J. Natn. Sci. Coun. Sri Lanka 1993 21(1) : 165-172

SHORT COMMUNICATION

PRELIMINARY STUDIES ON THE USE OF SYNTHETIC OLIGONUCLEOTIDE PROBES BASED ON CONSERVED PROTEIN SEQUENCES FOR IDENTIFYING GENES IN MOSQUITOES AND MALARIA PARASITES.

R. RAMASAMY AND MANTHRI S. RAMASAMY
Malaria and Vector Biology Laboratories, Division of Life Sciences, Institute of Fundamental Studies, Hantana Road, Kandy.

(Date of receipt : 22 January 1993)
(Date of acceptance : 23 April 1993)

INTRODUCTION

The isolation and sequencing of particular genes is a necessary step in determining the amino acid sequence of the corresponding proteins and the regulation of their synthesis. In one approach to isolating genes, limited amino acid sequence is used to deduce the corresponding coding sequence. Oligonucleotides of this sequence are then chemically synthesized, radioactively labelled and used as hybridization probes to detect target coding sequences in a clone bank. The redundancy of the genetic code poses a limitation to this approach and probe sequences are usually based on regions of protein sequence that are rich in amino acids with a restricted number of codons eg. methionine and tryptophan. To cover all codon possibilities, it is sometimes necessary to synthesize probes containing a mixture of different individual nucleotides. One way to overcome this problem is to use a base analogue that can pair with any one of the four natural bases at ambiguous positions with or without forming hydrogen bonds. Inosine, which is found in the 5' or "wobble" position of tRNA anticodons, forms base pairs with A, C and U in mRNA. 2'-deoxyinosine is therefore an useful insert at ambiguous codon positions in synthetic DNA probes.[1]

Proteases in malaria parasites are needed for degrading of haemoglobin, protein processing,[2] invasion of host erythrocytes[3] and immune evasion. Parasite proteases may therefore be useful targets for the rational design of inhibitory drugs.

Analogues of insect hormones may modulate the physiology of insects and are therefore potentially useful as agents for controlling insect vectors of disease. Many insect hormones are peptides, and in some cases are structurally related to vertebrate peptide hormones, indicating a common ancestral gene. Examples are some of the prothoracicotropic hormones (PTTH) of *Bombyx mori* (silkworm) that stimulate ecdysone release from the prothoracic glands. The 4400 m.wt PTTH-II is structurally homologous to the insulin A chain and insulin - like growth factor 1[4]. Such conserved

sequences may be used to produce oligonucleotides to identify homologous neuropeptide coding genes in insects such as mosquitoes that are of medical importance.

We report here on experiments to test the feasibility of using synthetic oligonucleotides to determine the presence of genes coding for a 26 kDa *P. falciparum* protein, serine protease, cathepsin D like aspartyl protease, cysteine protease and a PTTH-II-like peptide in the human malaria parasite *Plasmodium falciparum* Welch and a laboratory vector of human malaria, *Anopheles farauti* Laveran.

METHODS AND MATERIALS

Protein sequence and oligonucleotide synthesis

Conserved amino acid sequences were selected for probe synthesis by examining published sequences of proteins for suitable sequences with minimal codon multiplicity. The temperature at which 50% of hybrids dissociate, Td, was calculated using the formula :

$$Td(^0C) = 4(G+C) + 2(A+T).$$

(a) Insulin like insect neuropeptide[4]

amino acid sequence	Gly -	Ile -	Val -	Asp -	Glu -	Cys - Cys
predicted coding sequence 5′ - 3′		AUU -	GUX -	GAU -	GAA -	UGU - UGU
		C		C	G	C C
		A				
probe synthesized 3′ - 5′		TAI -	CTI -	CTI -	CTT -	ACI - ACI
			C		C	

Calculated Td = 37^0C

(b) Cysteine protease[5]

amino acid sequence	Pro -	Val -	Lys -	Asn -	Gln
predicted coding sequence 5′ - 3′	CCX -	GUX -	AAA -	AAU -	CAA
			G	C	G
probe synthesized 3′ - 5′	GGI -	CAI	TTT	TTI	GTT
	C	C	C		C

Calculated Td = 33^0C

(c) Cathepsin D (aspartyl protease)[6]

amino acid sequence	Val - Val - Phe - Asp - Thr - Gly
predicted coding sequence	GUX - GUX - UUU - GAU - ACX - GGX
5′ - 3′	C C
probe synthesized	CAI - CAI - AAI - CTI - TGI - CC
3′ - 5′	C C C

Calculated Td = 36^0C

(d) Serine protease[7]

amino acid sequence	Thr - Ala - Ala - His - Cys
predicted coding sequence	ACX - GCX - GCX - CAU - UGU
5′ - 3′	C C
probe synthesized	TGI - CGI - CGI - GTA - TCA
3′ - 5′	C C C G G

Calculated Td = 40^0C

(e) 26 kDa *P. falciparum* antigen[8]

amino acid sequence	Gly - Glu - Asn - Ala - Phe - Asp - Pro
predicted coding sequence	GGX - GAA - AAU - GCX - UUU - GAU - CCX
5′ - 3′	G C C C
probe synthesized	CCI - CTT - TTI - CGI - AAI - CTI - GG
3′ - 5′	C C C

Calculated Td = 44^0C

Synthesis of oligonucleotides was performed on an automated synthesizer (Applied Biosystems, CA), and cleavage and deprotection in concentrated NH_4OH at 60^0C for 16h. The oligonucleotides were purified by gel filtration on Sephadex G-50 columns using 10mM triethylammonium bicarbonate buffer pH 7. Fractions absorbing strongly at 260 nm were pooled and freeze dried.

Preparation of DNA

Parasite DNA was obtained from *in vitro* cultures of *P. falciparum* FC27 isolate. 0.5 ml schizonts at > 50% parasitaemia prepared by gelatine sedimentation[8] were used for DNA preparation. Three ml of *An. farauti* pupae obtained from a laboratory colony[9] were used to prepare mosquito DNA.

Parasites and pupae were resuspended in 1 - 2 ml TE (10 mM Tris.Cl, 1 mM EDTA pH 7.4) to which was added 20 - 50 ml of a solution containing 0.5 M EDTA pH 8, 100 μg ml^{-1} proteinase K and 0.5% sarkosyl (Sigma, MO) and incubated for 3h at 50^0C to digest protein. The preparations were extracted with phenol (containing

0.1% 8-hydroxyquinoline as a free radical scavenger) by standard procedures[10] three times and the aqueous layer with DNA dialysed against several changes of a buffer containing 50 mM Tris.Cl, 10 mM EDTA, 10 mM NaCl, pH 8. The DNA was then treated for 3h at 37^{0}C with 100 μg ml^{-1} of bovine pancreatic ribonuclease A (Boehringer, CA) that had been pre-heated to 100^{0}C to destroy contaminating DNAase.

The DNA was then re-extracted with phenol twice and once with water saturated ether to remove residual phenol. The DNA was precipitated by adding 2 vols of ice cold ethanol and incubated for 16h at -70^{0}C. The precipitates collected by centrifugation were resuspended in TE and dialysed against TE at 4^{0}C. The concentrations of the resulting mosquito and parasite DNA were 0.175 mg/ml and 0.075 mg/ml respectively.

Restriction enzyme digestion and electrophoresis

10 μg of mosquito and parasite DNA were completely digested with 10 units of the following restriction enzymes Bam H1, Dra 1, Eco R1, Hind III, Pst 1 and Rsa 1 (BRL, Gaithersberg, MD) by standard procedures[10]. Digested DNA was separated by electrophoresis on 0.7% agarose gels using Hind III digest as a size marker and the separated DNA, denatured and transferred to nylon membrane by Southern blotting[10].

Radiolabelling of probes

Stock solutions of oligonucleotides were made in TE. Aliquots were labelled with ^{32}P at the 5' termini using ^{32}P-ATP (Amersham, UK) and the enzyme T4 polynucleotide kinase (Pharmacia, Sweden)[10]. Ten - twenty picomoles of oligonucleotides were radiolabelled at a given time.

Dot blots

10, 1 and 0.1 μg aliquots of the DNA, diluted in TE, were heated at 95^{0}C for 5min to denature the DNA and then cooled on ice for 5min. An 8 x 11 cm piece of Hybond N nylon membrane (Amersham, UK) was pre-wetted in 0.1 M Tris pH 7.4, and samples of mosquito and parasite DNA were blotted onto the membrane using a dot-blot apparatus (Biorad, CA). The DNA in the blots were again denatured in 1.5M NaCl, 0.5M NaOH for 1 min followed by neutralization in 1.5M NaCl, 1M Tris pH 8 for 1 min. Dried membranes were UV irradiated for 3 min to cross-link the DNA to nylon. The membranes were treated with pre-hybridizing solution[10] containing 5 x Denhard's solution,[10] 0.1% sodium dodecyl sulphate, 6mM EDTA, 0.9M NaCl, 20mM Tris pH 8 and 10% dextran sulphate at 25^{0}C for 4h. 100 μl of solution was used per sq. cm of

membrane. For hybridization, the pre-hybridization solution was changed once and the labelled oligonucleotides added directly to the pre-hybridization solution. Hybridization was performed at 25^{0}C for 16h. The membranes were then washed 3x at 25^{0}C for 20 min each in 6x concentrated saline sodium citrate[10] (6xSSC) containing 0.1% SDS followed by three washes at 25^{0}C of 10min each in 2xSSC, 0.1% SDS. The membranes were dried, wrapped in thin plastic and autoradiographed with X-ray film overnight.

Southern blots

Southern blots were hybridized with ^{32}P labelled oligonucleotides in an identical manner to dot blots.

RESULTS

The results of dot blots with the different labelled oligonucleotides are presented in Figure 1. The probes directed against the 26 kDa *P. falciparum* antigen, the insulin - like insect neuropeptide, cysteine protease and serine protease hybridised to sequences present in both parasite and mosquito DNA. The latter three oligonucleotides appeared to react more strongly with mosquito DNA rather than

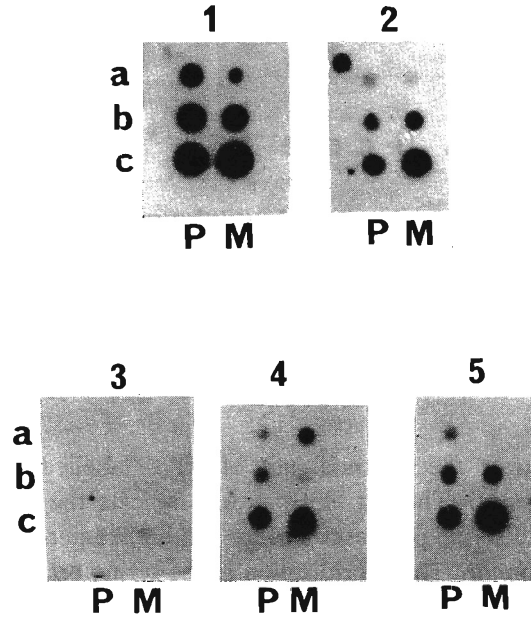

Figure 1: Autoradiograph of dot blots of An. farauti and P. falciparum DNA probed with oligonucleotides. a - 0.1 μg; b - 1 μg; c - 10 μg of mosquito (M) and parasite (P) DNA were reacted with ^{32}P - labelled probes for: 1 - 26 kDa P. falciparum antigen, 2 -insulin like neuropeptide, 3 - cathepsin D, 4 - serine protease and 5 - cysteine protease.

parasite DNA. The aspartyl protease showed a weak reaction with mosquito but not parasite DNA.

In subsequent Southern blot analysis of *P.falciparum* DNA, the serine protease specific probe hybridized to multiple fragments in many restriction enzyme digests (Figure 2).

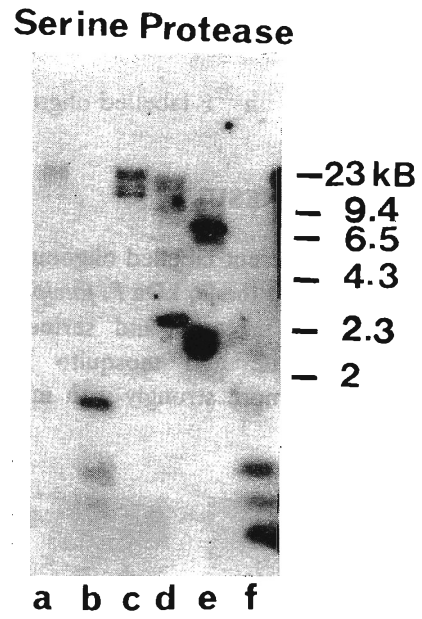

Figure 2. **Autoradiograph of a Southern blot of restriction enzyme digests of P. falciparum DNA reacted with the ^{32}P - labelled probe for serine protease. The DNA was digested with the following restriction enzymes: a) Bam III, b) Dra 1, c) Eco R1, d) Hind III, e) Pst 1 and f) Rsa 1. The migration positions of fragments of lambda DNA digested with Hind III and used as molecular weight markers are indicated in kilobases (kß).**

DISCUSSION

The results demonstrate the potential use of ^{32}P - labelled oligonucleotides for identifying specific genes. The dot blot patterns indicate the probable presence of genes for cysteine and serine proteases and for insulin like peptides or proteins in *P. falciparum* and *An. farauti*. The hybridization of the serine protease probe to several but discrete DNA fragments after restriction enzyme digestion may be due to the presence of different genes with the corresponding conserved sequences. This is similar to what is observed in higher organisms where genes for several different serine proteases and serine protease - related proteins are present in the genome. However, hybridization of the probes to irrelevant coding or non-coding sequences in mosquito or parasite DNA, due to probe redundancy, cannot be excluded.

False-positive hybridization can be minimized by increasing the stringency of the hybridization and washing conditions eg. by using higher temperatures. Since the codon usage frequencies for *P.falciparum* have been determined recently[11], it is now also possible to synthesize less-redundant oligonocleotide probes that can further reduce false-positive hybridization.

The function of the cytoplasmic 26 kDa antigen in *Plasmodium*[8] is not known. The presence of homologous sequences in mosquitoes suggests that this may be a gene conserved across different organisms or that the probe sequence is very degenerate.[1]

Polymerase chain reaction based amplification of the genes using conserved oligonucleotide sequences may be used to isolate the corresponding genes as the next step. An alternative procedure is to use the DNA probes to isolate clones from a DNA library prepared by digesting DNA with appropriate restriction enzymes. From the data presented, Dra 1, Hind III and Rsa 1 libraries of *P. falciparum* genomic DNA may be suitable for isolating genes for serine proteases. With both procedures, a number of irrelevant coding or non-coding sequences homologous to the probes might be isolated together with the relevant genes.

Acknowledgements

We thank C. Reed and M. Sands for technical assistance. Experimental work was carried out at the Queensland Institute for Medical Research, Brisbane, Australia.

References

1. Ohtsuka E., Matsuki S., Ikehara M., Takahashi Y. & Matsuhara K. (1985). An alternative approach to deoxyoligonucleotides as hybridization probes by insertion of deoxyinosine at ambiguous codon posiitons. *The Journal of Biological Chemistry* **260**: 2605-2608.

2. Lyon J.A., Geller R.H., Haynes J.D., Chulay J.D. & Weber J.L. (1986). Epitope map and processing scheme for the 195,000-dalton surface glycoprotein of *Plasmodium falciparum* merozoites deduced from cloned overlapping segments of the gene. *Proceedings of the National Academy of Sciences, USA* **83**: 2989-2993.

3. Dluzewski A.R., Rangachari K., Wilson R.J.M. & Gratzer W.B. (1986). *Plasmodium falciparum*: Protease inhibitors and inhibition of erythrocyte invasion. *Experimental Parasitology* **62**: 416-422.

4. Nagasawa H., Kataoka H., Isogai A., Tamura S., Suzuki A., Ishizaki H., Mizoguchi A. Fujiwara Y. & Suzuki A. (1984). Amino terminal amino acid sequence of the silkworm prothoracicotropic hormone; homology with insulin. *Science* **226**: 1344-1345.

5. Rogers J.C., Dean D. & Heck G.R. (1985). Aleurain: a barley thiol protease closely related to mammalian cathepsin, H. *Proceedings of the National Academy of Sciences USA* **82**: 6512-6516.

6. Faust P.L., Kornfeld S. & Chirgwin J.M. (1985). Cloning and sequence analysis of cDNA for human cathepsin D. *Proceedings of the National Academy of Sciences USA* **82**: 4910-4914.

7. Hartley B.S., Brown J.R., Kauffman D.L. & Smillie L.B. (1965). Evolutionary similarities between pancreatic proteolytic enzymes. *Nature* **207**: 1157-1159.

8. Ramasamy R., Simpson R.J., Dexter A., Keeghan M., Reed C., Bushell G., Ingram L., Henderson H., Moloney M.B., Moritz R.L., Rubira M.R. & Kidson C. (1988). Isolation and partial characterization of a 26 kDa antigen from *Plasmodium falciparum* recognized by an inhibitory monoclonal antibody. *Molecular and Biochemical Parasitology* **29**: 125-132.

9. Ramasamy M.S. & Ramasamy R. (1990). Effect of anti-mosquito antibodies on the infectivity of the rodent malaria parasite *Plasmodium berghei* to *Anopheles farauti*. *Medical and Veterinary Entomology* **4**: 161-166.

10. Sambrook J., Fritsch E.F. & Maniatis T. (1989). Molecular cloning - a laboratory manual. Cold Spring Harbour Laboratory Press, New York.

11. Hyde J.E., Kelly S.L., Holloway S.P., Snewin V.A. & Sims P.F.G. (1989). A general approach to isolating *Plasmodium falciparum* genes using non-redundant oligonucleotides inferred from protein sequences of other organisms. *Molecular and Biochemical Parasitology* **32**: 247-262.