# CHAPTER 2

## *Identification of target proteins and the design of antisense oligodeoxynucleotides*

### *2.1. Introduction*

#### 2.1.1 General

The aim of the work described in this chapter was to select a target protein for the inhibition of *P. falciparum,* with an antisense ODN targeted to the area surrounding the translation initiation codon as well as ODNs based on computational predictions of the mRNA secondary structure.

The design of antisense ODNs can be divided into several steps, each of which is important to the successful application of the end product. These steps include choosing the target protein, the targeted nucleotide sequence and type of ODN modification. Since the efficacy of antisense ODNs is largely dependent on their successful hybridization to mRNA, accessibility of the target sequence is of paramount importance (Lima and Crooke, 1997; Phillips and Zhang, 2000).

The aim of the study largely influences the choice of the target protein. If the antisense ODN is to be applied for therapeutic use, it is important that the target protein play a vital role in the organism. In many cases however, antisense studies are performed to validate the protein as a potential drug target. It is also possible that the protein has already been validated but that its specific inhibition, using conventional methods is problematic.

Another common aim of antisense studies is to determine the function of a particular protein or enzyme in a cell or metabolic pathway. For application in the current projects in our laboratory the most applicable aims would be to utilize the antisense technique in the validation of a protein as a potential drug candidate and/ or to determine the function of a particular protein or enzyme in a cell or metabolic pathway, akin to knock down applications.

Once the target protein has been identified, the secondary or higher order structures of the mRNA target sequence(s) must be determined since it is not a linear molecule (Smith *et al*, 2000; Sohail and Southern, 2000). There are several complex secondary and tertiary structures along the length of mRNA, which can have profound effects on antisense ODN hybridization to the target. Thermal denaturation studies performed in the 1950's, showed that between 40 and 60% of the bases in an RNA molecule are stacked and paired. The most common secondary structures include hairpin loops; internal loops; bulges and bifurcations (Figure 2.1). These secondary structures may associate with each other to form tertiary structures. An example of such tertiary structure formation is that of a pseudoknot. The latter forms when the single bases of two hairpin loops interact with each other (Riesner, 1997). Triple-stranded structures on the other hand, result from the interaction between the stems of two hairpin loops.

These different structures are distributed along the length of the mRNA molecule and influence the accessibility of these regions to antisense ODN hybridization. The secondary structures illustrated in Figure 2.1 represent certain functional signals (Riesner,

1997; Toulme, *et. al*., 1996).

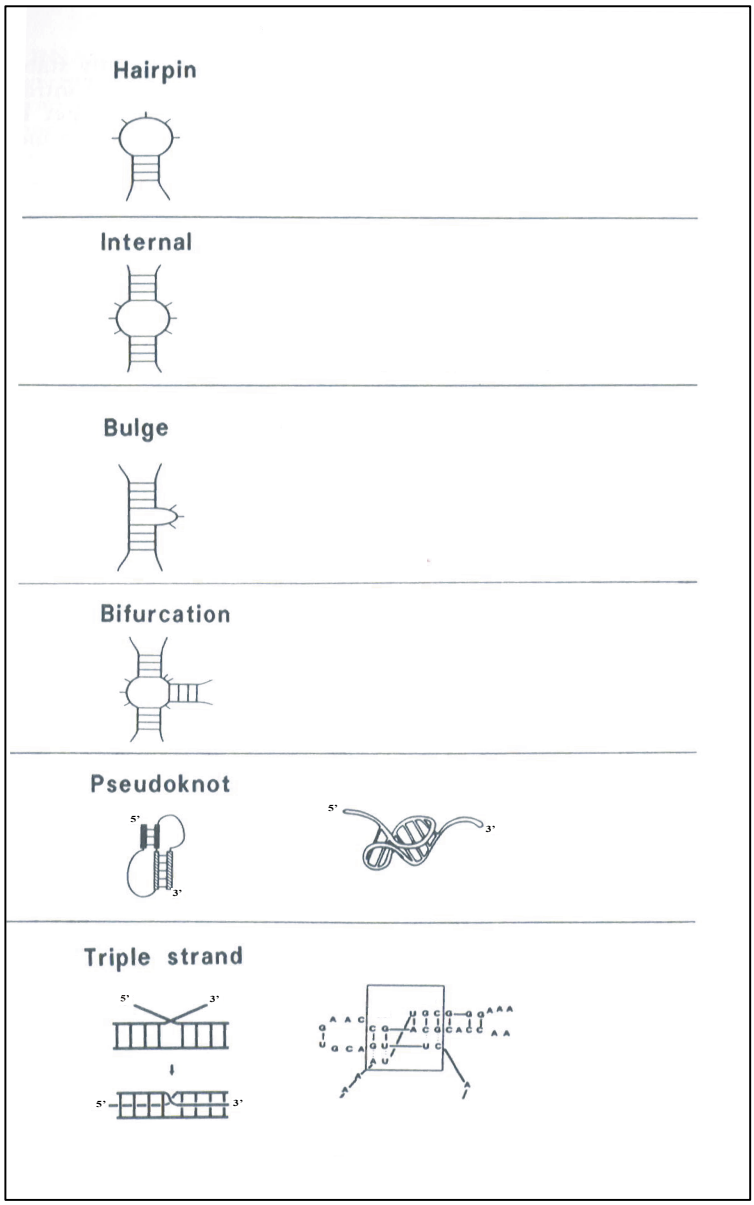



Figure 2.1: Graphic portrayal of the most common secondary and tertiary structure formations found in RNA. The hairpin loop, internal bulge, bulge and bifurcated loop are all examples of secondary structure formations, while the pseudoknot and triple strands are examples of tertiary structures. (Adapted from Riesner, 1997).

Hairpin loops are involved in the regulation of mRNA, which includes aspects such as stability, frame shifting and binding of specific proteins. Bulged loops serve as binding sites for proteins and are able to change their conformation after such binding has taken place. The tertiary structures too have specific functions such as pausing translation and frame shifting. From the nature of the tertiary structure formation it becomes clear that structures such as pseudoknots and triple strands are needed to define the relative orientation of distant helical regions (Riesner, 1997; Toulme *et. al*., 1996).

### 2.1.2. Target protein selection

Since the aim of the designed antisense ODN is to inhibit the *in vitro* proliferation of the malaria parasite, the target protein must fulfill certain criteria. Firstly, the protein must be vital for the continued survival of the parasite and secondly, the nucleotide sequence must be known.

Tubulin is a globular protein and ubiquitous to eukaryotic cells (Downing and Nogales, 1998). It is known to play a central role in cell division, cell movement and intra-cellular transport as a major component of microtubules (Bell, 1998). Inhibition of tubulin production and/or function therefore, has been shown to result in inhibition of cell proliferation and cell death (Downing and Nogales, 1998). Several anticancer drugs target tubulin proteins and thereby lead to the apoptosis of these cancer cells. Since cancer cells are rapidly dividing it is equally relevant and important to investigate the effect of tubulin inhibitors on the rapidly dividing malaria parasites and in so doing consider tubulin as a suitable target to inhibit parasite proliferation

The tubulin protein family can be divided into three subclasses, namely alpha (α), beta (β) and gamma (γ). The γ-tubulins can be found primarily at the microtubule organising centres (MTOCs), which act as the points of self-assembly in the cell, for example the spindle poles. The α- and β-tubulins, each with a molecular mass of 55 kDa, form heterodimers. The resultant 110 kDa proteins are the functional units that form the polar protofilaments. Staggered arrangement of thirteen protofilaments yields a helical arrangement of tubulin heterodimers to form the microtubule (Jordan and Wilson, 2004). Numerous ligands have been shown to bind to the tubulin component of microtubules in particular, of which several have proven anticancer properties (Downing and Nogales, 1998).

Most anticancer drugs such as colchicine, vincristine, vinblastine, taxotere and taxol interact with the β-tubulin subunit (Figure 2.2).

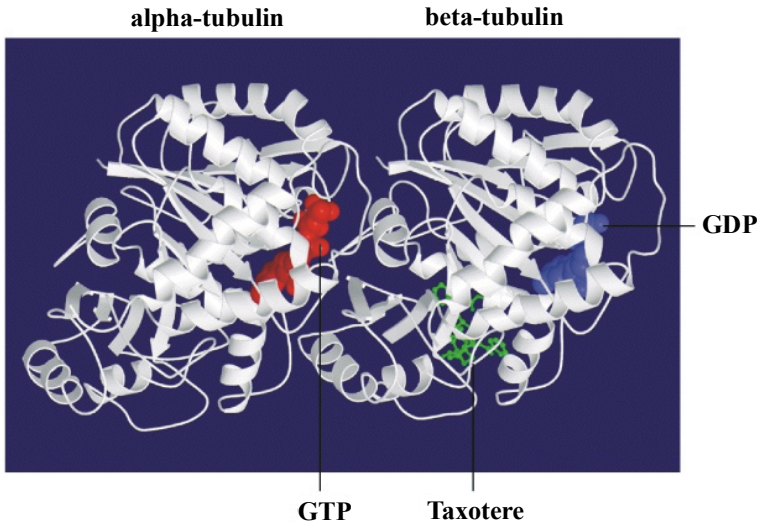



Figure 2.2: Illustration showing the spatial arrangement of α- and β-tubulin in the heterodimers that are assembled into microtubules. The binding site of taxotere, a microtubule inhibitor is indicated on the β-tubulin subunit (Downing and Nogales , 1998).

Jan *et al*. (2000) however discovered a new anticancer drug (COBRA-1) that exercises its action by binding to a cavity in α-tubulin. This drug has led to the destruction of microtubule organization and eventually apoptosis.

The effects of numerous known microtubule inhibitors on *P. falciparum* proliferation were investigated and the results are shown in Table 2.1.

Table 2.1: Inhibition of proliferation of *P.falciparum* in cultured erythrocytes and mammalian cells by microtubule inhibitors (Bell, 1998).

| **Compound** | **$IC_{50}$ (*P. falciparum*, µM)** | **$IC_{50}$ (mammalian cell, µM)** |
|---|---|---|
| Colchicine | 13 | $3.2x10^{-3}$ |
| Vinblastine | $2.8x10^{-2}$ | $6x10^{-4}$ |
| Vincristine | $6x10^{-3}$ | $2.2x10^{-3}$ |
| Taxol | $7.1x10^{-2}$ | $2.2x10^{-3}$ |
| Taxotere | $3.1x10^{-3}$ | $3.9x10^{-2}$ |

All the above compounds bind to β-tubulin but have different modes of action. Colchicine binds to a single site of an unassembled β-tubulin and thereby blocks or inhibits tubulin polymerization. Vinblastine and vincristine cause depolymeriztion of the already formed microtubule while taxol and taxotere stabilize the assembly of tubulin into microtubules to the extent that mitosis is disrupted and eventually leads to apoptotic cell death (Bell, 1998; www.dentistry.leeds.ac.uk/biochem). It is important to note that all

the above drugs affect the proliferation of mammalian cells as is evident from the fact that these agents are also employed as anticancer agents. From the information provided it becomes evident that tubulin proteins are suitable targets for the inhibition of *P. falciparum* proliferation, if this can be achieved in such a way as to specifically target the parasitic tubulin. This is problematic since Table 2.1 shows that all the compounds tested, with the exception of taxotere, are more effective inhibitors of mammalian cell proliferation than of *P. falciparum* proliferation.

Since the malaria parasite has five different known types of tubulin, namely α-I, α-II, β, γ and a putative δ tubulin it is necessary to consider which of the above would be the most suitable target. α-II tubulin is expressed predominantly in the sexual stages of the parasite lifecycle, whereas α-I and β tubulin are expressed in both the sexual and asexual stages (Delves *et. al*., 1990). It was decided to target α-I-tubulin since it is proposed as the binding and regulatory site for the key glycolytic enzymes: phosphofructokinase, glyceraldehyde-3-phosphate dehydrogenase and aldolase in the parasite (Itin *et. al*., 1993). This is particularly important due to the high glycolytic activity of infected erythrocytes as reflected by the 50 to 100-fold increase in the uptake of glucose (Hess *et. al*., 1995).

### 2.1.3. Target sequence selection

Several methods have been described to determine accessible regions of the mRNA molecule in order to improve the targeting of antisense ODNs. A brief description of the most commonly used methods to select mRNA target sequences is given below.

2.1.3.1. The random shotgun approach or gene walking: In this method, several ODNs (usually containing between 50 and 100 nucleotides) are targeted to various regions of a mRNA molecule (Smith *et al.*, 2000; Sohail and Southern, 2000). Typically only between 2 and 5% of the ODNs are successful (Sohail and Southern, 2000). Stull *et al.* (1996) measured the apparent dissociation constants of 32 phosphodiester and 5 phosphorothioate antisense ODNs, targeted against different regions of murine TNF-mRNA. These studies showed that a slight change in the target site of the mRNA has a profound influence on the ODN affinity for the site. Results such as these emphasize the complexity of successful target site selection. This is obviously an expensive and time-consuming method to identify possible target sites (Akhatar, 1998; Sohail and Southern, 2000).

2.1.3.2. Predictions of mRNA secondary structures: There are two ways in which mRNA secondary structure can be predicted to aid the identification of antisense target sequences. The first one utilizes computer-based predictions of the mRNA secondary structure in order to identify areas free of intramolecular hydrogen bonds. Three programs are commonly used for the prediction of RNA secondary structure namely, *Mfold*, *LinAll* and RNA*fold*. *Mfold* is used for the prediction of optimal and sub-optimal structures of linear or circular sequences at 37 °C. *LinAll* predicts optimal and sub-optimal secondary structures of linear or circular sequences at different temperatures. Finally RNA*fold* predicts only the optimal secondary structure at a given temperature (Riesner, 1997). The optimal structure refers to the structure having the minimum ΔG

(free energy) value. The sub-optimal structures are those that are more than 20% higher than the ΔG value of the optimal structure (Matthews *et. al*., 1999).

One of the drawbacks of the computer prediction techniques is that higher order interactions such as those with cellular proteins as well as possible tertiary structures that can influence ODN hybridization are unknown and therefore cannot be taken into account (Smith *et al.*, 2000; Sohail and Southern, 2000). The predicted secondary RNA structures therefore offer the researcher only a first approximation and not a comprehensive picture of all possible RNA structures.

The second method utilizes a phylogenic approach that is based on the principle that sequences with similar function acquire homologous secondary structures (Michel and Westhof, 1990, Sohail and Southern, 2000). This approach has been particularly successful in predicting the folding of large mRNA molecules. The drawback of this method lies in the fact that it is of limited use when only a small number of comparative nucleotide sequences are available or if the functional significance of the folded structure is unknown.

2.1.3.3. Oligonucleotide scanning arrays: A set of oligonucleotides, with each one being complementary to a section of the target sequence, is synthesized directly on the solid surface (glass or polypropylene) on which hybridization studies are conducted (Sohail and Southern, 2000). Radiolabelled transcripts are then hybridized to the arrays of complementary ODNs at physiological temperature. Hybridization patterns are obtained

by exposure of the arrays to a phosphor storage screen and scanning in a phosphorimager. This method is costly but offers the accuracy vital to the identification of the optimal ODN hybridization site.

2.1.3.4. Oligomer library / Ribonuclease-H-digestion: RNase-H, which specifically cleaves RNA in a DNA:RNA heteroduplex, is used in combination with a random or semi-random library. A random library consists of a complete set of all possible ODNs of a defined length (usually 10-12 mers). Semi-random libraries consist of ODNs (typically 11-mers) that have one constrained (i.e. non-variable) nucleotide at an internal position (Ho *et al*., 1996). To obtain accurate results, four libraries are synthesized each with a different constrained nucleotide. The size of each library is dependent on the length of the transcript under investigation. The target mRNA is usually transcribed *in vitro*, end-labeled and mixed with a random or semi-random library. Those library components complementary to accessible mRNA regions are identified by electrophoresis after cleavage of the DNA oligomer:RNA heteroduplexes by RNase-H. The problem with this particular method is that it is difficult to accurately determine the precise cleavage site and hence identify the precise location of the ODN hybridization site. This is significant since the shift of a single base or a single base length difference can have significant effects on the antisense activity of the ODN (Sohail and Southern, 2000). This method is therefore not as accurate as the oligomer scanning arrays but can be applied as an initial screening for RNA accessible sites (Milner *et. al*., 1997).

From the above it is evident that there is no single, reliable or inexpensive method for accurately predicting the most successful antisense target sites on mRNA and that

selection is still largely based on trial and error. Fortuitously, the area surrounding the translation initiation codon can be often used as a starting point due to the number of successful results reported in the literature using this region as a target site (Phillips and Zhang, 2000).

### 2.1.4. Oligonucleotide modification selection

The next important decision to be taken when designing antisense ODNs is the choice of modification. These may serve several purposes, the most important being protection against nuclease degradation and improving cell membrane permeation. Phosphodiester bonds are particularly susceptible to 3' exonuclease degradation (Uhlmann *et al*, 2000) and the hydrophobic core of cell membranes does not readily facilitate the uptake of the polyanionic ODNs into cells. Several types of chemical modifications to protect the ODN against enzymatic degradation have already being discussed in Chapter 1. The addition of a stable hairpin loop to either the 3' or 5' terminus of ODNs have to date only been applied and tested in conjunction with phosphorothioate modifications (Barker *et. al*., 1998). No other information is available except that this modification was shown to improve the efficacy by about 20% compared to the corresponding unmodified phosphorothiate ODN. Since this structural modification is the most simple and cost-effective, hairpin loops were added to all the ODNs used in this study to investigate their stability and uptake into parasitized erythrocytes.

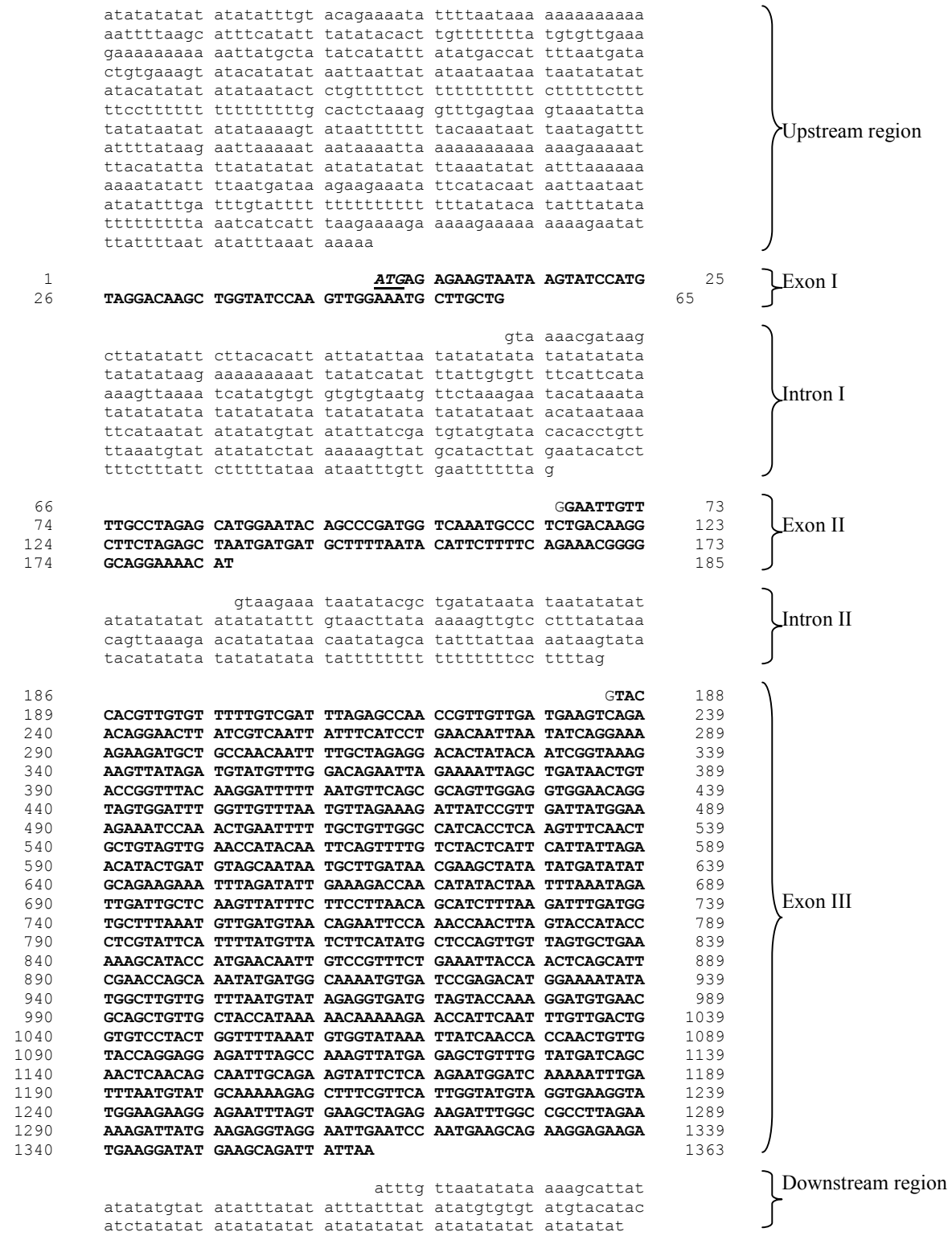



Figure 2.3: Complete gene sequence of *P. falciparum* α-I-tubulin. The translation initiation codon is underlined and the exons are indicated in bold letters (GenBank accession number: X15979).

## *2.2. Materials and Methods*

### 2.2.1. **Materials**

The ODNs were purchased from Roche Diagnostics.

### 2.2.2. **Methods**

#### 2.2.2.1. α-I-tubulin sequences

The complete sequence of *Plasmodium falciparum* α-I-tubulin gene (Genbank accession number: X15979; Holloway, 1989) is given in Figure 2.3.

The α-I-tubulin DNA sequence contains two introns, which were removed prior to analysis of the secondary structure of the mRNA by *Mfold* (Matthews *et. al.*, 1999 and Zuker, 1989).

#### 2.2.2.2. The RNA folding program, *Mfold*

*Mfold* was used to predict RNA secondary structure by energy minimization. Michael Zuker and co-workers at the Institute for Biomedical Computing, Washington University, USA wrote this program, and it was downloaded (version 2.3) from the Zuker homepage at http://bioinfo.math.rpi.edu/~zukerm/home.html on a Silicon Graphics Workstation.

##### 2.2.2.2.1. Drawing and evaluation of secondary mRNA structures

i) RNA Structure 3.5

This program utilizes the ct (connect) file format generated by the *Mfold* program to produce visual representations of the mRNA structures. The program was downloaded from the homepage of Doug Turner (sub-heading: The Turner Group) at http://www.chem.rochester.edu/Faculty/Turner.html.

RNA Structure 3.5 offers several features that allows a thorough investigation of the folded mRNA molecules. The two features used in these investigations, found under the *File* option, were *Draw* and *Oligo Walk*. The *Draw* function allows for the entry of an *Mfold* ct file and utilizes this information to generate a visual representation of the required RNA molecules. These can be printed as a hard copy or exported to Corel Draw or any word processing program.

The *Oligo Walk* function was extremely important in these investigations as it identified optimal ODN hybridization sites (a length of 27-mer at a concentration of 0.5 µM was specified) in each of the RNA structures predicted by *Mfold*. Each RNA structure is "walked" one base at time, and the overall ΔG of hybridization for each hybridization site is calculated. After conclusion of the *Oligo Walk* the energetically most favourable hybridization site (minimum ΔG) is identified. It is important to note that the overall hybridization energy took the energy required to break any local secondary structures into account.

ii) RNAdraw

The visual representations of the RNA molecules produced by this program were compared to those obtained with RNA Structure 3.5. This program, written by Ole Matzura, was downloaded from the RNAdraw homepage at: http://rnadraw.base8.sel. This program also utilizes the ct files generated by the *Mfold* program to produce visual representations of the folded RNA molecules.

All mentioned web addresses were last visited and confirmed in April, 2005.

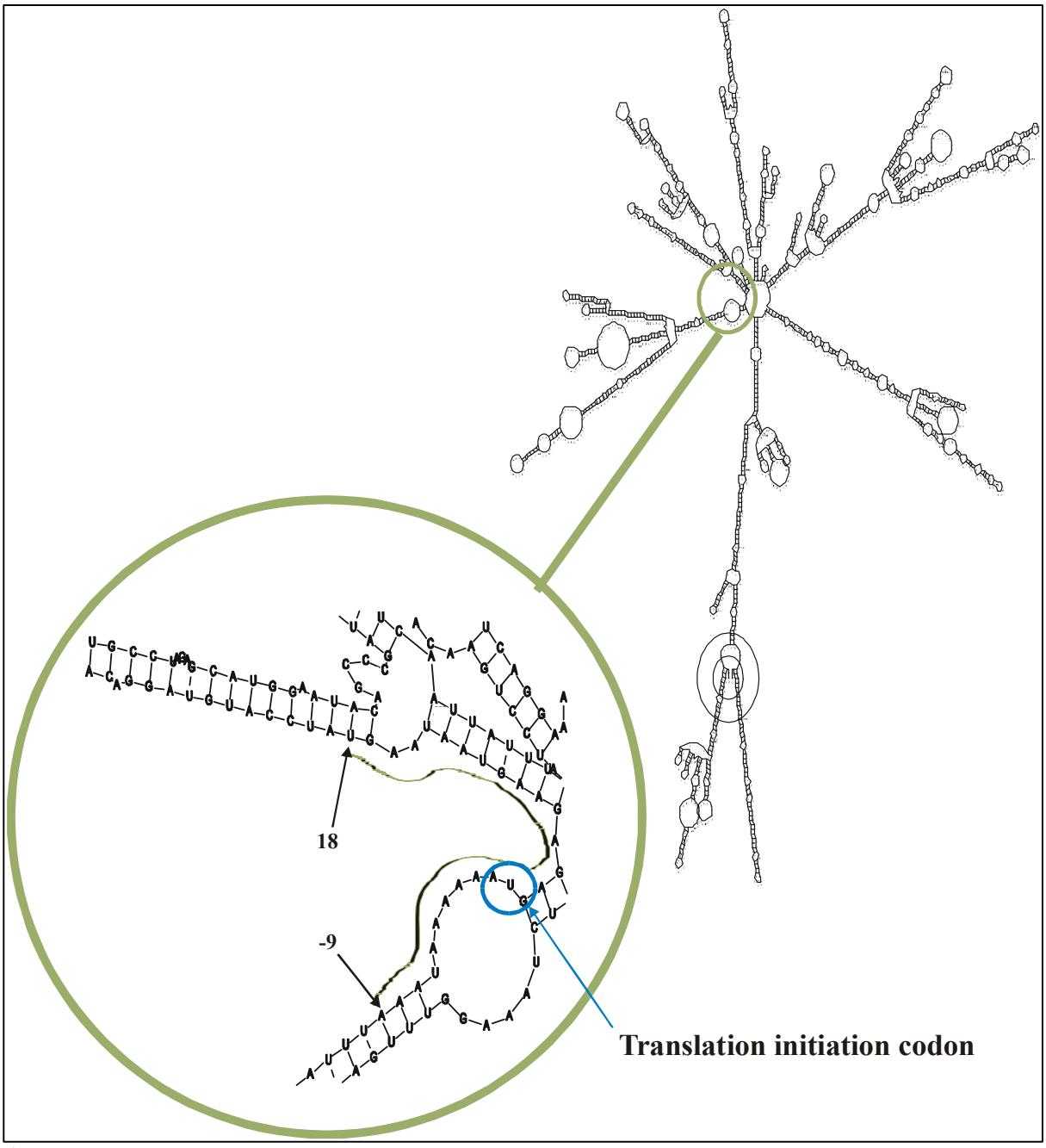
## *2.3 Results*

### 2.3.1. Selection of target mRNA sequences

Three antisense ODNs, targeted against different regions of the mature α-I-tubulin mRNA sequence of *P. falciparum*, were designed as well as the appropriate controls. One targeted the region around the translation initiation codon, while the selection of the other two was based on mRNA secondary structure information (Refer to Tables 2.4 and 2.5 in section 2.3.1.3 for sequences).

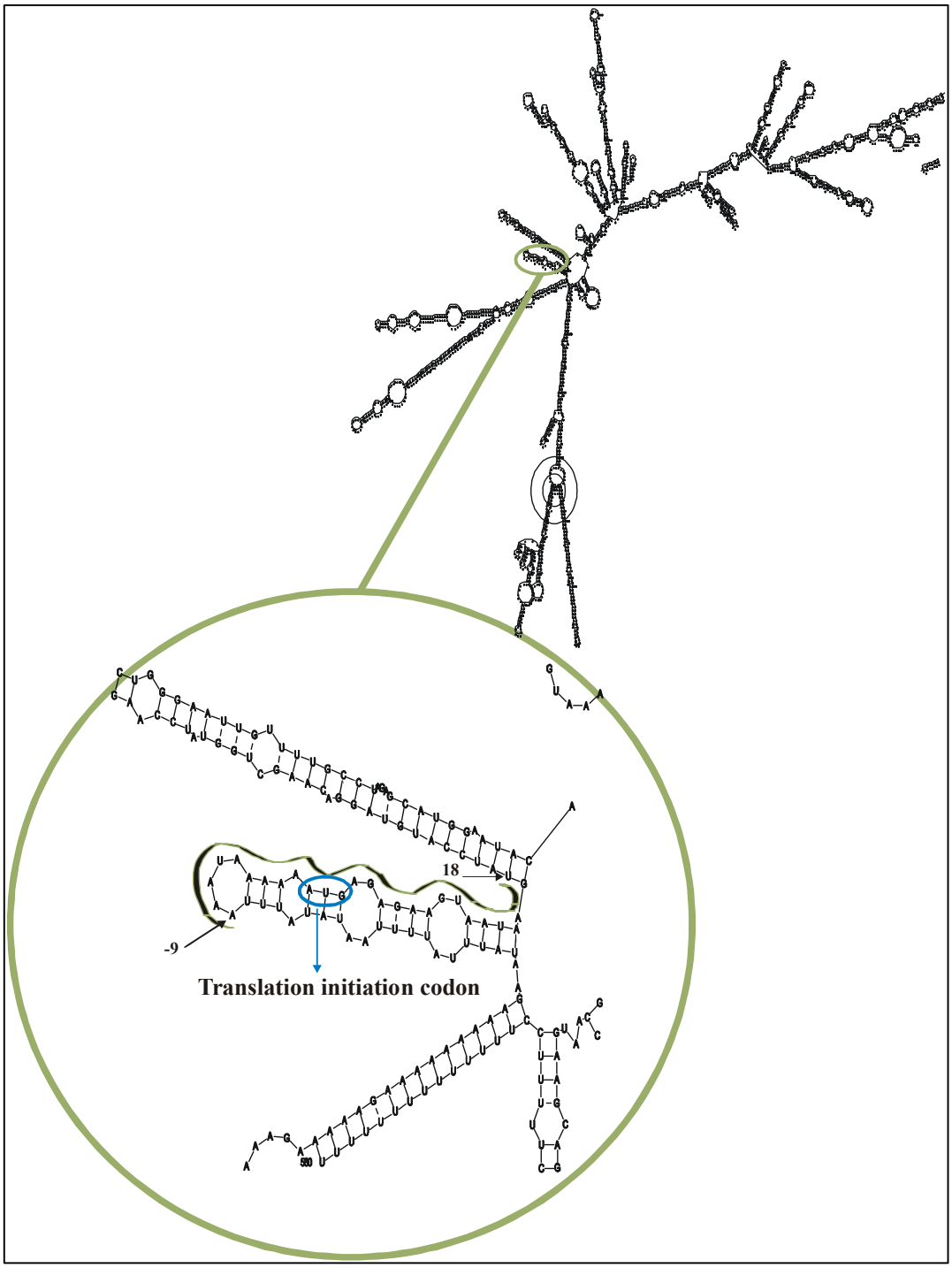
#### 2.3.1.1. Secondary structure prediction of *P. falciparum* mature α-I-tubulin mRNA

The optimal and sub-optimal foldings of *P. falciparum* α-I-tubulin mature mRNA were computed with the aid of the program, *Mfold* (Zuker, 1989). The optimal structure, i.e. the structure having the most favourable (minimum) ΔG value (–417.1 Kcal/mol), is shown in Figure 2.4A. The 17 sub-optimal structures, predicted to have ΔG values within 20% of that of the optimal structure, were generated in the same way. An example of such a structure is shown in Figure 2.4.B.




**A**

Figure 2.4: A: The predicted optimal structure of *P. falciparum* mature α-I-tubulin mRNA. The location of AS-1 is shown in green. Refer to Table 2.5 for all ODN sequences. These structures were generated by the program RNA Structure 3.5.




**B**

Figure 2.4 B: One example of the 17 sub-optimal structures of *P. falciparum* mature α-I-tubulin mRNA. The location of AS-1 is in green. Refer to Table 2.5 for all ODN sequences. These structures were generated by the program RNA Structure 3.5.

Significant deviations in the secondary structures were evident after comparison of all the predicted mRNA secondary structures. Literature reports have indicated that the most successful antisense ODNs are those that target single-stranded hairpin loop regions i.e. mRNA regions involved in minimal local structure formation (Bruice and Lima 1997; Lima *et. al*., 1997; Vickers *et. al*., 2000; Francois *et. al*., 1994). It was therefore important to analyse the optimal as well as the 17 predicted sub-optimal structures in order to identify the best possible antisense ODN hybridization site. This entailed inspection of all single stranded regions in an attempt to identify common or shared sequences predicted to be available for hybridization to an antisense ODN.

Possible ODN hybridization sites fitting the stated criteria were identified from the optimal and 17 predicted sub-optimal mRNA secondary structures as described below.

2.3.1.2 Identification of accessible hybridization sites and design of antisense ODNs.

The software program RNA Structure 3.5 was used to identify accessible binding sites of a 27-mer ODN on the mature mRNA secondary structures. The first base of the hybridization site for each of the 18 predicted mRNA secondary structures (one optimal and 17 sub-optimal structures), is given in Figure 2.5.

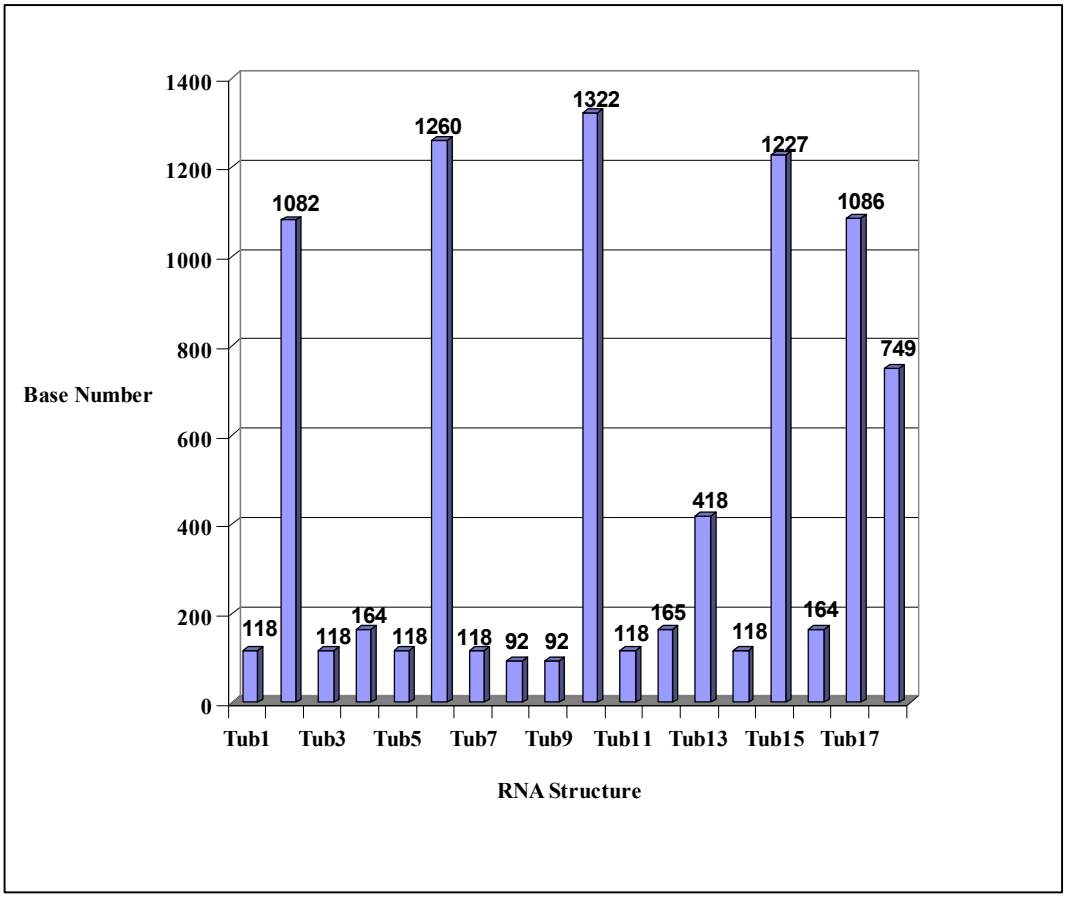



Figure 2.5 : Histogram depicting the first base of the mature mRNA sequence in each of the secondary structures of *P. falciparum* mature α-I-tubulin mRNA predicted by the program RNA Structure 3.5 to be optimal for hybridization to a 27-mer antisense ODN AS-1. Tub1 being the optimal structure.

The mRNA sequence starting at nucleotide 92 was predicted to be the most accessible hybridization site for a 27 mer antisense ODN in structures Tub 8 and 9. The sequence starting at nucleotide 118 was predicted to be the most accessible hybridization site for structures Tub 1, 3, 5, 7, 11 and 14 while the sequence commencing from nucleotide 164 was predicted to be the most accessible for the predicted structures of Tub 4 and 16, respectively.

The number of times a particular hybridization site is predicted to be accessible to the ODN in each of the different mRNA secondary structures, is given as % prevalence in Table 2.2.

Table 2.2: The percentage prevalence of the first base in each of the mature mRNA structures predicted by the RNA Structure 3.5 program to be the most accessible hybridization site for a 27-mer antisense ODN.

| First Base Number | % Prevelance |
|---|---|
| 92 | 11.10% |
| 118 | 33% |
| 164 | 11.10% |
| 165 | 5.60% |
| 418 | 5.60% |
| 749 | 5.60% |
| 1082 | 5.60% |
| 1086 | 5.60% |
| 1227 | 5.60% |
| 1260 | 5.60% |
| 1322 | 5.60% |

Three 27-mer antisense ODN binding sites starting at nucleotides 92, 118 and 164 of α-I-tubulin mRNA, are suggested as the most accessible in all the predicted mature mRNA secondary structures. Their sequences are shown in Table 2.3.

Table 2.3: Three *P. falciparum* α-I-tubulin mature mRNA sequences predicted to be accessible for hybridization of a 27-mer antisense ODN.

| mRNA Base Numbers | Sequence |
|---|---|
| 92 – 118 | 5’ AGC CCG AUG GUC AAA UGC CCU CUG ACA 3’ |
| 118 – 144 | 5’ AAG GCU UCU AGA GCU AAU GAU GAU GCU 3’ |
| 164 – 190 | 5’ AAA CGG GGG CAG GAA AAC AUG UAC CAC 3’ |

The secondary structure motifs of the three mRNA sequences displayed in Table 2.3 are shown in Figures 2.7 – 2.9. The position of the antisense ODN (AS-1) targeted against the region surrounding the translation initiation codon is also indicated.

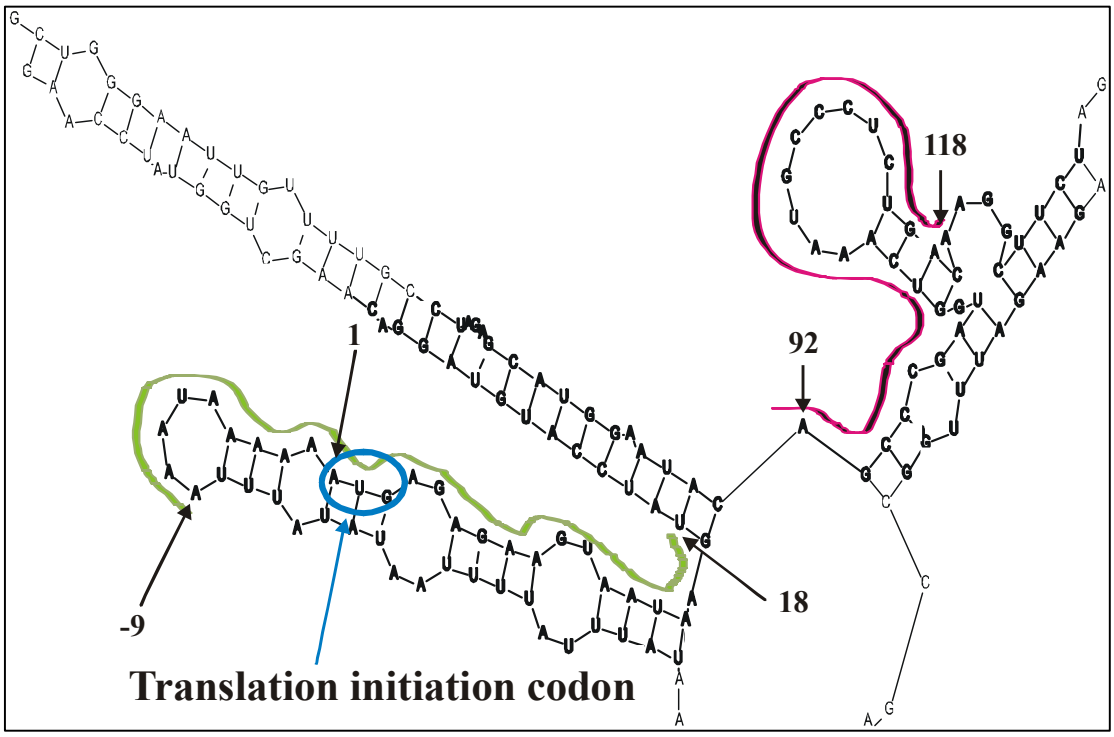
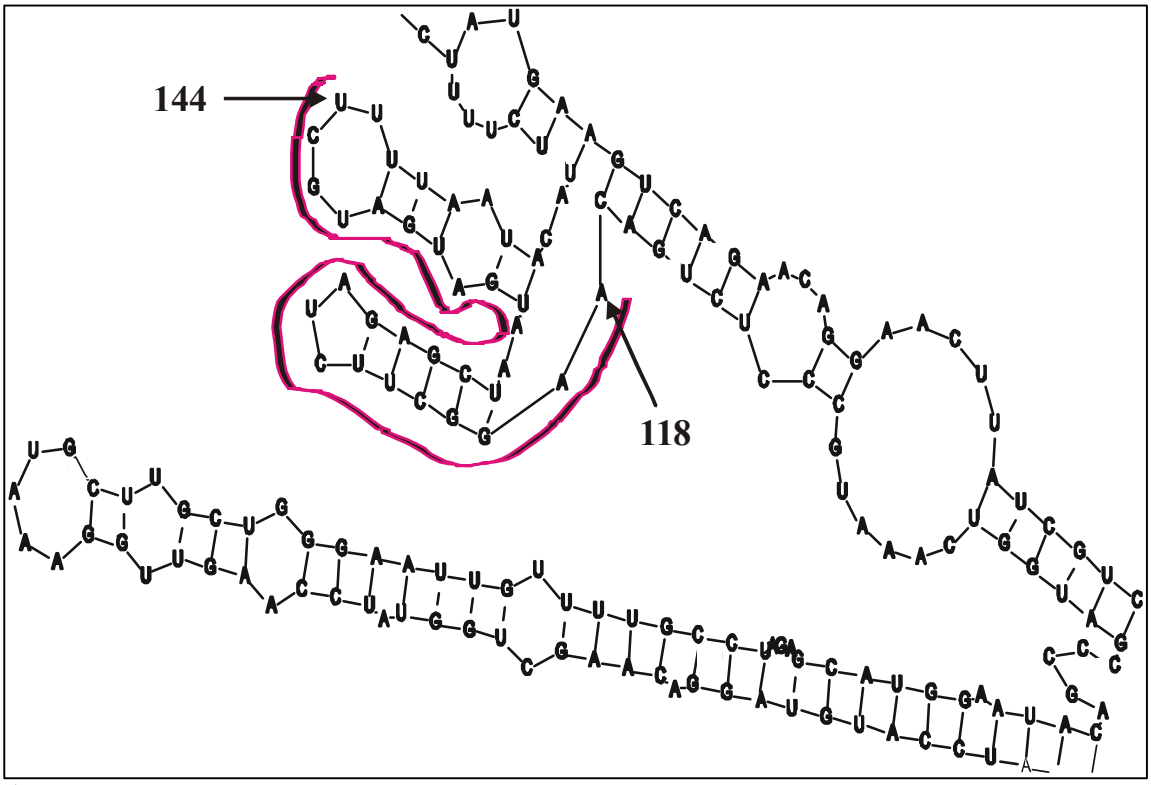



Figure 2.6: The mRNA secondary structure of nucleotides 92 to 118 (magenta area) and –9 to 18 (green area). The sequence 92 to 118 was predicted to be the optimal hybridization site for a 27-mer ODN in the predicted mRNA structures, Tub 8 and 9 (same secondary structure in the indicated regions). The sequence –9 to 18 is the region to which antisense ODN AS-1 is targeted and includes the initiation codon.

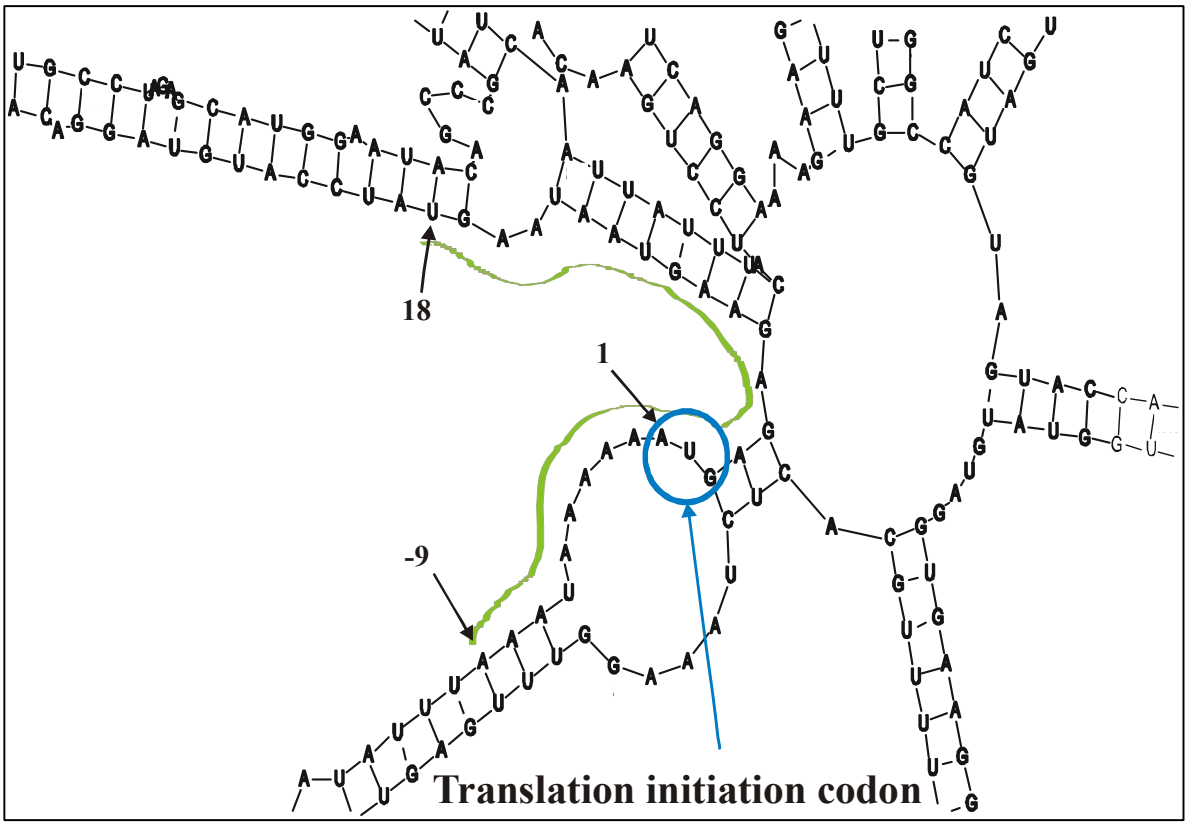
Part of an internal and a hairpin loop structure are common to both the predicted secondary mRNA structures Tub 8 and 9 in the region of bases 92 – 118. The bases –9 to 18 span three internal loops as well as part of a hairpin loop (Figure 2.6).




A

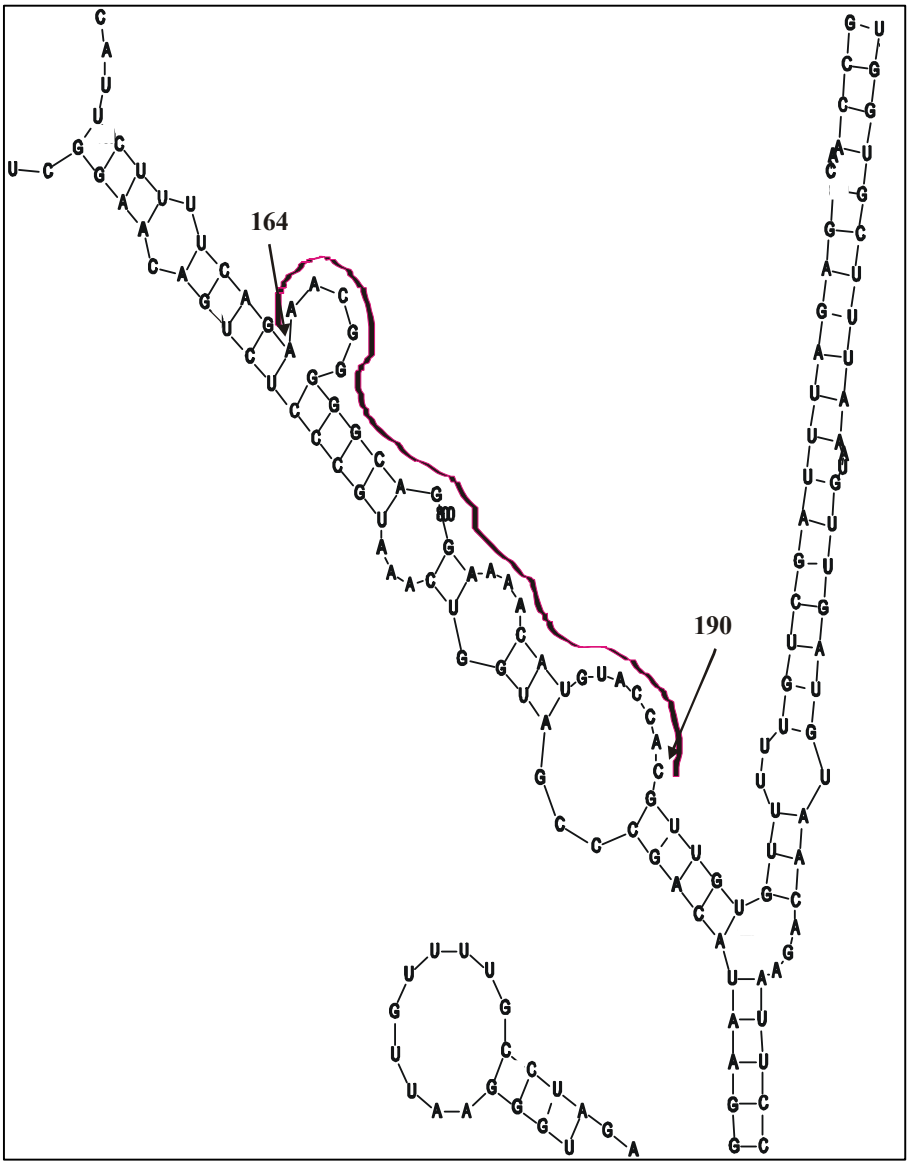
Figure 2.7 A: The optimal ODN hybridization sites in the predicted secondary mRNA structures of Tub 1, 3, 5, 7, 11 and 14: Nucleotides 118-144 (magenta area)(same secondary structure observed in the indicated region).

The nucleotide sequence 118 – 144 in the relevant mature mRNA structures where this sequence was predicted to be the optimal hybridization site, spans two hairpin loops and part of an internal loop (Fig 2.7 A). The bases –9 to 18 that were targeted by AS-1 include the initiation codon and span a bifurcation, an internal loop and part of a second bifurcation (Fig 2.7 B).



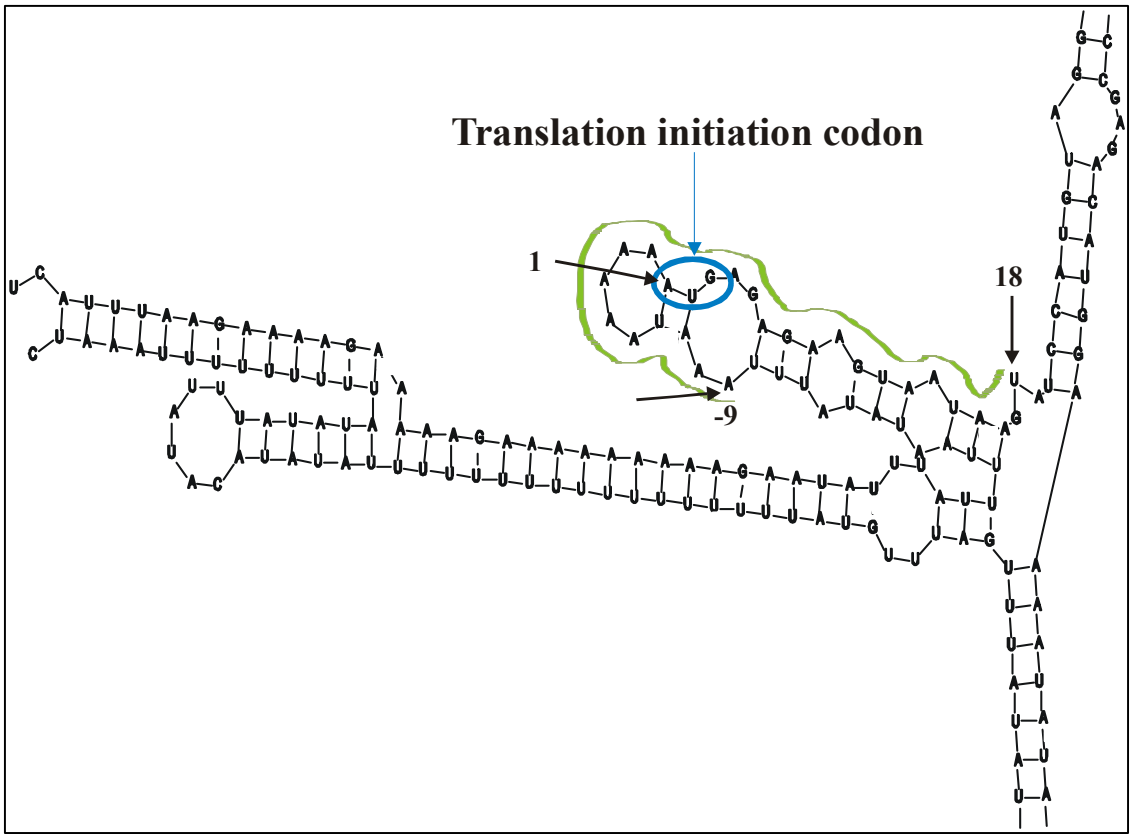

**B**

Figure 2.7: B: The optimal ODN hybridization sites in the predicted secondary mRNA structures of Tub 1, 3, 5, 7, 11 and 14: Nucleotides –9 to 18 (green area) is the region to which antisense ODN AS-1 is targeted and includes the initiation codon.




**A**

Figure 2.8: A : The most accessible ODN hybridization sites in the predicted secondary mRNA structures of Tub 4 and 16: Nucleotides 164-190 (magenta area)(same secondary structure observed in the indicated region).




**B**

Figure 2.8: B: The most accessible ODN hybridization sites in the predicted secondary mRNA structures of Tub 4 and 16: Nucleotides –9 to 18 (green area) is the region to which antisense ODN AS-1 is targeted and includes the initiation codon.

The predicted accessible nucleotide sequence 164 – 190 is shown to span one bulge and three internal loops (Fig 2.8 A). The sequence –9 to 18 targeted by AS-1 spans a hairpin loop, three internal loops and part of a bifurcation (Fig 2.8B).

Based on the criteria stated earlier on the properties of a ODN hybridization site and the results obtained from secondary structure predictions of α-1-tubulin mRNA, the most promising accessible binding areas for a 27-mer antisense ODN appear to include the

nucleotides 92 - 118 and 118 –144 (Figures 2.6 and 2.7A and Table 2.3). Both of these sequences span hairpin loop structures in the relevant predicted mRNA secondary structures. The nucleotides 164 – 190 do not span a hairpin loop structure in the relevant predicted mRNA structures and this sequence was therefore not included as a suitable antisense ODN binding site.

2.3.1.3. Design of an antisense ODN targeted against the translation initiation codon

The antisense ODN targeted against the translation initiation codon region, along with the sense and mismatch control sequences are shown in Table 2.4.

Table 2.4: Sequences and Tm values of the antisense ODN (AS-1), targeted against *P. falciparum* alpha-I-tubulin initiation codon region as well as the sense and mismatch control sequences. The translation initiation codon is underlined.

| **ODN Name Nt** | **SEQUENCE** | **Tm[1]** |
|---|---|---|
| Antisense ODN (AS-1) | 5' ACT TAT TAC TTC TCT CAT TTT TTA TTT 3’ | 64.1 °C |
| Mismatch ODN (MS-1) | 5' TTT ATT TTT TAC TCT CTT CAT TAT TCA 3’ | 64.1 °C |
| Sense ODN (S-1) | 5’ AAA TAA AAA ATG AGA GAA GTA ATA AGT 3’ | 64.1 °C |

[1]81.5 + 16.6log1 + 0.41x(%GC) – (675/L), L=27 (%GC method) (Howley, 1979).

Fifteen bases downstream, nine bases upstream from and including the initiation codon were targeted by the antisense ODN. The mismatch control is the inverse of the antisense

sequence to ensure a similar base composition between the antisense and mismatch control ODNs (Francois *et. al*., 2000).

The 3' terminal of the antisense ODN, as well as the sense and mismatch control sequences were structurally modified by the addition of a 15 base self-complementary 3'-loop in an attempt to protect the phosphodiester ODNs against 3'exonuclease degradation (Barker *et. al*, 1998). The modified ODNs i.e. the ODNs containing this loop, are referred to as ASL-1, MSL-1, SL-1, ASL-2, MSL-2, ASL-3 and MSL-3, respectively. All ODNs were investigated for sequence specificity against the *P. falciparum* and human sequences available at the time of ODN design (1997/1998) as well as retrospectively (April 2005) and are shown and discussed in Chapter 4.

The full sequences of the ODNs including the chosen loop modification, are shown in Table 2.5.

Table 2.5: Sequences and Tm values of the structurally modified antisense ODN, ASL-1 targeted against *P. falciparum* α-I-tubulin initiation codon (ODN: ASL-1, together with the sense and mismatch control sequences) and *P. falciparum* α-I-tubulin internal sequences (ASL-2 and ASL-3, together with mismatch control sequences). The 3'-loop region is shown in bold and the translation initiation codon is underlined.

| ODN Name | **SEQUENCE** | **Tm**[1] |
|---|---|---|
| Antisense ODN (ASL-1) | 5' ACT TAT TAC TTC TCT CAT TTT TTA TTT **GCG CGC TTT GCG CGC** 3' | 64.1 °C |
| Mismatch ODN (MSL-1) | 5' TTT ATT TTT TAC TCT CTT CAT TAT TCA **GCG CGC TTT GCG CGC** 3' | 64.1 °C |
| Sense ODN nt -9 – 18 (SL-1) | 5' AAA TAA AAA ATG AGA GAA GTA ATA AGT **GCG CGC TTT GCG CGC 3'** | 64.1 °C |
| Antisense ODN (nt 92 – 118) (ASL – 2) | 5' TGT CAG AGG GCA TTT GAC CAT CGG GCT **GCG CGC TTT GCG CGC** 3' | 79.2 °C |
| Mismatch ODN (MSL-2) | 5' TCG GGC TAC CAG TTT ACG GGA GAC TGT **GCG CGC TTT GCG CGC** 3' | 79.2 °C |
| Antisense ODN (nt 118 – 144) (ASL – 3) | 5' AGC ATC ATC ATT AGC TCT AGA AGC CTT **GCG CGC TTT GCG CGC** 3' | 73.2 °C |
| Mismatch ODN (MSL-3) | 5' TTC CGA AGA TCT CGA TTA CTA CTA CGA **GCG CGC TTT GCG CGC** 3' | 73.2 °C |

[1]81.5 + 16.6log1 + 0.41x (%GC) – (675/L), L= 27 (% GC method) **(**Howley, 1979**).** The Tm of the hairpin loop alone (GCG CGC TTT GCG CGC) was calculated to be 69.3 °C with a ΔG value of –40.3 kcal/mol.

Analysis with the Oligo 4.1 Primer Analysis Software (National Biosciences Inc.) indicated no dimer or loop formations, other than the 15-base self-complementary 3' terminal loop.

### *2.4 Discussion*

The choice of *P. falciparum* α-I-tubulin mRNA as the target for antisense ODNs in these studies was based on the central role that microtubules, of which tubulin is the major component, play in vital cellular functions such as cell division, to name but one (Jordan and Wilson, 2004). The α-I-tubulin protein, in particular is known to have regulatory functions in parasite glycolysis (Itin *et. al*., 1993). The choice of tubulin as target protein was further supported by the fact that numerous anticancer agents, inhibiting microtubule functioning were also effective inhibitors of *P. falciparum* proliferation (Bell, 1998). The problem with this approach however is the specific inhibition of the parasite tubulin without adversely affecting the mammalian protein. This aim is complicated by the fact that there is an 83% amino acid identity between the mammalian and *P. falciparum* α-tubulin proteins (Holloway *et al*., 1989). Antisense ODNs, that target the mRNA and thus the translation of the protein target and not the protein itself, are advantageous in this particular instance due to the 80% A + T rich *P. falciparum* genome as opposed to the 60% G + C rich human genome (Gardner *et al*., 2002). This difference in nucleotide composition enables the specific targeting of the *P. falciparum* α-I-tubulin mRNA.

The design of three phosphodiester antisense ODNs targeting different regions of *P. falciparum* α-I-tubulin mature mRNA is described in this chapter. RNA secondary

structures have been shown to affect the hybridization ability and hence the efficacy of antisense ODNs (Crooke, 2000; Smith *et al*, 2000; Sohail and Southern, 2000; Vickers *et al*, 2000). Two different approaches were used in the design of the ODNs. In the first approach the ODN was targeted to regions adjacent to and including the translation initiation codon. This strategy was chosen due to the results of the majority of studies reporting successful antisense-mediated inhibition by ODNs targeted to this region (Crooke, 2000; Francois *et al*., 1994; Sohail and Southern, 2000). In the second approach the predicted secondary structures of the mature mRNA were used in the design of two further ODNs targeted to the predicted accessible sites for ODN hybridization. Inclusion of the latter strategy was important since mRNA is known to have secondary and tertiary structures rendering certain sites inaccessible to ODN hybridization (Ho *et al*., 1996).

Several groups have reported the utility of computer-generated secondary structures to predict ODN hybridization sites (Cirino *et al*, 1997; Lima *et al*, 1992; Thierry *et. al*., 1993) whilst others have criticized this approach (Smith *et al*, 2000; Sohail and Southern, 2000). The points of criticism were aimed at two aspects of computer predictions. The first one being that tertiary structures are not taken into account when determining accessible antisense ODN target sites with computer algorithms used to predict the secondary mRNA structures. The second criticism is that large RNA molecules probably do not reach one global minimum energy structure, i.e. one optimal structure (Smith *et al*, 2000; Sohail and Southern, 2000). A structural hierarchy exists in the unfolding and folding process of RNA molecules, with secondary structures forming first (Leontis and Santa-Lucia, 1997). Tertiary structures then ensue from interactions between these

secondary structures and the reverse process takes place when the RNA molecules unfold. This structural hierarchy therefore refers to the preferential unfolding of the less stable tertiary structures compared to the more stable secondary structures. Earlier observations along these same lines led Tinoco and co-workers (1973) to postulate that the use of computer algorithms to predict secondary mRNA structure could indeed be useful in the identification of hybridization sites. Furthermore, most of the computer algorithms used to predict RNA secondary structures can generate optimal and sub-optimal structures, thereby fully acknowledging the fact that no single optimal RNA structure is continually present (Leontis and Santa-Lucia, 1997).

Computer generated RNA structures therefore appear to be useful due to the saving in time and cost, which is considerable compared with the random shot-gun or gene walking approach to identify successful RNA target sites for antisense ODNs. The number of possible antisense ODNs to be tested is narrowed down significantly. For example, calculations showed that 1981 oligonucleotides should be tested individually in order to identify the best 20-mer antisense ODN to target a 2-kilobase mRNA molecule when using the random shot gun approach (Sohail and Southern, 2000).

Of the three possible programs available for secondary structure predictions, the *Mfold* program allows the determination of optimal and sub-optimal secondary structures at a physiological temperature of 37 °C. The optimal as well as sub-optimal secondary structures of α-I-tubulin mature mRNA were determined. This was deemed necessary due firstly, to the fact that RNA molecules are dynamic structures and secondly due to the

fact that slight differences in the free energy parameters may have profound effects on the secondary structure of the RNA molecule. It is therefore important to utilize methods that allow an approximation of these parameters. *Mfold* utilizes three methods to achieve these goals.

Firstly an empirical method is used, which is based on the free energies of optical melting experiments for representative sequences. A second method bases the assignment of stability (or free energy) to unpaired regions on the frequency with which sequences occur in a specific motif in a database of phylogenetic structures. The third method varies the free energy parameters so as to optimize the accuracy of the computer-generated sequences against their phylogenetically determined structures (Riesner, 1997). The *Mfold* program is therefore capable of predicting optimal and sub-optimal structures for comparison in order to determine which regions remain the same upon slight changes of these free energy parameters (Riesner, 1997).

Eighteen secondary structures of the mature mRNA sequence of α-I-tubulin were predicted by *Mfold*, one optimal structure and 17 sub-optimal structures, the latter having ΔG values within 20% of the ΔG value of the optimal structure. Three potential (accessible) target sequences were identified (Figs 2.6 – 2.8 and Table 2.3) that incorporated nucleotides 92-118; 118-144 and 164-190). The identification of these three sequences was primarily based on the percentage prevalence, i.e. the number of structures in which the relevant sequence was predicted to be the most accessible hybridization site (Smith *et al*, 2000).

The sequence 118-144 was the most prevalent (33%) with sequences 92 -118 and 164-190 being accessible in 11% of the structures (Table 2.2). All the other sequences were found to be accessible in only 5.6% of the predicted structures and therefore not included in the list of potential target sites. The sequence 118-144 included 2 hairpin loops and part of an internal loop while the sequence 92-118 contained a hairpin loop and an internal loop. Sequence 164-190 on the other hand included 1 internal bulge and 3 internal loops. Sequences 92-118 and 118-144 were chosen as targets for antisense ODNs due to the presence of hairpin loop structures, which were shown to be the most effective targets for antisense ODNs and needed to initiate heteroduplex formation (Eguchi *et. al*., 1991; Frauendorf and Engels, 1996; Southern *et. al*., 1994; Southern *et. al*., 1999; Toulme *et. al*., 1996; Lima *et al.,* 1992; Thierry *et al.,* 1993). Studies using yeast tRNA which is structurally well characterized in solution, revealed that stable secondary structures prevent heteroduplex formation (Sohail and Southern, 2000).

The choice of sequence –9 to 18 targeted by AS-1 was based on the premise that this region is less obstructed by secondary structure formation due to the presence of amongst others, the ribosome binding site and the translation initiation codon (Sohail and Southern, 2000). Four of the predicted structures (Tub 4, 8, 9 and 16) contained a hairpin loop along with internal loops in this translation initiation region. This means that 22% of the predicted *P. falciparum* α-I-tubulin mRNA secondary structures contained a hairpin loop structure in this region. This compares favourably with sequences 92-118 and 118-144 in so far as each of these also contained at least 1 hairpin loop structure.

All ODNs were 27 nucleotides long, an acceptable length for specificity since a 17-mer ODN should already find a unique target within any genome (Francois *et al*, 1994). There is no fixed maximum limit for antisense ODN length, but generally the hybridization regions of the ODNs are in the range of 15-25 nucleotides (Phillips and Zhang, 2000; Riesner, 1997). However, since efficient cellular uptake is a prerequisite for all antisense ODNs and decreases with increasing ODN length, the latter should be kept as short as is possible while still allowing the achievement of sequence specific inhibition of target protein translation (Crooke, 2000). Furthermore, since ODN modifications and cell lines also have an influence on cellular ODN uptake, the decision of the length of ODN is largely an empirical one. In this particular study the choice of ODN length was also influenced by the high A + T content of the *P. falciparum* genome, a characteristic that decreases the Tm of any ODN targeted against the parasite DNA or RNA (Gardner, 2001). The lower Tm could in turn negatively influence the hybridization efficacy of the ODN. For this reason a longer 27-mer ODN with a correspondingly higher Tm was decided upon rather than a shorter, 15-mer ODN for example.

The inclusion of the correct control sequences is also extremely important in the design of antisense experiments. The selected mismatch ODN sequences were the inverse of their respective antisense sequences. This ensured a difference only in the base sequence and not the composition, which is regarded as a standard requirement for evaluation of antisense ODN results (Francois *et al*, 1994).

The effect of the self-stabilizing loop structural modification on the stability of phosphodiester ODNs in culture medium, targeted against *P. falciparum* mRNA has not yet been investigated. The stability of the three ODNs designed in this chapter was therefore investigated in culture medium supplemented with serum and the results are reported and discussed in the next chapter.