

UNIVERSITÀ DEGLI STUDI DI MILANO - BICOCCA
Dipartimento di Biotecnologie e Bioscienze
Dottorato di ricerca in Biotecnologie Industriali - XXVIII ciclo



**Development of a Software Application for
Loop-Mediated Isothermal Amplification (LAMP)
Primer Design**

Coordinatore: Prof. Marco Vanoni

Tutor: Prof. Luca De Gioia

Chiara Montrasio
Matr. 775146

Anno Accademico 2014/2015

Dottorato in Biotecnologie Industriali, XXVIII ciclo

Chiara Montrasio

Matr. 775146

Coordinatore: Prof. Marco Vanoni

Tutor: Prof. Luca De Gioia



Il lavoro presentato in questa tesi è stato realizzato presso i laboratori di Diagnostica Molecolare DiaSorin, in collaborazione con il Dott. Luciano Ruo Rui e sotto la supervisione della Dr.ssa Giulia Minnucci e del Dott. Francesco Colotta.

Il materiale contenuto nella seguente tesi è strettamente confidenziale. E' stata inoltre presentata richiesta di embargo tesi per un periodo di 3 anni dalla data di conseguimento del titolo.

The logo for DiaSorin, consisting of the word 'DiaSorin' in a white, serif font, centered within a dark blue rectangular background.

INDEX

RIASSUNTO	5
INTRODUCTION	15
1. LOOP-MEDIATED ISOTHERMAL AMPLIFICATION	15
1.1 - LAMP principle	16
1.2 - Loop primers	20
1.3 - Detection of LAMP amplification products	22
1.4 - RT-Q-LAMP	23
2. PRIMER DESIGN	26
2.1 - Primer design: crucial step in a LAMP assay development	26
2.2 - LAMP primer design: key factors	27
2.2.1 - Key factors for individual primers design	29
2.2.2 - Key factors for primer sets design	33
3. DNA THERMODYNAMICS IN PRIMER DESIGN	36
3.1 - Solution equilibrium and calculation of the amount bound	36
3.2 - The meaning of ΔH° , ΔS° and ΔG_T° parameters	37
3.3 - Nearest-Neighbor (NN) model	40
3.4 - Prediction of the Melting Temperature (T_m)	43
3.5 - Multi-state model for T_m prediction	44

4. LAMP PRIMER DESIGN: CURRENT OPTIONS	46
4.1 - Primer Explorer V4	46
4.2 - Visual OMP	49
4.3 - Other software for LAMP primer design: LAMP designer	51
AIM OF THE THESIS	53
MATERIALS AND METHODS	57
1. THE ALGORITHM FOR PRIMER DESIGN: BASIC SETTINGS	57
2. LAMPZERO PRIMER DESIGN: DESIGN MODE OVERVIEW	60
3. ASSISTED PRIMER DESIGN MODE	62
4. AUTOMATIC PRIMER DESIGN MODE	65
4.1 - Automatic primer design: selection of the primer candidates	67
4.2 - Automatic primer design: selection of the primer set candidates	70
5. CALCULATION OF THE CONCENTRATION CURVES AND DETERMINATION OF THE MELTING TEMPERATURES	73
6. SCORE CALCULATION ALGORITHM	76
6.1 - Score calculation algorithm for individual primers	77
6.2 - Score calculation algorithm for primer sets	79

7. T _m & DIMERS PREDICTIONS: LAMPZERO vs VISUAL OMP	81
8. CHARACTERISTICS AND REQUIREMENTS	83
RESULTS	85
1. SOFTWARE PROTOTYPE BUILDING	86
2. VERIFICATION OF LAMPZERO PREDICTIONS: COMPARISON WITH VISUAL OMP	88
3. LAMPZERO PRIMER DESIGN: DEFINING DESIGN CRITERIA AND REACTION CONDITIONS	89
3.1 - Target selection	89
3.2 - Solution conditions	91
3.3 - Design criteria	92
4. LAMPZERO ASSISTED PRIMER DESIGN	96
4.1 - Dumbbell visualization	99
4.2 - Reports	102
5. LAMPZERO AUTOMATIC PRIMER DESIGN	103
6. PRIMERS DIMERS EVALUATION IN MULTIPLEX SYSTEMS	108
DISCUSSION	113
REFERENCES	125
PATENT	131

RIASSUNTO

LAMP (*Loop-mediated isothermal AMPlification*) è una tecnica innovativa per l'amplificazione e la rilevazione di acidi nucleici. A differenza della PCR, che è caratterizzata dall'alternanza di fasi (denaturazione, appaiamento dei primer, estensione) a differenti temperature, LAMP avviene a temperatura costante, comunemente tra 63°C e 65°C, grazie all'utilizzo di una DNA polimerasi dotata di attività di *strand displacement*, ed è caratterizzata dall'uso di un totale di 6 primer: 2 primer interni (FIP e BIP, costituiti rispettivamente da due distinte regioni nucleotidiche), 2 primer esterni (F3 e B3) e 2 *loop* primer (LF ed LB), che fungono da catalizzatori della reazione di amplificazione. Di conseguenza, il disegno dei primer per un saggio LAMP richiede la selezione di 8 differenti sequenze nucleotidiche all'interno della regione genomica d'interesse.

Una selezione accurata dei primer, delle loro temperature di melting (T_m), nonché delle loro concentrazioni, risulta essere di fondamentale importanza, affinché il loro appaiamento sulla sequenza bersaglio sia finemente "orchestrato" alla temperatura costante di reazione: per questa ragione, la fase di disegno dei primer è cruciale nello sviluppo di un saggio LAMP.

Attualmente, i metodi a disposizione per il disegno di primer LAMP sono piuttosto limitati e si basano principalmente sull'utilizzo di Primer Explorer V4 (Eiken Chemical Co.), un software gratuito per il disegno automatico di primer specifici per la tecnologia LAMP.

Tuttavia, esso presenta diverse limitazioni. In primo luogo, non fornisce alcuna opzione di disegno manuale sulla sequenza d'interesse, caratteristica che impedisce di apportare modifiche ai primer selezionati dal software e di riposizionarli nel punto desiderato della regione genomica, ad esempio in corrispondenza di una mutazione. Inoltre, il disegno dei *loop* primer non viene effettuato in maniera simultanea al disegno dei 4 primer base, ma richiede un passaggio aggiuntivo, che può impedire verosimilmente di identificare tutte le possibili combinazioni di primer. In aggiunta, Primer Explorer non tiene in considerazione alcuna eventuale modifica alle condizioni sperimentali di un saggio LAMP (temperatura, concentrazione salina, ecc.), fornendoci calcoli delle temperature di melting poco accurati. Infine, esso non garantisce la protezione della proprietà intellettuale, essendo un software gratuito e disponibile tramite Internet.

Il disegno automatico di Primer Explorer può essere implementato mediante Visual OMP™ (DNA Software, CEO J. SantaLucia Jr.), un software per il disegno "*in silico*" di primer e sonde per PCR. Esso fornisce dei valori di T_m più affidabili, calcolati tenendo in considerazione le effettive condizioni sperimentali della reazione. Inoltre, permette di valutare la formazione di possibili strutture secondarie tra i primer presenti in soluzione, permettendoci di discriminare con maggiore precisione i migliori set di primer automaticamente generati da Primer Explorer. Tuttavia, si tratta di un software molto costoso, specifico per la PCR e che non tiene conto della complessità del metodo LAMP.

Pertanto, attualmente non esiste uno strumento efficace e veloce per la valutazione "*in silico*" dei set di primer specifici per LAMP. Di conseguenza,

la selezione dei migliori set di primer viene effettuata sperimentalmente, attraverso una procedura piuttosto lunga e laboriosa che consiste nel testare all'interno della reazione i vari set di primer e le loro possibili combinazioni.

Lo scopo di questa tesi è stato quello di sviluppare un software unico per il disegno dei primer, realizzato "su misura" in base ai requisiti della tecnologia LAMP sviluppata da DiaSorin. La creazione di un tale strumento può garantire l'esistenza di un sistema di disegno di primer esclusivo di proprietà aziendale, che assicuri la proprietà intellettuale, possa essere ulteriormente implementato in base alle nostre richieste e, soprattutto, possa dare un contributo significativo ad uno sviluppo più efficiente dei saggi LAMP.

Al fine di sviluppare tale software, abbiamo preso in considerazione una serie di regole che devono essere seguite per disegnare correttamente primer: alcune di esse sono regole generali, pertanto valide anche per PCR, mentre altre sono specifiche per la tecnologia LAMP.

In particolare, i criteri principali di selezione utilizzati dal software, per individuare i singoli primer data una sequenza bersaglio, includono:

- lunghezza: la lunghezza ottimale dei primer è generalmente 18-22 paia di basi. Assieme alla composizione in basi, essa influenza direttamente la temperatura di melting (T_m) del primer.
- contenuto di GC: corrisponde alla percentuale di nucleotidi Guanina e Citosina presenti nella sequenza del primer e dovrebbe aggirarsi attorno al 40-60%. Questo parametro influisce sul calcolo della temperatura di melting (T_m) e sulla lunghezza del primer.

- temperatura di melting (T_m): corrisponde alla temperatura alla quale il 50% delle molecole di DNA si trova in forma di doppia elica stabile ed il restante 50% in forma di singola elica e rappresenta un indice della stabilità dell'eteroduplex a DNA formato da primer e template. Essa è influenzata da vari fattori, quali la lunghezza del primer, il contenuto dei vari nucleotidi che lo compongono, ma anche le condizioni sperimentali di reazione (ad esempio la concentrazione dei sali e dei primer stessi). In base ai parametri forniti dal manuale di Primer Explorer, la T_m dovrebbe corrispondere a 65°C ($64 - 66^\circ\text{C}$) per le regioni F1c e B1c dei primer interni e per i *loop* primer e a 60°C ($59 - 61^\circ\text{C}$) per le regioni F2 e B2 dei primer interni e per i primer esterni F3 e B3.
- legame non specifico: i primer dovrebbero legarsi solo alla sequenza target, per garantire un'elevata specificità alla reazione di amplificazione. Di conseguenza, è importante evitare possibili regioni di omologia totale o parziale tra il primer e altre regioni genomiche al fine di evitare la formazione di prodotti aspecifici di reazione.
- stabilità dell'estremità del primer: la porzione 3' del primer rappresenta il punto di partenza per la reazione enzimatica di polimerizzazione del DNA e quindi deve avere un certo grado di stabilità termodinamica. In particolare, le estremità 3' delle regioni F2/B2, F3/B3, LF/LB e le estremità 5' delle regioni F1c/B1c sono diseguate in modo che l'energia libera (ΔG calcolato per le ultime 6 paia di basi del primer) sia inferiore a -4kcal/mol : un valore di ΔG

negativo, infatti, garantirà una maggiore stabilità di legame del primer al target in corrispondenza del punto di inizio della reazione di polimerizzazione.

- strutture secondarie intramolecolari (*hairpin*): le strutture ad *hairpin* (o forcina) si formano per l'interazione intramolecolare tra i nucleotidi del singolo primer. E' importante che la loro formazione venga evitata, dal momento che possono influenzare negativamente la resa della reazione di amplificazione.

E' importante sottolineare come la costruzione di un buon set di primer si basi anche sull'identificazione di una combinazione ottimale di più primer, che possano agire in maniera sincronizzata sulla sequenza bersaglio. Per questo motivo, vi sono ulteriori criteri di selezione utilizzati dal software per valutare i set di primer, che comprendono:

- distanza tra i primer: Primer Explorer suggerisce i criteri di distanza ottimale tra i primer da tenere in considerazione durante la fase di disegno. In particolare, la distanza tra il 5' della regione F2 e il 5' della regione B2 (ossia la lunghezza dell'amplicone) dovrebbe essere compresa tra 120 e 160 paia di basi; la distanza tra F2 ed F3 e fra B2 e B3 dovrebbe corrispondere a 0-60 paia di basi; la lunghezza delle regioni che costituiscono i *loop* dovrebbe essere compresa tra 40 e 60 paia di basi.
- bilanciamento delle temperature di melting: affinché una reazione LAMP sia efficiente, è importante non solo che i singoli primer abbiano le corrette temperature di melting, ma anche che esse

siano opportunamente bilanciate tra le varie coppie di primer *forward* e *backward* (F3/B3, F2/B2, F1c/B1c, LF/LB).

- strutture secondarie intermolecolari (omodimeri ed eterodimeri): esse si formano per l'interazione intermolecolare tra due molecole di primer dello stesso tipo (omodimeri) o di tipo diverso (eterodimeri) nelle regioni di omologia di sequenza. E' importante che la loro formazione venga evitata, dal momento che possono compromettere la resa e la specificità della reazione di amplificazione.

Considerando i fattori chiave per il disegno di primer LAMP, ci siamo focalizzati sul calcolo delle temperature di melting, sulla valutazione della stabilità delle estremità dei primer e sulla formazione di strutture secondarie.

A questo scopo, abbiamo preso in considerazione il valore di ΔG , che è un indice della spontaneità termodinamica di una reazione e influenza questi tre importanti parametri per il disegno dei primer. Per questa ragione, un calcolo accurato dei valori di ΔG è estremamente importante, al fine di ottenere delle predizioni ottimali.

ΔG viene normalmente calcolato attraverso il metodo di *Nearest-Neighbor* e questi valori si trovano nelle pubblicazioni di riferimento (ΔG , ΔS , ΔH , a 37°C e 1M NaCl). Tali parametri termodinamici, circa 500, relativi a tutte le possibili combinazioni nucleotidiche (coppie Watson-Crick, *mismatch*, *loop*, ecc.) sono stati inseriti nel database del nostro software,

assieme a formule per calcolare i valori di ΔG a diverse temperature e concentrazioni saline.

Inoltre il sistema è stato integrato con un algoritmo di ibridazione che, partendo dal database dei parametri termodinamici corretto, utilizza il calcolo della ΔG per individuare le strutture di DNA più stabili tra le moltissime combinazioni possibili. Questa integrazione permette al software di calcolare efficacemente le temperature di melting dei primer nelle diverse condizioni di reazione (mediante il calcolo della ΔG del dimero primer-target), nonché di predire l'eventuale formazione di strutture secondarie (mediante il calcolo della ΔG di *hairpin* e dimeri).

L'efficacia del nostro sistema è stata verificata mediante test di comparazione con Visual OMP™, da noi utilizzato come sistema di riferimento per la predizione delle temperature di melting e dei dimeri di primer. Un totale di 20 set di primer sono stati valutati con entrambi i sistemi: per ciascun set sono state calcolate ΔG , ΔH e T_m di 27 combinazioni di primer (omodimeri, eterodimeri e eteroduplex primer-target) e 6 monomeri (*hairpin*).

Il confronto tra i due sistemi ci ha permesso di stabilire come essi siano comparabili nella predizione delle temperature di melting, nonché delle strutture secondarie dei primer.

Abbiamo sviluppato un sistema di disegno assistito e un sistema di disegno automatico.

Il disegno assistito aiuta efficacemente l'operatore durante la selezione manuale dei primer lungo la sequenza di interesse. Esso consente di

ottenere in tempo reale le caratteristiche dei primer disegnati sul target (lunghezza, % GC, Tm, omo o eterodimeri, ecc), le quali vengono marcate in verde (se corrette) o in rosso (se errate) sulla base dei criteri di accettazione inseriti nel sistema.

Il sistema di disegno automatico è in grado di selezionare i migliori set di primer fra tutti quelli automaticamente generati dal software e viene effettuato mediante due fasi principali:

- una prima fase in cui il sistema valuta svariate migliaia di primer individualmente e attribuisce loro un punteggio;

- una seconda fase, in cui i primer che hanno superato la prima selezione vengono combinati fra di loro a generare fino a diverse migliaia di set di primer, classificati mediante un punteggio che permette di identificare i migliori.

La definizione di criteri di accettazione affidabili è, quindi, fondamentale per la selezione dei set di primer ottimali da parte del sistema. I criteri e le modalità attraverso cui viene attribuito il punteggio sono in fase di valutazione.

I set di primer ottenuti con la modalità automatica possono essere trasferiti nella modalità assistita, per essere modificati secondo le necessità dell'utente (sistema di disegno combinato).

Quindi, dopo aver selezionato la sequenza di interesse, le condizioni di reazione e i criteri per il disegno dei primer, il ricercatore può scegliere di effettuare il disegno dei primer attraverso tre differenti modalità: assistita, automatica o combinata.

Inoltre, il software è stato implementato attraverso le considerazioni derivanti dall'esperienza maturata da DiaSorin nello sviluppo della tecnologia LAMP.

Abbiamo sviluppato uno strumento per la rappresentazione grafica dei *dumbbell*, strutture la cui formazione rappresenta il punto di partenza della reazione di amplificazione esponenziale isoterma caratteristica della tecnica LAMP. La predizione della loro struttura, ottenuta mediante l'algoritmo di *folding*, è importante nella fase di disegno e progettazione dei primer, in quanto ci permette di valutare parametri quali la dimensione dei *dumbbell*, la loro simmetria, la dimensione dei *loop*, la posizione ottimale dei *loop* primer sulla sequenza bersaglio, nonché la formazione di *loop* imprevisti ed indesiderati.

In aggiunta, abbiamo sviluppato uno strumento per valutare la formazione di dimeri di primer in saggi multiplex, in cui più trascritti di interesse sono amplificati da diversi set di primer in una singola reazione. Esso è in grado di calcolare i principali parametri termodinamici (ΔG , ΔS , ΔH) nonché le temperature di melting delle strutture secondarie (*hairpin*, omodimeri, eterodimeri) che si possono generare fra le molecole di acidi nucleici presenti all'interno della miscela di reazione. Queste strutture vengono visualizzate graficamente, per consentire una migliore valutazione del grado di appaiamento esistente. Inoltre, il sistema fornisce informazioni riguardanti l'estensibilità dei dimeri di primer, ossia la possibilità che essi vengano allungati mediante una reazione di polimerizzazione a partire dalla loro estremità 3', permettendo così di individuare i primer potenzialmente più pericolosi.

Questo strumento è in grado di simulare con maggiore precisione le reali condizioni di interazione dei primer in una reazione multiplex, consentendo di effettuare uno screening preliminare degli oligonucleotidi durante la fase di disegno. Ciò risulta essere particolarmente importante, se si considera che attualmente non esistono altri sistemi di predizione in grado di eseguire tale funzione.

In conclusione, la presente tesi illustra lo sviluppo di un nuovo software, chiamato "LAMPZero", per il disegno specifico di primer applicabile alla tecnologia LAMP.

Esso unisce all'accuratezza dei calcoli dei parametri termodinamici utili per il disegno dei primer, l'efficacia di uno strumento che consente molteplici opzioni di disegno (automatico, manuale, combinato) e la flessibilità di un programma altamente personalizzabile in base alle esigenze dell'utilizzatore. Un altro importante punto di forza è legato alla possibilità di apportare ulteriori miglioramenti al software, implementandolo con le nostre considerazioni derivanti dall'esperienza maturata nello sviluppo della tecnologia LAMP.

Questi valori aggiunti del software da noi sviluppato hanno permesso di superare le limitazioni dei sistemi di disegno precedentemente utilizzati, che rendevano la fase di progettazione dei primer piuttosto laboriosa.

L'innovazione tecnica legata alle funzionalità del software LAMPZero è stata pertanto tutelata mediante richiesta di brevetto.

INTRODUCTION

1. LOOP-MEDIATED ISOTHERMAL AMPLIFICATION

The Loop-mediated isothermal AMPlification (LAMP) technology is an innovative non-PCR based nucleic acid amplification method performed using a strand-displacement polymerase under isothermal conditions. Both amplification and detection of nucleic acid sequences can be completed in a single step by incubating the mixture of sample, primers, DNA polymerase at a constant temperature, allowing the use of simple and cost effective reaction equipments (Notomi et al., 2000).

The amplification efficiency of LAMP is very high and the reaction proceeds rapidly as there is no need for initial heat denaturation of the template DNA or thermal cycling.

LAMP has been explored so far in hundreds of papers for the detection of DNA of many different organisms, such as bacteria (Nakao et al., 2010; Yang et al., 2009; Aryan et al., 2009), viruses (Kaneko et al., 2005; Yoda et al., 2009; Hong et al., 2004; Curtis et al., 2009), fungi (Sun et al., 2009; Uemura et al., 2008), parasites (Chen et al., 2009; Zhang et al., 2009; Nkouawa et al., 2009) and it is also suitable for the analysis of human DNA (Minnucci et al., 2012). Thanks to its peculiar characteristics, LAMP has shown to be an easy and rapid diagnostic tool in molecular medicine.

Moreover a Reverse Transcription (RT) LAMP has been developed for the detection and amplification of RNA by introducing a reverse

transcriptase in the reaction mix, expanding the fields of applications for this technique (Parida et al., 2004; Ushio et al., 2005).

Thus, LAMP constitutes an attractive alternative to PCR for sequence detection, with a sensitivity and quantitative performance comparable to PCR, allowing different detection methods, simplex and multiplex reaction applications in a fast, easy and potentially cheap molecular diagnostics platform (Parida et al., 2008).

1.1 - LAMP principle

LAMP is an isothermal DNA amplification method which relies on the use of a thermostable DNA polymerase with strand displacement activity (*Bst* polymerase from *Bacillus stearothermophilus*) and four primers specifically designed to recognize six distinct regions on the target gene, consisting in a pair of outer primers, F3 and B3, and a pair of inner primers, FIP (Forward Inner Primer) and BIP (Backward Inner Primer).

Both inner primers contain two distinct sequences (F2 + F1c for the FIP primer and B2 + B1c for the BIP primer), corresponding to the sense and antisense sequences of the target DNA, one for the initial priming and the other for the subsequent self-priming (Figure 1).

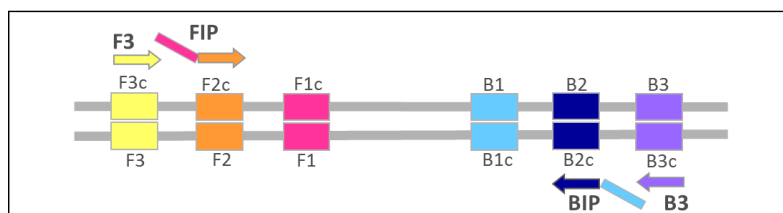


Figure 1. Schematic representation of the four primers, recognizing six distinct nucleotide sequences on the target gene, required for a basic LAMP reaction: two internal primers (FIP, BIP) and two external primers (F3, B3).

The mechanism of the LAMP amplification reaction includes three steps (Figure 2):

- I. production of starting structure;
- II. cycling amplification;
- III. elongation and recycling.

All four primers are operative in the initial steps of the reaction, but in the later cycling steps only the inner primers are required for strand displacement DNA synthesis.

At a constant reaction temperature, inner primer F1B hybridizes to F2c in the target DNA and initiates complementary strand synthesis (structure 1). Outer primer F3 hybridizes to F3c in the target and initiates strand displacement of the new DNA chain (structure 2), releasing a FIP-linked complementary strand, which forms a looped-out structure at one end (structure 4). This single stranded DNA serves as template for BIP-initiated DNA synthesis and subsequent B3-primed strand displacement DNA synthesis leading to the production of a dumbbell form DNA which is quickly converted to a stem-loop DNA (structure 6).

The dumbbell serves as the starting structure for LAMP exponential amplification, the second step of the LAMP reaction. The stem-loop DNA structure contains a double stem-loop which is the starting point for the whole LAMP amplification process as it promotes an amplification from its self-annealed 3' terminus and from a newly annealed internal primer (FIP or BIP) (structure 7). During amplification, FIP hybridizes to the loop in the stem-loop DNA and primes strand displacement DNA synthesis, generating as an intermediate one gapped stem loop DNA with an additional inverted

copy of the target sequence in the stem, and a loop formed at the opposite end via the BIP sequence (structure 8). Subsequent self-primed strand displacement DNA synthesis yields one complementary structure of the original stem-loop DNA and one gap repaired stem-loop DNA (structure 10) with a stem elongated to twice as long and a loop at the opposite end (structure 9).

Both of these products then serve as templates for BIP-primed strand displacement in the subsequent cycles and in the elongation and recycling step. Thus, in LAMP the target sequence is amplified 3-fold every half cycle. Amplification proceeds promoting itself, each strand being displaced by elongation of the new loops formed (Notomi et al., 2000).

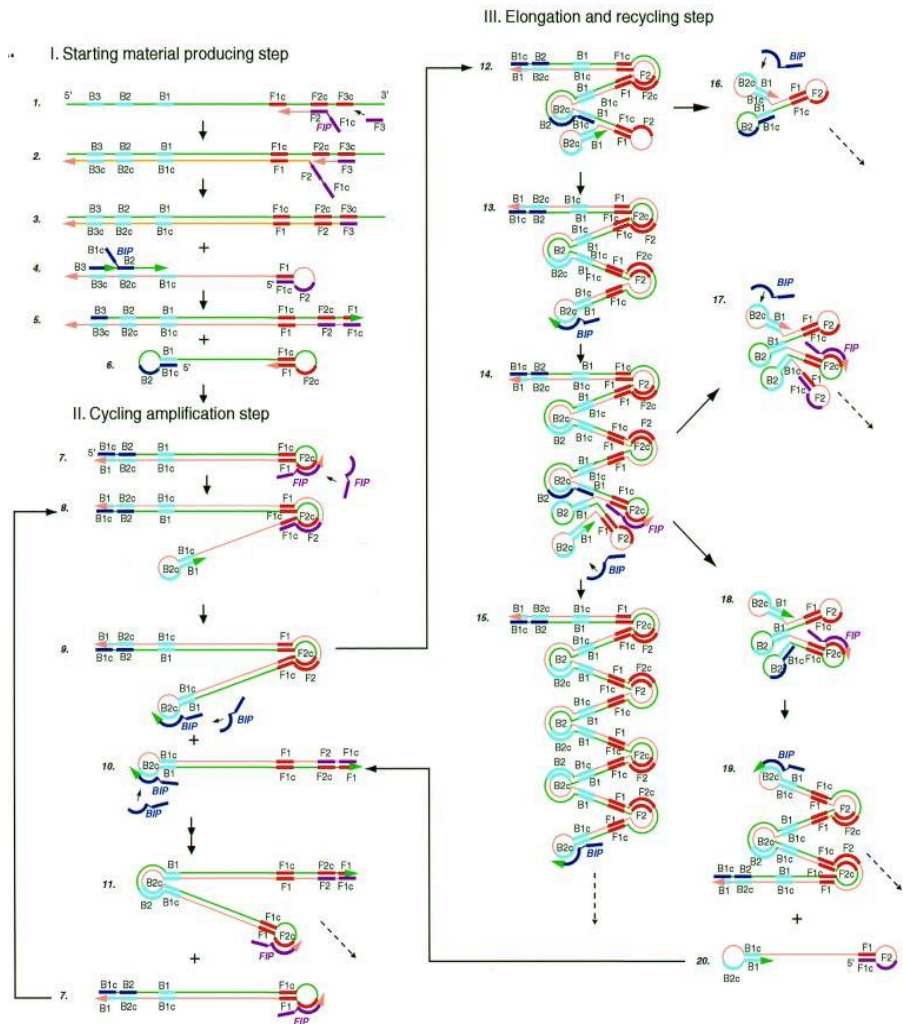


Figure 2. Schematic representation of the steps of a LAMP reaction. This figure shows the process that starts from primer FIP. DNA synthesis can also begin from primer BIP.

The final product is a mixture of stem-loop DNA with various stem length and cauliflower-like structures with multiple loops formed by annealing between alternately inverted repeats of the target sequence in the same strand (Figure 3).

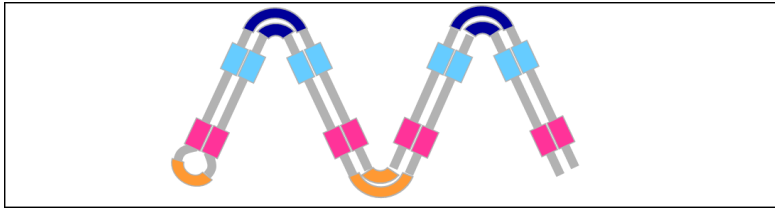


Figure 3. Schematic representation of inverted repeats concatamers that are the main amplification products of a LAMP reaction.

1.2 - Loop primers

In order to accelerate the LAMP reaction, it is possible to introduce additional primers, called Loop Primers.

The Loop Primers LF (Loop Primer Forward) or LB (Loop Primer Backward) contain sequences complementary to the single stranded loop region (either between the F1 and F2 regions or between the B1 and B2 regions) on the 5' end of the dumbbell-like structure, providing increased number of starting points for DNA synthesis.

In particular, two different types of dumbbell-like structures are generated during a LAMP reaction:

- i. a dumbbell on which the FIP primer and the LB primer (if present) hybridize: the LB primer anneals on the stem-loop structure generated by the self-annealing of the BIP primer, while the FIP

primer hybridizes to the complementary sequence on the second stem-loop structure of the dumbbell;

- ii. a dumbbell on which the BIP primer and the LF primer (if present) hybridize: the LF primer anneals on the stem-loop structure generated by the self-annealing of the FIP primer, while the BIP primer hybridizes to the complementary sequence on the second stem-loop structure of the dumbbell.

As a consequence, each loop primer will hybridize to the stem-loop that is not hybridized by the inner primer and will prime strand displacement DNA synthesis (Figure 4).

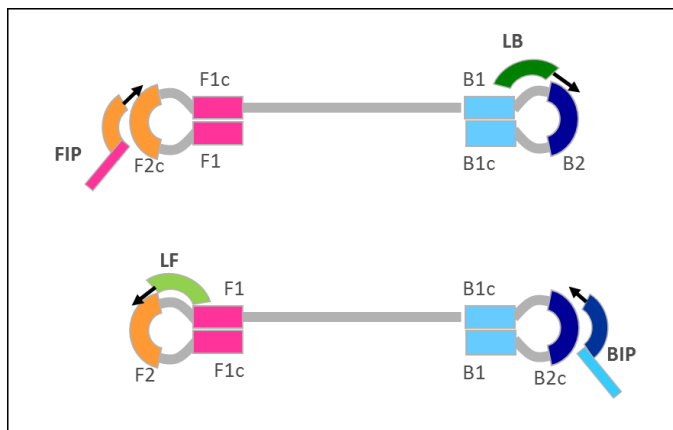


Figure 4. Schematic representation of the two stem loop structures (also called “dumbbell” structures) generated during a LAMP reaction: FIP and LB primers will hybridize on one dumbbell type, while BIP and LF primers will hybridize on the other dumbbell type.

The loop primers act as catalyzers of the amplification reaction, since their presence increases the rate of DNA production and therefore the

sensitivity of the method, considerably decreasing the reaction time of the original LAMP method (Nagamine et al., 2002).

By these peculiar amplification dynamics and exploiting the continuous displacement by *Bst* polymerase of the already produced DNA strands with no need for denaturation steps, it is possible to amplify the initial DNA amount up to 10^{10} times in less than 60 minutes with very high specificity, due to the many independent target-recognition events.

1.3 - Detection of LAMP amplification products

Two principal methods have been widely used to detect LAMP amplification products and distinguish between positive and negative samples: turbidimetry or fluorescence.

Turbidimetry detection is made possible by the high amplification efficiency of LAMP reaction (Mori et al., 2004; Mori et al., 2001). This detection method consists in the measurement of turbidity of the reaction mix, due to the precipitation of magnesium pyrophosphate. This insoluble salt is formed by the interaction between Mg_{2+} present in solution and the inorganic pyrophosphate produced by the incorporation of dNTPs in the DNA growing strands. Thus, the amount of magnesium pyrophosphate is proportional to the amount of amplified DNA. The turbidity is visible by the naked eye and is measurable on a turbidimeter as a transmittance signal either at end-point for a qualitative analysis or in real-time, allowing quantitative applications.

Another approach for the detection of LAMP reaction products is represented by the use of fluorescent intercalating dyes that bind double-

stranded DNA: they allow a real-time direct visualisation of the isothermal amplification by measuring the increase in fluorescence throughout the reaction correlated to the exponential synthesis of LAMP amplicons. The detection step has then to be followed by an annealing analysis to confirm the reaction product.

However, turbidimetry and fluorescent intercalating dyes are not very specific detection approaches, because they allow to detect double-stranded DNA without easy and rapid discrimination between specific and aspecific products.

Other fluorescence approaches are being applied to LAMP technique, such as molecular beacons or quenching probes, relying on sequence-specific detection of a desired LAMP product. The use of fluorescent-labelled target-specific probes results in increased specificity and sensitivity. In addition, the use of different fluorophores specific for distinct targets allows to potentially amplify many sequences simultaneously.

1.4 - RT-Q-LAMP

The 'traditional' LAMP technique initially described by Notomi was an isothermal method alternative to PCR for the rapid amplification of DNA, based on the use of 4 to 6 primers and a DNA polymerase with strand displacement activity (Notomi et al., 2000). The amplification reaction could be monitored by measuring either the turbidity or the fluorescence of intercalating dyes: however, in both cases, a single gene of interest was generally detected in the reaction mix.

Throughout the years, however, this method has been modified and implemented.

RT-Q-LAMP (“Reverse Transcriptase-Qualitative-LAMP”) technique, developed by DiaSorin, has represented a further improvement of the ‘traditional’ LAMP reaction (Spinelli et al., 2014). The insertion of fluorescent probes and a new polymerase with both reverse-transcriptase and DNA polymerase activities has allowed the amplification and the real-time detection of multiple target genes in a single tube, with a single enzyme, at a single temperature.

These implementations entail several advantages compared to a ‘traditional’ LAMP.

First, the use of a polymerase having both reverse-transcriptase and DNA polymerase activities allows the rapid and direct amplification of the target RNA included in the reaction mixture in a single-step procedure, avoiding further reverse transcription step, usually long and laborious.

Furthermore, the use of labelled oligonucleotide probes specific for the transcripts of interest has allowed the development of a LAMP multiplex technology, in which more than one target gene is amplified within the same reaction tube with very high efficiency. For example, it is possible to perform duplex and triplex reactions, in which respectively two or three genes of interest are amplified at the same time and real-time monitored, by the use of fluorophores emitting at different wavelengths revealed by dedicated fluorescence channels.

Finally, the possibility to simultaneously amplify an internal control within the same reaction tube represents a further advantage of this

technique. The internal control consists in a housekeeping gene (ex: GUS β , ABL, etc.) and its amplification allows to verify the proper extraction of RNA from cells, the quality and integrity of the extracted RNA, as well as the correct reaction conditions (buffer, enzyme, temperature, etc.).

Compared to the 'traditional' LAMP method, RT-Q-LAMP is a even more rapid and efficient technique in the detection of target sequences: the steps of reverse-transcription, amplification and monitoring of the signal are combined in a single one, resulting in an overall simplification of the diagnostic procedure.

2. PRIMER DESIGN

Design of primers and probes is one of the most crucial factors affecting the success and quality of molecular diagnostic techniques, in particular for quantitative assays, since an accurate and reliable quantification depends on using efficient primers and probes.

If the primer design phase is an important step in the development of a 'standard' PCR assay, based on the use of only two different primers (and one probe, if considering quantitative assay), we can easily understand how this phase becomes crucial in the development of a LAMP assay, that requires a coordinated hybridization of the numerous primers on the target without a dedicated annealing step.

2.1 - Primer design: crucial step in a LAMP assay development

Since the commonly used application of LAMP requires a total of six primers (with FIP and BIP primers encompassing two primer binding sites each), primer design for LAMP assays requires the selection of eight different regions of a target nucleic acid sequence.

In addition to that, loop primers have significant restrictions on their positioning, since LF and LB must be positioned strictly between the F2 and F1 sites of the FIP primer and the B2 and B1 sites of the BIP primer, respectively.

Therefore, LAMP primer design can be extremely challenging, especially when targeting polymorphic markers, sequences containing complex

secondary structure, but also when designing on a specific region to detect translocations or mutations (Gandelman et al., 2011).

It should also be considered that LAMP assays take place at a constant temperature, in contrast to PCR that occurs by repeated cycles of three different steps (denaturation, annealing, extension) each one taking place at a specific reaction temperature. This aspect introduces further complexity in the phase of primer design: an accurate selection of the primers melting temperatures (T_m) is important for successful reactions and essential so that the annealing of the numerous primers on the target could be finely orchestrated at the proper reaction temperature.

Moreover, significant care must be taken in order to avoid dimers formation among the primers used in the reaction, especially for FIP and BIP primers that are generally about 40 nucleotides long. In order to do this, it may be necessary that the primer design software could make predictions about secondary structures formation that are the most precise and closest as possible to LAMP experimental conditions.

For all these reasons, it is not difficult to infer that the primer design is a crucial step in a LAMP assay development and the use of an accurate primer design tool, able to guide the researcher in this phase, is of extreme importance.

2.2 - LAMP primer design: key factors

The design of LAMP primers is based on the six regions in the target sequence designated from the 5' end as F3, F2, F1, B1, B2 and B3.

Forward Inner Primer (FIP) consists of the F2 sequence (at its 3' end) that is complementary to the F2c region, and the same sequence as F1c region at its 5' end. Backward Inner Primer (BIP) consists of the B2 sequence (at its 3' end) that is complementary to the B2c region, and the same sequence as B1c region at its 5' end.

Loop primer Forward (LF) is designed using the complementary strand corresponding to the region between F1 and F2, while Loop primer Backward (LB) is designed using the complementary strand corresponding to the region between B1 and B2 (Figure 5) (Notomi et al., 2000).

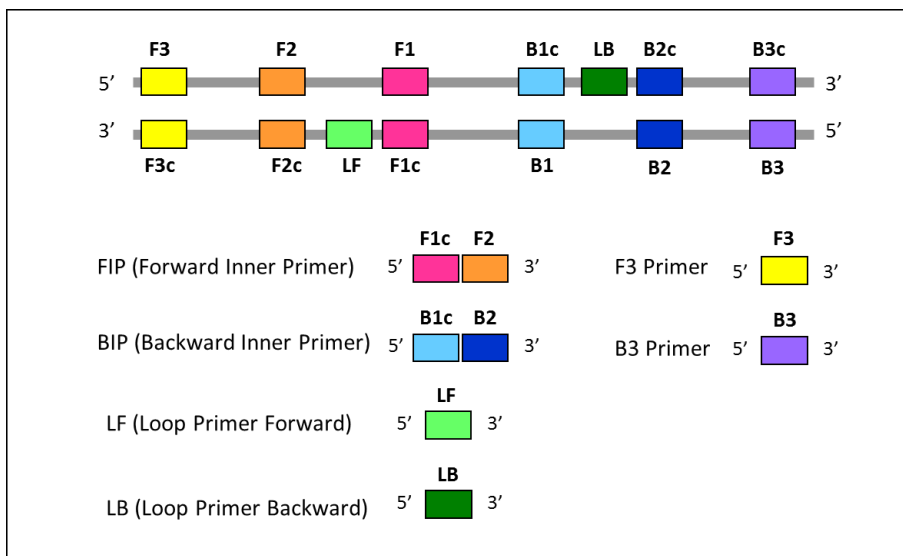


Figure 5. The design of LAMP primers is based on the six regions in the target gene, designated from the 5'-end as F3, F2, F1, B1, B2 and B3. FIP and BIP primers encompass two primer binding sites each. Furthermore two additional regions are required for the design of loop primers (LF and LB).

There are several important design criteria to be considered during the design phase, in order to potentially select the best primers ensuring specific amplification with high yield (Rodriguez et al., 2015; Alvarez-Fernandez, 2013; Rychlik and Rhoads, 1989; Hillier and Green, 1991; Hyndman et al., 1996).

The selection process has to be performed at first for each individual primer and then also for the combination of the primers (primer set), since all the primers have to work well together on the target, as if they were a 'synchronized team'.

2.2.1 - Key factors for individual primers design

The key factors for individual primers design include:

- ***Length***

It is generally accepted that the optimal length of primers is 18-22 base pairs (bp). This length is long enough for adequate specificity and short enough for primers to bind easily to the template at the annealing temperature.

The primer length directly influences the Melting Temperature (T_m) of the primer. For this reason, depending on the base pairs composition of the target sequence, the length of the primer can be shorter (in the presence of GC rich region) or longer (in the presence of AT rich region) than the optimal one, in order to obtain T_m values that are compatible with the temperature of a LAMP reaction (generally 63-65°C).

- **GC content**

The primer GC content corresponds to the number of G's and C's nucleotides in the primer as a percentage of the total bases and should range between 40-60%. In particular, primers with GC content between 50% and 60% tend to be relatively good primers. The GC content directly affects the calculation of the primer Melting Temperature (T_m) and the range for primer length.

Another parameter than can be correlated to the GC content is the GC clamp, that is the presence of consecutive G or C bases within the last six bases from the 3'end of primers. The GC clamp facilitates specific binding at the 3'end due to the stronger bond of G and C bases. However, more than 3 consecutive G's or C's at the 3'end of the primer should be avoided.

- **Non specific binding (NSB)**

Optimal primers should hybridize only to the target sequence, avoiding the formation of alternative amplification products, that could compromise the specificity of the reaction. For this reason, it is necessary to avoid regions of homology during the design phase. To test specificity, primers are commonly designed and then BLASTed with the whole genome of interest. Moreover, it is also important to identify and avoid regions of partial homology between the 3'end of the primer and the template, in order to potentially decrease the risk of primer non specific binding.

- ***Stability at the end of each primers***

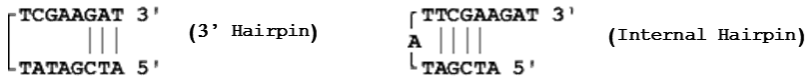
The end of the primers (intended as the last 6bp of the terminal portion) serves as the starting point of the DNA synthesis and thus must have certain degree of stability. The 3'ends of F2/B2, F3/B3 and LF/LB and the 5'end of F1c/B1c (corresponding to the 3'end of F1/B1) are designed so that the change in free energy (ΔG value) is -4kcal/mol or less. Negative ΔG values ensure higher thermodynamic stability of the primer-binding site, promoting primer annealing and DNA polymerization from the 3'ends. However, the more negative the ΔG is, the more often primer will anneal to the template, increasing the risk of false priming.

- ***Primer secondary structures (hairpins)***

It is important that primers, particularly the inner primers FIP and BIP, are designed to avoid the formation of secondary structures. In fact, the presence of primer secondary structures produced by intermolecular or intramolecular interactions can lead to poor or no yield of the product, adversely affecting primer template annealing and amplification and greatly reducing the availability of primers for the reaction.

Hairpins are formed by intramolecular interaction within the primer and should be avoided. The stability of hairpins is commonly represented by the ΔG value, the energy required to break the secondary structure. Highly negative values for ΔG indicate stable, undesirable hairpins. Presence of hairpins at the 3'end most adversely affects the reaction.

A 3'end hairpin with a ΔG of -2 kcal/mol and an internal hairpin with a ΔG of -3 kcal/mol are generally tolerated.



- **Melting Temperature (T_m)**

Primer Melting Temperature (T_m) by definition is the temperature at which one half of the DNA duplex will dissociate to become single stranded. The T_m of a primer is an index of the stability of a primer-template DNA duplex and influences the sensitivity and the specificity of the amplification.

T_m is calculated using the Nearest-Neighbor method, which is currently considered to be the approximation method that gives the closest value to the actual one, and it is affected both by primer length and GC content, but also by experimental conditions such as salt and oligo concentrations. Therefore the most accurate predictions of T_m are calculated taking into consideration also the experimental conditions of a LAMP reaction.

Based on Primer Explorer V4 manual (https://primerexplorer.jp/e/v4_manual/), the T_m for each LAMP region is designed to be about 65°C (64 - 66°C) for F1c and B1c, about 60°C (59 - 61°C) for F2, B2, F3, and B3, and about 65°C (64 - 66°C) for the loop primers.

Primers with higher melting temperatures (above 66°C) have a tendency for secondary annealing.

2.2.2 - Key factors for primer sets design

The key factors for primer sets design include:

- **Distance between primers**

The primers are designed so that the distance from the end of F2 to the end of B2 is between 120 and 160 bases. This region corresponds to the LAMP amplicon (the region amplified by the LAMP method) and is strictly correlated to the size of the dumbbell structure. The primers are also designed so that the distance from the 5' end of F2 to the 5' end of F1 (and from the 5' end of B2 to the 5' end of B1) is between 40 and 60 bases: these are the regions that form the loops, where the loop primers hybridize.

The distance between F2 and F3 should be between 0 and 60 bases (Figure 6) (https://primerexplorer.jp/e/v4_manual/).

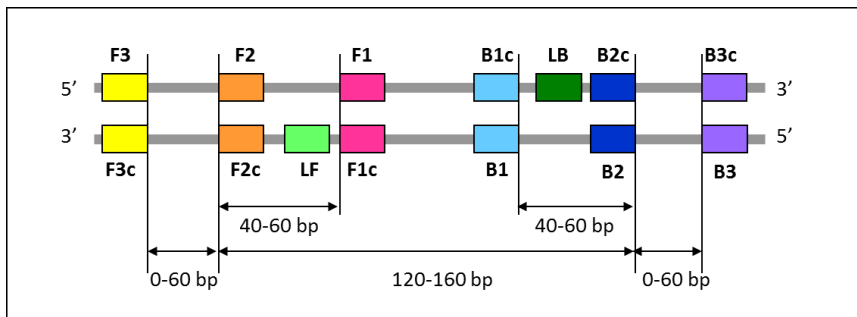


Figure 6. Distance between primers of a LAMP primer design, as suggested by Primer Explorer V4 manual.

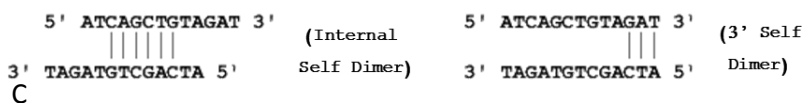
- **Primer secondary structures (self dimers and cross dimers)**

Generally a large amount of primers are used in a LAMP reaction compared to the amount of target gene. When primers form intermolecular dimers much more readily than hybridizing to target DNA, they reduce the product yield. This is particularly problematic when the homology occurs at the 3'end of the primer. The 3'end will extend readily leading to primer dimer artefacts.

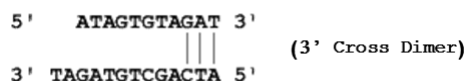
For this reason, as previously discussed, it is important that primers, particularly the inner primers FIP and BIP, do not form secondary structures and that the primer 3'ends are not complementary.

There are two types of intermolecular interactions able to generate secondary structures formation: self dimers and cross dimers.

Self dimers are formed when two copies of the primer bind to each other, where the primer is homologous to itself. A 3'end self dimer with a ΔG of -5 kcal/mol and an internal self dimer with a ΔG of -6 kcal/mol is generally tolerated.



Cross dimers are formed by intermolecular interaction between different primers, where they are homologous. A 3'end cross dimer with a ΔG of -5 kcal/mol and an internal cross dimer with a ΔG of -6 kcal/mol is generally tolerated.



- ***T_m balance***

The Melting temperature (T_m) is an important parameter for primer design. When considering a primer set, also the T_m balance is fundamental: in particular, the two primers of each primer pair (such as F3/B3, F2/B2, F1c/B1c, LF/LB) should have closely matched T_m for maximizing LAMP product yield.

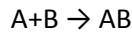
3. DNA THERMODYNAMICS IN PRIMER DESIGN

The physical principles of DNA hybridization and folding are very important for designing optimal PCR reactions. These models, studied for PCR, can be applied also to other amplification techniques where nucleic acids are involved, such as LAMP reactions.

In the next section, the background for understanding DNA thermodynamics will be provided.

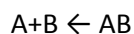
3.1 - Solution equilibrium and calculation of the amount bound

The process of duplex hybridization for a forward bimolecular reaction is given by:

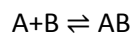


where A and B imply strands A and B in the random coil state and AB implies the ordered AB duplex state. This is called the two-state approximation (it is assumed that there are no intermediate states).

The reverse reaction is unimolecular and is given by:



If enough time elapses, the forward and reverse reaction rates will be equal and equilibrium will be achieved as shown in the following equation:



The equilibrium constant, K, for the reaction is given by the law of mass action:

$$K = \frac{[AB]}{[A][B]} \quad (1)$$

The equilibrium constant is independent of the total strand concentrations, [A_{tot}] and [B_{tot}], which is why it is called a “constant.” However, K depends strongly on temperature, salt concentration, pH, [DMSO] and other environmental variables.

Even though K is a constant, if you change the total concentration of one or both of the strands, [A_{tot}] and/or [B_{tot}], then the system will respond to re-achieve the equilibrium ratio, which in turn means that the individual concentrations [A], [B], and [AB] will change. Simply stated, the more the strand A added, the more [AB] will increase.

The equation shows that if the equilibrium constant and total strand concentration are known, then the amount of primer bound to target [AB], which is the quantity that matters for hybridization in PCR, can be computed (SantaLucia, 2007).

The equilibrium constant and its changes related to temperature can be predicted, as shown in the next paragraph.

3.2 - The meaning of ΔH° , ΔS° and ΔG_T° parameters

In thermodynamics, considering a PCR reaction, the “system” is defined as the contents of the test tube that contains the nucleic acid strands, solvent, buffer, salts and all the other chemicals. To determine whether a process is spontaneous and to determine the equilibrium, we only need to determine the changes in certain properties of the system (namely, ΔH° , ΔS° and ΔG_T°).

The process of reaching equilibrium results in the release of heat from the system to the surroundings, when strands change from the random coil state to the duplex state. At constant pressure, this change in heat of the system is called the enthalpy change, ΔH . The naught symbol, $^\circ$, is added (e.g. ΔH°) to indicate that the energy values given are for the idealized “standard state”, which simply means that the energy change refers to the amount of energy that would be released if a scientist could prepare each species in 1M concentration (i.e., $[A] = 1M$, $[B] = 1M$, and $[AB] = 1M$, which is a non-equilibrium condition), mix them, and then allow them to come to equilibrium.

The more heat that is released from the reaction to the surroundings, the more disordered the surroundings become and thus the more favourable the reaction is (because of the second law of thermodynamics). As a result of the hybridization reaction, the amount of order in the system also changes (a duplex is more ordered than a random coil because of conformational entropy). In addition, solvent molecules and counterions bind differently to the duplexes and random coils. These effects are all accounted for in the entropy change of the system, ΔS° .

The ΔH° and ΔS° are combined to give the Gibb’s free energy change for going from random coil to duplex:

$$\Delta G^\circ_T = \Delta H^\circ - T\Delta S^\circ \quad (2)$$

where T is the Kelvin temperature, ΔH° is given in kcal/mol, and ΔS° is given in cal/mol K.

Importantly, there is a relationship between the Gibb's free energy change at temperature T and the equilibrium constant at temperature T:

$$\Delta G_T^\circ = -RT \times \ln(K) \quad (3)$$

where R is the gas constant (equals 1.9872 cal/mol K).

This equation provides the critical link that allows for the equilibrium constant to be computed from ΔG_T° . Next, ΔG° can be computed at any temperature T, if ΔH° and ΔS° are known by using Equation 2. Thus, if ΔG° is known at all temperatures, then K can be computed at all temperatures and thus the concentrations of all species can be computed at all temperatures as described in Equation 1.

All these statements mean that given ΔH° and ΔS° , we can compute the concentration distribution for all species at all temperatures, as illustrated in Figure 7 (SantaLucia, 2007).

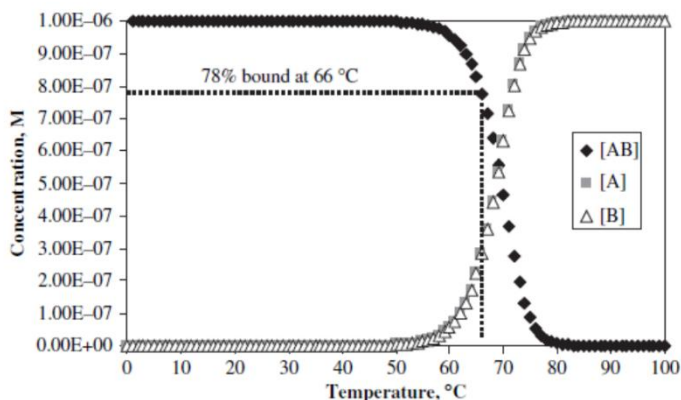


Figure 7. Simulation of the hybridization profile for a simple non-self-complementary two-state transition, given only ΔH° and ΔS° and using Eq. 1, as described in the text. Note that the percent bound can be determined at any temperature. The random coil concentrations of strands A (squares) and B (triangles) are overlaid.

ΔH° and ΔS° values are accurately predicted from the strand sequences involved in the duplex by applying the Nearest-Neighbor (NN) model.

3.3 - Nearest-Neighbor (NN) model

The NN model for nucleic acids assumes that the stability of a given base pair depends on the identity and orientation of neighboring base pairs (SantaLucia, 1998). Instead of treating a DNA helix as a string of interactions between base pairs, the Nearest-Neighbor model considers a DNA helix as a string of interactions between 'neighboring' base pairs.

The 10 Watson-Crick NN dimer duplexes are represented with a slash separating strands in antiparallel orientation (e.g., AC/TG means 5'-AC-3' Watson-Crick base-paired with 3'-TG-5'). For oligonucleotide duplexes, additional parameters for the initiation of duplex formation are introduced.

The total ΔG_{37}° is given by:

$$\Delta G (\text{total}) = \Delta G (\text{initiation}) + \Delta G (\text{symmetry}) + \Delta G (\text{propagation})$$

where ΔG (symmetry) equals +0.43 kcal/mol, if the duplex is self-complementary, and zero, if it is non-self-complementary, and ΔG (propagation) corresponds to the sum of the standard free-energy changes for the 10 possible Watson-Crick NNs.

An example of ΔG calculation is illustrated in Figure 8:

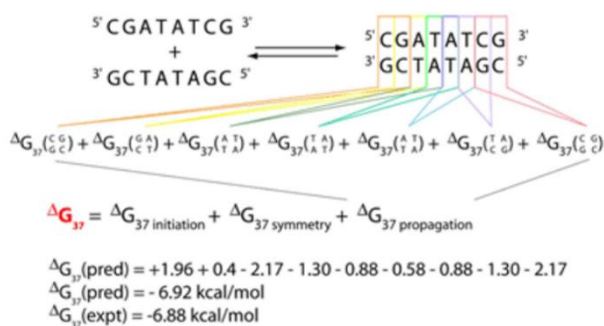


Figure 8: Example of ΔG_{37}° calculation for the sequence CGATATCG·CGATATCG using the Nearest-Neighbor (NN) parameters. The duplex CGATATCG·CGATATCG is constituted by seven NN dimers and is self-complementary (thus ΔG° (symmetry) is +0.4).

Thermodynamic parameters (ΔG° , ΔH° and ΔS°) for each NN Watson-Crick pair are published in reference documentations (Table 1) (Allawi and SantaLucia, 1997). ΔH° and ΔS° parameters are analogously calculated (as for ΔG calculation) from the values in Table 1. The ΔG_{37}° can also be calculated from ΔH° and ΔS° parameters by using the equation 2 ($\Delta G_T^\circ = \Delta H^\circ - T\Delta S^\circ$).

Propagation sequence	ΔH° (kcal mol ⁻¹)	ΔS° (e.u.)	ΔG_{37}° (kcal mol ⁻¹)
AA/TT	-7.6	-21.3	-1.00
AT/TA	-7.2	-20.4	-0.88
TA/AT	-7.2	-21.3	-0.58
CA/GT	-8.5	-22.7	-1.45
GT/CA	-8.4	-22.4	-1.44
CT/GA	-7.8	-21.0	-1.28
GA/CT	-8.2	-22.2	-1.30
CG/GC	-10.6	-27.2	-2.17
GC/CG	-9.8	-24.4	-2.24
GG/CC	-8.0	-19.9	-1.84
Initiation	+0.2	-5.7	+1.96
Terminal AT penalty	+2.2	+6.9	+0.05
Symmetry correction	0.0	-1.4	+0.43

Table 1. Nearest-Neighbor thermodynamic parameters for DNA Watson-Crick pairs in 1M NaCl₂. ΔH° and ΔS° are assumed to be temperature independent.

In addition to the NN predictions of Watson–Crick base-paired duplexes, J. SantaLucia’s laboratory has published the empirical equations that allow the NN model to be extended to include salt dependence (SantaLucia, 1998), terminal dangling ends (Bommarito et al., 2000) and all possible internal and terminal mismatches (Peyret et al., 1999; Allawi and SantaLucia, 1997; Watkins et al., 2011) (Figure 9).

The availability of the dangling-end parameters is important because in PCR the short primers bind to the longer target DNA and the first unpaired nucleotides of the target sequence next to the 5’ and 3’ of the primer-binding site contributes significantly to binding (in some instances, a dangling end can contribute as much as a full AT base pair).

The mismatch parameters allow for the T_m of mutagenic primers and for the specificity of hybridization to be computed.

The availability of the salt dependence allows for the accurate prediction of thermodynamics under a wide variety of solution conditions that occur in biological assays.

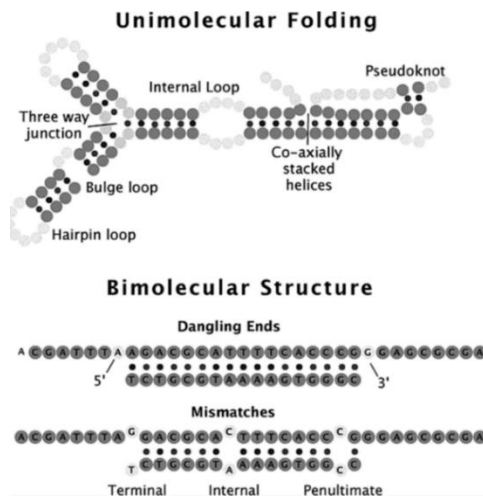


Figure 9. Structural motifs that occur in folded DNA (top) and in bimolecular duplexes (bottom).

3.4 - Prediction of the Melting Temperature (T_m)

Melting temperature and thermodynamic parameters (ΔG° , ΔH° and ΔS°) are related to each other. Therefore, when the thermodynamic parameters of a given nucleic acid sequence are known, the melting temperature (T_m) can be predicted.

T_m is defined as the temperature at which half of the strands are in the double-helical state and half are in the “random-coil” state and is calculated through the formula:

$$T_m(^{\circ}\text{C}) = \{\Delta H / \Delta S + R \ln(C)\} - 273.15$$

where ΔH and ΔS are the sum of all the ΔH and ΔS values of the NN dimer duplexes, R is the gas constant (1.987 cal/K·mol) and C is the total oligonucleotide strand concentration (for self-complementary oligonucleotide duplexes). For non-self-complementary molecules, C is replaced by $C/4$ if the strands are in equal concentration or by $(C_A - C_B/2)$ if the strands are at different concentrations, where C_A and C_B are the concentrations of the more concentrated and less concentrated strands, respectively (SantaLucia, 1998).

For more accurate T_m prediction, an additional salt correction is applied to the ΔS parameter, since the NN reference values are calculated from DNA melting studies in 1M NaCl (default condition):

$$\Delta S (\text{salt correction}) = \Delta S (1\text{M NaCl}) + 0.368 \times N \times \ln([\text{Na}^+])$$

where N is the number of nucleotide pairs in the primer (i.e. primer length - 1), $[Na^+]$ is salt equivalent in mM.

The T_m equation apply only to “two-state transitions” (i.e., the molecules that form only random coil and duplex states), and they do not apply to transitions that involve intermediate partially folded or hybridized structures (SantaLucia, 2007).

However, natural polymers with heterogeneous sequences, such as DNA usually melt with many stable intermediate states (non-two-state).

3.5 - Multi-state model for T_m prediction

The two-state model, based on the principle to calculate the amount bound for a two-state transition (refer to paragraph: “Solution equilibrium and calculation of the amount bound”), can be inaccurate because there are often many different equilibria that can compete with the desired equilibrium (Figure 10).

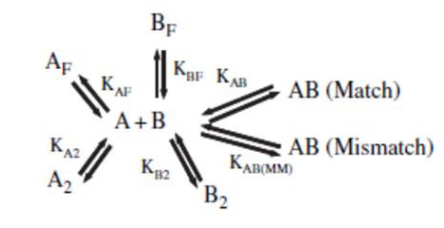


Figure 10. Structural seven-state model for hybridization (AB match) with competing equilibria for unimolecular folding (AF and BF), homodimers (A2 and B2), and mismatch hybridization (AB mismatch). The presence of the competing equilibria will decrease the concentration of AB (match).

In addition to target secondary structure folding, other structural species can also form, such as folded primer, mismatch hybridization, primer homodimers and heterodimers (especially in a LAMP reaction where numerous primers are present).

It is desirable to compute the concentrations of all the species for such a multi-state system. This can be accomplished by generalizing the approach described above for the two-state case (equation 1) and calculating all the species concentrations at equilibrium.

The ΔH and ΔS for each of the reactions in Figure 10 can be predicted using the NN model (SantaLucia and Hicks, 2004). Thus, we can compute ΔG at all temperatures and thus all the K 's at all temperatures, and those can be solved to give the concentrations of all species at all temperatures.

The "effective T_m ", defined as the temperature at which half the strands form a desired species, is calculated taking into consideration the effective concentration of all the species in solution. The usefulness of the effective T_m is that it is related to what would be observed experimentally in a non-two-state model and therefore it is more appropriate to describe thermodynamics occurring in biological systems (SantaLucia, 2007).

4. LAMP PRIMER DESIGN: CURRENT OPTIONS

In this section of the Introduction, an overview of the software applications to be used for LAMP primer design will be provided.

DiaSorin principal option for LAMP primer design consisted in the use of Primer Explorer V4 in combination to Visual OMP. It is worth highlighting that Visual OMP is a primer design system specific for PCR: however, due to its accuracy, it is used as reference system for thermodynamic predictions regarding nucleic acids.

4.1 - Primer Explorer V4

Primer Explorer V4 (PE4) is a free primer design software specific for LAMP technique, supplied by Eiken Chemical Co (<https://primerexplorer.jp/e/>). Once provided the target sequence and a number of design parameters, this program processes and displays the primer set that are considered the most efficient by its algorithm.

In particular, the design of a complete LAMP primer set proceeds through two subsequent steps: in the first step, the software selects the basic LAMP primers (FIP, BIP, F3 and B3) based on the 6 regions required by the LAMP method. Then, since the loop primers are not an essential requirement for LAMP, they are selected in a second and separate execution using the primer information file of the selected basic LAMP primers.

Primer Explorer provides two different design modes: an 'Easy Mode' that eliminates the need to change parameters and displays five primer sets that are likely to have high amplification efficiency and an 'Expert Mode'

that is designed for primer set customization, allowing the user to change parameters and to specify the number of primer sets to be designed.

Being a specific software for LAMP method, Primer Explorer V4 takes into consideration most of the key factors for LAMP primer design, previously discussed. For melting temperatures (T_m) calculation, it uses the "Nearest-Neighbor" method and fixed experimental conditions (oligo concentration at 0.1 μ M, sodium ion concentration at 50 mM, magnesium ion concentration at 4 mM).

When the target sequence is loaded, Primer Explorer determines automatically its GC content. In the case of a normal sequence ($45 < \text{GC} \% < 60$), the primer design conditions are entered as a default setting. If the sequence is classified as AT rich ($\text{GC}\% < 45$) or GC rich ($\text{GC} \% > 60$), then the primer design conditions (T_m , length, GC content) are automatically modified as shown in Table 2:

	T_m ($^{\circ}$ C)	Length (mer)	GC content (%)
AT rich	>55	18-25	<45
GC rich	<68	15-22	>60

Table 2. Primer design conditions (T_m , length and GC content) automatically selected by Primer Explorer software, based on the GC content (AT rich or GC rich) of the target sequence.

When conducting the primer design, the primers are designed for FIP-BIP and F3 and B3 in the entire target region. Next, for each FIP-BIP region, F3 and B3 are selected to form a primer set. The generation of primer sets, which consists in a combination of FIP-BIP with the F3 and B3, begins at the

5'end and proceeds until the 3'end is reached. Each FIP- BIP can form primer set with a maximum of three combinations of F3-B3.

Moreover PE4 have some additional features, such as the possibility to take the location of mutation into account, choosing either between the primer design function that does not include the mutation point or the function that allow the mutation to be included in the 5'end, internally or in the 3'end.

However, Primer Explorer presents some important limitations. The need to perform an additional design session for loop primers selection can prevent optimal primer combinations from being identified, providing results that may not be satisfactory. In this case, the user has to repeat the design of the 'basic' LAMP primer set by varying the design parameters and the procedure can become inefficient and time-consuming.

Then, since T_m is affected by experimental conditions such as salt and oligo concentration, the use of the "Nearest-Neighbor" method for T_m calculation with fixed concentrations may provide poor T_m predictions, preventing to make simulations with the real concentrations used in a LAMP reaction.

In addition, Primer Explorer does not allow possible adjustment to the primer sets automatically generated by the system, thus having significant restrictions in primers positioning on the target.

Finally, primer dimers cannot be analysed and last, intellectual property is not ensured, since PE4 is a public internet software application.

Thus, the primer sets that results from the processing of Primer Explorer require to be tested and confirmed by using another software application

able to provide more reliable "in silico" predictions of melting temperatures (T_m) and secondary structures formation: Visual OMP.

4.2 - Visual OMP

Visual OMP™ (Visual Oligonucleotide Modelling Platform) is a software application developed by DNA Software (www.dnasoftware.com; CEO J.SantaLucia Jr.) for the analysis and design of oligonucleotide sequences for use with the PCR technique. It provides researchers with an "in silico" laboratory environment for designing PCR and microarray based experiments (Royce et al., 2003; SantaLucia and Hicks, D. 2004).

Visual OMP is a multiplex primer design system that can analyze multiple primers and probes at the same time using a design algorithm that queries each oligo design against every other oligo in the system to check for intended and unintended cross- and mishybridization. At this purpose, the algorithm simulates all secondary structures for target (optimal and suboptimal), primers (optimal and suboptimal), homodimers, heterodimers and calculates the energy (ΔG) of all these possible structures by using a thermodynamic database.

Accurate values for T_m , ΔG , percent bound and concentration for all species at the defined experimental conditions can be provided. In addition, target and oligo structures can be displayed much faster than with other available products.

The software is also integrated with a BLAST design function that allows to select primers unique to a target and against a specified genome, preventing mishybridization artefacts in the assay.

Visual OMP uses the following formula for the calculation of T_m:

$$T_m = \Delta H^\circ \times 1000 / (\Delta S^\circ + R \times \ln(C_T/x)) - 273.15$$

The T_m varies in function of the concentration (C_T). The concentration is generally calculated considering a "two-state model", that is assuming the presence of only two species in the solution: the single-stranded DNA and the primer, both hybridized or not.

This is an approximation, since the species that are present in the solution are numerous (i.e. all other primers and their combinations). For this reason, Visual OMP optimizes the T_m calculation considering the effective concentration of all the species in solution ("multi-state model"), obtaining a so-called "effective T_m".

Based on our experience, Visual OMP is a useful tool to verify and implement the primer sets generated by the automatic design of Primer Explorer V4, because it provides accurate T_m calculations and secondary structures predictions, based on the effective experimental conditions used in a LAMP reaction.

However, it is a licensed and very expensive tool. Moreover, being specific for PCR, it does not take into account the complexity of the primer design for the LAMP method, not providing any design output of the primer sets and turning out to be laborious.

4.3 – Other software for LAMP primer design: LAMP Designer

LAMP Designer (Optigene) is a new software application to specifically design primers for Loop-Mediated Isothermal Amplification assays (<http://www.optigene.co.uk/lamp-designer/>).

LAMP Designer retrieves batches of gene sequences directly into the program from on-line sources such as GenBank. The application can perform a BLAST search to avoid cross homologies while designing the primers. Once designed, a verification BLAST can be performed to verify primers' specificity.

LAMP Designer incorporates a sequence view with primers marked on the sequence and calculates melting temperatures using Nearest-Neighbor method and SantaLucia values. In addition, the most stable cross homologies between the primers designed (based on ΔG calculation) are displayed by the program.

This software application allows to modify solution conditions (such as reaction temperature and salt concentration) and primer design parameters (such as length, T_m and distances), in order to meet the experimental requirements of the user.

However, this tool presents several limitations for DiaSorin LAMP primer design applications. Primer concentrations, which strongly affects T_m and secondary structures evaluation, cannot be set, thus preventing to obtain accurate predictions.

The analysis of secondary structures formation between different primer sets (multiplex reactions) is not allowed.

As with Primer Explorer, also with LAMP Designer the user cannot modify or adjust the primer sets automatically generated by the system, having significant restrictions in primers positioning on the target.

Additionally, it is a licensed software application developed by Optigene, therefore intellectual property is not ensured.

AIM OF THE THESIS

Loop-mediated isothermal AMPLification (LAMP) is an innovative technique for the amplification and the detection of nucleic acids. The commonly used application of LAMP requires a total of six primers: four basic primers (F3, B3, FIP, BIP) annealing to the target region in a peculiar way, that allows the formation of an initial stem-loop structure, and two loop primers (LF, LB) able to accelerate the amplification reaction. Since FIP and BIP primers encompass two primer binding sites each, the selection of eight separate regions of a target nucleic acid sequence is required.

An accurate selection of the primers, in order to obtain their synchronized hybridization to the target DNA, is a very crucial step for the efficiency of LAMP reactions. Therefore, the phase of primer design is critical for successful LAMP assays and certainly more complex and challenging than PCR.

Since LAMP is an emerging amplification technique, LAMP primer design systems are still rather limited and mainly consist of one principal tool, Primer Explorer 4, a free software developed by Eiken Chemical Co.

However, this tool has several limitations: in particular, it does not allow manual design options, in order to make possible modifications on the primer sequences automatically generated by the system; it gives poor estimations of primers melting temperatures (T_m); it does not provide any information regarding primers secondary structures formation.

For these reasons, the primer sets resulting from the Primer Explorer automatic design require a verification and implementation step,

performed by using Visual OMP™ (DNA Software, CEO J.SantaLucia Jr.), an efficacious tool for *in silico* PCR specific primer design, that ensures better T_m estimations and secondary structures predictions.

As a consequence, the phase of primer design is laborious and time-consuming and the selection of the candidate primer set has to be done by numerous trial-and-error laboratory experimentations.

In this contest, the need for an effective LAMP primer design system has emerged and become urgent, due to the substantial lack of efficient tools for the *in silico* evaluation of LAMP primer sets.

The aim of my thesis was the development of a primer design software application tailored to DiaSorin LAMP requisites.

This project has been carried out thanks to the collaboration with Luciano Ruo Rui, a colleague with expertise in Information Technology and Software Engineering.

The first main goal of my PhD project was to develop an assisted design software function that could help and guide the operator in the manual selection of LAMP primers on the target sequence, providing accurate information on primer melting temperature and secondary structures formation.

The second main goal consisted in the development of an automatic design software module that, given a target sequence and defined the acceptance criteria, could automatically generate a list of candidate primer sets, to be eventually moved to the assisted design for any possible adjustment.

The creation of such a tool can guarantee the existence of a DiaSorin LAMP primer design system, in order to ensure intellectual property, be implemented with possible future improvements and, above all, make a significant contribution to an easier and more efficient development of LAMP assays.

MATERIALS AND METHODS

DiaSorin LAMP primer design system (named LAMPZero) is a software application that allows a researcher to design specific LAMP primer sets, considering a target DNA sequence, specific reaction conditions and design criteria.

In the following section of my thesis, the methods underlying the functions of LAMPZero will be described. In particular, detailed explanations and flowcharts of the mechanisms used by the system to design and analyse nucleic acids molecules (primers) will be provided.

1. THE ALGORITHM FOR PRIMER DESIGN: BASIC SETTINGS

When setting up a new LAMP primer design experiment, three are the main steps to be considered:

- 1) Select a target (DNA sequence on which the primers will be designed);
- 2) Define the reaction conditions (experimental conditions that will be used in the LAMP assay);
- 3) Establish the design criteria (rules to be followed by the system to provide primers, according to the user's needs).

As represented in the flowchart of Figure 11, these preliminary stages (100, 104, 106) of the LAMP primer design method have been included into the calculation algorithm of the software application (1).

The target sequence can be selected from a database of DNA sequences (102) included in the system.

The reaction conditions consist of: the assay temperature (108), the concentration of monovalent cations (110), the concentration of divalent cations (112) and the concentrations of the primers in solution (114). Optionally, the reaction conditions also comprise the percentage of glycerol (115). All these parameters can be modified, based on the real conditions used for each experiment.

The design criteria include: the melting temperature T_m ($^{\circ}\text{C}$, minimum and maximum value) (116), the CG content (percent, minimum and maximum value) (118), the threshold of stability of 3' and 5' ends (ΔG , minimum value) (120), the threshold of stability of the primers dimers (ΔG , maximum value) (122), distances between primers (base pairs, minimum and maximum value) (124), primer length (base pairs, minimum and maximum value) (126) and the number of bases of non-specific binding, NSB (base pairs, maximum value)(128). Also design criteria, as well as reaction conditions, can be varied: values can be entered for each couple of forward and backward primers (F3 / B3, F2 / B2, F1c / B1c and LF / LB).

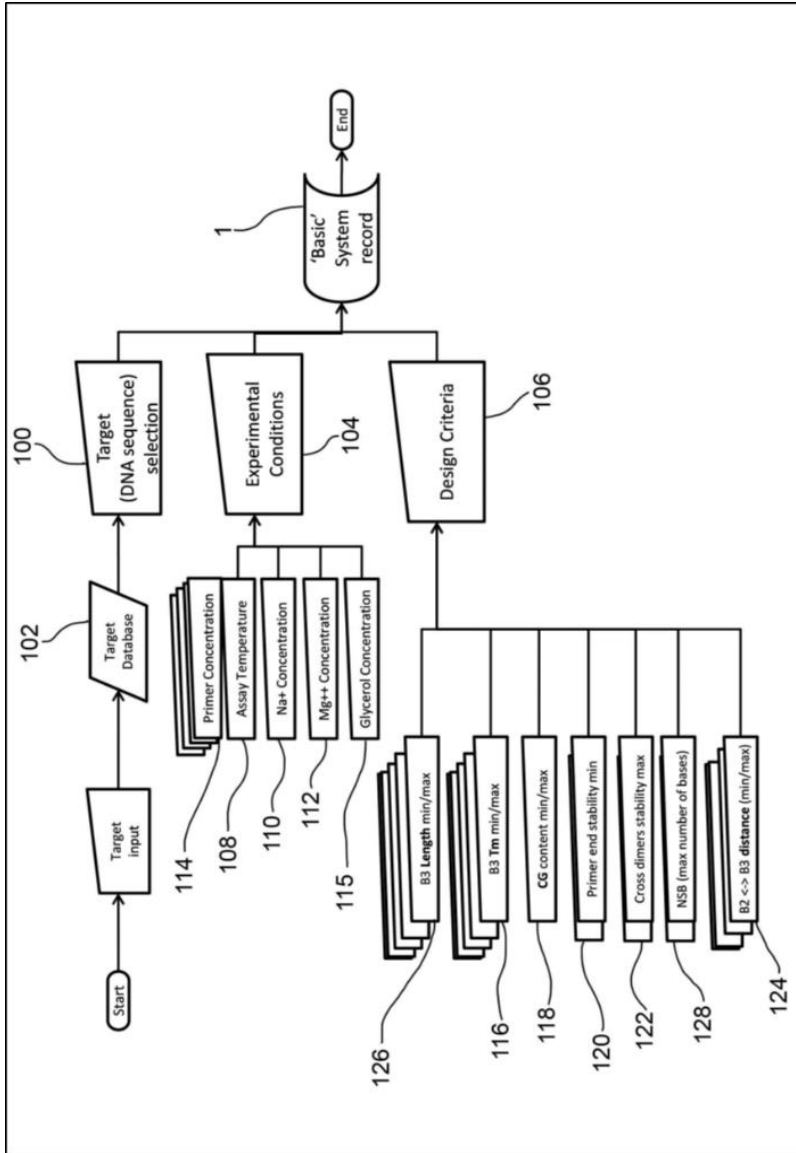


Figure 11. Flowchart representing the inclusion of the target sequence, the experimental conditions and the design criteria into the primer calculation algorithm of the software.

2. LAMPZERO PRIMER DESIGN: DESIGN MODE OVERVIEW

Once the sequence of the target DNA (100), the reaction conditions (104) and the design criteria (106) have been selected, the user can decide whether to perform an assisted primer design (4), in order to obtain a single primer set candidate (6), or perform an automatic primer design (2), in order to obtain several primer set candidates (8), as shown in Figure 12.

In the assisted design mode the user can select the primers directly on the target sequence. The design algorithm performs a calculation of the thermodynamic parameters of the primer sequence and of the secondary structures formed by the primer sequence with all the other sequences that are present in the LAMP reaction. The primer set candidate (6) obtained with the assisted design mode can be graphically represented by the dumbbell structure (5) for further evaluations made by the researcher.

In the automatic design mode, the design algorithm performs a scan along the selected target sequence, identifying the possible primers (based on the primer design criteria set by the user) and assigning them a score. The primers with a better score (primer "candidates") are then evaluated in combination, in order to detect the best "*in silico*" primer set by score.

Moreover, the primer sets obtained with the automatic design mode (2) can be moved to the assisted design mode (4), in order to be modified according to the user's needs. This design option is called 'combined mode'. In both the assisted and the automatic design modes, the primer set candidate(s) can be stored into a system record (9).

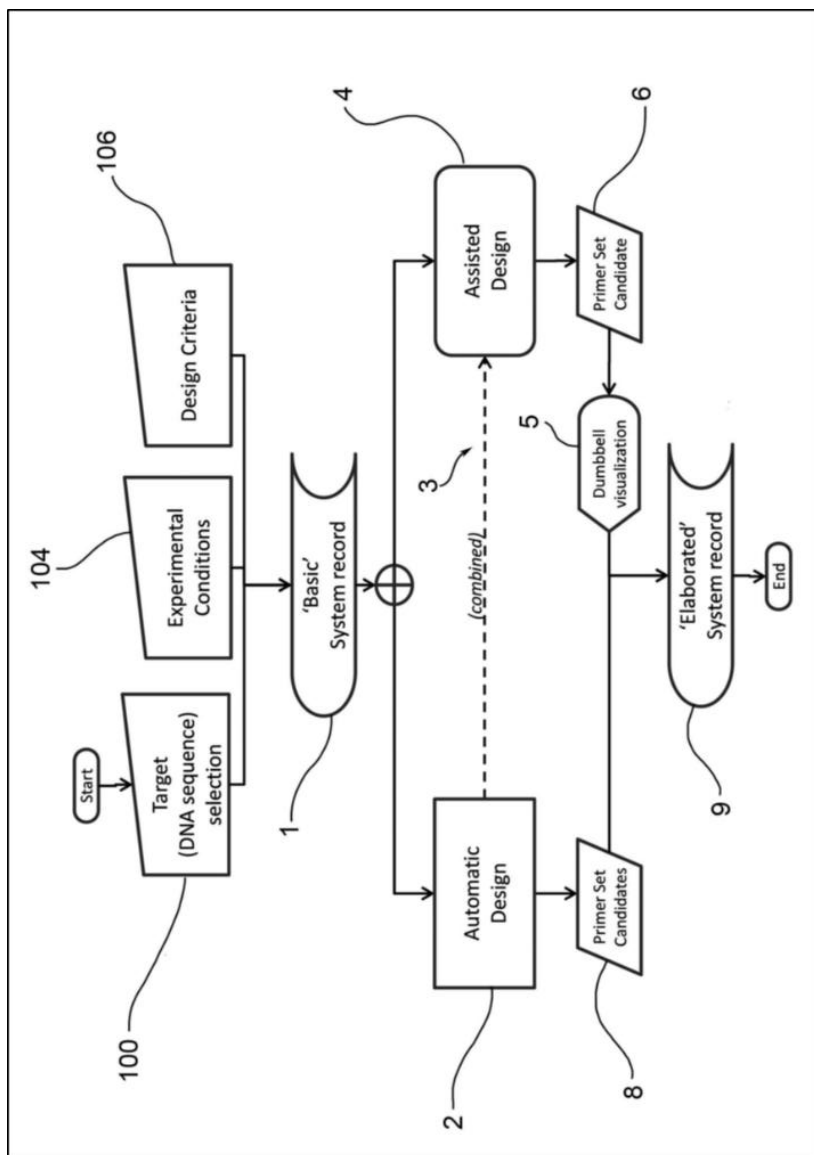


Figure 12. Flowchart representing the three primer design modules (automatic, assisted and combined) of LAMPZERO primer design software.

3. ASSISTED PRIMER DESIGN MODE

In the following paragraph a detailed description of the algorithm underlying the assisted primer design will be provided (Figure 13).

In the assisted design mode the user starts from a “basic” system record (1), containing the target DNA sequence (100), the reaction conditions (104) and the design criteria (106). The system allows to specify reaction conditions (200), such as temperature and salt concentrations, in order to perform an automatic update of the thermodynamic parameters database (202), which refers to an assay temperature of 37°C and concentrations of monovalent and divalent ions of 1M and 0M (201), respectively.

Through a graphical interface, the program allows the user to select the primers sequences directly on the text string of the target sequence (204).

The nucleotide sequence of the selected primer is then sent to the hybridization algorithm (206), which predicts the most stable dimers structures for all possible combinations of dimers (both for primer-target dimer (208) and for primer dimers (210)) through the calculation of the thermodynamic parameters ΔG , ΔH , ΔS . In other words, starting from the “corrected” database of thermodynamic parameters (202), the algorithm uses ΔG calculation in order to identify the most stable structure among all the numerous possible combinations.

ΔG value of the primer-target dimer allows the calculation of the primer melting temperature (T_m) (212).

The selection of the primer is also used by the program to calculate the stability (ΔG) of the ends of primer (214) and to predict and visualize the most stable hairpin (218) by applying the algorithm of folding (216).

The ΔG value of dimer and hairpin structures allows the calculation of the equilibrium constants K (220) and, consequently, of the concentrations of all species (structures) in solution (222). These concentrations will be used to calculate the effective primer-target T_m (224), as will be described in more detail with reference to the flowchart of Figure 17.

The algorithm provides real-time thermodynamic parameters aforementioned together with the GC content percentage and the possible non-specific bindings (226).

Conditional formatting clearly marks any discrepancies from the design criteria, regarding both individual primers and their combinations (primer dimers).

If the results of the selection process are not acceptable, the user can simply change the primer selection (228). The algorithm is repeated starting from the selection of one primer on the DNA target sequence (step 204) until completing the selection of the primer set (229).

Once the entire primer set has been selected, the algorithm of the assisted design mode graphically displays the dumbbells (230), the key structures of LAMP amplification technique, together with the hybridizing inner primers and loop primers (if included in the primer design). If the dumbbell is not accepted by the user, it is sufficient to modify primers selection to obtain a new one (232).

As will be further discussed in the next paragraphs (Figures 19 and 20), specific algorithms will also assign a score to each individual primer (234) and to the primer set (236) based on how the predicted parameters values are close to the corresponding target values for each design criteria.

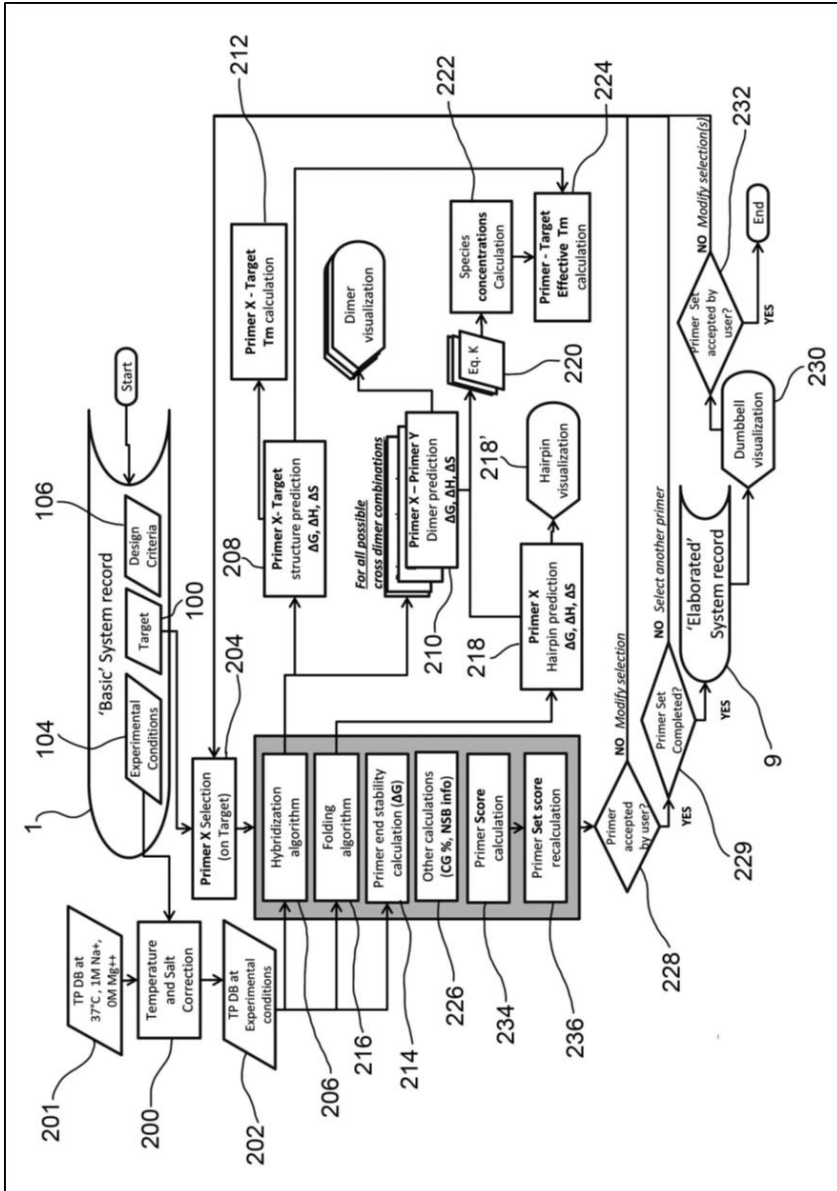


Figure 13. Flowchart representing the algorithm of the assisted primer design mode.

4. AUTOMATIC PRIMER DESIGN MODE

In the following paragraphs a detailed description of the algorithm used for the automatic primer design will be provided (Figure 14, 15 and 16).

An overview of the mechanism underlying the identification of the primer set candidates by the algorithm of the automatic design mode is represented in the flowchart of Figure 14.

In a first phase (300), the program scans the target sequence in order to identify all possible F3 primer sequences that meet the design criteria related to F3. The scan process is then performed for all the other primer types (F2, F1c, etc.) including, whether required, the loop primers (optional). At the end of the first phase, all possible primers divided by type are identified.

In a second phase (302), the algorithm identifies the possible internal primers (FIP and BIP) considering F2 + F1c and B2 + B1c combinations that meet the design criteria (such as the distance criteria).

Distance criteria are also applied in a third phase (303), where the system selects the possible combinations FIP + BIP, and in a fourth phase (304), where the combinations FIP + BIP are matched to the external primers (F3 and B3), thus obtaining the primer set candidates.

If the loop primers are required, the algorithm combines the primer set candidates to the loop primers based on the distance criteria (305).

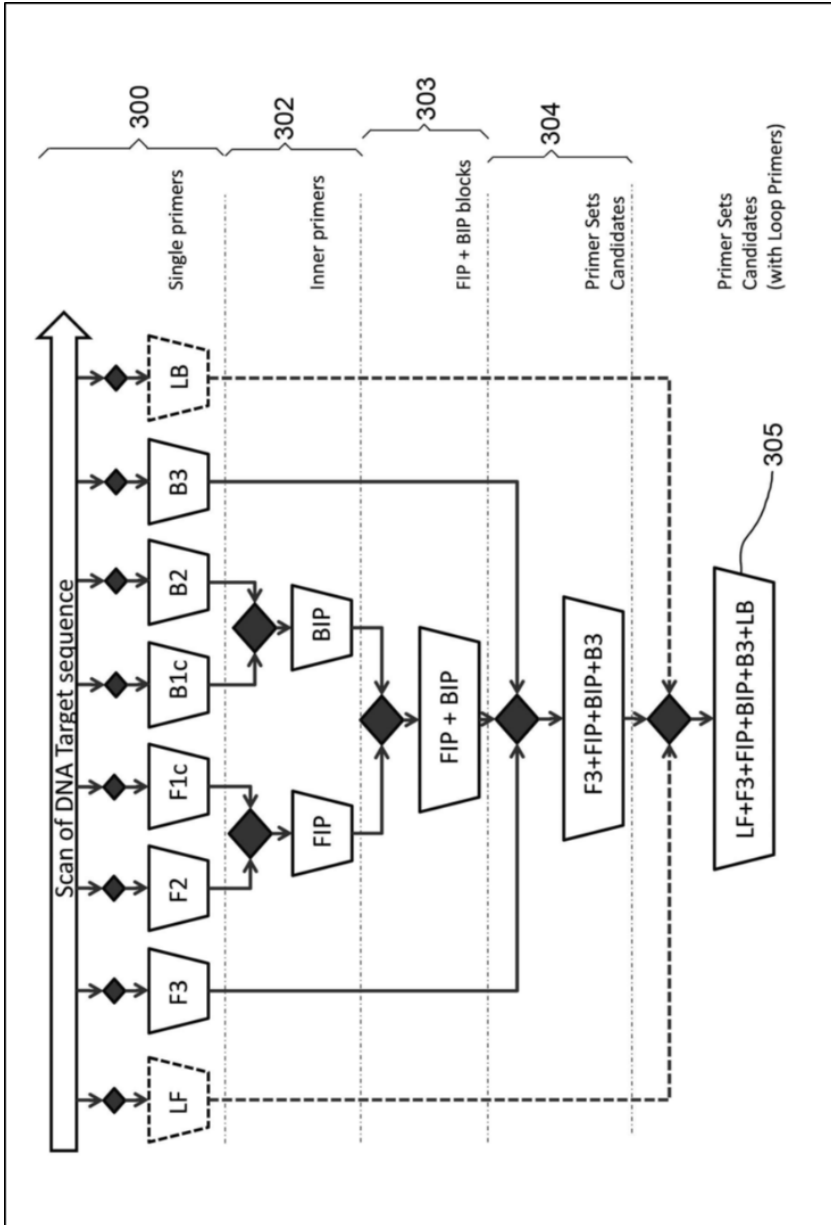


Figure 14. Flowchart representing the main processing steps of the design algorithm used in the automatic mode.

4.1 - Automatic primer design: selection of the primer candidates

The algorithm of the automatic primer design is described in details in Figure 15 and Figure 16.

In particular, the flowchart in Figure 15 represents the process by which the algorithm identifies individual primer candidates, while the flowchart in Figure 16 defines the process of selection of primer set candidates.

As in the assisted design mode (Figure 13), a system record (1) where the target DNA sequence (100), the reaction conditions (104) and the design criteria (106) were registered, is provided to the user also for the automatic design mode (Figure 15). The solution conditions, such as temperature and salt concentrations, can be specified (400) in order to perform an automatic update of the 'basic' thermodynamic parameters database (402), which refers to 37 °C and the concentrations of monovalent and divalent ions are 1M and 0M, respectively.

The user can perform an automatic design on the whole target sequence, but also on a selected region of the target (404). In both cases, the algorithm calculates which is the sequence range (406) to be used for the search of each individual primer (for example starting from F3 primer) according to the design distance criteria (124). Along this range specific for each primer type, all the sequences matching length criteria (126) for that specific primer are then identified (408). Then, sequences are verified against all other specific design criteria.

Sequences not passing the check are discarded. The first control (410) is related to the CG content (118). The subsequent controls (412, 414) are related to the stability at the primer ends (120) and to the melting

temperature (116), both calculated using the hybridization algorithm (206). The last control (416) is related to the presence of possible non-specific bindings (NSB) along the target sequence (128).

After these controls the system provides a series of primer candidates (418) to which is assigned a score (420), that will be as much higher as its parameters are close to the design criteria (106).

The selection process is repeated in order to obtain all the necessary primers to form a complete set, that includes internal primers (422), external primers (424) and, in case, loop primers (426).

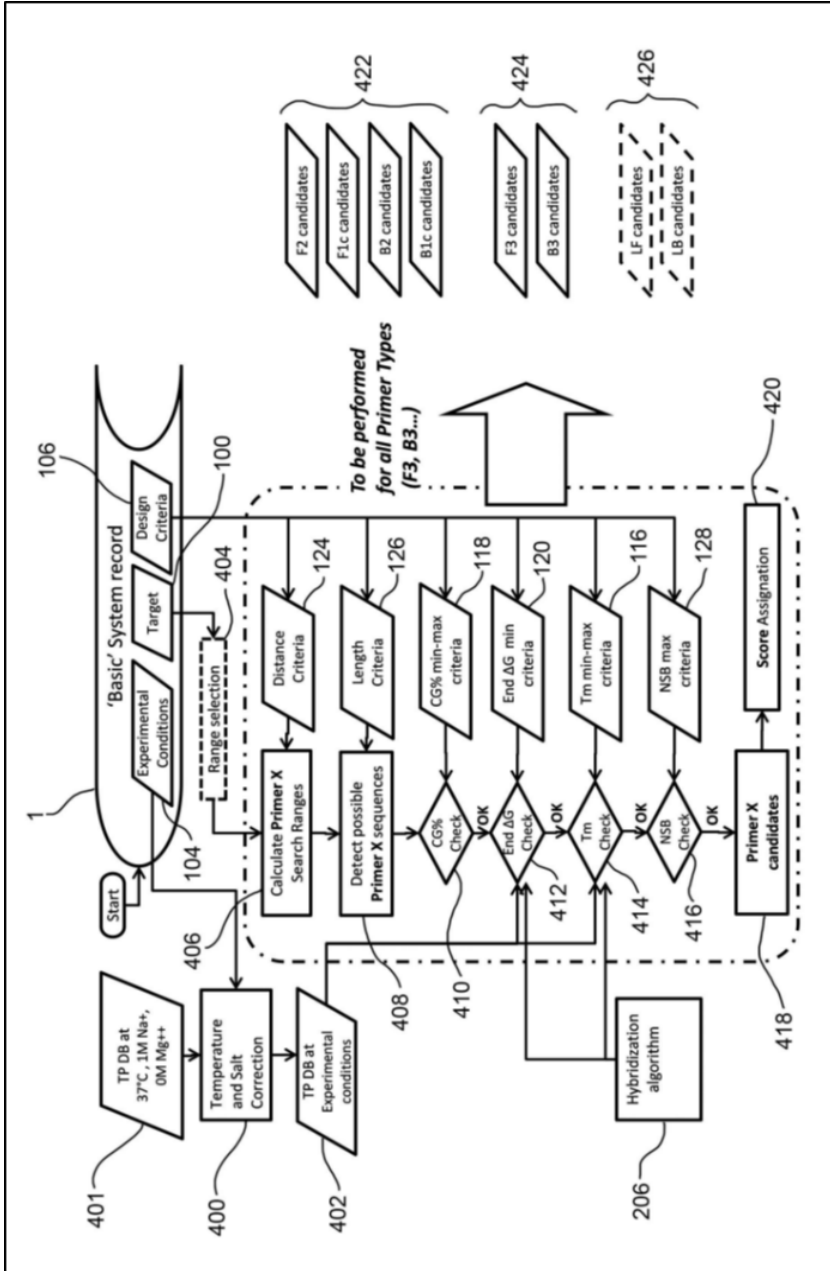


Figure 15. Flowchart representing the algorithm used for primers selection in the automatic design mode.

4.2 - Automatic primer design: selection of the primer set candidates

Figure 16 schematically represents the process of primer set selection, starting from the individual primer candidates (422, 424, 426).

The first selection phase (500) identifies all possible FIP primers (502) generated by the combination of the best (by score) F2 and F1c candidates, considering the design distance criteria (124) and their melting temperatures (as per Primer Explorer manual, the T_m of F1c sequence is suggested to be 5°C higher than the T_m of the F2 sequence).

A similar selection process (500') is performed for the BIP primer (504) generated by the combination of the best (by score) B2 and B1c candidates.

Then, all the possible combinations FIP + BIP are selected (506, 508) based on the design distance criteria (124).

Such combinations are then checked against the stability of primer dimers (510): this parameter is evaluated using the correspondent design criteria (122), the thermodynamic parameters database (402) "corrected" according to the reaction conditions (400) and the hybridization algorithm (206).

A preliminary score is then attributed to the FIP + BIP combinations (512) that are not discarded (if the ΔG value of the most stable dimer is higher than the limit indicated in the design criteria).

As previously performed for the internal primer (422), the algorithm evaluates the best (by score) F3 and B3 candidates (external primers, 424) and compares them (514) with the design distance criteria (124). A score to the selected F3-B3 primers couples is then assigned (516) based on the

balance of melting temperatures of the two primers (the closer, the better). The best F3-B3 couples (518) are then matched to the best FIP + BIP combinations (520) in accordance with the design distance criteria, in order to get the primer set candidates (522).

The user may also indicate the number of F3-B3 couples to be matched to any combination of FIP + BIP primers, so as to achieve possible additional variants of the primer sets.

Whether the user requires also the loop primers, both or one of them (426), the algorithm evaluates the best (by score) LF and/or LB candidates and compares them with the design distance criteria (524). A score is assigned to the LF and LB primer pairs thus selected (526) based on the balance of the T_m of the two primers (the closer, the better). Best LF-LB couples or single loop primers (528) are assigned to the primer sets previously selected (522) according to the design distance criteria (530).

Also in this case, the user may also indicate the number of LF-LB couples to be matched to each set, in order to achieve further design variants.

The resulting primer sets (522), with or without loop primers, are then evaluated for the acceptance criteria of primers dimers stability (532). The sets meeting this criterion (if the ΔG value of the most stable dimer is higher than the limit indicated in the design criteria) can then be identified as primer set candidates (534) and a score is assigned to them (536).

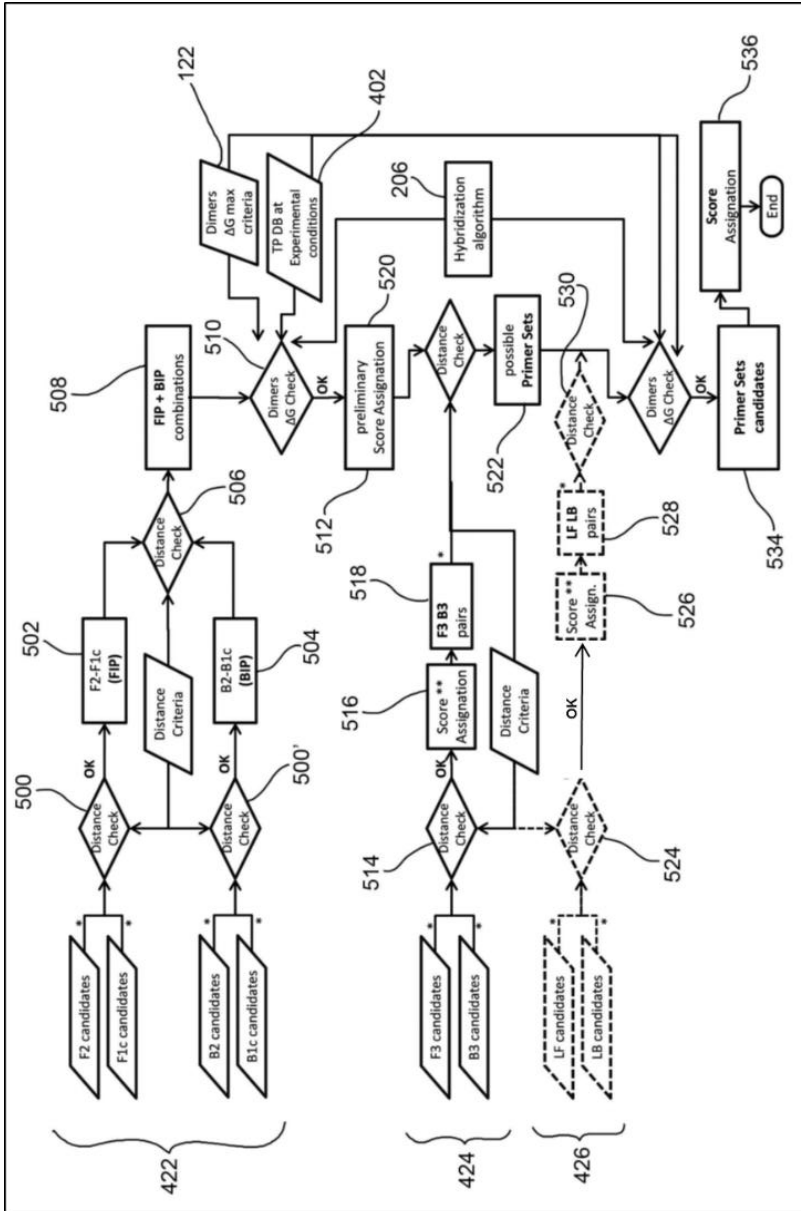


Figure 16. Flowchart representing the algorithm used for primer sets selection in the automatic design mode, starting from the primer candidates selected by the algorithm of figure 15.

5. CALCULATION OF THE CONCENTRATION CURVES AND DETERMINATION OF THE MELTING TEMPERATURES

The design algorithm is able to calculate and display all the secondary structures deriving from one or more primer sets, providing parameters such as ΔG , T_m , effective T_m , concentration and percentage in solution, as represented in Figure 17.

The user can select up to three different records (600, 600a, 600b), each one containing information about the primers and the target of each primer sets, in order to evaluate secondary structures formation not only for simplex (one set), but also for duplex and triplex reactions (two and three sets, respectively).

The algorithm processes (604) all possible combinations of primers dimers or primer-target (602a) and monomers, random coils or hairpins (602b) using the database of “corrected” thermodynamic parameters (202), according to the specified experimental conditions, such as temperature and salt concentrations (104).

The algorithm applies to all the possible combinations aforementioned the hybridization (206) and folding (216) algorithms, thus predicting the most stable secondary structures (dimers and hairpins) and providing their ΔG (606). Starting from ΔG values, it is possible to calculate the equilibrium constant K (608) of the species in solution (then calculating their concentrations) according to the formula:

$$\Delta G_T^\circ = -RT \times \ln(K)$$

Once the concentration of all the structures in solution is obtained (614), the algorithm is able to predict all structures concentration (hairpins & dimers as well as random coils) (616) at all values of the temperature scale. The value of the temperature scale in which the sequence of DNA target is 50% bound to the primer (i.e. the concentration of the heteroduplex target-primer will be equal to 50% of the target concentration) represents the effective melting temperature (618).

The calculation of the T_m often refers to the two state model, which is an approximation that assumes that in the reaction only two species are present : $A + B \rightleftharpoons AB$. The effective T_m is a most accurate parameter, because it is calculated considering the concentrations of all species in solution.

The process starts again until the calculation has been carried out for all the degrees of the temperature scale (for example from 10 to 100 ° C). Once the process is complete, the melting curves (620), which represent the trend of the structures concentration related to the increasing temperature, can be displayed.

6. SCORE CALCULATION ALGORITHM

Next paragraph provides an example of how the score to be assigned both to a primer and a primer set is calculated. This calculation is applicable to both the algorithm for the assisted primer design and the automatic primer design modes (Figures 18, 19 and 20).

Concerning each parameter associated to a primer, the score is directly proportional to the proximity of the value predicted by the design algorithm to a target value.

An example of calculation of the target value is shown in Figure 18. The target for the melting temperature (T_m) is considered as the middle value of the design criteria range (i.e. the mean value between $T_{m \text{ max}}$ and $T_{m \text{ min}}$) and the tolerance is the same interval divided by two. The example illustrates how a predicted melting temperature T_m (61.85) fairly close to the target value (62) produces a quite high score (85%).

704

$$T_{m \text{ target}} = (T_{m \text{ max}} + T_{m \text{ min}}) / 2$$

$$T_{m \text{ tolerance}} = (T_{m \text{ max}} - T_{m \text{ min}}) / 2$$

$$T_{m \text{ score}} = 1 - \frac{\text{Abs}(T_{m \text{ predicted}} - T_{m \text{ target}})}{T_{m \text{ tolerance}}} * 100$$

$$T_{m \text{ max}} = 61 \quad T_{m \text{ min}} = 63 \quad T_{m \text{ predicted}} = 61.85$$

$$T_{m \text{ score}} = 1 - \frac{\text{Abs}(61.85 - 62)}{1} * 100 = 85\%$$

Figure 18. Example of the score calculation for a primer or a primer set by the specific calculation algorithm.

6.1 - Score calculation algorithm for individual primers

In the score calculation to be assigned to a primer (Figure 19), the calculation algorithm compares the predicted value (700) of a parameter associated to a primer, for example the melting temperature value, with the correspondent design criterion (702) and calculates the score (704).

This score can be corrected according to a weight (706) configurable in the program (708) (e.g. it is possible to give a greater weight to the melting temperature T_m with respect to the CG content percentage).

With the mechanism described above, the attribution of the score is then carried out not only for the calculated values (such as the CG percentage or the number of NSB bases) but also for the predicted values of primers ends' stability (ΔG) and homodimers.

At the end, the algorithm calculates the score of the primer (710) as the sum of the individual parameters scores (709).

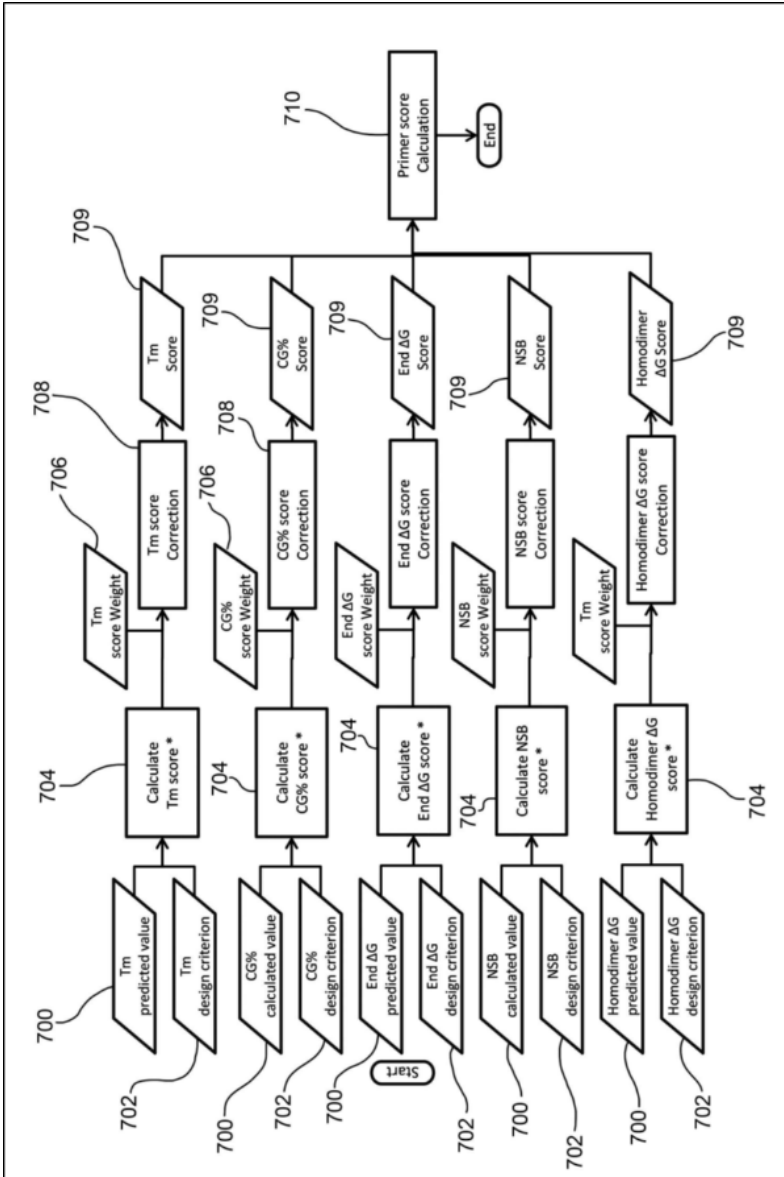


Figure 19. Flowchart of the calculation algorithm used to assign a score to individual primers.

6.2 - Score calculation algorithm for primer sets

The attribution of the score to a primer set is done with the calculation algorithm schematically represented in Figure 20.

In a first step (800), the calculation algorithm corrects the scores (710) of the individual primers considering the weight (801) attributed to them, allowing to get weighed scores of each individual primer (802).

In a second step (804), the calculation algorithm evaluates the balance of the melting temperature of each primers pairs (F3/B3, F2/B2, F1c/B1c, and potentially LF/LB), and assigns a score to that balance (805). The closer the values of the two primers of the pair are, the higher the score will be.

In a third step (806), the calculation algorithm calculates the stability score (as in figure 18) relative to the stability of the primers dimers using the ΔG value of the most stable primer dimer as the predicted value to be used. Also in this case, the calculation algorithm corrects the dimers stability scores considering a weight (807) attributed to them, thus obtaining weighed stability scores (808).

The calculation algorithm also calculates (810) an amplicon score relative to the amplicon length, that is the distance between the end of the F2 portion and the end of the B2 portion. Also in this case, the calculation is performed similarly to what described in figure 18. The calculation algorithm corrects the amplicon score by considering an amplicon length weight (811), thus obtaining a weighed amplicon score (812).

At the end, the algorithm calculates the score of the primer set as the contribution of the scores, possibly weighed, of each individual primer, of the balance, of the stability and possibly of the amplicon length (814).

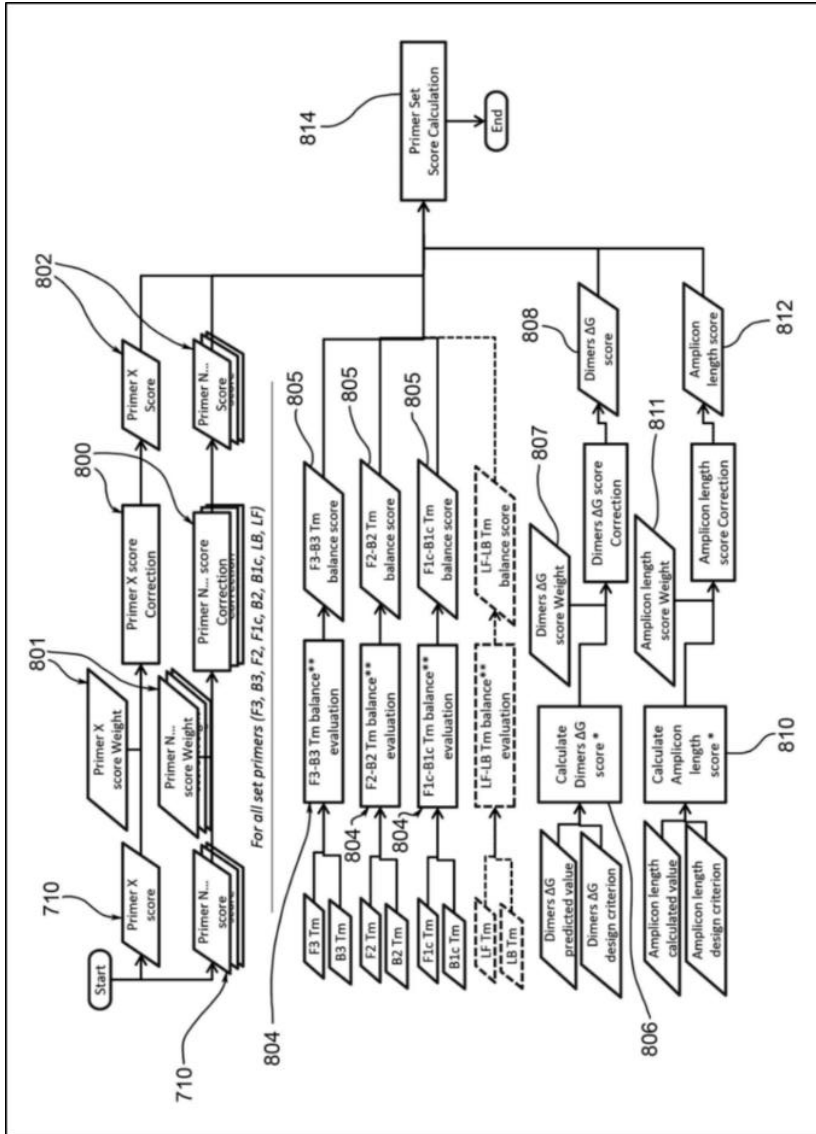


Figure 20. Flowchart of the calculation algorithm used to assign a score to primer sets, starting from individual primers scores calculated as shown in figure 19.

7. T_m & DIMERS PREDICTIONS: LAMPZERO vs VISUAL OMP

In order to compare their capability in predicting melting temperatures (T_m) and secondary structures, “*in silico*” experiments were performed in parallel with LAMPZero and VisualOMP™ software applications.

20 different LAMP primer sets, each one composed by 2 outer primers and 2 inner primers, were tested together with their target sequence. Inner primers FIP and BIP were analysed considering the two distinct sequences of F1c and F2 and B1c and B2, respectively.

The concentrations set were:

- 1.00E-07 for the double stranded target DNA
- 2.00E-07 for the outer primers F3 and B3
- 1.60E-06 for the inner primer sequences F2 and B2
- 1.60E-06 for the inner primer sequences F1c and B1c

The reaction conditions set were:

- Assay Temperature: 65°C
- Monovalent Cations Concentration [+]: 0.060M
- Divalent Cations Concentration [++]: 0.002M
- Glycerol Percentage: 2.50%

ΔG, ΔH and T_m on 27 dimers combinations (including homodimers, heterodimers and primer-target duplexes) and 6 monomers (hairpins) for each primer set were calculated using the two different software applications. For primer-target duplexes, only T_m values were considered.

The results obtained were compared and values > ±0.01 (±0.1 for T_m) were marked as discordant (Figure 21).

	V.OMP Results				LPD Results				Comparison Results		
	ΔG^*	ΔH	ΔS	T_m	ΔG^*	ΔH	ΔS	T_m	ΔG^*	ΔH	T_m
Monomers (Hairpins)											
F3	1.94	-4.4	-18.8	-38.6	1.94	-4.4	-18.9	-38.5	0	0	-0.1
B3	2.04	-26.9	-85.6	41.2	2.04	-26.9	-86.1	41.2	0	0	0
F2	2.6	-15	-51.9	15	2.6	-15	-52.4	14.9	0	0	0.1
F1c	2.91	-3.8	-19.9	-81.8	2.86	-15	-13.0	-156.8	0.05	-2.3	75
B2	2.88	-14	-49.8	7.2	2.85	-10.8	-40.6	-5.9	0.03	-3.2	13.1
B1c	1.83	-23.3	-74.2	40.3	1.83	-23.3	-74.8	40.3	0	0	0
Homodimers											
F3	1.3	-31.3	-96.5	-26.8	1.29	-31.3	-97.0	-26.7	0.01	0	-0.1
B3	0.17	-44.5	-132	0.2	0.17	-44.5	-132.9	0.2	0	0	0
F2	0.75	-19.9	-61	-46.1	0.75	-19.9	-61.4	-46.1	0	0	0
F1c	0.28	-32.1	-95.8	-10.6	0.28	-32.1	-96.3	-10.6	0	0	0
B2	-0.74	-55.8	-162.8	215	-0.75	-55.8	-163.8	216	0.01	0	-0.1
B1c	-1.29	-49.9	-143.8	19.9	-1.28	-49.9	-144.6	19.9	-0.01	0	0
Primer + Target											
target-B3	-8.97	-164	-458.4	61.8	-8.92	-163.9	-461	61.7	-0.05	-0.1	0.1
target-F1c	-8.93	-140	-387.7	64.9	-8.9	-140	-390	64.8	-0.03	0	0.1
target-B2	-7.38	-148.4	-417.1	61.4	-7.37	-148.4	-419.5	61.4	-0.01	0	0
Primer + Target (R&C)											
targetRevCompl-F3	-9.46	-148	-409.6	62.5	-9.43	-148	-412.2	62.5	-0.03	0	0
targetRevCompl-F2	-7.6	-156.7	-441.1	62	-7.59	-156.7	-443.6	62	-0.01	0	0
targetRevCompl-B1c	-10.04	-152.2	-420.5	67.4	-10.02	-152.2	-423	67.3	-0.02	0	0.1

Figure 21: Example of the Excel spreadsheet used for the comparison experiments between VisualOMP and LAMPZero. ΔG , ΔH and T_m values for all the possible dimers/monomers combinations were calculated using VisualOMP (light blue) or LAMPZero (light orange). The comparison results, obtained by the difference between the two values for each possible combination, were marked in green if concordant (ΔG , $\Delta H \leq \pm 0.01$; $T_m \leq \pm 0.1$) or in pink if discordant (ΔG , $\Delta H > \pm 0.01$; $T_m > \pm 0.1$).

8. CHARACTERISTICS AND REQUIREMENTS

Programming Language: MS Visual C# 2010

NET Framework: MS Framework 3.5

Client/Server System

- Database Format: MS SQL Server 2008
- Client Compatibility: (Microsoft) Windows XP, Windows Vista e Windows 7

RESULTS

My PhD project has been focused on the creation of a software application specific for LAMP primer design, in order to improve and make easier the very first phase of the development of a LAMP assay: the design and selection of a candidate primer set that could guarantee a proper amplification of the target DNA sequence with high sensitivity and specificity.

Although some software applications specific for LAMP primer design already exist, they often have several limitations: therefore, the selection of the candidate primer set is often time-consuming.

At this purpose, we have developed a LAMP primer design software application that overcomes limitations of current primer design tools, making the phase of design more effective.

As a first step, we have created a database of thermodynamic parameters and we have integrated it with the hybridization algorithm: this tool has enabled us to obtain primers melting temperatures calculation and prediction of molecular interactions at the different temperatures and solution conditions of the reaction.

Once created the engine of our prediction system, we have developed two main primer design modes: an assisted mode to guide and help the operator in the manual selection of primers on the target and an automatic mode that generates a series of primer sets, according to the acceptance criteria previously set by the operator.

The design system that has been developed allows a researcher to specifically design LAMP primer sets on a DNA target sequence through the use of a series of graphical interfaces, taking into consideration the desired reaction conditions and design criteria.

1. SOFTWARE PROTOTYPE BUILDING

The first step in the development of DiaSorin software application prototype consisted in taking into consideration the key factors specific for LAMP primer design.

As previously mentioned in the Introduction (refer to the section “Primer Design”), they include both factors specific for each individual primer and factors specific for the entire primer set. In particular, in the evaluation of the most important design criteria for individual primers, we have considered primer length, GC content, non specific binding, homodimers formation, stability at the end of each primer and melting temperature (T_m). In the analysis of the main design criteria for primer sets, we have considered distance between primers, T_m balance between each couple of primers and dimers formation.

All these key factors should be included into the software system in order to guarantee an efficient selection process of the best primers.

Among the key factors for LAMP primer design, we focused our attention on T_m calculation, primers end stability evaluation and secondary structures formation. The melting temperature (T_m) is an index of the stability of a primer-template DNA duplex and strongly affects the sensitivity and the specificity of the amplification. The stability of the end of

the primers is helpful for anchoring the primer onto the template for efficient priming. The evaluation of secondary structures formation is useful to determine possible non specific products formation.

In order to match these three key factors, we took into consideration the ΔG value, which is an index of the thermodynamic stability of a reaction and, in turn, affects these three important parameters. For this reason, an accurate calculation of ΔG is an extremely important point in the development of a primer design software.

ΔG is usually calculated through the Nearest-Neighbor method and these values (ΔG , ΔS , ΔH , at 37°C and 1M NaCl) are published in reference documentation. Such thermodynamic parameters have been entered into our software application database, together with formulas to calculate ΔG at different temperatures and salt concentrations and to determine T_m .

In addition, a hybridization algorithm, that predicts the most stable structures at certain temperatures and salt conditions, has also been integrated into our system.

The creation of such a database, containing more than 500 thermodynamic parameters records (ΔG , ΔH , ΔS) for all the possible nucleotide combinations (Watson and Crick pairs, internal and terminal mismatches, loops etc.), is the core of DiaSorin primer design system and has allowed us to calculate accurate T_m , as well as to describe primer thermodynamics.

This has represented the starting point for the development of an assisted primer design system that could efficiently help the operator in the 'manual' selection of the primers on the target.

2. VERIFICATION OF LAMPZERO PREDICTIONS: COMPARISON WITH VISUAL OMP

We have verified the accuracy of our software prototype performing some comparative studies with VisualOMP™ (DNA Software, CEO J.SantaLucia Jr.), our reference system for melting temperatures (T_m) and dimers prediction.

We have performed a total of 20 “*in silico*” experiments on 20 different primer sets. For each one, we have analyzed ΔG , ΔH & T_m on 27 primer combinations (including homodimers, heterodimers and primer-target duplexes) and 6 monomers (hairpins), in order to evaluate if the two systems could generate similar T_m and thermodynamics predictions.

The results collected in table 3 show that we have obtained corresponding results, both on dimers and hairpins evaluation, suggesting that our prediction of primers Melting temperatures and dimers could be comparable to the Visual OMP one.

The very few discordant results obtained on monomers do not represent a relevant issue, since these structures tend to be extremely unstable at the reaction temperature of a LAMP assay.

	CORRESPONDING RESULTS	DISCORDANT RESULTS	% DISCORDANCE
Hairpins Evaluation	349	11	0,03%
Dimers Evaluation	1374	6	0,004%

Table 3. Summary results of the comparison experiments between Visual OMP and LAMPZero software applications, concerning hairpins and dimers (primer-target, self dimers, cross dimers) evaluation.

3. LAMPZERO PRIMER DESIGN: DEFINING DESIGN CRITERIA AND REACTION CONDITIONS

As previously discussed, an accurate selection of primer melting temperatures (T_m) is important for successful reactions and essential for the proper annealing of the primers on the target at the constant reaction temperature. At this purpose, it is necessary that the primer design software provides predictions of the primers melting temperatures that are the closest to the LAMP experimental conditions.

Since the reaction conditions (such as temperature, salt concentrations, etc.) can affect the melting temperature of the primer-target DNA duplex, it is important, when designing new primers, to adequately set the design criteria, in order to obtain the most reliable T_m calculations.

When creating a new design experiment, three are the main steps to be considered:

1. Select a target
2. Define the solution conditions
3. Define the design criteria

Once target, reaction conditions and design criteria have been set, the user can decide to perform the primer design through the assisted mode, the automatic mode or with the combination of the two (combined mode).

3.1 - Target selection

When creating a new design experiment, a new ID (progressive number) will be attributed to it. User and creation date will be automatically added.

The system provides the possibility to enter comments and the type of assay linked to the design experiment, in order to specify all the information needed.

Moreover, it is possible to change the status of the experiment in “open”, whether there is the need to modify the experiments, or “closed”, whether the experiment must not be changed.

The very first step in the phase of primer design is represented by the selection of the target DNA sequence. It is possible to select the target sequence from the available list, previously created, or also add the sequence of a new target. Once selected, the target sequence will be assigned to the experiment (Figure 22).

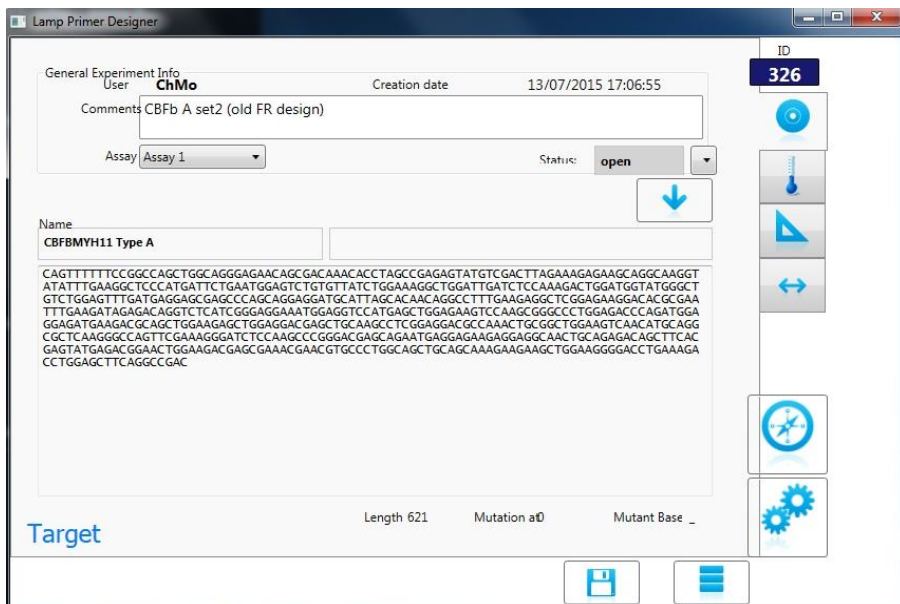


Figure 22. Screenshot representing the selection of the target sequence.

3.2 - Solution conditions

The second step in the phase of primer design is represented by the definition of the solution conditions (Figure 23).

In order to provide accurate thermodynamic parameters and therefore accurate calculation of melting temperatures and prediction of secondary structures, LAMPZero allows to define the specific experimental conditions.

In particular it is possible to change primers concentration, the monovalent and divalent cations concentration, glycerol percentage and the assay temperature, in contrast to Primer Explorer 4, where these parameters have predetermined fixed values, thus determining poor T_m calculations.

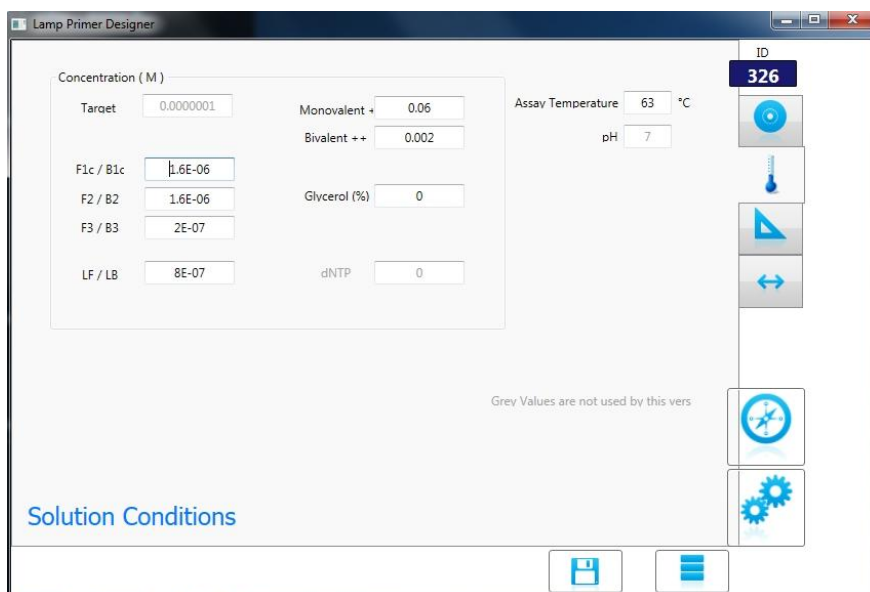


Figure 23. Screenshot of the solution conditions settings, regarding nucleic acid and salt concentrations. The percentage of glycerol and the assay temperature can also be set.

3.3 - Design criteria

The third step in the phase of primer design is represented by the definition of the design criteria, regarding primer length, melting temperature and ΔG threshold of primers end stability and of primer dimers formation (Figure 24).

In the assisted mode the design criteria are used as reference when comparing newly designed primers. In the automatic mode these criteria may be changed to reduce or increase the results generated by the system.

Based on the target GC content, the software assigns default values to primer length and T_m for each couple of primers (F1c/B1c; F2/B2; etc.): these values can further be modified by the user.

The software gives the possibility also to change primers end stability and dimers' stability default values.

Finally, another possible design criteria that can be used by the system as a selection factor is represented by the CG clamp, that is the number of consecutive C/G nucleotides at the end of the primer sequence (last 6 bases). The default value is set on 0: however it is possible to increase the setting, in order to reduce the selection to primers having more efficient priming abilities.

It is worth to highlight that the design criteria represent a 'barrier' in the automatic mode, since the sequences not respecting them are excluded from the eligible results. In the assisted mode, however, such sequences are eligible and can be distinguishable from the sequences meeting the design criteria through the use of conditional formatting.

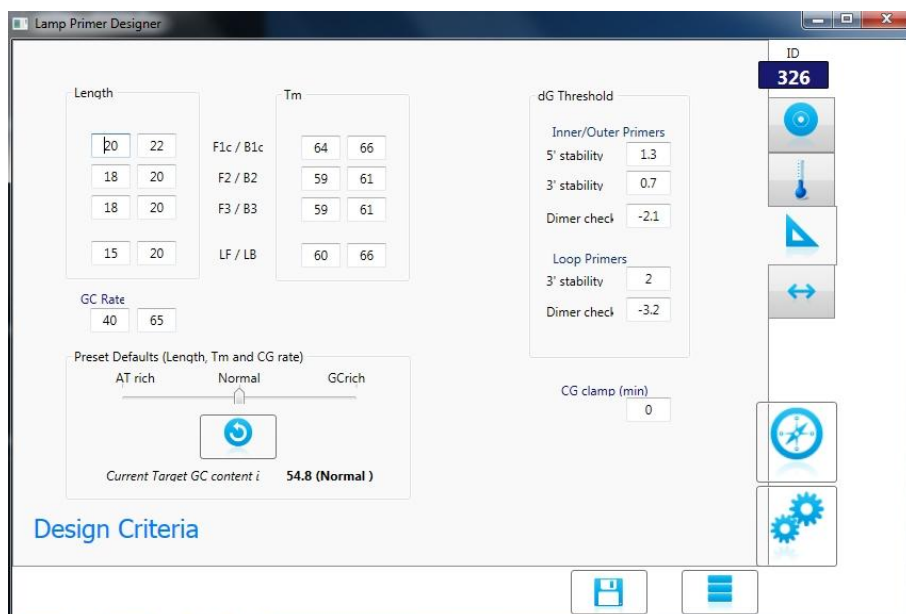


Figure 24. Screenshot of the design criteria settings regarding primer length, melting temperature (Tm) and ΔG threshold for primers end stability and dimers check evaluation.

Another important point in defining the design criteria is represented by the possibility to change distance criteria among the different primers, as shown in Figure 25.

As previously discussed in the introduction, one of the key factors in the LAMP primer design is represented by the correct distances between primers on the target (Figure 6): this 'standard' design mode allows to obtain symmetric dumbbell structures and optimal amplicon length (between 120 and 160 base pairs), theoretically ensuring best performances in a LAMP reaction.

However, during the primer selection phase there is often the need to obtain as many different design options as possible.

Unlike Primer Explorer 4, DiaSorin software application allows to set different distances for the two sides of the dumbbell (for example the distance between F1c and F2 or B1c and B2 primers, corresponding to the loop regions).

In addition, also the distance between the F1c and the B1c (corresponding to the single stranded region between the forward and the backward loop) can be modified, to generate LAMP amplicon of variable length.

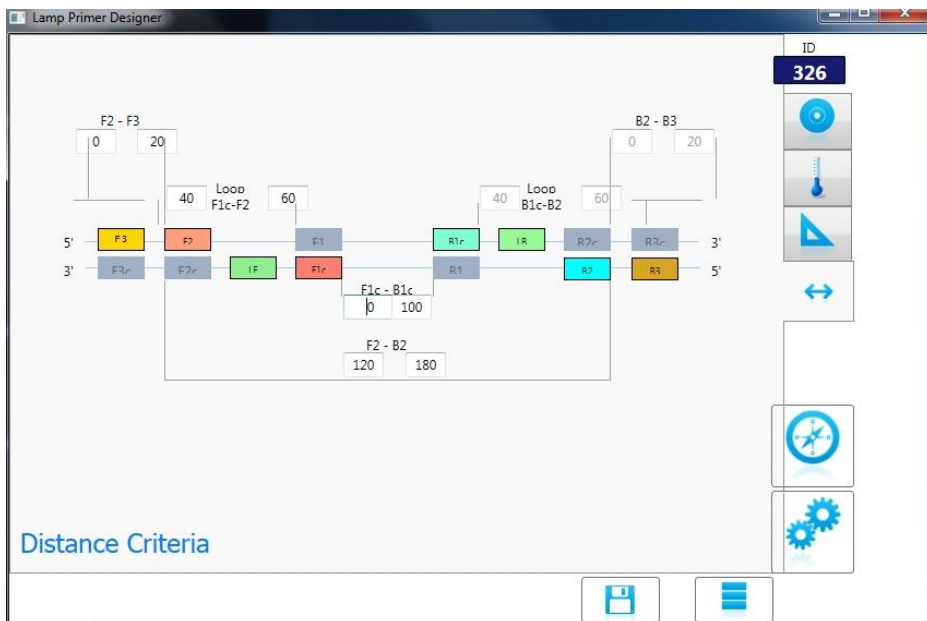


Figure 25. Screenshot of the distance criteria settings: the user can change distances between the different primers of the set, on both the forward and the backward sides.

This 'flexible' design mode could be extremely appropriate in some circumstances. For example, primer design in correspondence to a chromosomal translocation or a genetic mutation requires the precise localization of the primers in a well-defined region of the target in order to perform the molecular diagnosis. If this region is particularly rich in GC or AT nucleotides, is palindromic or homologous to other genomic regions, it could be advantageous, if possible, to modify primers positioning, to avoid the region where the design could be difficult.

4. LAMPZERO ASSISTED PRIMER DESIGN

The assisted design system has been developed in order to help the operator in the 'manual' design of LAMP primers, that consists in the direct selection of the primer sequences on the target. This design mode relies on the use of an algorithm that calculates thermodynamic parameters and predicts secondary structures of the primer sequence together with the other sequences present in the solution, thus providing useful information to guide the operator in the primer selection process.

As shown in Figure 26, the window of the assisted design is divided into three main sections: the design area, the primers data area and the dimers data area.

In the design area, the operator can design primers through direct selection on the target: the user can select the primer label (F3, F2, F1c, etc.), deciding on which primer to work, then select a sequence of bases on the target and adjust the sequence on both ends adding or removing bases.

Once the design phase on the primer has been finished, the current selection can be stored in the grid below (primers data) in the proper corresponding primer section.

Moreover it is possible to detect a past sequence, copied from a document (ex: Primer Explorer results page, word document, etc.), simply selecting the primer label and pasting the sequence in the dedicated space. The system looks for the sequence as it is (but also for its complement, reverse or reverse and complement) allowing to find it along the target. Once the selection is in place, the system calculates values considering the type of primer that was selected in the label.

In the primers data area, the operator can visualize all the main information regarding each primer that has been designed: its position along the target, length, sequence, CG%, primers end ΔG , CG clamp, concentration, T_m , information whether it has passed or not the acceptance criteria and its score.

The system gives also the possibility to adjust a primer sequence, comparing sequences of the same primer type, in order to possibly find a better primer. Once selected the primer label, the operator can use the test sequence space in order to design new modified primer and compare it to the same primer type already stored in the grid.

In the dimers data area, the operator can obtain information regarding secondary structures formation: in particular, by a double click on ΔG prediction values, pictures of heterodimers or hairpins can be visualized.

Modifications of the primer sequence determine changes of the predicted values (regarding both primers and dimers data).

During the design the user can obtain a real time calculation of primers and dimers data, that are marked in green or in red, whether they have matched or not the acceptance criteria inserted into the system. In this way the operator can be quickly and easily guided by the system through the difficult phase of manual primer design.

Design Area

CAAGTTGCCCTAATCAGAAATAAATTCATTGCAACGTTAAATACATCACAAAATGGCCGCTCAAAAACGGCTTGAACATATAATGACTGCCTCTCATTC
 TCTGTTTCATTTTCACTCCCATATTAAGGTGAAATATAATCGCTATGCCGAGATAGTGGCAACATATTGCTCTCGCATTTTGGCAACCCAAAT
 GGAGGCCAGACGAACGAGATACCATATGCAAAATGTAGCTAATTTTCTCCCATATATATAAGCCGAAAACGGCTGTTAAAAGTCCCAATAGATT
 GCCAATTAATGACTATATATTCTGAGATAAAATGACCAGATTAATATAAAGCCATGTTTTGTATTGCTTTTTTATACCCCTTACTCATATAT
 TTAAGGGGTACACTATATCTGTGAGAGATGTAAACAACAGTATAAATGCTCCGACTATATAAAGTATATATTTTTGATCAATCAAAAGCCG
 AGCGCATAGCCATCCGCTGCTGCTAGCAAGCTCAGATCTCAAGACTATAAATGAGAGGTTGAGATTAGCAGCATATAGACTGAGCG
 TGGAGCGAGTGCAGTTGGCCAGCTTAATGGCTTGCCTATTACGCTCCAAAACCCGCAAAATGCCATGCCACATTTTTAAACCATTTTCTCGATA
 TTTTTCACATTTTATACCTTTGTAAATCTATACGCTCTCAAAAAAATTTGCCATGCCACATTTTTTAACCATTCTCGATA

Primers Data

Solution Conditions: $t^m = 63.0$ Na[] = 6,000 M Mg[] = 2,000 M glycerol % = 0,0

Name	Pos	Len	5' Sequence	CG%	End dg	C1	Conc. M	Tm	i	Score	Info
F3	-	926	21 tccgtctgttcgtatGAACGT	47.6	-3.6	2	2E-07	92.2	▲	37.4	
B3	-	1129	22 gtpatagaaattacaAAGTTA	22.7	-1.8	1	2E-07	78.7	▲	26.3	
F2	-	958	20 gaactataaantcaGAAGG	30.0	-3.4	2	1.6E-06	81.1	▲	41.5	
F1c	-	1000	18 GGTCTCTGCTCTGAGAA	50.0	-5.5	3	1.6E-06	86.8	▲	35.2	
B2	-	1100	19 aaaaaaataggaaatATGG	21.1	-3.2	2	1.6E-06	80.1	▲	31.5	
B1c	-	1057	21 ACGTCCacaanccgcacaaa	47.6	-3.5	2	1.6E-06	94.1	▲	34.6	
LF	-						8E-07			0.0	
LB	-						8E-07			0.0	
Test	N/A						0			0.0	

Dimers Data

	B1	F2	F1c	B2	B1c	LP	LB	Test	homo	hpin
x1 >	-1.2	-2.4	-1.9	-2.0	-3.1				-6.2	-1.1
x2 >	-1.9	-1.8	-1.6	-1.2					-2.0	0.0
F2 >	-2.2	-1.5	-1.0						+0.1	-1.6
F1c >	-2.6	-2.5							-2.9	-1.4
B2 >	-2.0								+0.9	-1.9
B1c >									-2.5	-1.3
LP >										
LB >										
Test >										

Secondary Structures

Figure 26. Screenshot of the assisted design mode. In the design area the user can design, modify, delete the primers on the target sequence: the primers selected are highlighted through the use of different colours. A summary of the primers information (position, length, CG %, etc.) is reported in the primers data area: they are marked in green (if right) or in red (if wrong) based on the acceptance criteria inserted in the system. Secondary structures prediction (ΔG values) is shown in the dimers data area.

Moreover, the folding algorithm is able to predict possible unexpected loops, showing them in the graphical representation of the dumbbell. Once they are detected, the user can potentially decide to perform a new primer selection (Figure 28).

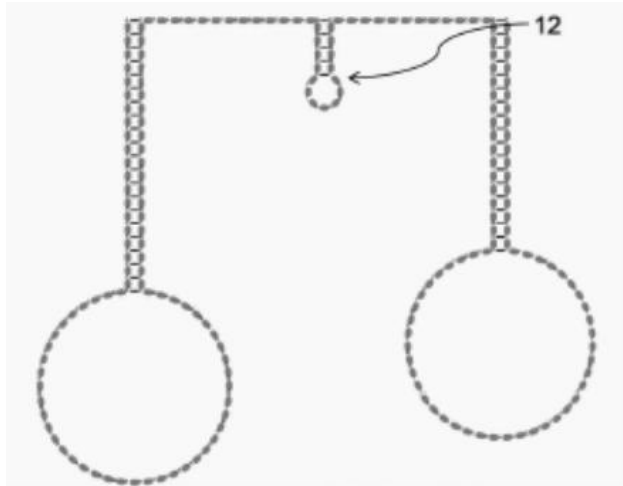


Figure 28. Example of an unexpected loop (12) formed by a LAMP primer set and visualized by the graphical representation of the dumbbell structure.

The visualization of the dumbbell structure can be important during the primer design phase, as it allows to make evaluations on the size of the dumbbell, the dimension of the loops, the symmetry of the dumbbell, the optimal positioning of the loop primers and eventually of the "stem primer", selected in the central part of the dumbbell ("stem region" area of the target, between F1c and B1c).

The size of the dumbbell is related to the amplicon length, that is, the region amplified by the LAMP method. Since there is a direct correlation

between the amplicon length and the amplification speed (i.e. small amplicon are generally amplified faster than larger ones), it is important to evaluate the size of the dumbbell, also to make comparisons among primer sets, potentially identifying the most appropriate based on the researcher's needs.

Another important parameter is represented by the dimension of the loops. Firstly, it can provide indication about the amplification of the dumbbell, since strongly asymmetric ones will unlikely be amplified by the LAMP method. Secondly, it can be useful to analyze the multiple options for loop primer design in that region.

Another important aspect of dumbbell visualization is linked to the possibility to display the loop primers on the dumbbell, allowing to more accurately evaluate their positioning on the loop regions, even by comparison with previous primer designs.

4.2 - Reports

The system generates a series of reports, useful to obtain all the complete information on individual primers and primer sets.

By double clicking on a row in the primers data area, the user can get the 'primer report'. This report contains information on possible failed acceptance criteria, non specific binding (NSB) map (to detect possible target locations where the primer is not designated to bind) and primer score calculation (the closer the values are to the acceptance criteria, the higher will be the score obtained by the primer, as explained in the Materials and Methods section, Figure 18).

The 'primer set report' contains summarized information of all the primers of the set, showing a list of the primers that failed (if any) the selection process since they didn't match the acceptance criteria. This report also highlights the most stable primer dimers, suggesting primers that could give possible specificity issues.

Moreover it shows the primer set score calculation: the partial scores calculated for each individual primer, for each pair of primers (T_m balance) and for dimers formation stability are then summed together in order to provide the final score for the primer set designed.

5. LAMPZERO AUTOMATIC PRIMER DESIGN

The automatic design system allows the selection of the best '*in silico*' primer sets among all those automatically generated by the software according to our acceptance criteria.

The algorithms of the software perform automatic design through two principal steps:

- a first step of primers selection, in which several thousands of primers are evaluated versus acceptance criteria, then a score is calculated for the accepted ones;
- a second step of primer sets selection, in which several thousands of possible sets are identified through the combination of the accepted primers. The primer sets are evaluated versus acceptance criteria and a calculated score is assigned to the accepted ones, highlighting the best candidates.

Therefore, the definition of proper acceptance criteria is a crucial step for the selection of the best primer sets by the system. The criteria used for the score attribution are still being evaluated.

In the automatic design, given a nucleotide target sequence, we can obtain a list of possible primer sets.

Before performing a search, the user can define the design parameters to be considered, such as selecting the number of F3-B3 pairs to be possibly matched with the found FIP-BIP combinations or selecting the number of loop primers to be possibly matched with the found primer sets (with the possibility to check only for LF or LB primer). If getting too much results, it could be useful to use lower values and viceversa (Figure 29).

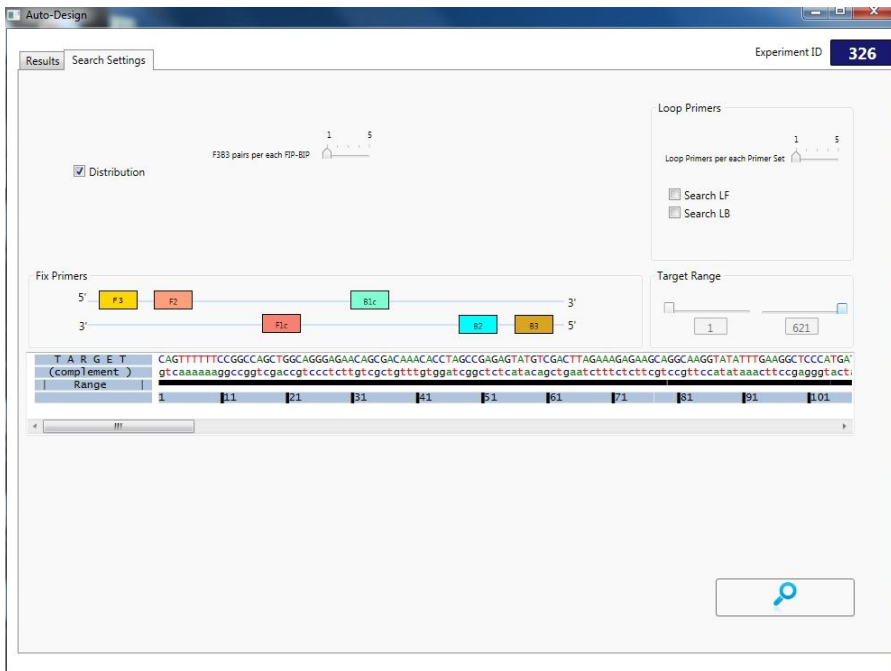


Figure 29. Screenshot of the search options to be set in the automatic primer design mode.

After setting search options, the system will find a number of possible primers which match the design criteria (Figure 30), following this order:

- Number of FIP
- Number of BIP
- FIP + BIP
- Primer sets
- Primer sets + loop primers

The system will calculate all the possible secondary structures (dimers and hairpins) of the found FIP+BIP couples and of the primer sets.

Cutoff settings allow to reduce the number of results, discarding the ones below a certain score.

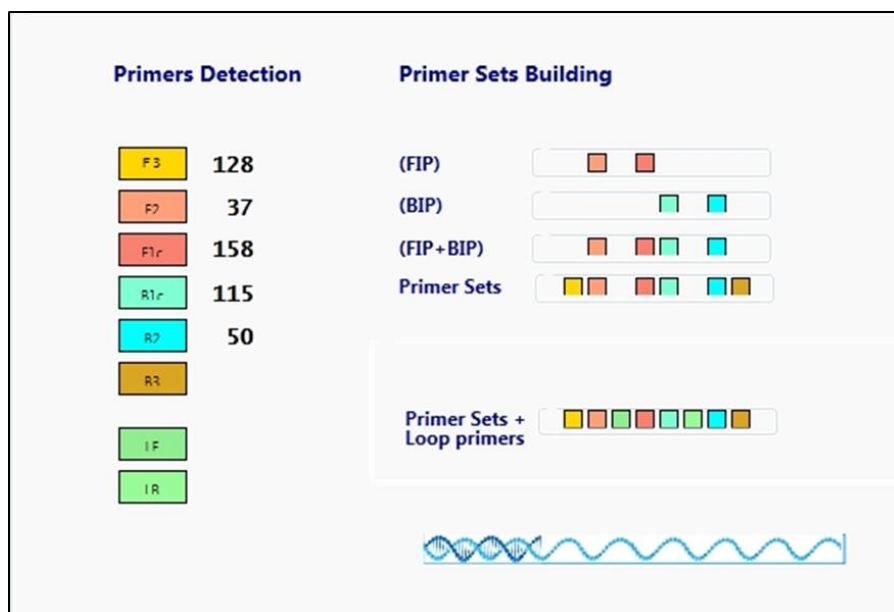


Figure 30. Screenshot of the primer search performed by the automatic design mode after the selection of the design settings. The system initially detects along the target all the possible primers matching the design criteria; then the primers are combined between each other for the primer sets building and finally loop primers are assigned to each primer set. The status of the search is showed by the double helical structure on the bottom right corner of the screen.

After the selection process, the results are displayed, sorted off by score (first the highest ones, than the lowest ones). With a click on a row, the related primer set is shown in the box.

Each set can be marked as favourite (a green rectangle will appear near the selected set) and moved to the assisted design for any possible adjustment or modification. It is also possible to limit the view to the favourite sets, especially when lots of primer sets are generated (Figure 31).

The screenshot shows the 'Auto-Design' software interface. The window title is 'Auto-Design' and the Experiment ID is 326. The main area displays a list of primer set candidates, sorted by score. The table has columns for primer ID, score, and sequence. The table is sorted by score, with the highest scores at the top. A search bar is visible at the top left, and a 'P.Set Temp ID' field is at the bottom left. A 'Favourites' checkbox is at the bottom right.

Primer ID	Score	Sequence
MP18	-1.0	55.0
MP18	-1.0	54.9
MP18	-1.4	54.7
MP18	-1.4	54.3
MP18	-1.4	54.2
MP18	-1.4	54.0
MP18	-0.8	53.9
MP18	-1.0	53.2
MP18	-1.0	53.0
MP18	-1.0	52.4
MP18	-1.0	52.1
MP18	-1.0	52.1
MP18	-1.4	52.0
MP18	-1.4	52.0
MP18	-1.4	51.8
MP18	-1.4	51.6
MP18	-1.0	51.5
MP18	-1.4	51.2
MP18	-1.4	51.2
MP18	-1.4	51.1
MP18	-1.4	51.1
MP18	-1.4	51.0
MP18	-1.4	51.0
MP18	-1.4	50.9
MP18	-1.4	50.8
MP18	-1.4	50.7
MP18	-1.4	50.5

Figure 31. Screenshot of a list of possible primer set candidates obtained with the automatic design mode.

Other features of the automatic design are represented by the possibility to design around a specific primer(s) called fixed primer(s) and to adjust the target range (the default is the whole target sequence).

It is possible to select the sequence that we want on the target and click on the label to fix the primer. Fixed primers will not be included in the search for new primers, so all the results will have as expected the same fixed primer. This feature could be useful when performing a new primer search or implementing an older one, not modifying a specific primer that proved to be efficient in experimental conditions.

Through the so-called 'combined' design mode, a primer set obtained with the automatic module can be moved to the assisted design, for any possible modification according to the user's needs.

6. PRIMERS DIMERS EVALUATION IN MULTIPLEX SYSTEMS

The first version of LAMP technology initially described by Notomi consisted of a simplex reaction amplifying a single gene of interest real time-monitored through turbidimetry.

Over the years, however, this method has been modified and implemented: for example, DiaSorin has developed a Q-LAMP technology able to guarantee the amplification and the detection of different target transcripts in a single reaction through the use of specific fluorescent probes.

In the case of multiplex reactions, where many primers at high concentration are included in the reaction mix, the probability to obtain template-independent primer interactions that give rise to non-specific products formation increases.

However, none of the existing LAMP primer design software applications was able to predict secondary structures formation in multiplex reactions.

For this reason, we have developed a tool that allows to evaluate not only the molecular interactions between primers of the same set (as in the assisted mode), but especially the formation of dimers between the primers of different sets.

This tool uses the software database in order to calculate thermodynamic parameters and then melting temperatures, while the UNAFold hybridization and folding algorithms {Mfold, Markham, NR Zuker, M.} are used for the evaluation and visualization of the secondary structures.

Before performing a search, the system allows to select the primer sets (up to a maximum of three) and the solution conditions (such as assay temperature and salt concentrations). Moreover it provides also the option to consider complete inner primers (e.g. FIP) instead of their components (e.g. F2 and F1c).

After the search, the user can obtain a list of the primer dimers predicted by the system, together with information regarding the type of structure formed (homodimer, heterodimer...), ΔG values, melting temperatures and their concentrations (Figure 32).

Sequence 1	Sequence 2	Structure	Delta G	Ext.	Delta H	Delta S	Tm	Eff DG	Eff Tm	Concentration	Order	Seq 1 pct	Seq 2 pct
71/TargetRC	71/B1c	Primers=Target	-10.0	99	-152.2	-420.5	67.3	-9.47	66	6.69E-11	4047	66.88	4.18
74/F3	74/B3	Heterodimer	0.1	0	-29.6	-87.7	-26.0	0.09		3.48E-17	5001	0.00	0.00
74/F3	74/F2	Heterodimer	0.2	1	-32.6	-96.8	-9.3	0.18		2.45E-16	5002	0.00	0.00
74/F3	74/F1c	Heterodimer	-0.2	0	-35.5	-104.5	-2.2	-0.13		3.88E-16	5003	0.00	0.00
74/F3	74/B2	Heterodimer	0.7	5	-34.5	-104.1	-9.0	0.91		8.20E-17	5004	0.00	0.00
74/F3	74/F3c	Heterodimer	0.3	0	-22.7	-68.0	-33.8	0.39		1.79E-16	5005	0.00	0.00
74/F3	71/F3	Heterodimer	0.5	0	-23.1	-69.7	-46.3	0.55		1.76E-17	5006	0.00	0.00
74/F3	71/B3	Heterodimer	0.3	0	-22.7	-68.0	-46.6	0.40		2.21E-17	5007	0.00	0.00
74/F3	71/F2	Heterodimer	0.3	0	-21.5	-64.2	-36.9	0.28		2.11E-16	5008	0.00	0.00
74/F3	71/F1c	Heterodimer	-0.6	0	-38.0	-110.7	3.7	-0.56		7.31E-16	5009	0.00	0.00
74/F3	71/B2	Heterodimer	-0.3	0	-38.5	-113.1	2.7	-0.29		4.89E-16	5010	0.00	0.00
74/F3	71/B1c	Heterodimer	-0.1	3	-41.2	-121.5	4.9	-0.02		3.30E-16	5011	0.00	0.00
74/B3	74/F2	Heterodimer	0.6	0	-28.1	-84.8	-20.9	0.59		1.33E-16	5013	0.00	0.00
74/B3	74/F1c	Heterodimer	-0.9	0	-33.1	-95.1	-1.3	-0.86		1.14E-15	5014	0.00	0.00
74/B3	74/B2	Heterodimer	0.7	0	-20.8	-63.6	-42.4	0.95		7.74E-17	5015	0.00	0.00
74/B3	74/B1c	Heterodimer	0.7	0	-27.1	-82.2	-23.8	0.72		1.10E-16	5016	0.00	0.00
74/B3	71/F3	Heterodimer	0.1	3	-23.2	-68.8	-43.1	0.14		3.25E-17	5017	0.00	0.00
74/B3	71/B3	Heterodimer	0.9	0	-19.3	-59.9	-63.0	1.01		8.84E-18	5018	0.00	0.00
74/B3	71/F2	Heterodimer	0.6	0	-19.9	-60.7	-45.2	0.65		1.21E-16	5019	0.00	0.00
74/B3	71/F1c	Heterodimer	-0.1	0	-34.0	-100.2	-5.5	-0.01		3.26E-16	5020	0.00	0.00
74/B3	71/B2	Heterodimer	-0.2	0	-36.4	-107.1	-1.0	-0.14		3.97E-16	5021	0.00	0.00
74/B3	71/B1c	Heterodimer	-0.7	13	-38.9	-112.8	5.5	-0.62		8.05E-16	5022	0.00	0.00
74/F2	74/F1c	Heterodimer	0.5	0	-37.1	-111.3	-6.6	0.60		1.05E-15	5024	0.00	0.00
74/F2	74/B2	Heterodimer	0.6	0	-29.3	-88.6	-21.5	0.90		6.75E-16	5025	0.00	0.00

Figure 32. Example of a screenshot showing the results obtained with the multiplex evaluation tool. The type of primers interacting (F3, B3, etc.), their corresponding set (ID74/ID71), the structure formed (primer-target, heterodimer, etc.), structure ΔG , extensibility and concentration are provided as basic information by the system.

The software also provides a graphic representation of the predicted secondary structures, helping the researcher in the evaluation of the primer dimers potentially most 'dangerous' based on the level of hybridization to each other (Figure 33).

In addition, the system is able to assess the primer dimers that could be extensible on their 3'end by the polymerization activity of the enzyme according to the complementary sequence, providing information on the number of base pairs of the extension (Figure 34).

In the analysis of secondary structures formation, ΔG and extensibility are two of the most important parameters to be evaluated: negative ΔG values and a high number of extensible bases at the 3' end will more likely favour non specific annealing and primer elongation events between two primers.

This tool for the evaluation of primer dimers formation in multiplex reactions allows to perform a preliminary screening of the oligonucleotides during the design phase. The researcher can delete, modify or redesign primers that are prone to form secondary structures between each other, in order to potentially prevent non specific products formation.

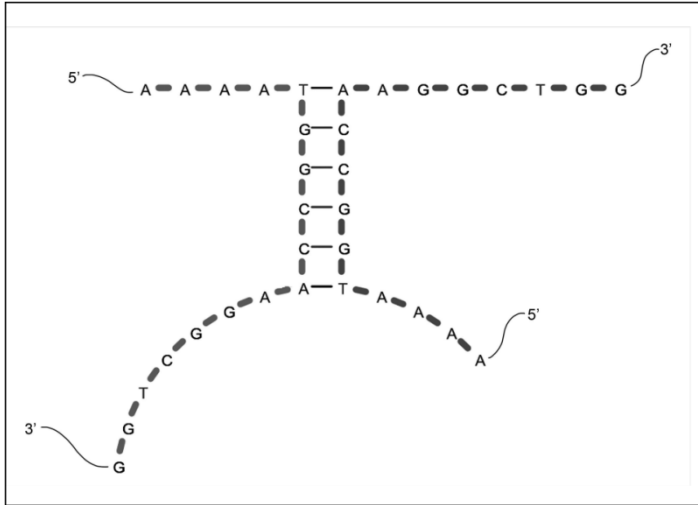


Figure 33. Example of an internal self dimer annealing for a total of 6 base pairs.

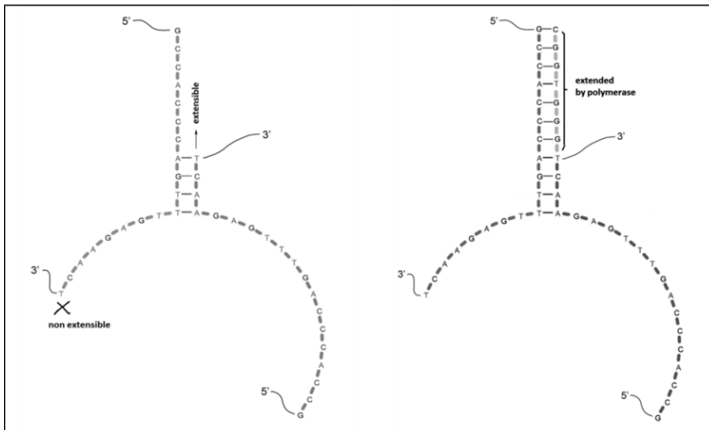


Figure 34. Example of a 7-base pairs extensible self dimer before (on the left) and after (on the right) the activity of the DNA polymerase enzyme.

DISCUSSION

Loop-mediated isothermal AMPlification (LAMP) is an innovative gene amplification technique alternative to PCR, whose main features consist of constant temperature (isothermal reaction), the employment of a strand-displacement polymerase and the use of 4-6 different primers.

The mechanism of the LAMP amplification begins with the annealing and the extension of the inner primers (FIP and BIP) on the target DNA; the newly synthesized strand is then displaced by the annealing and the extension of the outer primers (F3 and B3). The displaced product generates a “stem-loop structure”, called dumbbell, which represents the starting structure for LAMP exponential amplification.

The use of two additional catalytic primers, called Loop Primers (LF and LB), provides an increased number of starting points for DNA synthesis, thus increasing the efficiency of the amplification reaction.

Since the commonly used application of LAMP requires a total of six primers (with FIP and BIP primers encompassing two primer binding sites each), primer design for LAMP assays requires the selection of eight different regions of a target nucleic acid sequence.

It should also be considered that LAMP assays take place at a constant temperature, in contrast to PCR that occurs by repeated cycles of three different temperatures. This aspect introduces further complexity in the phase of primer design: primers melting temperatures (T_m) should be

accurately selected to promote the synchronized annealing of the numerous primers on the target at the proper reaction temperature.

Moreover, significant care must be taken in order to minimize dimers formation among the primers used in the reaction, especially for FIP and BIP primers that are generally about 40 nucleotides long.

For all the reasons previously explained, it is not difficult to infer that the primer design phase is a crucial step in a LAMP assay development and the use of an accurate primer design software application is of extreme importance.

Since LAMP is an emerging amplification technique, LAMP primer design systems are rather limited.

The most common method to specifically design primer sets for LAMP method is to use Primer Explorer V4 from Eiken Chemical Co. Ltd. (<http://primerexplorer.jp/e/index.html>). It is an Internet free software that can be used without any installation procedure. Based on the target sequence information and the primer design requirements provided by the user, this software automatically generates primer sets composed by the basic 4 primers (F3, FIP, BIP, B3) required by the LAMP method. Through a second additional step, it can also design the loop primers.

However this software has several limitations. First, it uses the "Nearest-Neighbor" method for T_m calculation, not considering any possible modification to the standard LAMP experimental conditions, such as reaction temperature or salt concentration variations: therefore it can give poor melting temperature (T_m) estimations.

Second, it does not provide manual design options, not allowing any possible adjustment to the primer sets automatically generated by the system and thus having significant restrictions in primers positioning on the target.

Third, loop primers cannot be designed in a single design step, instead requiring two serial executions, which can prevent more optimal primer combinations from being identified.

Fourth, it does not allow to evaluate primer dimers and last, it does not ensure intellectual property, being a public internet software application.

Due to the primer design restrictions described, the automatic design of Primer Explorer V4 has to be verified through the use of Visual OMPTM (DNA Software, CEO J.SantaLucia Jr.), an efficacious tool for in silico primer design specific for PCR, that provides better T_m estimations and secondary structures predictions, allowing to discriminate the candidate primer sets among those generated by Primer Explorer V4.

As a result, the phase of primer design is multi-step, laborious, time-consuming and the selection of the candidate primer set has to be done by numerous trial-and-error laboratory experimentations. In this contest, the need for an efficient LAMP primer design system has emerged and become urgent.

In order to overcome all the limitations of existing LAMP design programs and make the phase of primer design easier and more effective, we have developed a LAMP primer design software application tailored to DiaSorin LAMP requisites.

DiaSorin LAMP primer design software (also named LAMPZero) has been created to be a versatile tool for specific LAMP primer design. The processing steps of this system are outlined in Figure 35.

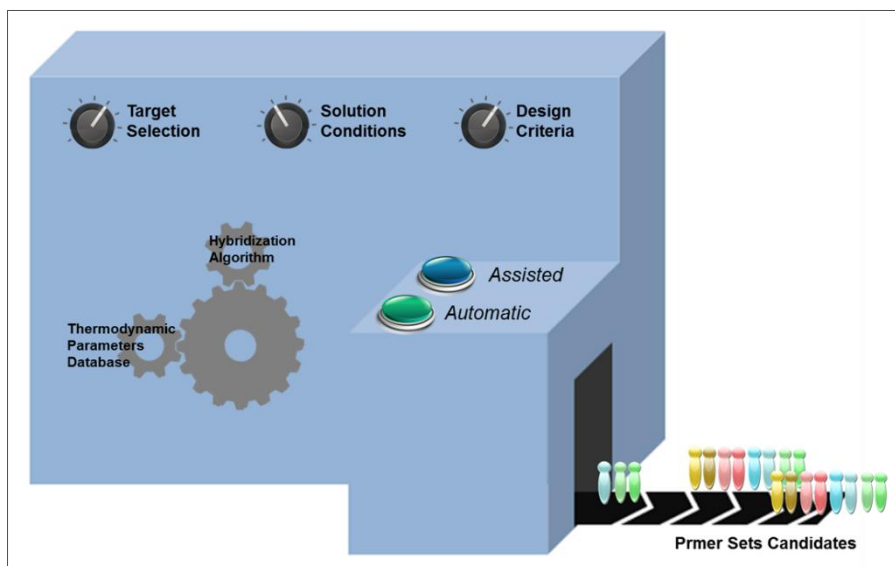


Figure 35. Schematic representation of LAMPZero primer design software application. The engine of the system consists of a database of thermodynamic parameters integrated with a hybridization algorithm. After the selection of a target sequence, the solution conditions and the design criteria, the user can select between an assisted or an automatic design mode, in order to obtain primer sets candidates.

The engine of this primer design system is represented by the integration of a database of thermodynamic parameters together with a hybridization algorithm derived from UNAFold {Mfold, Markham, NR Zuker, M.} (step 1). In contrast to Primer Explorer, that provides T_m calculation

through the Nearest-Neighbor method at reference conditions (37°C, 1M NaCl, 0M MgCl₂), the database of our system is able to calculate ΔG values and hence determine T_m at different temperatures and salt concentrations. In addition to that, the hybridization algorithm can efficiently provide information on primers thermodynamics, predicting the most stable structures at different reaction conditions.

This feature of the system is of extreme importance, because it allows to obtain efficient thermodynamics predictions about primers at the effective temperature and salt concentrations of a DiaSorin LAMP reaction. The predictions obtained with Primer Explorer V4 highly deviate from the real LAMP reaction conditions.

After the selection of a target sequence, the insertion of the solution conditions of the reaction and of the criteria of primer design, the system provides accurate primer T_m calculation and prediction of molecular interactions (step 2).

This implementation has represented a crucial step in the development of an efficient primer design software application able to generate primer thermodynamics predictions that could be as close as possible to the real LAMP experimental conditions.

The accuracy of these predictions has been demonstrated through “in silico” comparative tests against Visual OMP, our reference system in T_m calculation and cross dimers prediction: data obtained have revealed that calculations may be considered equivalent between the two systems, therefore demonstrating the efficiency of our software application.

Both an assisted and an automatic primer design systems have been developed (step 3).

Given a target nucleic acids sequence, the automatic primer design is able to generate LAMP primer sets through 2 principal steps: a first phase of individual primers selection and a second phase of primer sets selection, in which several thousands of possible sets are identified through the combination of the previously accepted individual primers. The definition of the acceptance criteria inserted into the system is extremely important for both the primers and the primer sets evaluation, in order to select the most promising candidates, calculate a score to be assigned to the accepted ones, hence highlighting the best candidates.

The assisted primer design, that guides the operator in the manual selection of primers on the target, can be used in order to adjust the primers designed in the automatic mode (combined design mode) but also to perform a LAMP primer design from scratch.

One of the main added value of LAMPZero compared to other LAMP primer design programs is represented by its **extreme flexibility**, that allows to overcome all the limitations previously experienced, being a tailored tool.

This system allows to easily switch from the automatic to the assisted design mode, in order to make possible adjustments to primer sequences based on the researcher's need: for example the user can modify primers positioning on the target in order to specifically design over a mutation or a

translocation site or to avoid possible primers dimers formation or non-specific bindings.

It also provides the possibility to set the design criteria to be used in the phase of design, thus allowing to implement the standard LAMP primer design mode, the one described by Notomi and traditionally used.

For example, one of the advantages of LAMPZero is represented by the possibility to modify the reference distances between primers defined as guidelines for LAMP technique. These series of restrictive rules are followed by PrimerExplorer software, that allows to design primers in a rather limited way. On the contrary, setting different distances between forward and backward sides of the primer set, as provided by LAMPZero, allows to obtain dumbbells of different size or shape, giving the chance to have greater number of available design alternatives to the standard one.

The definition of the acceptance criteria used for the primer design are extremely important, especially for the automatic design: in this design mode they represent a 'filter' for primer selection, since the sequences not meeting them, are excluded from the eligible results.

Other features of DiaSorin software application are represented by its **efficiency** and **accuracy**.

Unlike PrimerExplorer V4, that requires two serial executions for loop primer design, thus preventing possible optimal primer combinations from being identified, the automatic design mode of LAMPZero system allows the simultaneous design of loop primers, allowing to identify all the possible primer combinations based on the acceptance criteria.

Moreover, if needed, the system provides several options in order to make the automatic primer design even more rapid: for example it is possible to include only one loop primer in the automatic design or adjust the target range, in order to reduce the target sequence where the primer selection has to be performed. Another automatic design option consist in the possibility to design around a specific primer(s) called fixed primer(s), that will not be included in the new primer design: this feature could be useful when performing a new primer search or implementing an older one, not modifying a specific primer that proved to be efficient in experimental conditions.

All these design options make the primer design less time-consuming.

The accuracy of the system in providing primer melting temperatures and secondary structures formation is guaranteed by the use of two different algorithms. The hybridization algorithm allows to predict the conformation of the most stable primer-target structure among all the possible combinations, evaluating its ΔG , ΔH and ΔS values and then calculating the melting temperature. It also predicts the conformation of the most stable structures between primer candidates. The folding algorithm is used to predict the most stable conformation of the hairpin of the primer candidate.

The use of specific formulas that apply corrections to the thermodynamics parameters considering the different solution conditions (temperature, salt and primer concentration) allow to obtain reliable predictions, that could be as close as possible to the real experimental condition of each LAMP reaction.

In addition, it is possible to obtain data regarding non-specific binding (NSB) of the primer along the target, that could suggest the formation of possible aspecific products during the amplification reaction.

Another added value of LAMPZero, compared to PrimerExplorer, consists in its **innovations**, represented by the development of two important tools: one for the visualization of the dumbbell structure and one for primer dimers evaluation in multiplex systems.

The dumbbell structure represents the starting point of the exponential amplification characteristic of LAMP reactions: during the phase of primer design its visualization can be useful to make considerations and evaluations on the size of the dumbbell, the dimension of the loops, the shape of the dumbbell and the optimal positioning of the loop primers. In particular, the analysis of the size of the dumbbell is very important because it is related to the amplicon length (the region amplified by the LAMP method). Since there is a direct correlation between the amplicon length and the amplification speed (i.e. small amplicon are generally amplified faster than larger ones), it is important to evaluate the size of the dumbbell. For example, if the primer design is aimed to select a very efficient and sensitive primer set, dumbbells of small dimensions will be preferred; on the contrary, if the primer design is intended to select a primer set specific for an housekeeping gene that could act as a weak internal control of the amplification reaction, also dumbbells of larger size could be selected.

The evaluation of primer dimers is a fundamental feature when considering multiplex systems, in which the amplification and the detection of different target transcripts take place in a single reaction. The great number of oligonucleotides used (a maximum of 18 different primers for a triplex reaction) increases the probability to obtain dimers formation between primers of the different sets. For this reason, the development of a system that allows to evaluate the formation of dimers between different primer sets is important.

Currently, no other primer design software application allows the evaluation of secondary structures formation in multiplex reactions. Therefore our system provides an extreme advantage in the phase of LAMP primer design.

This tool uses the software database in order to calculate thermodynamic parameters and melting temperatures and the hybridization algorithm to evaluate and visualize primers secondary structures. In particular it provides ΔG values, in order to indicate the most problematic dimers structure (having highly negative ΔG) and information regarding extensibility, that is the possibility that a primer is extended on its 3' end by the polymerase activity of the enzyme.

So, it can be a valuable tool in order to perform a preliminary screening of the primers in the first phase of the design, allowing to delete, modify, or redesign primers that are potentially dangerous as prone to interact between each other.

The analysis of dimers formation between different primers sets can provide a more accurate simulation of the interactions that could take place

among oligonucleotides in a multiplex reaction. If the predictions of the "*in silico*" design get closer to the experimental reality of a LAMP reaction, the design phase and the subsequent experimental evaluation of primers could become simpler, faster and less laborious.

LAMPZero is a **protected design system**, having a private access that ensures intellectual property, unlike PrimerExplorer V4 that is a public internet software application. This is an important added value from an industrial point of view, ensuring the design phase to be preserved and avoiding to spread possible information about the primers used in our diagnostic assays.

Finally a crucial added value of DiaSorin primer design software is represented by the **possibility to be implemented** with future enhancements based on our experience in developing LAMP assays, being fully tailored for DiaSorin.

The importance of this feature can be understood taken into consideration the mechanism of score attribution in the automatic primer design. The definition of proper acceptance criteria is a crucial step for the selection of the best primer sets by the system and the criteria used for the score attribution are still being evaluated. For example, evidences derived from our experiments show the existence of a correlation between size of the dumbbell and amplification speed: in particular, small dumbbells provide usually faster amplification of the target than larger dumbbells. Therefore, the assessment of a score to the amplicon length could

potentially help discriminating the most sensitive primer sets among those designed by the software. It is worth to highlight that the amplification rate of a primer set is influenced not only by the amplicon dimension, but also by other parameters: the attribution of a score depending on the dumbbell size should not be a limit to the design, on the contrary it should be optional based on the needs of the researcher.

The possibility to further improve the automatic design through the adjustment of current score parameters for primer selection based on our LAMP experience will be investigated.

Explorative studies on the software through dedicated LAMP tests will be useful to compare “*in silico*” predictions with “*in vitro*” performances, in order to further unravel the mechanisms underlying LAMP amplification technique, thus implementing the functions of DiaSorin primer design system.

REFERENCES

Allawi, H. T. & SantaLucia, J., Jr. (1997). Thermodynamics and NMR of internal G-T mismatches in DNA. *Biochemistry* 36, 10581–10594.

Álvarez-Fernández R1. Explanatory chapter: PCR primer design. *Methods Enzymol.* 2013;529:1-21.

Aryan E, Makvandi M, Farajzadeh A, Huygen K, Bifani P, Mousavi SL, Fateh A, Jelodar A, Gouya MM, Romano M. A novel and more sensitive loop-mediated isothermal amplification assay targeting IS6110 for detection of *Mycobacterium tuberculosis* complex. *Microbiol Res.* 2009 Jun 8.

Bommarito, S., Peyret, N. & SantaLucia, J., Jr. (2000). Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res.* 28, 1929–1934.

Chen JH, Lu F, Lim CS, Kim JY, Ahn HJ, Suh IB, Takeo S, Tsuboi T, Sattabongkot J, Han ET. Detection of *Plasmodium vivax* infection in the Republic of Korea by loop-mediated isothermal amplification (LAMP). *Acta Trop.* 2009 Sep 22.

Curtis KA, Rudolph DL, Owen SM. Sequence-specific detection method for reverse transcription, loop-mediated isothermal amplification of HIV-1. *J Med Virol.* 2009 Jun;81(6):966-72.

Gandelman O, Jackson R, Kiddle G, Tisi L. Loop-mediated amplification accelerated by stem primers. *Int J Mol Sci.* 2011;12(12):9108-24.

Hillier, L. & Green, P. (1991). OSP: a computer program for choosing PCR and DNA sequencing primers. *PCR Methods Appl.* 1, 124–128.

Hong TC, Mai QL, Cuong DV, Parida M, Minekawa H, Notomi T, Hasebe F, Morita K. Development and evaluation of a novel loop-mediated isothermal amplification method for rapid detection of severe acute respiratory syndrome coronavirus. *J Clin Microbiol.* 2004 May; 42 (5):1956-61.

Hyndman, D., Cooper, A., Pruzinsky, S., Coad, D. & Mitsuhashi, M. (1996). Software to determine optimal oligonucleotide sequences based on hybridization simulation data. *Biotechniques* 20, 1090–1097.

Kaneko H, Iida T, Aoki K: Sensitive And Rapid Detection Of Herpes Simplex Virus And Varicella-Zoster Virus Dna By Loop-Mediated Isothermal Amplification. *J Clin Microbiol* (2005) 43(7): 3290-3296.

Minnucci G, Amicarelli G, Salmoiraghi S, Spinelli O, Guinea Montalvo ML, Giussani U, Adlerstein D, Rambaldi A. A novel, highly sensitive and rapid allele-specific loop-mediated amplification assay for the detection of the JAK2V617F mutation in chronic myeloproliferative neoplasms. *Haematologica.*2012;97(9):1394–1400.

Mori Y, Nagamine K, Tomita N, Notomi T: Detection of loop-mediated isothermal amplification reaction by turbidity derived from magnesium

pyrophosphate formation. *Biochem Biophys Res Commun* (2001) 289: 150-154.

Mori Y, Kitao M, Tomita N, Notomi T. Real-time turbidimetry of LAMP reaction for quantifying template DNA. *J Biochem Biophys Methods*. 2004 May 31;59(2):145-57.

Nagamine K, Hase T, Notomi T. Accelerated reaction by loop-mediated isothermal amplification using loop primers. *Mol Cell Probes*. 2002 Jun;16(3):223-9.

Nakao R, Stromdahl EY, Magona JW, Faburay B, Namangala B, Malele I, Inoue N, Geysen D, Kajino K, Jongejan F, Sugimoto C. Development of loop-mediated isothermal amplification (LAMP) assays for rapid detection of *Ehrlichia ruminantium*. *BMC Microbiol*. 2010;10:296.

Nkouawa A, Sako Y, Nakao M, Nakaya K, Ito A Loop-mediated isothermal amplification method for differentiation and rapid detection of *Taenia* species. *J Clin Microbiol*. 2009 Jan;47(1):168-74.

Notomi T, Okayama H, Masubuchi H, Yonekawa T, Watanabe K, Amino N, Hase T. Loop-mediated isothermal amplification of DNA. *Nucleic Acids Res*. 2000 Jun 15; 28(12):E63.

Parida M, Posadas G, Inoue S, Hasebe F, Morita K. Real-time reverse transcription loop-mediated isothermal amplification for rapid detection of West Nile virus. *J Clin Microbiol*. 2004 Jan;42(1):257-63.

Parida M, Sannarangaiah S, Dash PK, Rao PVL, Morita K. Loop mediated isothermal amplification (LAMP): a new generation of innovative gene amplification technique; perspectives in clinical diagnosis of infectious diseases. *Rev Med Virol.* 2008;18(6):407–21

Peyret, N., Seneviratne, P. A., Allawi, H. T. & SantaLucia, J., Jr. (1999). Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A-A, C-C, G-G, and T-T mismatches. *Biochemistry* 38, 3468–3477.

Rodríguez A1, Rodríguez M, Córdoba JJ, Andrade MJ. Design of primers and probes for quantitative real-time PCR methods. *Methods Mol Biol.* 2015;1275:31-56.

Royce, R. D., SantaLucia, J., Jr. & Hicks, D. A. Building an in silico laboratory for genomic assay design. *Pharm. Visions* 2003; 10–12.

Rychlik, W. & Rhoads, E. R. (1989). A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing, and in vitro amplification of DNA. *Nucleic Acids Res.* 17, 8543–8551.

SantaLucia, J., Jr. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U. S. A.* 95, 1460–1465.

SantaLucia, J., Jr. & Hicks, D. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* 2004; 33, 415–440.

SantaLucia J. Jr. Physical principles and visual-OMP software for optimal PCR design. *Methods Mol Biol.* 2007;402:3-34.

Spinelli O, Rambaldi A, Rigo F, Zanghì P, D'Agostini E, Amicarelli G, Colotta F, Divona M, Ciardi C, Coco FL, Minnucci G. Simple, rapid and accurate molecular diagnosis of acute promyelocytic leukemia by loop mediated amplification technology. *Oncoscience.* 2014 Dec 27;2(1):50-8

Sun J, Najafzadeh MJ, Vicente V, Xi L, de Hoog GS. Rapid detection of pathogenic fungi using loop-mediated isothermal amplification, exemplified by *Fonsecaea* agents of chromoblastomycosis. *J Microbiol Methods.* 2009 Oct 13.

Uemura N, Makimura K, Onozaki M, Otsuka Y, Shibuya Y, Yazaki H, Kikuchi Y, Abe S, Kudoh S. Development of a loop-mediated isothermal amplification method for diagnosing *Pneumocystis pneumonia*. *J Med Microbiol.* 2008 Jan;57(Pt 1):50-7.

Ushio M, Yui I, Yoshida N, Fujino M, Yonekawa T, Ota Y, Notomi T, Nakayama T. Detection of respiratory syncytial virus genome by subgroups-A, B specific reverse transcription loop-mediated isothermal amplification (RT-LAMP). *J Med Virol.* 2005 Sep;77(1):121-7.

Watkins NE Jr, Kennelly WJ, Tsay MJ, Tuin A, Swenson L, Lee HR, Morosyuk S, Hicks DA, Santalucia J Jr. Thermodynamic contributions of single internal rA-dA, rC-dC, rG-dG and rU-dT mismatches in RNA/DNA duplexes. *Nucleic Acids Res.* 2011 Mar;39(5):1894-902.

Yang W, Pin C, Haibing G, Yang C, Hui L, Qigai H. Loop-mediated isothermal amplification targeting the apxIVA gene for detection of *Actinobacillus pleuropneumoniae*. *FEMS Microbiol Lett.* 2009 Nov;300(1):83-9.

Yoda T, Suzuki Y, Yamazaki K, Sakon N, Kanki M, Kase T, Takahashi K, Inoue K. Application of a modified loop-mediated isothermal amplification kit for detecting Norovirus genogroups I and II. *J Med Virol.* 2009 Oct 25;81(12):2072-2078.

Zhang H, M Thekiso OM, Aboge GO, Kyan H, Yamagishi J, Inoue N, Nishikawa Y, Zakimi S, Xuan X. *Toxoplasma gondii*: Sensitive and rapid detection of infection by loop-mediated isothermal amplification (LAMP) method. *Exp Parasitol.* 2009 Jan 31.

PATENT

An Italian patent application (no. 102015000087649) concerning a software-based method for the design of LAMP primer sets to be used for the amplification of a nucleic acid target sequence was requested.

RINGRAZIAMENTI

Il primo doveroso e sincero ringraziamento è per Luciano, con il quale ho collaborato per lo svolgimento di questo progetto. Il suo aiuto nel corso di questi tre anni è stato immenso e preziosissimo e per questo lo ringrazio infinitamente.

Vorrei ringraziare Giulia, per avermi dimostrato fiducia, per aver creduto in me e per avermi dato l'opportunità di affrontare tante sfide, che mi hanno permesso di mettermi alla prova, di migliorare giorno dopo giorno e di crescere.

Vorrei ringraziare il Dott. Colotta per la sua grande disponibilità e per il tempo dedicatomi per la revisione di questa tesi.

Ringrazio di cuore i miei colleghi per le risate, i momenti di allegria e spensieratezza che mi hanno regalato durante questi anni, ma soprattutto per essermi stati vicini nei momenti difficili e per aver sempre cercato di strapparmi un sorriso, anche nelle "giornate no".

Infine, vorrei ringraziare la mia famiglia, che crede incondizionatamente in me e che è sempre al mio fianco.