux P, Tuffery P (2009) PEP-FOLD: an online resource for de novo peptide structure prediction. Nucleic Acids Res 37: W498-W503.

25. Maupetit J, Derreumaux P, Tufféry P (2010) A fast method for large-scale de novo peptide and miniprotein structure prediction. J Comput Chem 31: 726-738.

26. Thévenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, et al. (2012) PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res 40: W288-293.

27. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22: 195-201.

28. VLife Sciences Technologies VLifeMDS: Molecular Design Suite. (2010)

29. Kozakov D, Hall DR, Beglov D, Brenke R, Comeau SR, et al. (2010) Achieving reliability and high accuracy in automated protein docking: ClusPro, PIPER, SDU, and stability analysis in CAPRI rounds 13-19. Proteins 78: 3124-3130.

30. Kozakov D, Brenke R, Comeau SR, Vajda S (2006) PIPER: an FFT-based protein docking program with pairwise potentials. Proteins 65: 392-406.

31. Comeau SR, Gatchell DW, Vajda S, Camacho CJ (2004) ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics 20: 45-50.

32. Comeau SR, Gatchell DW, Vajda S, Camacho CJ (2004) ClusPro: a fully automated algorithm for protein-protein docking. Nucleic Acids Res 32: W96-99.

33. De Vries SJ, van Dijk M, Bonvin AM (2010) The HADDOCK web server for data-driven biomolecular docking. Nat Protoc 5: 883-897.

34. De Vries SJ, van Dijk AD, Krzeminski M, van Dijk M, Thureau A, et al. (2007) HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. Proteins 69: 726-733.

35. ENCODE Project Consortium, Bernstein BE, Birney E, Dunham I, Green ED, et al. (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489: 57-74.

36. Saliba KJ, Kirk K (1998) Uptake of an antiplasmodial protease inhibitor into Plasmodium falciparum-infected human erythrocytes via a parasite-induced pathway. Mol Biochem Parasitol 94: 297-301.

37. Morris MC, Depollier J, Mery J, Heitz F, Divita G (2001) A peptide carrier for the delivery of biologically active proteins into mammalian cells. Nat Biotechnol 19: 1173-1176.