UNIVERSITAT DE BARCELONA

2010

Application of Surface Plasmon Resonance and spectroscopic techniques for characterizing the interaction between human c-KIT G–quadruplex DNA and water soluble porphyrins

Sintayehu Manaye

Master's Thesis

UNIVERSITAT DE BARCELONA

Application of Surface Plasmon Resonance and spectroscopic techniques for characterizing the interaction between human c-kit G–quadruplex DNA and water soluble porphyrins

by

Sintayehu Manaye

A thesis submitted in partial fulfilment of the requirements for the degree of

European Master in Quality in Analytical Laboratories

University of Barcelona

2010

Certificate

This is to certify that the project entitled "*Application of Surface Plasmon Resonance and spectroscopic techniques for characterizing the interaction between human c-kit G-quadruplex DNA and water soluble porphyrins*" submitted by Sintyehu Manaye towards the partial fulfilment of European Master in Quality in Analytical Laboratories is based on his own effort under my supervision. To the best of my knowledge it hasn't been submitted to any other program.

Dr. Raimundo Gargallo Gomez	Dr. Ramon Compano
Project Supervisor	Head of Department

Table of contents

Dedication

To my mother - the wisest oak tree that ever stood;

she taught me to stand firm to the end, no matter

how many axes swung my way.

Acknowledgement

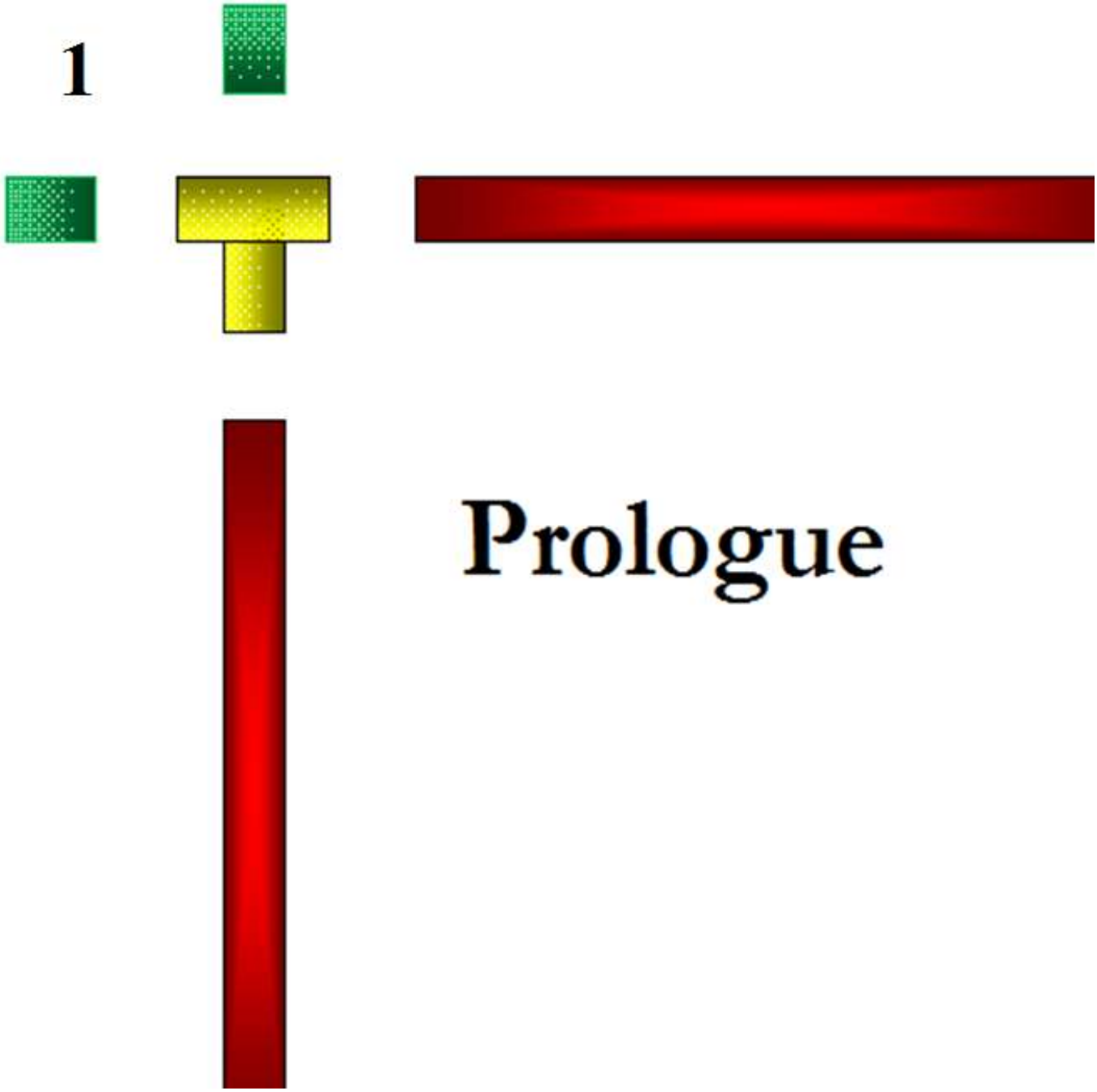
I would like to express my deep appreciation to

- European Commission for Education and Training for granting me the scholarship; an opportunity that I needed most
- My advisor, Dr. Raimundo Gargallo Gomez who not only took me seriously but also provided me with everything I need for my project and after. He has taken me one step further in my academic career via his encouragement, discussion, recommendation and very constructive feedback.
- My teachers, friends and relatives who have always been there for me.

Abstract

G-quadruplex DNA structures are hypothesized to be involved in regulation of telomere homeostasis and gene expression. The development of small molecules that could modulate the stability of G-quadruplex structures are of therapeutic interest in both cancer and aging. In this study, the interaction between human c-kit proto-oncogenic sequence 5'-d [CGG GCG GGC GCG AGG GAG GGG]-3', which can form an intramolecular G-quadruplex structure, and water soluble porphyrins 5,10,15,20-tetrakis-(N-methyl-4-pyridyl)-21,23H-porphyrin tetratosylate (TMPyP4) and N-methyl mesoporphyrin IX (NMM) was studied by UV-Vis and Circular Dichroism spectroscopy. Another sequence, 5'-d [CGG GTG GGT GTG TGG GTG GGG]-3', in which cytosine and adenine residues of c-kit gene sequence were systematically replaced with thymine was used to study the influence of loops on folding and binding interactions. Molecular absorption and CD data recorded along thermal denaturation, acid base titration and mole ratio experiments were analyzed by means of multivariate data analysis methods. Surface Plasmon Resonance (SPR) was used to probe the kinetics and affinity of molecular interaction. The results showed the formation of an interaction complex with a stoichiometry 1:1 (drug/DNA) and logarithm of the equilibrium constant equal to 6.7 ± 0.7 and 7.3± 0.5 for NMM and TMPyP4 respectively. It is proposed TMPyP4 binds to the longest loop in c-kitG1 blunt-end conformer where as NMM stacks at either end of the G-quartet. A small population of c-kitG1 exists as a dangling-end conformer which can form an interaction complex with a stoichiometry 2:1. Kinetic studies showed that the two conformers undergo very slow interconversion.

Keywords: G-quadruplex, NMM, TMPyP4, DNA-Drug interactions, conformational change, multivariate data analysis



1 Prologue

1. Systems Science and Multivariate Data Analysis

1.1. Systems Science and the 'arch of knowledge'

Chemical and physical quantities can be broadly classified into constants and variables. A constant is any chemical or physical quantity that is considered to be universal and identical in time. On the other hand, a variable can assume different values in time. The different values of a variable are called levels. A method of analyzing data on a single variable at a time is known as univariate analysis. In this method, a particular variable is selected and analyzed in isolation. Univariate data analysis is helpful to compare different classes in terms of their central tendency and variability with respect of a measured quantity(Goldstein and Shelly 1973).

A method of analysis concerned with the relationship between a pair of variables in a data set is called bivariate analysis. It is a method of considering two variables at a time, assuming all other variables are held constant. Bivariate data analysis is helpful to examine dependence relationship or correlation between the variables under investigation. Techniques of bivariate analysis include statistical hypothesis testing, analysis of variance, regression and correlation analysis. Univariate and bivariate data analyses can be considered as techniques of reductionist approach. Reductionist approach is based on an assumption that the whole is the sum of its parts.

Multivariate data analysis deals with analyzing two or more variables simultaneously in a data set. Bivariate analysis is the simplest form of multivariate analysis. Multivariate data analysis has two additional values. First, any chemical or physical setting in a natural environment is subjected to several variables. Measuring these variables is more reasonable than controlling them. Second, interactions among different variables which are important can only be observed in multivariate analysis. Some of the techniques of multivariate analysis include principal components analysis, partial least square discrimination analysis and extended canonical variate analysis. Multivariate data analysis is a holistic approach which is based on an assumption that the sum is more than the sum of its parts(Goldstein and Shelly 1973).

Systems science is an interdisciplinary field of science that deals with complex natural systems. It originated in Biology to describe the nature of ecosystems in ecological hierarchy pioneered by Carl Ludwig Bertalanffy in 1950s. Central to the systems science are the concept of system and emergent property. A system is a set of interacting entities making up an integrated whole. It is a configuration of parts connected and joined together by a web of relationships. An emergent property is the property of a complex system which arises from multiplicity of simple interactions. Systems science is dealing with all

major components of a system at the most possible fundamental level. A system has three aspects: structure which is the component entities (composition) of the system, function which is the process taking place in the system (input-processing-output) and relationship which refers to functional and structural interaction between components of a system.

1.2. Multivariate Data Analysis

Advances in science and technology made it possible to read and measure practically everything. Lack of data is hardly the problem these days. A set of techniques dedicated to the analysis of large data sets with more than one variable comprise multivariate data analysis. Several of these techniques were developed recently in part because they require the computational capabilities of modern computers. Different multivariate techniques have different applications which depend on the nature of the data set and the nature of the problem. The data table to be analyzed is constructed from several measurements collected on two sets of variables. In general, one variable constitutes rows and the other variable columns.

1.2.1. Principal Component Analysis (PCA)

This is the basis of all multivariate data analysis. PCA is a multivariate projection method designed to extract and display the systematic variation in a data matrix. It represents a data table as low dimensional planes such that the overview of the data can be obtained. Statistically PCA finds lines, planes or hyperplanes in K-dimensional space that approximate the data set as well as possible in the least square sense. The goal of PCA is to decompose a data table with correlated measurements into a new set of uncorrelated (i.e., orthogonal) variables. These variables are called principal components (eigenvectors). Each variable is also assigned a set of scores which correspond to its projection on the components.

The results of the PCA are often presented with graphs plotting the projections of the units onto the components, and the loadings of the variables. The importance of each component is expressed by the variance (i.e., eigenvalue) of its projections or by the proportion of the variance explained. In this context, PCA is interpreted as an orthogonal decomposition of the variance matrix of a data table. It is best suited to study one data structure. Prior to PCA, data are often pretreated in order to transform a data into a form suitable for analysis. In fact, pre-processing can make the difference between a useful model and no model at all.

1.2.2. Spectroscopic Techniques

Spectroscopy is the study of the interaction of electromagnetic radiation with matter. Spectroscopic techniques are used to determine the concentration or amount of a chemical species in a sample based on its response signal upon interaction with electromagnetic radiation. With the arrival of the current fast photo diode-array instrumentation that allows the recording of whole spectra, spectroscopic techniques have proven to be indispensable tools in analytical laboratories. There are several advantages for using multi-wavelength data. In contrast to univariate methods (selecting single wavelengths) multiple wavelength allows for the analysis and quantification of analytes in complex matrices without requiring previous separation. Spectroscopic methods are also suitable for studying chemical equilibria in solution. If the spectral responses of the components involved do not overlap, such analysis is, in general, trivial.

In most cases chemists wish to explain measurements by using a chemical model that comprises some values of parameters such as equilibrium constant to characterize a chemical system. Along with advanced instrumentation, computers acquired ever-increasing importance in analytical laboratories. Consequently, several computational approaches were developed for the evaluation of equilibrium data using multiple wavelengths. A different approach, soft modeling, has emerged as a consequence of the explosive growth of chemometrics in recent years. Nowadays, mathematical procedures used for determination of equilibrium constants from experimental data can be classified into hard-modeling and soft-modeling methods(Amigo, de Juan et al. 2006).

1.2.3. Hard Modeling

A model is a function that describes the matrix of the concentrations of all components in a system with respect to a process. The most classical approaches used absorbance recorded at one wavelength as input information to unravel kinetic mechanisms and to determine rate constants. This was replaced by the modeling of reaction mechanisms to multivariate data at multiple wavelengths. Hard modeling computational approaches are based on initial proposal of a model for chemical equilibrium. The first step in hard modeling is model definition. It refers to postulation of a set of species defined by their stoichiometric coefficients and equilibrium constants. Model definition relies on the knowledge and experience of the researcher. Hard modeling requires the fulfillment of the mass-balance equations and the mass-action law. Experimental data are then refined to the proposed model using a least squares curve fitting minimization(Daz-Cruz, Daz-Cruz et al. 2000).

In a general manner, the concentration profiles of the species involved in binding equilibrium of a DNA its ligand can be described as follows:

$$a(DNA) + b(drug) \leftrightharpoons DNA_a - drug_b \text{ complex} \quad \text{----- } K_{eq}$$

In a chemical model, the concentration profiles of all involved species are expressed as a function of two kinds of parameters, the total concentration of the components in equilibrium and the suitable equilibrium constants, K_{eq} (also denoted by the logarithmic pKi). Those are the parameters that are refined during nonlinear fitting. Hard-model constraint selects concentration profiles in matrix involved in the known reaction process as input for a nonlinear hard-modeling fit. Absorption data are governed by Beer–Lambert's law and the measurements are well described by a matrix equation:

$$\boldsymbol{D = CS^T + E}$$

D is a matrix, the rows of which are formed by the absorption spectra measured as a function of progress of the process. The columns of **D** are the absorption traces measured at different wavelengths. According to Beer–Lambert's law, this matrix can be decomposed into a product of a matrix **C** containing, column wise, the concentration profiles of the absorbing species and another matrix **S** containing, row wise, their molar absorptivities. The difference between the measurement **D** and its calculated representation CS forms a new matrix **E**; a collection of residuals. Nonlinear least squares hard-modeling algorithm consists of finding set of parameters for which the sum of all the squares over all the elements of the error matrix, **E**, is minimal. This crucial sum is a function of the measurement, **D**, the pre-defined model and the parameters. Examples of these hard-modeling methods include SUPERQUAD, LETAGROP-SPEFO, SPECFIT(Gampp, Maeder et al. 1985), SQUAD, EQUISPEC(Dyson, Kaderli et al. 1997) and HYPERQUAD(Gargallo, Tauler et al. 1996).

EQUISPEC computational program has four main components. These are defining a chemical model of an equilibrium, speciation calculation, refinement of the equilibrium constant and model selection. The chemical model defines the stoichiometric coefficients of all species present in solution. It is consisted of the set of reactants and complex species formed from them. Speciation calculation uses known analytical concentrations as input to calculate the distribution of all the proposed species in the predefined model. At the beginning, the logarithms of the free concentrations are refined with the added advantage of automatically imposing a non-negativity constraint on the free concentrations. Then equilibrium constant refinement process proceeds to find equilibrium constant values that give the best fit to the experimental data. EQUISPEC computes cumulative equilibrium constants for the overall equation. Once a refinement is completed the results will be checked to verify if the chosen model is acceptable.

Generally the combination of parameter errors, quality of fit, physical meaning of constants and distribution of residuals are considered in model selection. Parameter errors refer to relative uncertainties

of the equilibrium constant. Models where the percentage uncertainties in the best estimate of equilibrium constant are comparably high are less acceptable. Distribution of residuals is generally assumed to recapitulate the distribution of experimental errors. However, residuals are correlated whereas experimental errors aren´t. Correlation independent of the magnitudes of the residuals show some randomness. Physical constraints for example, non-negativity of concentration and molar absorptivity are also used in model selection.

The main advantage of hard-modeling is the significantly higher robustness compared to soft-modeling analyses. The number of variable parameters reflects the robustness in a simple and representative way. This approach has an excellent performance if the proposed model is appropriate and if all the variation related to the spectrometric response is linked to the components involved in the process. Results start worsening when some uncontrolled variation in the reaction occurs during the process or when inert absorbing interferences are present(Gemperline and Cash 2003).

The postulation of a theoretical physico-chemical model is very difficult because multiple processes can be involved simultaneously. The most difficult aspect of hard-modeling is determination of the correct model. The process of proposing several models and comparing results can be tedious. The choice of correct model is not so straightforward, and often not much guidance is available; consequently, relatively exhaustive trials have to be conducted in order to be able to choose appropriate model. In such cases, lack of a hard model makes any other type of approach of highest interest(Esteban, Arino et al. 2000). The process of determining the correct model, e.g. the correct mechanism, can be supported by prior soft modeling analyses of the data.

1.2.4. Soft Modeling

Soft-modeling approaches, like Evolving Factor Analysis (EFA), orthogonal projection approach (OPA), simple-to-use interactive self-modeling mixture analysis (SIMPLISMA), and multivariate curve resolution (MCR) are meant to describe processes without using explicitly the underlying chemical model linked to them. Soft-modeling techniques work without any assumption of a chemical model and do not have the requirement of the compliance of the mass-action law.

Within this group, Multivariate Curve Resolution by Alternating Least Squares optimization (MCR-ALS) method has been successfully applied to spectrometric titrations. This soft-modeling procedure has three main steps. These are detection of number of components, initial estimations of parameters and iterative optimization of estimated parameters by ALS. The first basic goal of MCR-ALS is to decompose mathematically the experimental data matrix into a product of two abstract orthogonal matrices for a pre-

selected number of components. The number of chemical contributions of the data matrix is chosen to minimize residual data variance, leaving in it, if possible, only the experimental error. Many methods have been proposed for selection of the number of components. By far the most common is singular value decomposition (SVD). First, a semi-log plot of the singular values as a function of number of components is used for visual inspection. Visual inspection of this plot helps in specifying the number of components to include in the MCR analysis. Plus, the magnitude of residuals expressed as a percentage of lack-of-fit in the original data matrix, after a particular number of principal components, is calculated(Esteban, Arino et al. 2000).

Once the number of components is estimated by SVD, data structure can be analyzed by using evolving EFA. In MCR–ALS method, EFA can provide an initial estimation of how concentration profiles of these components change along the experiment. EFA method is based on evaluation of magnitude of singular values of the submatrices built up by adding successively, one by one, all the rows of the original data matrix in forward and backward direction.

From initial estimations of concentration profiles or individual spectra, a constrained alternating least squares (ALS) optimization is started to recover correct set of concentration profiles and pure individual spectra. This recovery is based on an assumption that instrumental responses of chemical contributions are bilinear and can be expressed within the matrix equation. Least square solutions obtained in this way are pure mathematical solutions without any physical meaning. Therefore some physical constrains are integrated into ALS optimization procedure at each stage of the iterative optimization. The constraints usually applied in spectroscopic studies are: a) non-negativity of concentrations, b) non-negativity of the signals, c) unimodality of the concentrations, d) unimodality of the signals, and e) closure (total analytical concentration remains constant)(Jaumot, Escaja et al. 2002).

The main drawback of soft modelling techniques is the fact that they don't provide estimates of process parameters. In addition, there is ambiguity associated with the raw data decomposition. This could be partially or completely suppressed by means of the addition of several soft-modeling constraints or by the simultaneous analysis of sets of experiments. Soft-modelling techniques model all the absorbing contributions in the original measurement, i.e., those contributing to the process of interest and those which do not.

Nevertheless, Soft-modelling approaches can give preliminary insight into the chemical processes in the data set and guide the choice of chemical models that can subsequently be fitted to the hard models. MCR-ALS is a flexible method that takes advantage of known chemical and mathematical information about the

data set through the use of constraints and allows for the simultaneous analysis of several data matrices without strong requirements related to the mathematical data structure(Gemperline and Cash 2003).

- **General Objectives:** The general assumptions of this study were

    - G-quadruplex DNA structures might be involved in regulation of gene expression
    - G-quadruplex DNA structures can be biological targets for drug designing

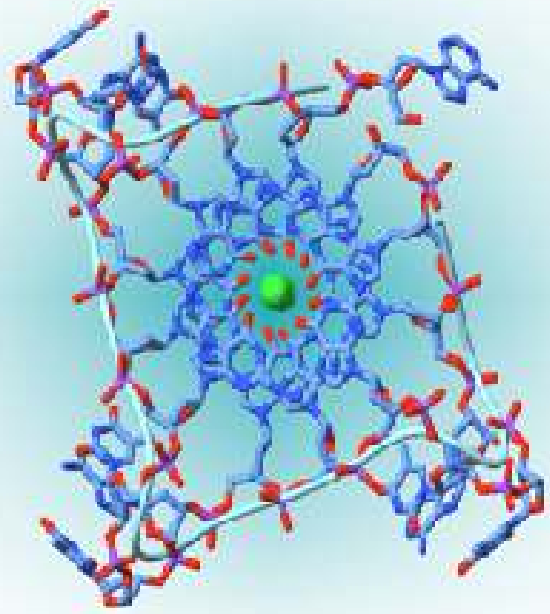
- **Specific objectives:** The specific objectives of this study were

    - to measure thermodynamic parameters of DNA folding
    - to determine binding stoichiometry of DNA-drug interactions
    - to measure binding affinities of drugs to DNA
    - to propose binding mechanism consistent with available experimental data (NMR structure, CD, UV-Vis spectroscopy and SPR)
    - to measure kinetic and thermodynamic parameters for DNA drug interactions

In this study, MCR-ALS methodology was used to study the equilibrium between the different conformations of G-quadruplex forming DNA sequences. The application of this method allowed the identification and resolution different conformations. The pure CD and UV spectra of each species were unambiguously resolved and thermodynamic data describing different equilibria were obtained.

2

G-Quadruplex DNA



2. Introduction to G-Quadruplex DNA

2.1. Cancer and Aging

Cancer is a major life-threatening disease and the second greatest cause of mortality in European Union after cardiovascular diseases. It is also a major public health problem in the United States and many other parts of the world. As it is predominantly a disease of old age this will continue to increase as the population ages. Despite significant progress in the fields of early diagnosis and standard treatments including radiotherapy and chemotherapy, the heterogeneity of mutations and development of drug resistance are some of the clinical challenges that remain to be overcome(Black, Bray et al. 1997).

The cellular basis of cancer is uncontrolled growth of immortalized cells. It is due to failure of the systems that control cell growth and proliferation. In healthy individuals, there is a happy balance between cell birth and cell death. When cell cycle regulation is lost by genetic damage, this happy balance is offset. Genetic aberrations that lead to cancer are often classified into two types. These are proto-oncogenic mutations and tumor suppressor gene mutations(Gazdar 1994).

Proto-oncogenes are genes analogous to time bombs. If they aren´t activated by genetic damage, they will not have any pathogenesis. However, when they are activated by some genetic changes, proto-oncogenes promote active proliferation due to increased hyper functional gene products. Proto-oncogenes are said to undergo change of function mutations. Tumor suppressor genes are active genes that encode proteins involved in cell cycle regulation. Tumor suppressor gene products include DNA repair proteins and cell cycle check point proteins. These genes are analogous to body guards. They are indispensable genes for the healthy balance of cell birth and cell death mentioned earlier. When tumor suppressor genes undergo loss of function mutation, cell cycle runs out of control and cancer prevails(Duquette, Huber et al. 2007).

In Molecular Biology, it is known that the RNA primer at the 5' end of the lagging strand can't be replaced by DNA. The extreme ends of linear DNA aren't replicated by the common replication machinery. As a result, each round of DNA replication would cause shortening of one of the two daughter DNA molecules producing loss of genes at the ends of chromosomes. This is called End Replication Problem(Levy, Allsopp et al. 1992). Cells solve end replication problems in a variety of ways. Eukaryotic cells solve this problem via the use of telomeres; the natural ends of linear chromosomes.

Telomeres are combinations of telomeric DNA and telomeric proteins. Telomeric DNA is consisted of several thousands of tandemly repeated species specific G rich 3' ending sequences. They have multiple functions in genome homeostasis. Cumulative telomeric erosion is a limiting factor in the happy balance of

cell birth and cell death. It elicits a signal for the onset of cellular senescence, a stage in which cells cease to divide. Extensive telomeric shortening leads to end to end chromosomal binding causing chromosomal instability, cellular senescence and cell death(Blackburn 1997).

These unusually G rich sequences localized at the ends of chromosomes are synthesized by a novel mechanism. The enzyme that catalyzes the synthesis of telomeric DNA is called telomerase(Yu, Bradley et al. 1990). Healthy cells have very low telomerase activity. They can divide a limited number of times. On the contrary, germ cells and nearly all cancer cells have active telomerase which gives them the ability of unlimited proliferation referred as immortality. Aging correlates with proliferative capacity and telomere length. In higher animals loss of telomerase function in somatic cells is believed to be the molecular basis of aging. Several lines of evidences suggest that cellular senescence evolved as a failsafe mechanism to prevent cancer.

A substantial portion of telomeres is maintained in single stranded form. The G rich overhang strands could potentially fold into a stable structure, known as G quartet, in which guanine bases forms strong hoogsteen base pairs that further associate to form cyclic tetramers. G quartet arrangement is consisted of four guanine bases in a square co-planar array linked by cyclic hydrogen bonding pattern. Utilization of both the N1 and N2 of one face with the O6 and N7 of the second face on guanosine yields eight hydrogen bonds per planar G-quartet as shown in figure 2.1. Hoogsteen faces are critical in the formation and stabilization of tetrads(Mergny and Maurizot 2001).

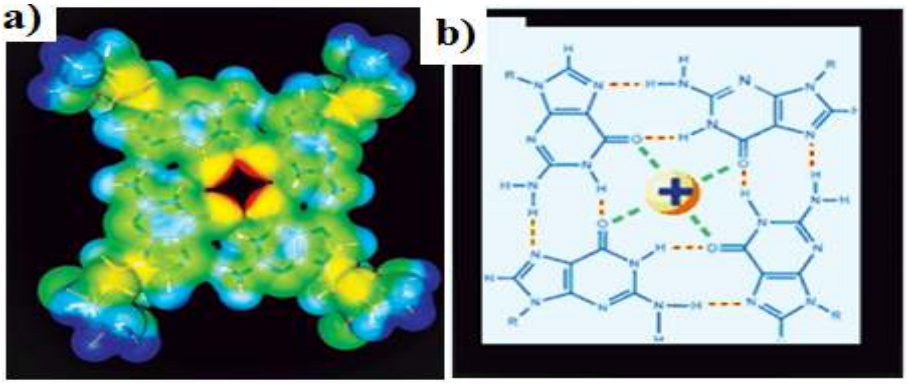


Figure 2-1 a) Space filling model of guanine quartet b) A pocket center in G-quartet which is lined by electronegative carbonyl oxygen is coordinated to a metal ion.

2.2. G-quadruplex DNA

Structural flexibility is a key for nucleic acids to function. That explains why nucleic acids are structurally diverse molecules. They can assume several distinct structures that vary with medium properties and

presence of alkali metal ions. There are three well established major conformational states of double helical DNA motifs. These are DNA forms A, B and Z. These forms have very different structural features. B DNA, the biologically predominant form of DNA is formed in the presence of alkali metals such as Na^+ when the relative humidity is as high as 90%. When the relative humidity is reduced to 75%, B DNA undergoes a reversible conformational change to the A form. Z-form is one of the more unusual polymorph of duplex DNA. Furthermore, two polynucleotide strands can associate into a three stranded arrangement termed a triplex by the inclusion of a third strand into the major groove of duplex DNA.

The association of four polynucleotide strands termed a quadruplex (figure 2.2) only adds to the structural diversity seen in nucleic acid structure. Guanine polynucleotides have been known to form G-quartets in solution especially in the presence of K^+. Interest in the structural arrangements of G-quadruplexes was ignited by the fact that human telomeric sequences could form such structures, at least *in vitro*.

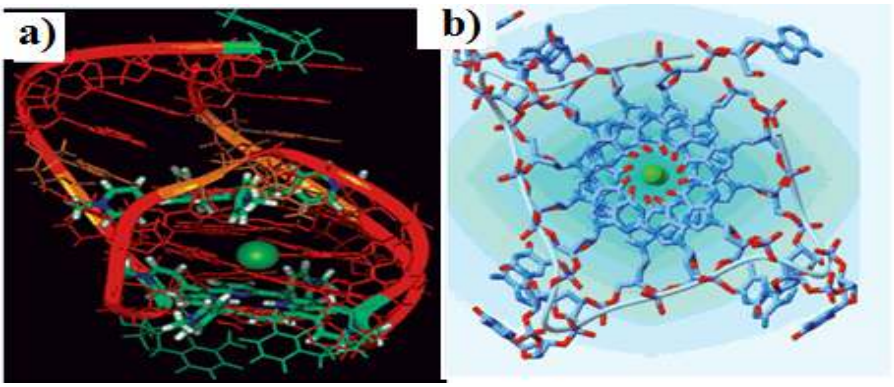


Figure 2-2 Different G-quadruplex structures view from different angles (a) side- view and (b) top-view. These PDB figures were adopted from [1] and [2] respectively.

G-quadruplexes are very interesting molecular objects. G-quadruplex structures have attracted a large number of research attentions. This is due to potential application in drug designing for cancer chemotherapy, antiviral activity and unique structure for nanotechnology(Michaela, ccaron et al. 2006).

2.2.1. G-quadruplex Structures and Topology

All G-quadruplex structures comprise a stack of two or more G-quartets, [tetrads], linked by the phosphodiester backbone and stabilized by specific monovalent ion. The stacked quartets constitute the G-tetrad core; the sugar phosphate backbones constitute columns and linkers connecting the G-stretches constitute loops. G-tetrad core, columns and loops are structural hall marks of G-quadruplex DNA. G-quadruplex structures can be unimolecular, bimolecular or tetramolecular depending on the number of polynucleotide strands associated. Unimolecular (foldback) structures have been studied in great detail for

practical reasons. They are the most likely forms to exist *in vivo*. Their physical chemistry is much easier to analyze (structures are independent of concentration, faster and concentration independent kinetics, reversible concentration-independent thermodynamics). The detailed conformation is easier to determine for unimolecular G-quadruplexes. In foldback structures, guanine repeats in the linear sequence form the G-quartets (represented as blocks in figure 2-3) whereas other nucleotides constitute the loops (represented by curved arrows). For molecules that comprise three loops with contiguous G-quartet strands six different types of folds have been experimentally determined(Lane, Chaires et al. 2008).

G-quadruplex columns have different orientations giving rise to parallel, antiparallel and mixed structures. There are different loop arrangements within these structures. Arrangements include lateral loops, double chain reversal loops and diagonal loops shown in the figure 2-3 below.

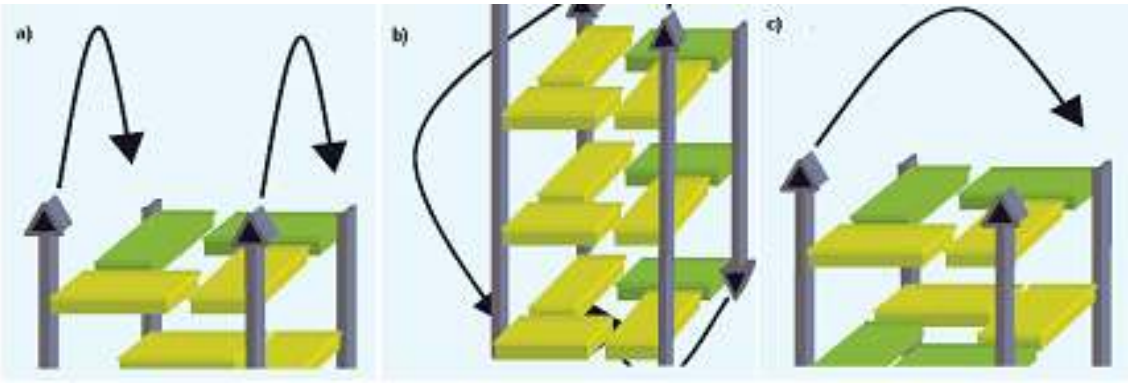


Figure 2-3 Schematic diagrams showing potential linking loop arrangements for unimolecular quadruplexes. (a) Lateral loops (b) Double chain reversal loops (c) Diagonal loops.

The number (length) and type (sequence) of the linking nucleotides impose certain topological constraints on the folded quadruplexes. Loop length exerts unfavourable strain on the G-quartets, which is itself structure dependent. However, even a single connecting nucleotide, forming an external connecting loop, can accommodate a stack of up to three guanine tetrads (Lane, Chaires et al. 2008).

2.2.2. Thermodynamics and Kinetics

Understanding the rules that govern the formation and stability of G-quadruplex DNA are still under investigation. The formation of quadruplex structures, whatever their type, is clearly enthalpy driven, with an enthalpy per quartet of $^{-}15$ to $^{-}25$ kcal mol^{-1}. This favorable enthalpy is partially compensated by unfavorable entropy of formation. The range of values for entropy changes depends on the stoichiometry of strand association. Besides, some loop structures are compact with ordered structure showing features such as base stacking, intraloop hydrogen bonding and extensive van der Waals' interactions affecting entropy changes. The closest thymine in the loop can stack over the outermost G-quartet associated by

hydrogen bonding. These observations support the important role played by the nature and length of loops in quadruplex stability(Lane, Chaires et al. 2008).

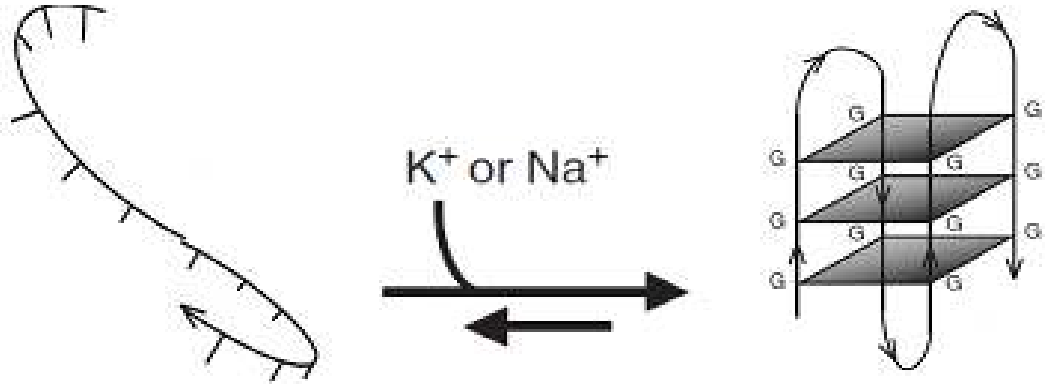


Figure 2-4 Solution equilibrium involving unimolecular G-quadruplex DNA molecule in the presence of monovalent ions.

The major reason for determining the kinetics of formation and dissociation of quadruplexes is that it is the only unambiguous means to obtain information about the mechanism of a complex reaction pathway. Under equilibrium conditions, it is not possible to detect intermediates unless they become sufficiently populated. Association and dissociation rate constants which characterize G-quadruplex formation kinetics are determined from the rate of forward and reverse reactions in G-quadruplex folding equilibrium (shown in figure 2-4 above). Kinetic data analysis is complicated by the existence of several quadruplex conformations in equilibrium. It was shown that the dissociation (unfolding) rate constant of G-quadruplexes was extremely small (in the orders of $10^{-5}s^{-1}$range), and that the association (folding) was characterized by negative apparent activation energy. The formation of 'simple' monomolecular quadruplexes as driven by either K^+ or Na^+ is kinetically complex. A further level of complexity is added by the very different paths of formation. As a result, the thermodynamically most stable structure is not necessarily the kinetically favored(Lane, Chaires et al. 2008).

2.2.3. G-quadruplex Stabilizing Factors

The most destabilizing component in nucleic acids at neutral pH is the very unfavorable electrostatic interactions between the oxygen atoms in the phosphodiester bonds, which bear a formal charge of –1. This is exacerbated by the arrangement of the guanine O6 carbonyl groups central to the G-quartet. This unfavorable energy needs to be offset by all other favorable interactions. Stabilizing interactions include hydrogen bonding, π-π stacking, hydration, specific ion binding and counter ion effect. Enthalpy change from intramolecular H-bond formation is relatively small. As a result, the contribution of H-bonding to the free energy of stabilization is thought to be small. Base stacking is stabilized through the combination of

hydrophobic, electrostatic and van der Waals interactions. The π-π stacking interactions between quartets are likely to account for a substantial part of the net stabilization free energy.

G-quadruplex formation absolutely requires the participation of cations. The negatively charged cavities located between the G-tetrads need to be stabilized by the coordination of cations [Figure 2-5]. Unlike other known DNA structures, G-quartets interact directly with dehydrated cations via inner sphere coordination. It isn´t surprising the cavity on the G-quadruplex binds cations selectively. G-quadruplexes are remarkably specific with respect to the cation type because Li^+ does not stabilize G-quadruplex Na^+, K^+, Sr^{2+}do so very efficiently. The selection of a suitable cation, based on size and charge, dramatically determines the overall stability of the final folded quadruplex. Multiple cation coordination geometries are possible and observed in quadruplexes. As the specific ion binding to the unfolded state is likely to be small, the stabilization due to specific ion binding of approximately one monovalent ion per quartet should be substantial.

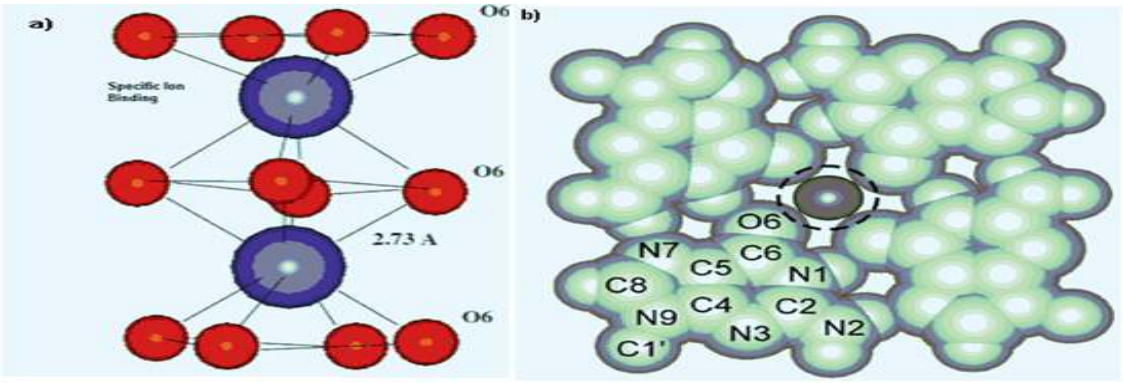


Figure 2-5 Specific-ion binding between G-tetrad a) lateral views shown with twist between bases b) top view of a Na+ coordinated in the plane of G-quartet the dashed circle represents ionic radius of K+ which is too large to coordinate in-plane. Adapted from Parkinson, G. N., (2006). "Fundamentals of G-quadruplex Structures". In Neidle, S. and Balasubramanian, S. (Ed.), Quadruplex Nucleic Acids (pp.9). Cambridge: The Royal Society of Chemistry Press.

DNA is a polyanion and cations should screen repulsion of DNA phosphates for the DNA strands to come close to each other in the quadruplex. Non-specific ion binding includes both counter ion condensation and Debye–Huckel counterion atmosphere (shielding). Such ions interact as outer sphere complexes (i.e. the species retain their hydration coordination). Ion condensation depends on charge density whereas Debye–Huckel counterion screening depends on the net charge. Unimolecular G-quadruplex structures are known to condense fewer ions than the unfolded state. Consequently, the net charge on G-quadruplex is higher than that of the strand state. That is the reason why increasing ionic strength favors G-quadruplex formation. Hydration is a process by which water combines with a molecule in a definite ratio as integral part of that molecule. The formation of small unimolecular G-quadruplex structure is accompanied with the release of ca. 13 waters per mole of quartet folding. Thus dehydration

would stabilize quadruplexes compared with duplexes, which are more hydrated than the single strand. Decreasing water activity has favorable electrostatics for G-quadruplex folding.

2.2.4. Do quadruplex structures actually exist *in vivo*?

Thermodynamically, telomeres are the most likely place to find G-quadruplexes *in vivo*. The best evidence came from the indirect immunostaining technique on isolated nuclei from ciliates. It showed the existence of antiparallel G-quadruplex only in the vegetative state. About 21 proteins including enzymes have been identified that bind, unwind, and cleave G quadruplex DNA i*n vitro* with great affinity and specificity. Eukaryotic cells contain many G-rich regions capable of forming quadruplex DNA at a surprisingly high frequency. Among these are repetitive and functionally essential chromosomal domains including the rDNA and the immunoglobulin heavy chain switch regions of higher vertebrates. Cells have well established homeostatic mechanisms for the regulation of water, glucose and salts in the blood. The activity of water in the cellular environment can be changed by the action of proteins and other biomolecules. If it can be shown that 10–20 water molecules are released per mole quartet formed, dehydration would help to stabilize G-quadruplex formation in the cellular milieu(Lipps and Rhodes 2009).

On the contrary, of the 21 proteins discovered to bind G-quadruplexes, only one quadruplex was a fold-back structure. Quadruplex binding ligands typically have rather modest affinity and selectivity even for the small numbers of structures tested, and do not always bind in the manner expected or designed. The slow unfolding suggests that such structures would be long-living – days, weeks or even more. If G-quadruplex DNAs formed *in vivo* they compromise genetic stability. So far there is no clear answer for the formation of G-quadruplex DNA inside the cell(Lane, Chaires et al. 2008).

2.2.5. Intragene G-quadruplex Structures

There is now a growing interest in non-telomeric human sequences and other genomes that can form (or perhaps can be induced to form) quadruplexes. It is hypothesized that these quadruplexes may play a role in the regulation of gene expression. A DNA sequence which controls the transcription of an associated gene is called a promoter. Gene promoters tend to be modular in architecture and can be very complex if they are differentially regulated in different tissue types, during different stages of development. Guanine rich sequences are present in the promoter region of key oncogenes, and one can observe an induction of duplex to G-quadruplex transition in the presence of a quadruplex ligand and/or KCl. The regulatory potential of G-quadruplexes towards cancer cell growth is strongly substantiated by their possible formation in the promoter regions of several human oncogenes such as *c-myc*(Onyshchenko, Gaynutdinov

et al. 2009), *k-ras* (Cogoi and Xodo 2006), *bcl-2*(del Toro, Bucek et al. 2009), *c-kit*(Bucek, Jaumot et al. 2009), or *RET* (Guo, Pourpak et al. 2007)oncogenes.

2.3. *C-kit* proto-oncogene

Research in *c-kit* started in 1986 when Evelyn E. Zuckerman, William D. Hardy and their associates isolated a new acute transforming feline retrovirus, the Hardy-Zuckerman 4 feline Sarcoma Virus (HZ4-FeSV) from a cat. The genome of this newly isolated virus contained new oncogene designated v-kit and had a structure 5' Δ-gag-kit- Δ pol- Δ env 3'. The gag-kit (group specific antigen kit) was known to specify a polyprotein and hence the name kit. The researchers speculated the viral v-kit gene to be transduction product of cellular *c-kit* sequence. They also predicted the gag-kit to display partial homology with tyrosine specific protein kinases(Besmer, Murphy et al. 1986).

2.3.1. Proto-oncogene KIT

KIT (Accession Number 31413) is now known to be a human proto-oncogene coding for a 145-160 kDa membrane-bound glycoprotein. This glycoprotein belongs to a family of growth factor receptors with tyrosine kinase activity (*c-kit*, OMIM*164920). KIT gene is about 7.08 kbases mapped to loci 4q11-q12 on the template strand (+, W). The promoter sequence of KIT proto-oncogene is situated -102 to -197 upstream of the transcription initiation site. The *c-kit* gene is expressed by and is critical for the development of mast cells, melanocytes, and hematopoetic stem cells(Hemesath, Price et al. 1998). A new interest in the structure and function of KIT gene is ignited due to the discovery of conserved quadruplex forming motifs in this promoter (Hsu, Varnai et al. 2009). The sequence from -140 to -160 upstream of transcription initiation site contains a G-rich sequence called *c-kitG1*(Rankin, Reszka et al. 2005). Its sequence is 5'-d [CGG GCG GGC GCG AGG GAG GGG]-3'. *C-kitG4* is a mutated sequence in which only guanine bases of the wild type *c-kitG1* are maintained whereas all other adenine and cytosine residues are replaced by thymine. The NMR solution structure [Figure 2-6] of G-quadruplex formed by *c-kitG1* sequence was published very recently.

The formation of G-quadruplex structure *in vivo* by proto-oncogenic sequences is more controversial than their telomeric partners. This is because in gene promoters, duplex–quadruplex competition is of special interest [Figure 2-7]. Experimentally, at physiological pH and temperature the duplex absolutely wins over the quadruplex. Thermodynamically, in the absence of any stabilization of the quadruplex, intra-genic G-quadruplex formation does not overcome the very unfavourable strand separation energy. It is clear that such structure will not spontaneously form without some help. The energy cost might be offset by the formation of a stable structure, i-motif in the complementary strand or ligand binding. Another alternative

is that supercoiling stress generated by transcription might supply this energy (even transiently). A binding free energy of 20 kcal mol^{-1} is equivalent to a dissociation constant of 10^{-14} M(Lane, Chaires et al. 2008).

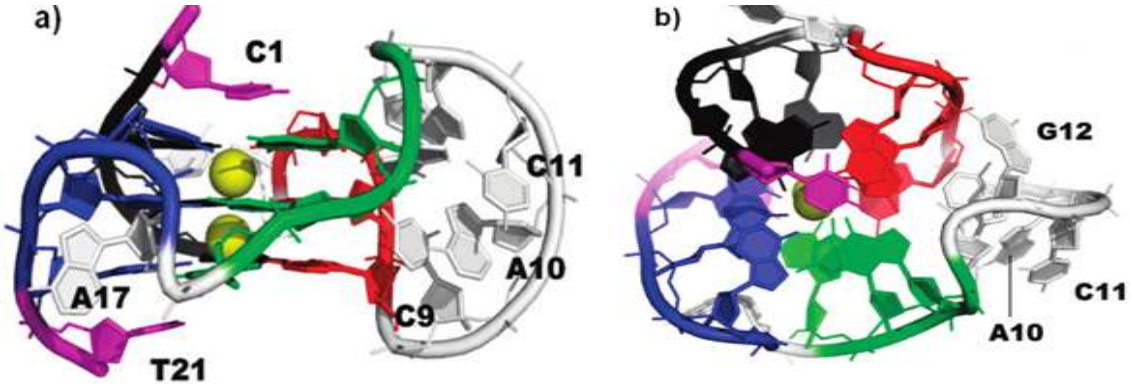


Figure 2-6 Cartoon representation of the structural model of *c-kitG1* NDB ID: 2KQG a) side view b) top view; the bases are labelled to highlight base stacking in L2, i.e., C9-C11, and the compaction of L3, i.e., A17.

Selective binding of a promoter oncogenic quadruplex by a small molecule may overcome the energy barrier and stabilize G-quadruplex structure. Stabilization of oncogenic G-rich sequences and the single-stranded ends of telomeric DNA into G-uadruplexes is potentially of therapeutic utility. A growing number of groups are targeting G-quadruplex DNA with small molecules hoping to inhibit cancer growth according to two distinct mechanisms. First, the over-expression of oncogenes like *c-myc, c-kit*, and *k-ras* might be inhibited by promoter deactivation. The second, more extensively studied mechanism is the inhibition of telomerase, a ribonucleoprotein complex that catalyzes the 3' extension of telomeric DNA. Inhibiting telomerase activity is the primary area anti-telomerase anticancer therapy. This can be supplemented by specific downregulation of the expression of a particular oncogene.

2.3.2. Gastrointestinal Stromal Tumors (GIST)

Gastrointestinal Stromal Tumors (GISTs) are relatively uncommon mesenchymal neoplasms that arise in the wall of the stomach, small intestine, colon, and other sites within the abdominal cavity(Hirota, Isozaki et al. 1998). Constitutive activation of the KIT receptor tyrosine kinase is a central pathogenetic event in most GISTs and generally results from oncogenic point mutations which can involve either extracellular or cytoplasmic domains of the receptor. Oncogenic mutations enable the KIT receptor to phosphorylate various substrate proteins, leading to activation of signal transduction cascades which regulate cell proliferation, apoptosis, chemotaxis, and adhesion. The drug imatinib has recently been licensed for clinical use in the treatment of GIST; its mode of action involves inhibition of *c-kit* kinase enzyme.

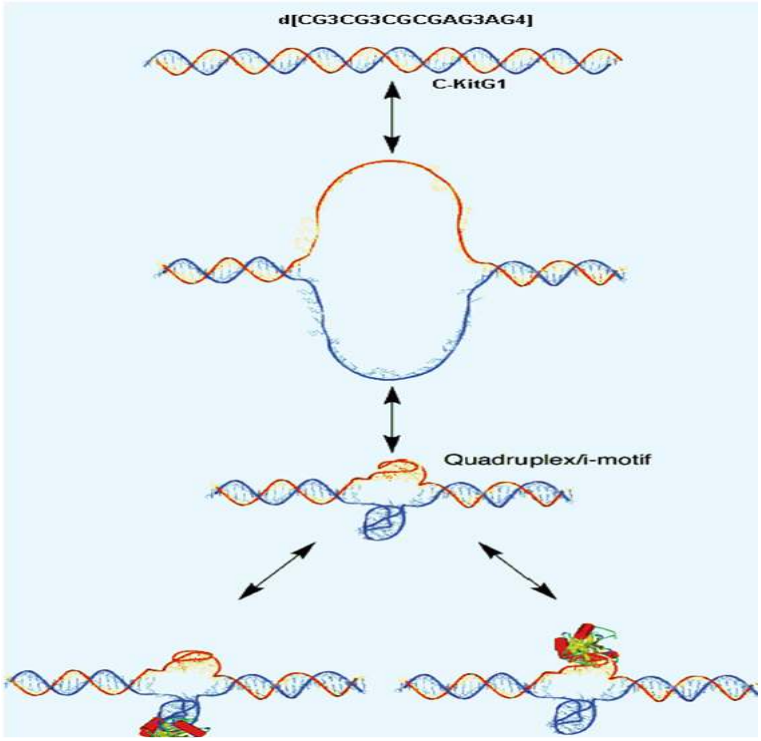


Figure 2-7 Steps in the formation of an internal G-quadruplex structure, clearly the formation of intragenic G-quadruplex faces fierce competition with the 'mighty' duplex DNA.

Newer kinase inhibitors have been developed, such as the compound sunitinib; however, the heterogeneity of kinase mutations and development of resistance are clinical challenges. A small molecule that inhibits *c-kit* expression at the transcriptional level would inspire the exploration of quadruplex-based therapeutic approaches to address GIST and many other cancer types(Hsu, Varnai et al. 2009).

2.4. Anticancer Drug Development

Novel therapeutic approaches based on understanding the molecular biology of cellular targets and chemical structures of target drugs are needed to reduce the high incidence of cancer. Molecular biology has a deep impact on drug discovery by simplifying the process of identifying a potential therapeutic target and its role in health and disease. Target validation is provision of objective evidence that the molecular target identified is critically involved in a disease process. Targets could potentially be enzymes, receptosr, ion channels, transporters, DNAs, RNA and ribosomal targets. When the biology of disease mechanism is well characterized as in the case of GIST, target validation is no longer a limiting step.

Analytical chemistry guides pharmacology by developing and validating *in vitro* assays. Assay development involves building chemical tests that measure the amount of a particular target molecule in a given matrix and characterizing its interactions. The binding affinity of a drug to its target, stoichiometry of binding and the presence of associated conformational changes are all determined by *in vitro* analytical assays. Moreover,

analytical techniques play indispensable role to investigate thermodynamics and chemical kinetics of the drug with the identified target.

Why G quadruplex?

Numerous pharmacological agents currently used in cancer chemotherapy bind to DNA non-specifically (e.g. cisplatin, mitomycin C(Lambert, Armstrong et al. 2009), chloroethyl nitrosoureas(Ru-Gang, Li-Jiao et al. 2007), daunomycin, etc.). B form DNA isn't the only target for drugs (Luedtke 2009). Other conformations may have a role in specific stretches of DNA. There are possibilities that certain groups of DNA sequences may *in vivo* adopt conformations other than B-DNA under the influence of environmental factors. Selection methods that yield nucleic acids structures that are capable of highly specific binding to small molecules are required to provide additional examples of recognition by structure-specific motifs.

The significant structural differences between G-quadruplex DNA and duplex DNA make quadruplex DNA a very attractive target for highly selective, structure-specific drug design. The potential formation of intragenic G-quadruplex structures which might be considered cancer-specific makes G-quadruplex DNA structures even more attractive targets. Nevertheless, so far direct proof for the formation and *in vivo* existence of G-quadruplex is still sparse. It is also important to note that G-quadruplexes are highly dynamic structures and losses in DNA conformational entropy upon ligand binding might be an important factor in determining specificity. Development of small molecules that specifically bind to a particular DNA secondary structure may improve cancer-specific targeting and decrease the side effects associated with chemotherapeutic treatments. Small molecule mediated stabilization of the G-quadruplex structure can effectively inhibit telomerase activity, and when applied to cells, G-quadruplex ligands can initiate apoptosis or replicative senescence. Porphyrins are known to bind to and stabilize different types of G-quadruplexes and, in some cases, to facilitate G-quadruplex formation(Luedtke 2009).

2.5. Water Soluble Porphyrins

Porphyrins are aromatic heterocyclic macrocycles characterized by the presence of four modified pyrrole subunits interconnected at their alpha carbon through methine bridges(Biesaga, Pyrzynska et al. 2000) . Porphyrins are known to play important roles in many fundamental biological processes, e.g. photosynthesis, metal coordination chemistry, biological redox reactions and oxygen transport processes. Since porphyrins possess large extended π electron system and have high stability, they are finding use in advanced materials and as components in organic metals, molecular wires and other devices for electronics.

2.5.1. Physical Properties

Many water soluble porphyrins dimerize readily in aqueous solution. The propensity of a porphyrin to dimerize in aqueous solution can be estimated by recording the optical spectrum of the solution as a function of the concentration of added salt. Analysis of the data in terms of the Debye–Hückel formalism gives an estimate of the extent of dimerization as a function of ionic strength. Aggregations can be modulated by changing the medium properties such as temperature, pH, ionic strength and salt concentration (Monsu Scolaro, Castriciano et al. 2002). Porphyrins bearing charged groups such as NMM were observed to aggregate due to screening of the repulsive interactions between side-groups by counterions (Dixon and Steullet 1998). Dimerization may affect binding of the cationic porphyrins to their targets, e.g., DNA.

2.5.2. Interaction with G-quadruplex DNA

The ease of synthesis of porphyrins and the ability to create modified porphyrins with different ring systems and substituents as well as their solubility makes porphyrins desirable to work with from a drug design standpoint. These compounds are excellent models for the design of new anti-cancer, anti-viral and anti-parasitic compounds with high affinity and selectivity. They can be artificially synthesized, can be extracted from natural sources and chemically modified very easily. Many water soluble porphyrins are commercially available. They can be monitored spectroscopically for they have intense absorption Soret band in the visible region. They have fluorescence excitation and emission spectra for monitoring by fluorimetry.

Porphyrins are known to produce photoinduced cytotoxicity. Upon excitation, singlet oxygen is generated, resulting in indiscriminant cell damage. This feature of porphyrins has been exploited advantageously through the use of photodynamic therapy for treatment of some cancers including lung, esophageal and bladder cancers. The phototoxicity of porphyrins would prevent them from being used as telomerase inhibitors or *c-myc* repressors.

Meso-5, 10, 15, 20-**T**etrakis-(N-**m**ethyl-4-**p***y*ridyl)**p**orphine (TMPyP4) is a cationic porphyrin. TMPyP4 is commercially available as sulfate salt of unsaturated hydrocarbon. While a large number of G-quadruplex ligands have been reported in the literature, the cationic porphyrin TMPyP4 is the most extensively studied to date (Gaynutdinov, Neumann et al. 2008). TMPyP4 has been shown experimentally to bind to human telomeric DNA with a 2:1 stoichiometry. Molecular modelling with human telomeric DNA suggests that one TMPyP4 molecule is stacked externally on each of the terminal quartets, with the positively charged groups oriented toward the sugar-phosphate DNA backbone. TMPyP4 also binds to and stabilizes the G-rich strand of the intramolecular *c-myc* sequence with a 2:1 stoichiometry, with each molecule stacked

externally on the G-quartet surface, in a similar manner as TMPyP4 binds to the human telomere. Subsequent X-ray crystallography and NMR studies have shown that TMPyP4 can bind to G-quadruplex DNA at many different positions, including the terminal G-tetrads, and the loops, grooves, and phosphodiester backbone.

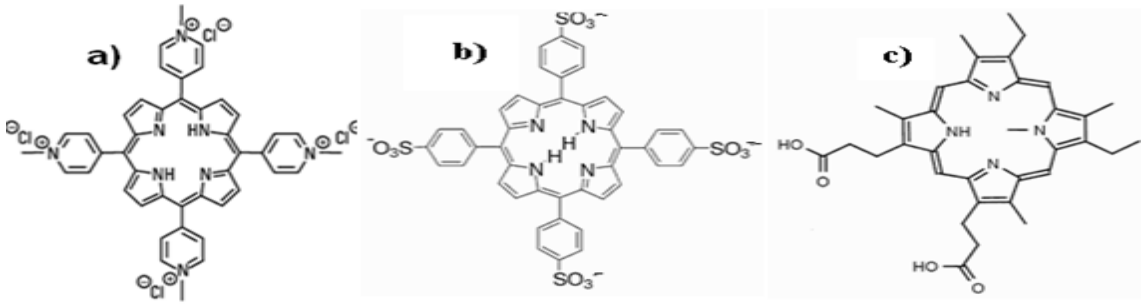


Figure 2-8 Structure of water soluble porphyrins a) cationic porphyrin TMPyP4 b) anionic porphyrin TPPS and c) anionic porphyrin NMM.

TMPyP4 has exhibited some promising anticancer activities *in vivo (*Mikami-Terao, Akiyama et al. 2009*)*. It has been shown to have a dramatic effect on gene expression, and it down-regulates transcription of both *c-myc* and the catalytic subunit of telomerase. When TMPyP4 binds to *c-myc*, it induces a conformational change in the DNA. This conformational change of the DNA renders other accessory protein unable to recognize G-quadruplex. TMPyP4 inhibits both telomerase (IC_{50} ≅ 0.7–10 μM) and Taq DNA polymerase (IC_{50} ≅ 2 μM). Telomerase inhibition by TMPyP4 is relatively insensitive to metal coordination by the porphyrin, but is highly dependent on the groups at the meso positions. Extensive substitution of the meso positions with groups other than pyridinium failed to generate compounds with improved activities, but the resulting structure–activity relationships (SARs) demonstrated that base stacking and charge-charge interactions are important for porphyrin-DNA binding.

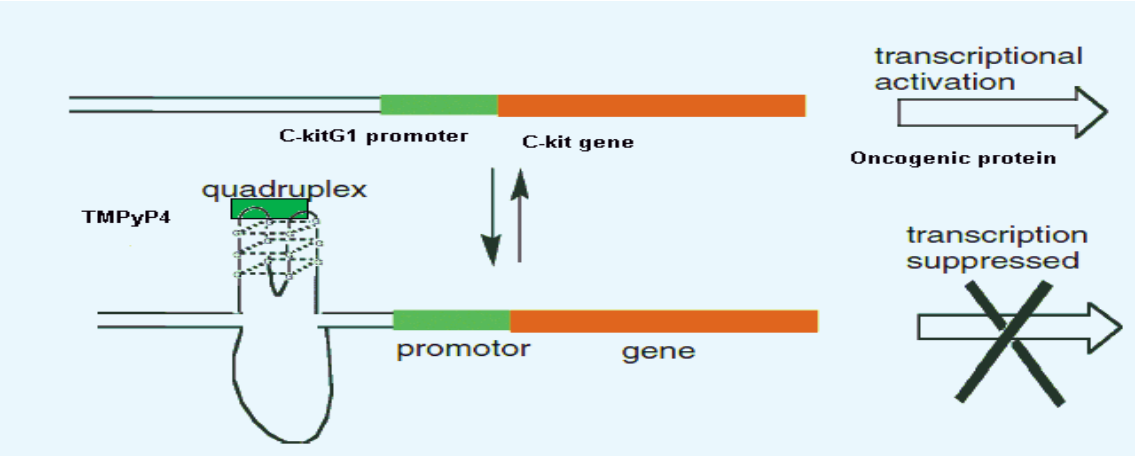


Figure 2-9 A possible model for G-quadruplex formation and small molecule stabilization in the promoter region of *c-kit* gene; subsequent inhibition of transcription is also shown.

Despite the effective quadruplex-binding and enzyme inhibition exhibited by TMPyP4, this compound has several drawbacks that would prevent it from being used therapeutically. TMPyP4 has very poor DNA specificity. It binds to single stranded DNA, duplex, triplex, G-quadruplex and bulk genomic DNA with similar affinities (Kd ≅ 200 nM), and therefore it cannot be considered a structure-selective ligand (Luedtke 2009). This is also important because indiscriminate duplex binding is associated with cytotoxicity. TMPyP4 also binds preferentially to intermolecular quadruplexes over the biologically relevant intramolecular quadruplexes. This intermolecular binding has been demonstrated through the TMPyP4-induced formation of anaphase bridges in sea urchins. Binding to intermolecular quadruplexes not only results in compound loss due to binding of non-targeted sequences, but is undesirable because it can result in end-to-end fusion of chromosomes. Also, the binding of TMPyP4 to the human telomere and *c-myc* sequence is not as strong as some of the newer quadruplex-interactive compounds which have been discovered. The porphyrin core of TMPyP4 is smaller (10.1 Å) than the size of the G-quartet (13.2 Å) which results in the molecule binding with an offset from the center of the quartet. TMPyP4 only overlaps with two out of four of the guanines in the terminal G-quartets. The binding of TMPyP4 to the quadruplex could potentially be improved if it were modified to increase the contact surface between the compound and the G-quartets, in order to maximize stacking interactions.

N-Methyl Meso porphyrin IX (NMM) is commercially available as a mixture of four regioisomers. Regioisomersism is type of isomerism produced as a result of the different outcomes of chemical reactions in which there are different orientations or sites to choose from. For meso-porphyrins, there are four unsymmetrical central N atoms, when methylation reactions between unsymmetrical central nitrogen atoms occur, four outcomes are possible. This anionic porphyrin has excellent G-quadruplex specificity, but relatively poor G-quadruplex affinity (Kd ≅ 2−10 μM)(Luedtke 2009). Changes in the fluorescence properties of NMM can be used to detect G-quadruplex DNA binding *in vitro*. NMM binds G-quadruplex DNA, and specifically inhibits its unwinding by RecQ-family helicases . NMM inhibited telomerase activity in the HEK293 cell extract. Recently, NMM was found to up-regulate genes having promoters with a high potential for G-quadruplex formation, and it also suppressed rDNA activity in yeast.

Despite the fact that NMM is known as very quadruplex-selective ligand, it is counterbalanced by a very low affinity. It is also an inhibitor of heme synthesis. In animal cells, N-methyl mesoporphyrin IX acts specifically to block iron insertion into protoporphyrin IX by ferrochetalase(Beale and Chen 1983). It is found to undergo dimerization even at very low concentration.

Meso-5, 10, 15, 20-tetrakis (4-sulfotophenyl) porphirine (TPPS) is a tetra negative anionic porphyrin. There are several reports that TPPS interacts with proteins in aqueous solution. Due to high electrostatic

repulsion with the negatively charged sugar phosphate backbone of nucleic acids, it is considered practically non-interactive to G-quadruplex DNA. TPPS can serve as a negative control to study structure dependent DNA drug interactions(Li and Tong 1994).

Porphyrins aren't the only compounds that recognize G-quadruplex structures. In fact the strongest and the most specific telomerase inhibitor ever reported of any G-quadruplex interactive small molecule is a natural product called telomestatin. Telomestatin is isolated from *Streptomyces anulatusn 3533-SV4*. A total synthesis of telomestatin has recently been reported, but little is known about the physicochemical and structural features important for its DNA binding and pharmacological activities. No SARs have been reported for telomestatin, due, in part, to the extreme difficulties associated with the synthesis of its highly constrained macrocyclic core. Synthetic analogs of telomestatin containing six or seven oxazole units have recently been prepared and show good G-quadruplex affinity and specificity. Telomestatin has molecular dimensions similar to those of G-tetrad DNA and can bind to various G-quadruplexes with modest affinity (Kd ≅ 30 nM). The ring system is similar in size to the G-quartet structure which increases π-π stacking, electrostatic and hydrogen-bonding interactions as compared to TMPyP4. Telomestatin exhibits good selectivity for intramolecular versus intermolecular G-quadruplex structures, and it has a 70-fold lower affinity for duplex DNA. Telomestatin is one of the strongest and most specific inhibitors of telomerase reported to date (IC_{50} ≅ 1 μM). It does not inhibit other DNA polymerases or reverse transcriptase likely due to the close match between the size and shape of the macrocyclic structure of telomestatin to that of the G-quartet. Modeling studies of telomestatin suggest that it binds to the ends of the basket conformation of the intramolecular human telomeric sequence with a 2:1 stoichiometry. Telomestatin is also able to induce and stabilize G-quadruplexes in the absence of added monovalent cations, which is a unique characteristic among small molecules(Luedtke 2009).

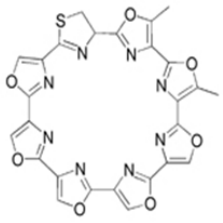


Figure 2-10 Structure of telomestatin, a natural product with the strongest affinity to G-quadruplex structure.

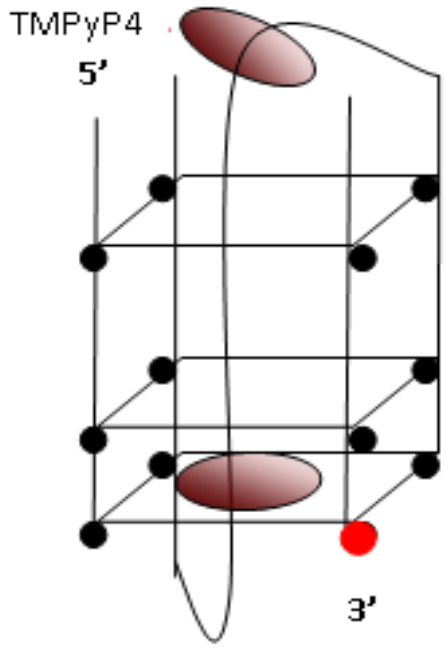
Since telomestatin is a natural product, it is very difficult to obtain, especially in quantities needed for widespread therapeutic use. It has complicated ring system which is difficult to synthesize, and at present would be cost-prohibitive to produce commercially. Also, since telomestatin is an uncharged molecule, it has low solubility in water, which is not only undesirable from a therapeutic standpoint, but also makes it

difficult to study. On the basis of modeling and comparative analysis of the binding of TMPyP4 and telomestatin to G-quadruplex, a series of metal-substituted porphyrins and expanded porphyrins, can be designed and synthesized. The large ring systems are expected to overlap the G-quartet more effectively than the ring system of TMPyP4, and hoped to improve binding and selectivity. The substitution of metals in the porphyrin cores can eliminate the photocytotoxicity associated with porphyrins through rapid quenching of singlet oxygen.

In the present study, thermally induced unfolding of the structure formed by a guanine-rich region *c-kit* was studied by means of circular dichroism and molecular absorption spectroscopies. The affinity of G-quadruplex DNA for three drugs was determined using Surface Plasmon Resonance. In addition, DNA drug interaction was investigated as a function of pH value and temperature. Multivariate data analyses methods based on both hard and soft modelling were used to allow accurate quantification of the various species present in different conditions. In this work an attempt was made to systematically investigate and classify the binding modes of TMPyP4 and NMM. TMPyP4 and NMM have different net peripheral charge, core charge density and ring substituents. The dramatic variation in binding affinities, large red-shift in the absorption bands and binding mechanisms of these small molecules to the quadruplex DNA were compared to understand the effect of structural modifications of substituents and peripheral charge on molecular recognition.

3

Results and Discussion



3. Biophysical characterization of target drugs using molecular absorption and circular dichroism spectroscopy

3.1. Results and Discussion

3.1.1. Test for dimerization of porphyrin drugs

Overlaying data at different concentrations but identical absorbance [figure 3-1] has allowed identification of possible aggregation in NMM but not in TMPyP4 and TPPS. The molecular absorption spectra of TMPyP4 and TPPS were the same at different concentrations showing no evidence of dimerization. On the other hand, the Soret band of NMM was different at different concentrations. At low total NMM concentration, the absorbance maximum of NMM solution was at 394 nm whereas at high total concentration the absorbance maximum was shifted to 379 nm. This showed dimerization of NMM. Dimerization produced blue shift in the Soret band which is referred to as hypsochromic shift (shorter wavelength) and hypochromicity (reduced absorptivity).

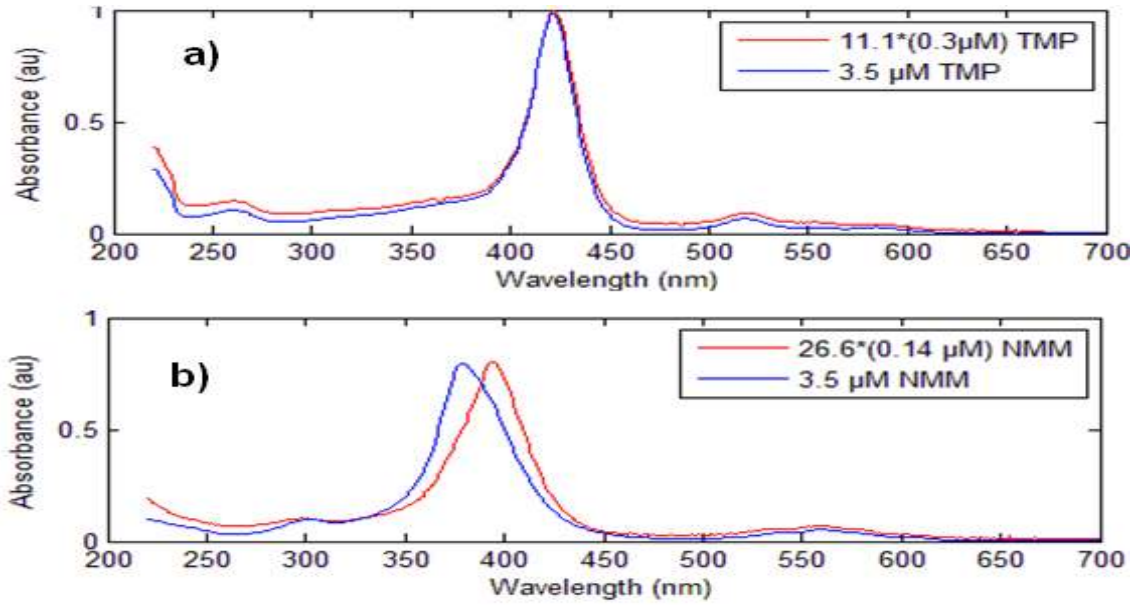


Figure 3-1 Results from monomer dimer association test for TMPyP4 (a) and NMM (b) in 0.1 M KCl, pH 7.22, buffer at 25°C. Absorbance at low concentration was scaled.

Estimation of fractional populations of the various monomer-dimer species present in NMM mixture was made possible by multivariate method based on soft MCR-ALS modeling using two components. The results are illustrated in figure 3-2 and table 3-1. Well resolved spectra were assigned to monomer and dimer states. Results showed that only at very low concentrations, below 1 μM, it is practical to assume the monomer as the only species present. At all other concentrations, NMM monomer and dimer coexist, higher concentrations favouring the latter.

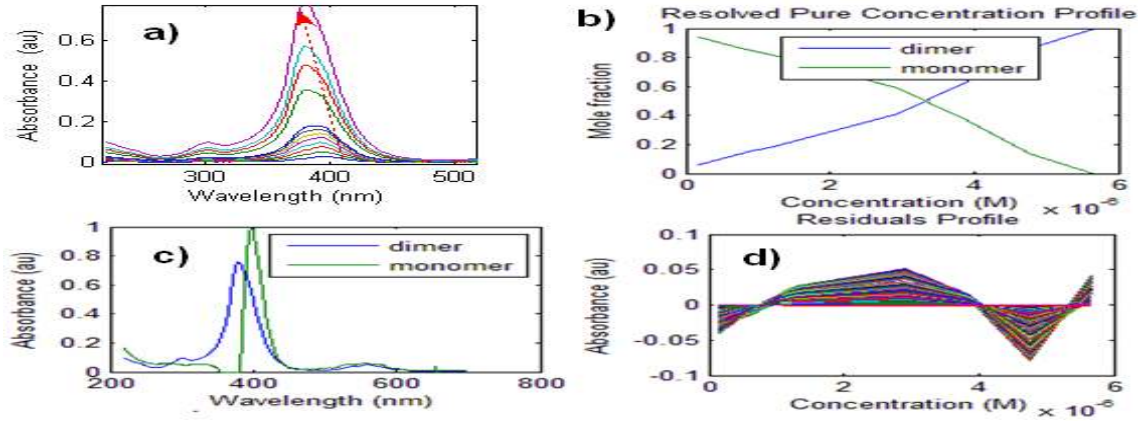


Figure: 3-2 Results from multivariate quantification of various species present in total NMM for the concentration range from 0.1 μM to 3.5 μM. a) Experimental molecular absorbance data. b) Resolved concentration profiles for the two optically active species present as mixture. c) Resolved pure molecular absorption profiles of the dimer and monomer. Absorbance for the monomer was multiplied by a factor of 26.6 as before. (d) Unexplained absorbance. Experimental conditions are pH 7.2 at 25°C in QSB.

Fitting error in percentage observed for the proposed components was found to be high (6.27%). This could be due to high uncertainty associated with estimated concentration of the monomer. The monomer is the dominant species at concentrations too low to measure accurately by molecular absorption spectroscopy.

Why does NMM dimerize whereas TPPS and TMPyP4 don´t?

Porphyrins have a well-extended π conjugate system. They are expected to form self-aggregated molecular complexes through van der Waals interactions. The major forces driving porphyrin aggregations are π-π stacking interactions which are intrinsic properties of porphyrin core. Van der Waals interactions are influenced by polarizability and surface area of contact. However, several arguments indicate these π-π stacking interactions are not enough to account for the whole story. An additional driving force is, therefore, required, and it is assumed to be solvophobic association. Porphyrin aggregation in water is generally similar to hydrophobic interaction of nucleic acids in being enthalpically driven and entropically opposed(Margalit and Rotenberg 1984).

When a molecule is dissolved, a cavity has to be created in the solvent to accommodate the molecule. The formation of the cavity is an increase in the surface area of the solvent and therefore requires input of energy, mainly enthalpic, to work against the surface tension of the solvent. When two such molecules, already dissolved, associate, two cavities are joined into one, reducing, therefore, the surface area of the

solvent and releasing some of the energy invested in creating the individual cavities. It should be noted that this solvophobic association is not restricted to aqueous solutions, but is expected to be strongest there, since the surface tension of water is exceptionally high compared with that of other common solvents. Thus a decrease in the dimerization constant from water to pure solvents of lower surface tension, and at the extreme, no detectable dimerization is predicted based on solvophobic association. Solvophobic association depends on the nature of solute, the nature of the solvent, the nature of cosolvents and medium properties.

Self dimerization of charged porphyrins requires π-π stacking and solvophobic interactions overcome electrostatic repulsive forces. Localized charge density on porphyrin core is very important indicator of stacking. This charge density depends on electrostatic interaction of charged substituents and dipole-dipole interactions. TPPS and TMPyP4 carry four negative and four positive charges respectively with large populations of charged forms at neutral pH. NMM carry two negative charges with lower population of charged forms than both TMPyP4 and TPPS at neutral pH. As a result, NMM is more hydrophobic, and has higher core electron density than both TPPS and TMPyP4. Thus electrostatic repulsive forces among NMM monomers have lower magnitude than that of both TPPS and TMPyP4. Subsequently, NMM dimerizes whereas TPPS and TMPyP4 don't. TPPS has higher core electron density than TMPyP4. As a result TPPS has been known to aggregate spontaneously in water at high ionic strength to reduce electrostatic repulsion. TMPyP4 has the lowest core electron density. The positive charges of the peripheral pyridinium groups delocalized on the porphine ring seem to prohibit the formation of the TMPyP4 dimer by overcoming π-π stacking and solvophobic interactions. In the case of NMM, however, the London's dispersion force between the porphine rings overcomes the electrostatic repulsion, leading to the dimer formation(Kano, Minamizono et al. 1997).

3.1.2. Determination of absorption coefficients

The concentration of NMM was determined from the concentration of the stock solution based on gravimetric and volumetric measurements. Concentrations determined in this manner are only approximate. The most accurate method to determine concentration is spectroscopic method. However, the true absorption coefficient of N-methyl mesoporphyrin IX which takes dimerization into account has not yet been established. Here the absorption contributions of the two species are dissected for the first time using multivariate curve resolution. The dissected values agreed very well with values reported in literatures. The values of absorption coefficient obtained are provided in table 3-1.

Why NMM has Soret band at shorter wavelength than TMPyP4 and TPPS?

NMM has lower molecular dimension compared to both TPPS and TMPyP4. The core electron charge density, which is responsible for the Soret band, is the function of delocalized electrons and surface area of delocalization. In cationic TMPyP4, N-methyl pyridyl groups delocalize the core electron density and diminish it. The amount of delocalized electron cloud is higher in TPPS than in NMM. However, NMM has the smallest molecular radius and two negative charges giving it a high electron density and hence high energy of excitation.

Ser. No.	Drug	Wavelength at A_{max} (nm)	Slope $(M\ cm)^{-1}$	SE(slope)	R^2	Reported Molar Absorptivity $(M\ cm)^{-}$
1	TMPyP4	422	283,178	3736	0.998	226, 000(Ren and Chaires 1999)
2	NMM Monomer	394	116,281	1458	0.999	(Ren and Chaires 1999)
	NMM Dimer	379	85,336	1859	0.995	
3	TPPS	413	446,975	14460	0.996	418,000 (Kojima, Koda et al. 2001)

Table 3-1 Molar absorption coefficients of water soluble porphyrins in QSB (7.22 pH, 0.15 M ionic strength) at 25°C.

3.1.3. Acid-base properties of porphyrin drugs

The acidity constants (pKa values) of the three drugs were determined by spectroscopic acid- base titrations using a constant amount of monomeric porphyrin. For this study, the titration data was examined for protonation of the free base [figure 3-4]. To evaluate pH-absorbance spectra, the data matrix was fitted by multivariate curve fitting programs based on hard-modeling (EQUISPEC) and soft modeling (MCR-ALS) approaches. The results are shown in figure 3-3. The concentration of monomeric NMM was determined as percentage of total concentration. At the concentration level used in this experiment, 70% of total NMM concentration was in the form of monomer. The acidity constants of all equilibria were estimated using an established simple chemical equilibrium model. The pKa values of TMPyP4, TPPS and NMM as determined by EQUISPEC program are as shown in the table 3-2. Experimentally determined value of pKa for TMPyP4 was significantly different from reported value. This could be due to the fact that the pH meter used in this study was calibrated at pH 4 and 9. There was high uncertainty in the measured values outside the calibration range particularly at very low pH. At neutral pH, where the interaction of water soluble porphyrins with quadruplex forming DNA is to be investigated, all the three drugs are essentially in free base form. Therefore, there are only two protons associated with central nitrogen atoms in the porphyrin ring.

Why TPPS is more basic than NMM?

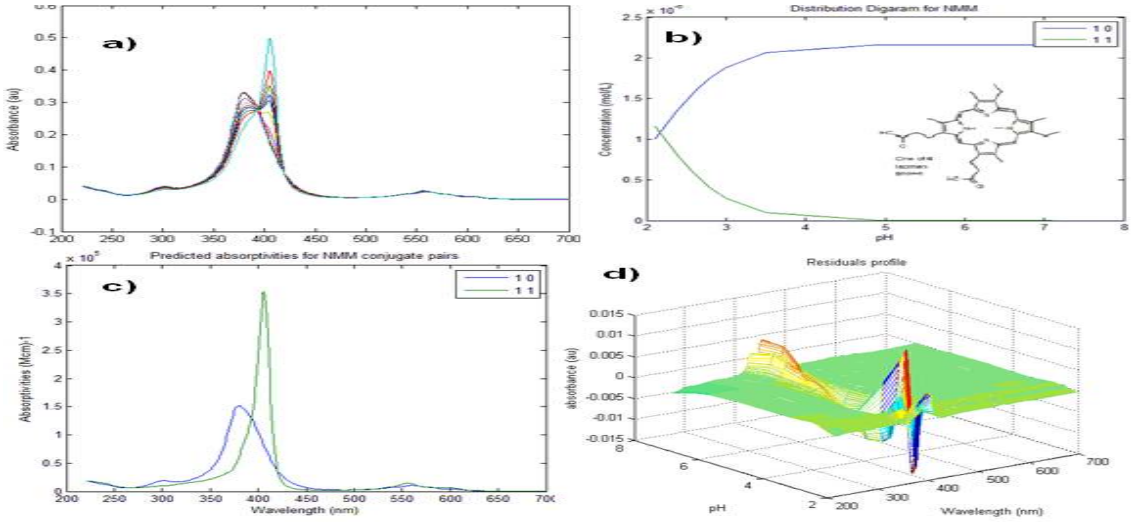


Figure 3-3 Acid-base properties of NMM. a) Original absorption spectral data at pH values ranging from 1 to 8. b) EQUISPEC resolved concentration profiles for the two optically active conjugate pairs present along the titration). c) Predicted molecular absorptivity. d) Un-explained variance. Experimental conditions are stated in the text.

Ser. No.	Drug	pKa	Standard error of pKa	Residual error	Reported pKa (Peter 2003)	Predicted absorptivity of free base (M cm)$^{-1}$	Predicted absorptivity of monocation (M cm)$^{-1}$
1	TMPyP4	1.07	0.03	0.0403	1.4	238,000	283,000
2	NMM	2.17	0.07	0.0281	2.2	152,000	353, 000
3	TPPS	4.62	0.02	0.0574	4.7	244,000	206,000

Table 3-2 Acid dissociation constants of water soluble porphyrins, different parameters that were used to evaluate the quality of fit during EQUISPEC curve fitting procedure are shown.

Covalently, mono-N-alkyated porphyrins are more basic than their non-alkylated partners(Weinraub, Peretz et al. 1982). Alkylation deforms the central nitrogen network and perturbs solvation pattern. This makes the central nitrogen atoms more accessible to protons. The core porphyrin nucleus is relatively planar in free bases. Protonation normally requires some degree of deformation. Porphyrins having electron donating group at the beta position have high proton affinity. NMM has alkyl substituents on alpha and beta positions which have electron donating effect. Proton affinity is mainly the function of net core electron charge rather than electron density. Delocalized electrons on meso substituents could be resonance conjugated with the central π system. Tetranegative TPPS has very high proton affinity due to conjugation of anionic substituents without delocalizing central π system whereas tetrapositive TMPyP4 has the lowest proton affinity due to effective delocalization of central π system and electrostatic repulsive forces with protons.

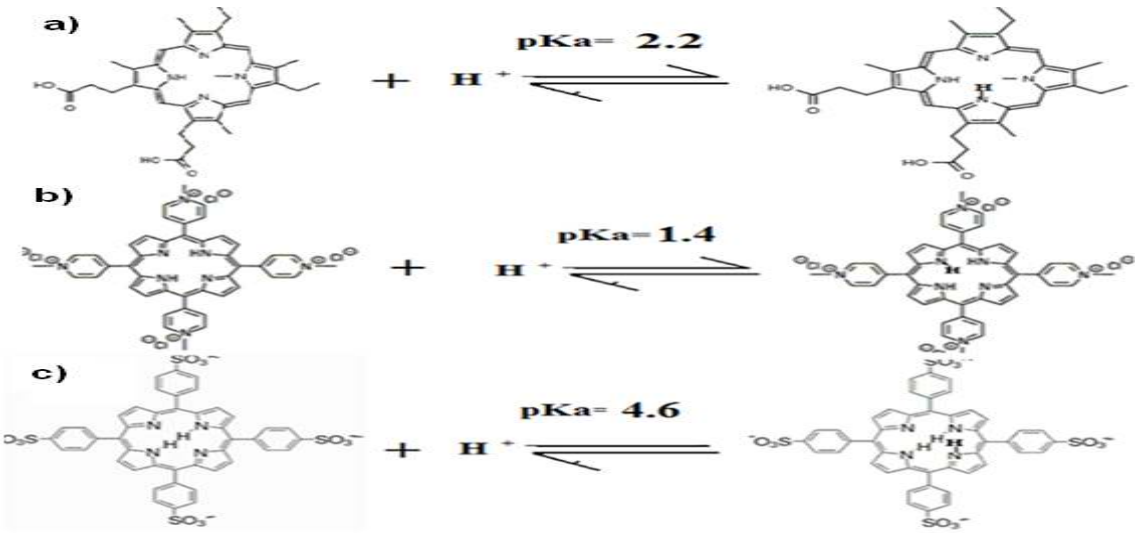


Figure 3-4 Acid base equilibria of (a) NMM (b) TMPyP4 and (c) TPPS. Note that NMM has only two protonated isomers whereas both TMPyP4 and TPPS have three isomers.

3.1.4. CD spectra of porphyrin drugs

Circular Dichroism (CD) is observed when optically active chemical species absorb left and right circularly polarized light slightly differently. As can be seen in figure 3-5, both NMM and TMPyP4 don't show CD signal between 220 nm and 500 nm as both drugs don't have chiral secondary structures.

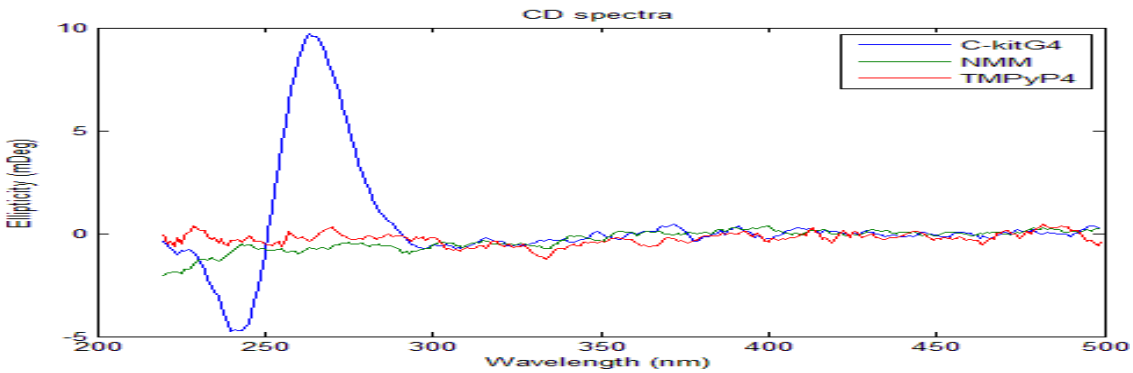


Figure 3-5 Circular dichroism spectra of *c-kitG4* (Quadruplex forming DNA), NMM and TMPyP4 in QSB pH 7.22 at 25°C.

3.1.5. Fluorescence spectra of NMM

Florescence spectroscopy indicated the intensity of fluorescence signal didn´t increase with increasing total concentration of NMM. This agreed with previous reports of dimerization of NMM where the monomer was fluorescent and the dimer wasn´t(Margalit, Shaklai et al. 1983). At this particular concentration (3.2 to 6.2 μM), increase in total amount of NMM in the solution decreased the fluorescent signal due to the formation of more dimers.

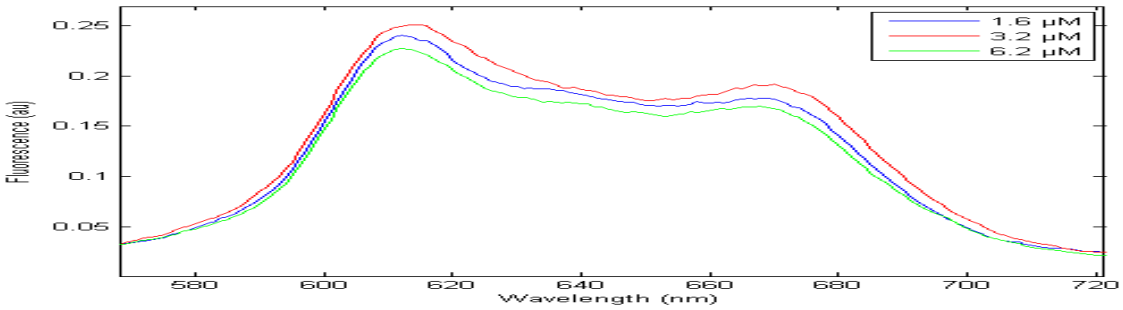


Figure 3-6 Fluorescence spectra of N-methyl mesoporphyrin IX in QSB containing 0.15 M KCl, pH 7.22 at 25°C.

3.1.6. Thermal Denaturation of porphyrin drugs

Spectroscopic thermal denaturation experiments of TMPyP4 and TPPS in QSB demonstrated that the Soret band was substantially diminished upon increasing temperature above 60°C [Figure 3-7]. This is interpreted as spectroscopic evidence for the formation of alkali metal porphyrin complexes. QSB contains mainly potassium ions (0.147M) and trace amounts of sodium and magnesium ions. Complexation reactions of most metals with porphyrins are often endothermic due to the release of free base protons. Magnesium is also reported to form stable metalloporphyrin complex.

Application of MCR-ALS for the resolution and quantification of different chemical species gave reasonable fit for proposed two components. These components are interpreted to be the free base and the alkali metal complex. The absorptivity of the metalloporphyrin complex wasn't reliably determined as the data matrix was insufficient. The red shift is only slightly observed and there is every reason to expect a new band at higher wavelength at elevated temperatures or higher alkali metal concentrations.

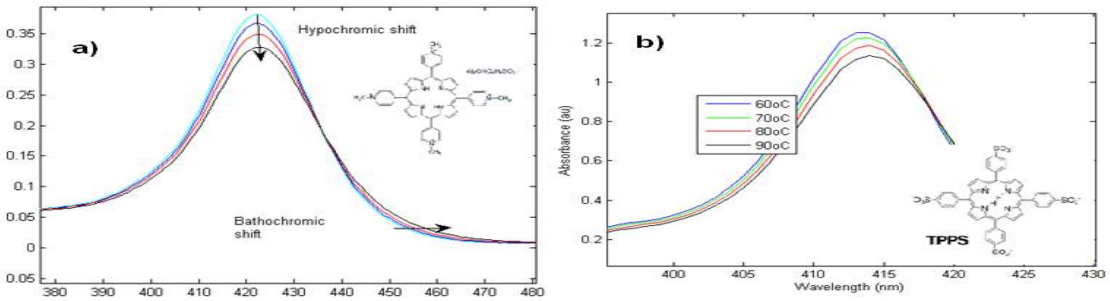


Figure 3-7 Change in absorbance spectra of NMM and TPPS at different temperatures during melting experiments in QSB. Effect of increasing temperature on absorbance spectra of NMM (a) and TPPS (b)

Hypochromicity at the Soret band of the free porphyrin base and red shift with increasing temperature might be due to formation of metalloporphyrin. Spectroscopic thermal denaturation experiment of NMM

clearly showed the melting of dimer into the monomeric form [3-8, a]. Moreover, as shown in [3-8, b] thermal denaturation of NMM required three components (including the dimer) to explain the observed spectroscopic data. Percentage fitting error of the observed was still high (3.78%) due to incomplete process.

Ser. No.	Drug	Number of components	Residual error	Variance explained	Predicted absorptivity of free base (M cm)$^{-1}$	Predicted absorptivity of metalloporphyrin (M cm)$^{-1}$
1	NMM	3	0.0080	99.8575	245,000	270,000
2	TMPyP4	2	0.0016	99.968	249,000	197,000
3	TPPS	2	0.0068	99.8758	392,000	366,000

Table 3-3 MCR-ALS predicted absorptivities and measures of goodness of fit for specified number of components as optically active species present along melting of porphyrin drugs in QSB.

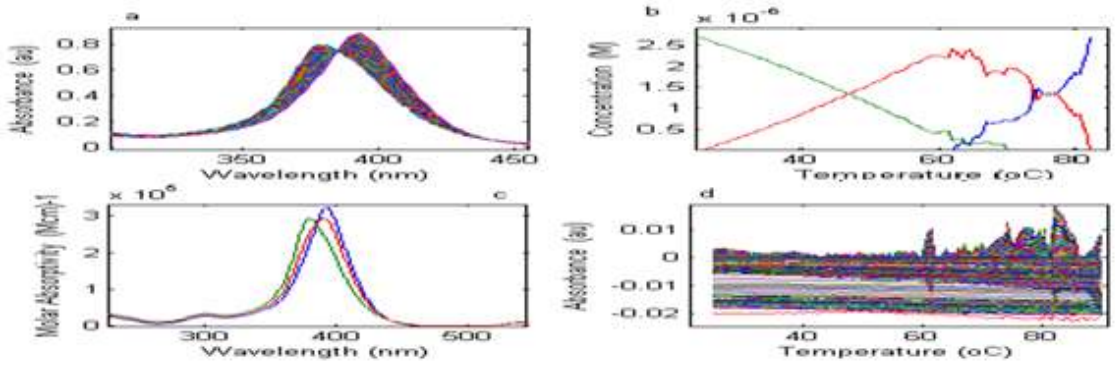


Figure 3-8 a) Experimental absorbance data obtained for thermal denaturation of NMM. b) MCR-ALS resolved pure concentration profiles c) The corresponding pure absorptivities. (d) Residual profile shows unexplained variance by three optically active species present along melting of NMM.

3.2. Conclusion

At sufficiently low concentrations used in the present study, only NMM aggregates to form dimer. NMM monomer and dimer coexist even at low concentration ranges; higher concentrations favour the latter. NMM is more hydrophobic; carry less net charge and has higher core charge density than both TPPS and TMPyP4. This could be the structural reason for its ability to dimerize. The Soret band of the free base showed red shift from NMM to TPPS and TMPyP4. This is interpreted to follow reduction in the core electron density of parent porphyrin ring due to ring substituents.

TMPyP4 has the least proton affinity whereas TPPS has the highest among the three. Alkylation in NMM deforms the central nitrogen network and makes the central nitrogen atoms more accessible to protons.

TPPS has very high proton affinity due to conjugation of anionic substituents to porphyrin ring enhancing electron charge accumulation.

Spectroscopic thermal denaturation experiments of TMPyP4 and TPPS in QSB demonstrated evidence for the formation of alkali metal porphyrin complexes at temperature above 60°C. Melting of NMM confirmed the existence of dimer which melted in to monomer at elevated temperatures.

At conditions relevant to biological activities namely, physiological pH, temperature and ionic strength, all the three porphyrins studied are found to be practically in free base form. Only NMM forms dimers and none of the three drugs form metalloporphyrin complexes. Most of the porphyrins are therefore essentially in the form of one species. Subsequent research to characterize binding interactions of these porphyrins with biological targets can be considered.

4. Biophysical characterization of G-quadruplex DNA targets using molecular absorption and circular dichroism spectroscopy

The three dimensional structure of *c-kitG1* sequence has been determined by solution state NMR spectroscopy. The result provided a clear evidence for structural polymorphism and the presence of at least two structurally similar but dynamically distinct populations (Hsu, Varnai et al. 2009). This sequence forms a parallel quadruplex in K^+ solution with internal dynamics that are modulated by loop composition. The c-*kit* quadruplex is structurally heterogeneous due to the four consecutive guanines at the 3' end, each of which could participate in G-tetrad formation [figure 4-1].

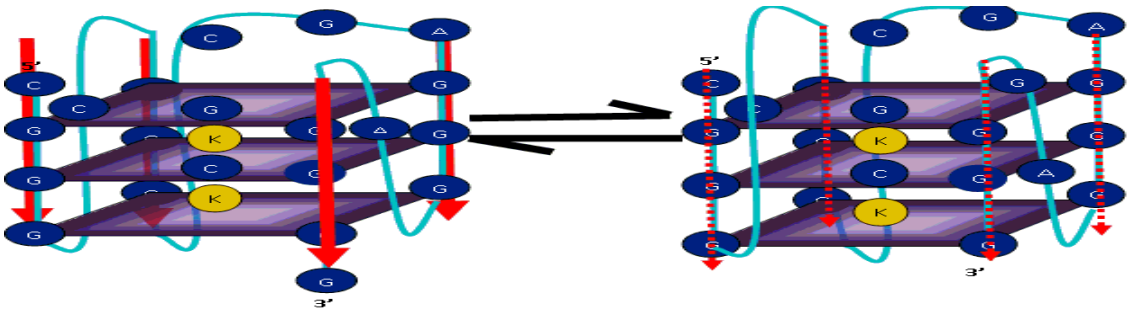


Figure 4-1 Schematic representation for two conformational isomers of *c-kitG1* quadruplex in QSB. Dangling-end conformer (left) has a flanking nucleotide and blunt-end conformer (right) has no flanking nucleotide.

The presence of potassium ion dependent conformation that is likely to be antiparallel has also been postulated. Folding of intramolecular G-quadruplex requires all favorable interactions to offset

electrostatic repulsions in phosphate backbone. These favorable stabilizations include non-specific ion interaction, specific site ion binding and hydration.

Ser. No.	Sample	Sequence	MW (Da)	ε (M cm) $^{-1}$	Tm (°C)
1	c-kitG1	CG3CG3CGCGAG3AG4	6659	222,750	75
2	c-kitG4	TG3TG3TGTGTG3TG4	6698	214, 650	76

Table: 4-1 Characteristic of wild type c-kitG1 sequence and its loop variant c-kitG4. The given absorptivites are and melting temperature are estimated by adjusted nearest-neighbour calculation method.

4.1. Results and Discussion

Thermal denaturation for both sequences [Figure 4-2] shows significant hypochromicity at 295 nm upon melting which is considered to be the hallmark of G-quadruplex structures. The mid- point of such rapid transition was assigned to be the melting temperature (T_m) of G-quadruplex DNA. Several melting experiments were carried out at different DNA concentrations. In all experiments, there was no significant change in melting temperature of *c-kitG1* and transition of *c-kitG4* was still incomplete. This was an evident for a monomolecular transition and higher melting temperature of *c-kitG4*. *C-kitG1* showed a sharp hypochromicity between 65°C and 85°C [Figure 4-2, b] and an additional small region of hypochromicity around 39°C. Here the melting temperature of c-*kitG1* was determined to be 74°C.

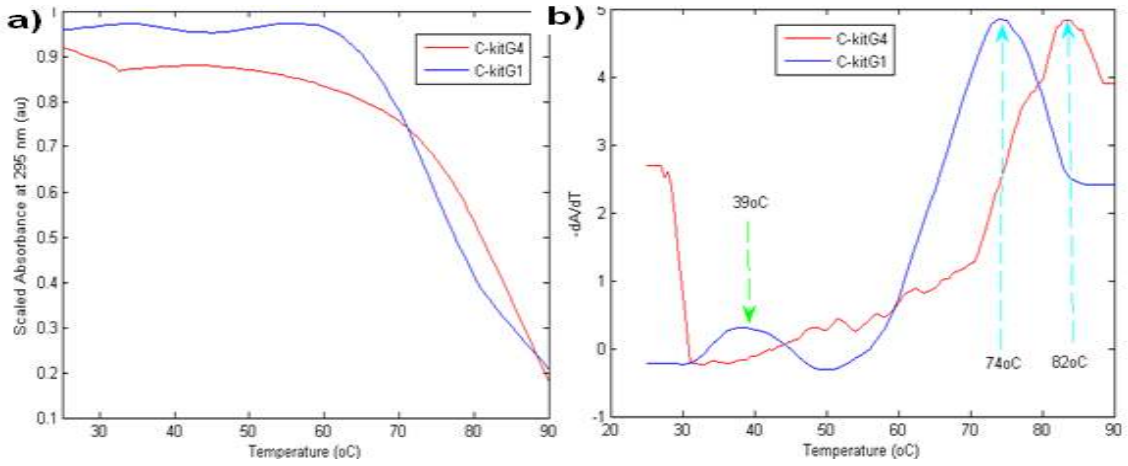


Figure 4-2 Thermal denaturation curve for *c-kitG1* and *c-kitG4* in QSB. a) Tracing scaled absorbance at 295 nm is plotted against temperature b) the negative of the first derivative of relative absorbance was smoothened and monitored at 295 nm.

The melting curve obtained for the modified sequence *c-kitG4* exhibited a single melting transition at 295 nm. *C-kitG4* showed incomplete hyperchromicity over the temperature range studied. It is likely to show clear cooperative melting at elevated temperature ranges. The melting temperature for *c-kitG4* was

estimated to be 82°C. This value is only approximate. Additional experiments at higher temperature ranges should be considered to get conclusive transitions.

4.1.1. Determination of Thermodynamic Parameters

The percentage fitting error of observed data was 1.61% for MCR-ALS decomposition of the experimental spectra obtained during melting of *c-kitG1* for a proposed two species. There was very little gain in information by adding a third component in the model. By looking at this fitting error, it seemed reasonable to explain the melting process with only two components.

Thermodynamic parameters for quadruplex unfolding were estimated from melting curve data using the two MCR-ALS resolved components. However fitting the two component concentration profiles to van't Hoff equation showed a funnel shaped tendency of residuals [Figure 4-4] with increasing reciprocal of absolute temperature. This is particularly true when more components are required to describe a model.

The inclusion of a third component eliminated this trend. For three components, the residuals were scattered around the mean value of zero [Figure 4-4]. This is a characteristic of a reasonable model. Consequently, the inclusion of an intermediate species was considered more plausible than the direct melting of *c-kitG1*. In general, the data recorded in this melting experiment were well fitted when three components, i.e. two transitions were considered. The first transition located at intermediate temperatures, actually modelled the small conformational change at the loop. The second, transition, located at high temperature and very sharp, was related to the cooperative unfolding of the G-tetrads.

MCR-ALS output [Figure 4-3] shows melting of the G-quadruplex conformation through an intermediate structure. At low temperature, most of the population of the oligonucleotide is in folded native conformation (green line) with very small amount of intermediate structure (red line). The native structure is proposed to be the blunt end conformer and the intermediate is proposed to be dangling end conformer. At high temperature, the random coil (blue line) is practically the only species present. The absorptivity of the native conformation is slightly lower than the random coil. However, the absorbance spectra and molar absorptivity of the intermediate is very similar to that of the native structure. According to MCR-ALS, at pH 7.2 in QSB melting temperature of *c-kitG1* quadruplex to form random coil was 75°C. The small hypochromicity at 39°C can be explained by an intermediate in which one of the three loops is deformed. Loop 2 is particularly susceptible loop as it is the longest and with the highest degree of freedom. This result is in line with NMR studies which concluded loop composition play important role in internal dynamics of c-*kit* G-quadruplex structure. The point can be further confirmed by systematically changing this loop and analyzing melting curves.

The enthalpy and entropy for the main transition were obtained by van't Hoff analysis [Figure 4-5] of melting data assuming a valid two state transitions in a narrow temperature range near the melting temperature. This assumption is further supported by a finite entropy difference between the three states (G-quadruplex, intermediate and random coil). Two-state approximation can be extended for an n-state system only if the entropy differences between the n-states are finite. In this case the van't Hoff enthalpy and the calorimetric enthalpy of the n-state system can have good agreement(Bakk and Metzler 2004). Thermodynamic data are given in table 4-2.

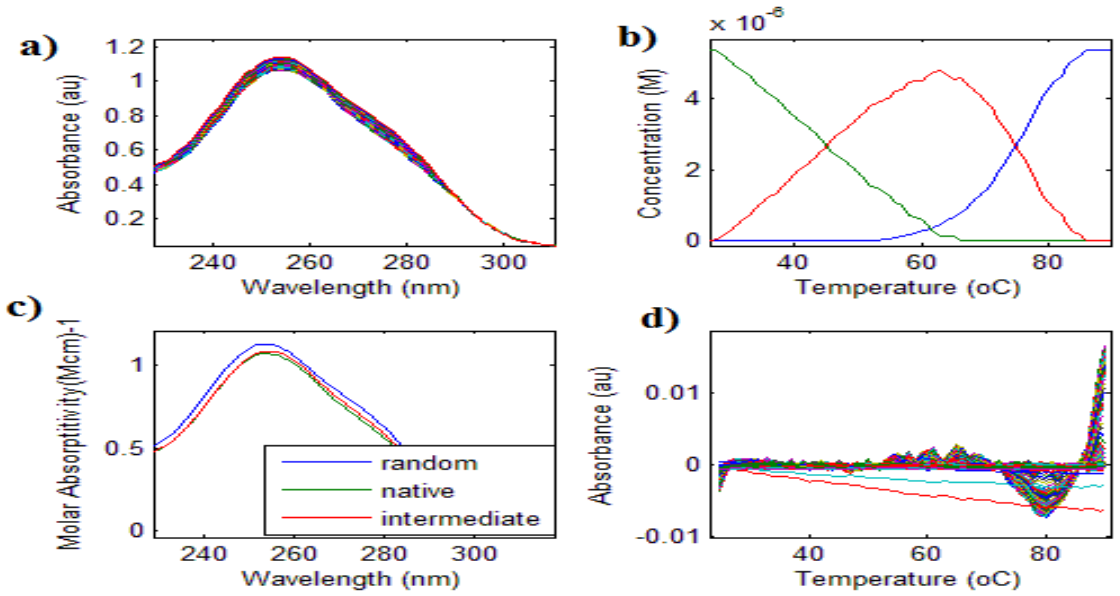


Figure 4-3 Multivariate quantification of various species present during melting of *c-kitG1*. a) Experimental absorbance data for melting of *c-kitG1* in QSB b) MCR-ALS resolved concentration profiles for the three optically active species present as mixture c) the corresponding resolved molecular absorptivity d) Unexplained absorbance. Experimental conditions are given in the text.

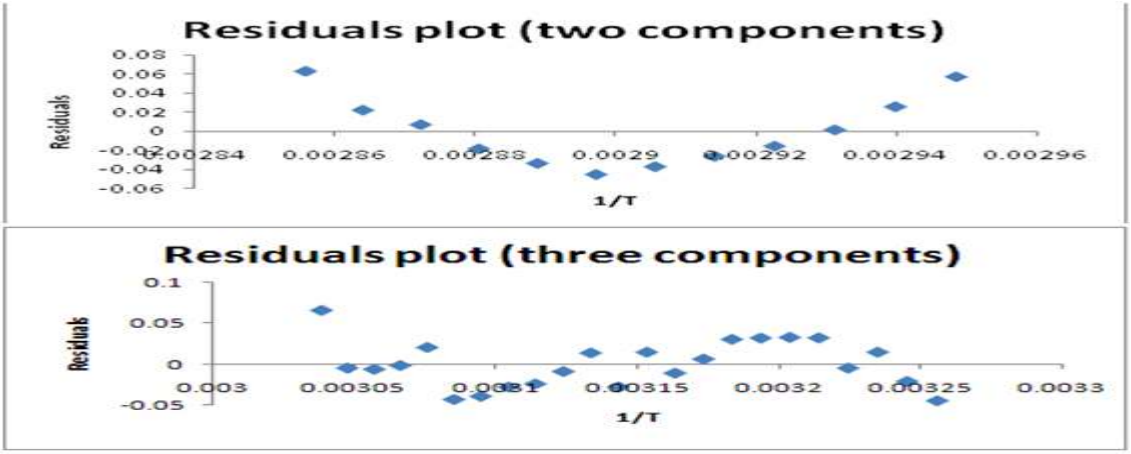


Figure 4-4: Residuals plot for linear regression modeling of thermodynamic parameters using van´t Hoff equation during thermal denaturation of *c-kitG1*in QSB based on two and three components model.

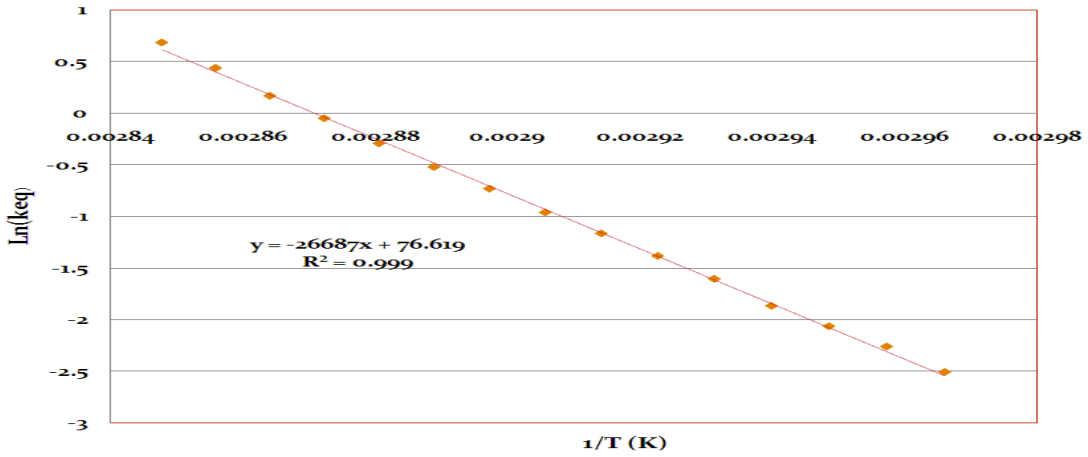


Figure 4-5 Linear regression modeling for van't Hoff representation of *c-kitG1* melting. Equilibrium constants were calculated from MCR-ALS resolved concentration profiles for three components

Once again, the experimental data for melting of *c-kitG4* were well fitted when three components were considered. MCR-ALS output for thermal denaturation of *c-kitG4* using three spectroscopically distinguishable species [Figure 4-6] shows direct melting of the G-quadruplex conformation. It is shown that a dominant G-quadruplex conformation (red line) and a somewhat minor conformational population (green line) essentially coexist at low temperatures. The dominant conformer is assumed to be the dangling end conformer and the minor population is assumed to be the blunt end conformer.

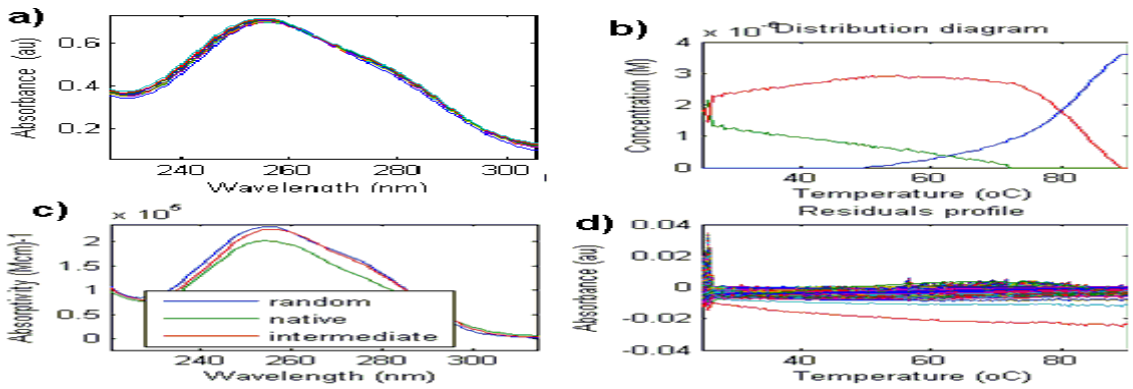


Figure 4-6 Multivariate quantification of various species present during melting of *c-kitG4*. a) Experimental absorbance data for melting of *c-kitG4*in QSB. b) MCR-ALS resolved concentration profiles for the three optically active species present as mixture. c) Resolved molecular absorptivity and unexplained absorbance (d). Experimental conditions are pH, 7.2 in QSB at 25°C.

Melting of *c-kitG4* showed some interesting difference from that of *c-kitG1*. As would be expected, the loops of *c-kitG4* were already unfolded even at low temperature and showed no clear transition. These altered loops without secondary structure were thought to be equivalent to the intermediate species formed during melting of *c-kitG1*. According to MCR-ALS, at pH 7.2 in QSB melting temperature of the *c-kitG4*

quadruplex to form random coil was 80°C. Melting temperature of *c-kitG4* was found to be higher than *c-kitG1*. This result is consistent with previously reported single base substitution on the loops. This implied that if loop 2 in *c-kitG1* wasn´t folded into some conformation, it would destabilize the folded G-tetrad structure.

Ser No.	Sequence	ΔH^o (kcal/mol)	ΔS^o (cal/Kmol)	ΔG^o (kcal/mol)	T_m (°C)
1	5'-$CG_3CG_3CGCGAG_3AG_4$	53.1 ± 0.5	152.4 ± 1.4	5.8 ± 0.6	75 ± 1
2	5'-$TG_3TG_3TGTGTG_3TG_4$	48.0 ± 1.2	136.0 ± 3.4	5.8 ± 1.6	82 ± 2
3	5'-$AG_3(TTAGGG)_3$	66.2	186.5	4.9	61.8
4	5'-$GGG(TTAGGG)_3$	77.5	202	7.3	69.3

Table 4-2 Thermodynamic parameters for thermal denaturation of intramolecular G-quadruplex forming sequences. *C-kitG1* and *c-kitG4* (1 and 2) are in pH 7.2 QSB at 37°C whereas telomere sequences (3 and 4) are in 100 mM potassium buffer at 37°C(Lane, Chaires et al. 2008).

Thermodynamic values agree with those previously described (Lane, Chaires et al. 2008) for similar structures in potassium solution. The free-energy change at 310K varied from stable (4.9 kcal mol^{-1}) to very stable (5.8 kcal mol^{-1}). Both enthalpy and entropy changes are positive in the context of G-quadruplex unfolding. The results for *c-kitG4* are not comfortably conclusive. Uncertainties associated with melting of this sequence are relatively large obscuring comparison of stability with other quadruplexes. One possible origin of high uncertainty could be very high melting temperature of *c-kitG4* and incomplete transition even at the highest temperature considered. In general, melting was accompanied by increase in entropy from a very ordered structure of quadruplex into a random coil and the process was endothermic. Folding, the reverse process, required favorable free energy changes resulting from favorable enthalpy. High enthalpy changes are noteworthy as they are the main driving forces for G-quadruplex folding.

Why *c-kitG4* has higher melting temperature than *c-kitG1*?

The most destabilizing component in nucleic acids at neutral pH is the very unfavorable electrostatic repulsion between the oxygen atoms in the phosphodiester backbone, which bear a formal charge of –1. It has been known for a long time that the carbonyl oxygen of cytosine has the ability to form three hydrogen bonds in water. The strong hydrogen-bonding of this carbonyl oxygen has also been associated with a greater formal charge and stronger dipole moment on this functional group than that of other pyrimidines(Jeffrey and Kinoshita 1963). Cytosine residues exacerbated unfavorable electrostatic interaction in the backbone. Searches for potential G- quadruplex-forming sequences in the human genome revealed that the most common linker between the G-tracts is single base A or T and not C. Solution state NMR structure of *c-kitG1* sequence showed C5 of loop 1 adopted a wide range of

orientations indicating significant internal mobility (Hsu, Varnai et al. 2009). Besides, the size of the base may also be significant factor. Some researchers found purine bases in the loop (AAA) produced less stable G-quadruplexes than pyrimidines (TTT and CCC)(Phillip, Tom et al. 2007). Adenine in the loops via increased steric hindrance with the phosphate backbone may limit the stability. Substituting cytosine and adenine in *c-kitG1* with thymines in *c-kitG4* could have reduced formal charge density and steric clash. The result was reduction of structural constraint providing higher melting temperature for *c-kitG4*.

4.1.2. Effect of environmental factors on G-quadruplex structures

In 1991, Prive et al. noted DNA is built from five structural elements rather than three: not only familiar bases, sugars and phosphates, but also ordered waters and bound counterions. This was to emphasize the major theme that water and ions which are ubiquitous in all nucleic acid environments have profound effect on the conformation and function of nucleic acids. There are many reasons to investigate the effect of environmental factors on the stability of G-quadruplex structures. Ions modulate the interaction of DNA with other charged DNA binding partners such as drugs. Nucleic acids are highly charged polyanions. They strongly interact with cations and there is great potential for regulation of biological processes by changes in ion concentrations. Cations play a central role in promoting G-quadruplex structures. The following results were obtained to study the effect of specific ion binding, non-specific ionic interaction and hydration on G-quadruplex structure.

G-quadruplexes exhibit characteristic CD spectra depending on the topology of the structure (parallel or antiparallel). A maximum positive absorbance at around 260 nm with a corresponding negative maximum at 240 is a characteristic feature of parallel G-quadruplex structures. The positive and negative ellipticity spectral profiles for both *c-kitG1* G1 and *c-kitG4* [figure 4-7, b] exhibited characteristic signature spectra for authentic parallel G-quadruplexes. These observations were consistent with the formation of parallel G-quadruplex reported by other researchers.

Thermal denaturation experiment of *c-kit* G1 was done in QSB. Experimental CD spectra are given in figure 4-7a, for selected temperatures. Every effort was made to quantify the different species present during melting using simultaneous multivariate analysis of merged CD and absorbance data [figure 4-8]. The distribution diagram didn´t suggest a clear transition of predicted species, probably due to high noise inherent to CD spectral profile and lack of sufficient data points. Nevertheless, the resolved pure spectral profile seemed to show an interesting spectral behavior for the predicted three spectroscopically active species. As shown in figure 4-8, there was a significant increase in the intensity of negative ellipticity in going from the native structure to the intermediate. The raw data also showed an increase in negative ellipticity in the temperature range 25°C to 50°C.

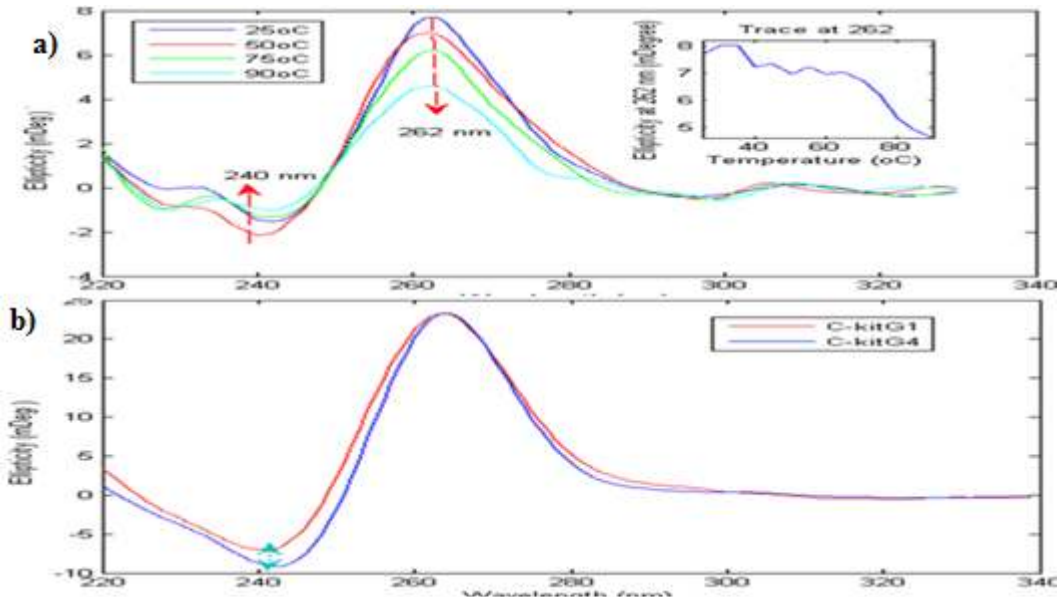


Figure 4-7 a) Experimental circular dichroic spectra of *c-kitG1* in QSB. Inset shows ellipticity trace at 262 nm. b) Overlaid CD spectra of C-kitG1 and C-kitG4 at 25°C in QSB. The arrow shows increased intensity of negative ellipticity in *c-kitG4*.

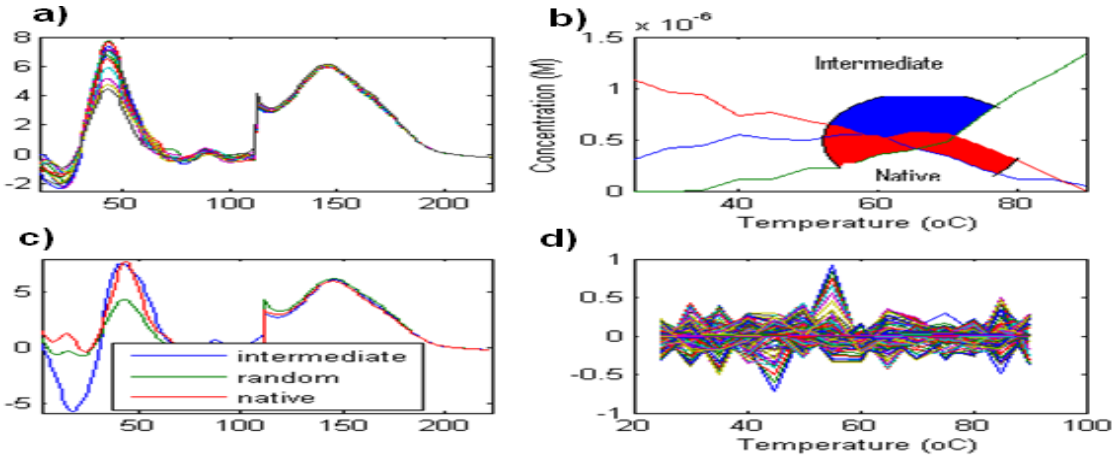


Figure 4-8 Multivariate quantification of various species based on simultaneous analysis of absorbance and CD spectra during thermal denaturation of *c-kitG1*. a) Merged experimental CD and absorbance data. b) Resolved concentration profiles for the three optically active species present as mixture the circle shows where the distribution diagram fails. c) Resolved pure spectral profiles. d) Unexplained variance. Experimental conditions are pH, 7.2 in QSB.

What causes increase in negative ellipticity in the temperature range 25°C to 50°C?

Both the raw data [figure 4-7, a] and resolved pure spectra profile [figure 4-8, c] suggested increase in negative ellipticity in the temperature range where the intermediate is expected to be dominant population. This could be interpreted in terms of an intermediate species with a deformed loop which affected the rotation angle between G-tetrads in a way that increased negative ellipticity and decreased positive

ellipticity. This agreed perfectly with the observed CD spectra of *c-kitG4* [figure 4-7, b] in which the loops are systematically altered to disrupt any ordered structure. The intermediate in thermal denaturation history of *c-kitG1* (native sequence) showed the same optical properties as the folded *c-kitG4* (loop mutated sequence). Such optical recapitulation was by no means merely fortuitous.

Ethanol stabilized G-quadruplex structure and the melting temperature was elevated in the presence of only 5% ethanol. Multivariate curve resolution [figure 4-9c] showed similar distribution diagram for different species during melting of *c-kitG1* with or without ethanol. These findings were consistent with previous reports of ethanol being better inducer of G-quadruplex than even K^+ ions.

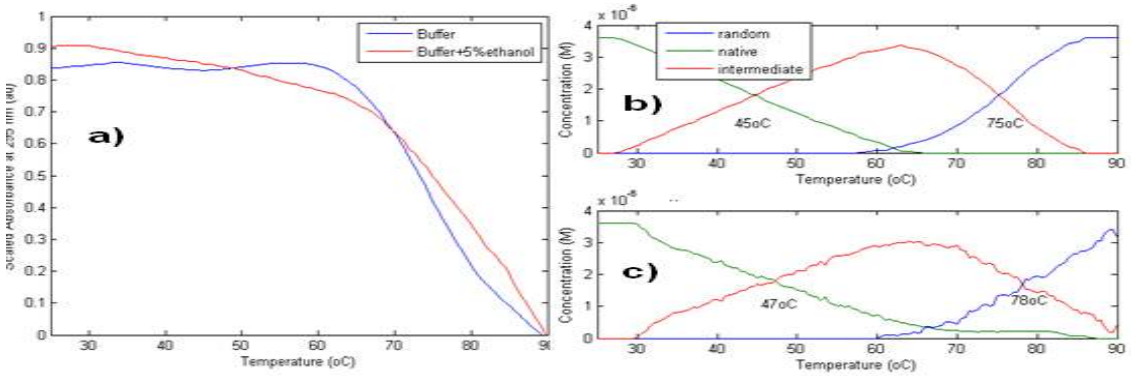


Figure 4-9 Thermal denaturation curves for *c-kitG1* in QSB with and without 5% ethanol. a) Tracing scaled absorbance at 295 nm is plotted against temperature. Results obtained from multivariate quantification of various species in QSB without ethanol (b) and with 5% ethanol (c).

How does ethanol enhance intramolecular G-quadruplex formation?

Several lines of mechanisms have been proposed. One possible mechanism is ethanol reduces water activity to increase G-quadruplex stability. Both G-quadruplex and single stranded oligonucleotide are heavily hydrated; the latter is even more hydrated. As a result, quadruplex folding is associated with the release of water molecules. Water molecules hydrating the G-tetrad might have the activity of reducing electron charge density of the quartet thereby diminish the extent of π-π stacking interactions. Replacing some of these waters by ethanol which has electron donating property could decrease water activity resulting in strong stacking. Adding ethanol may also shift the equilibrium towards formation of G-tetrads by reducing water of hydration. Furthermore, cosolvents such as ethanol not only decrease the water activity, but also the bulk dielectric constant(Michaela, ccaron et al. 2006). Reducing dielectric constant improves the interaction between cations in the cavity and phosphate backbone. This will effectively diminish phosphate-phosphate repulsive interaction that destabilizes the G-quadruplex structure most.

It has been suggested, in the absence of K^+ or Na^+, and only macrocation counterions present, the fraction of the folded state is small. Thermal denaturation of *c-kitG4* in Milli-Q water [Figure 4-10a] shows

cooperative unfolding as evidence for the presence of G-quadruplex. This G-quadruplex structure was found to have very low melting temperature.

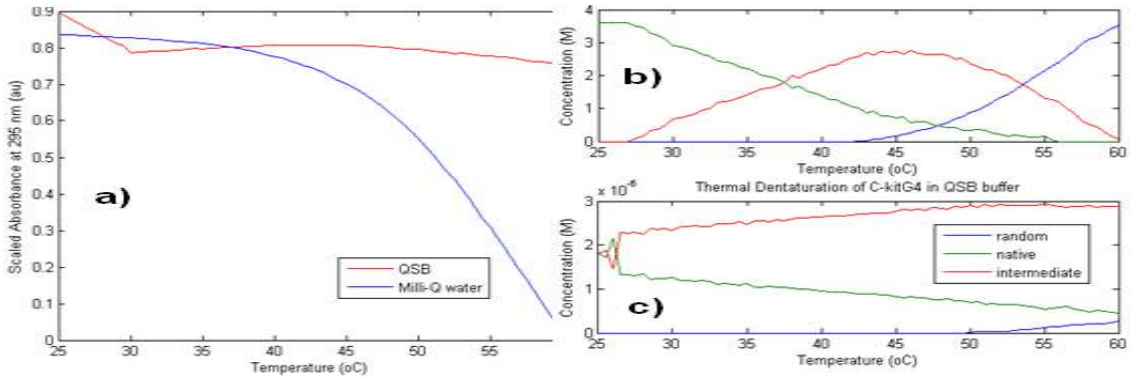


Figure 4-10 Thermal denaturation curves for *c-kitG4* in QSB and in Milli-Q water. Tracing scaled absorbance at 295 nm is plotted against temperature (a). Results obtained from multivariate quantification of various species in Milli-Q (b) and in QSB (c).

The formation of G-quadruplex in Milli-Q water was independently confirmed by characteristic CD spectra for parallel G-quadruplex [cf. figure 4-11] obtained for the same sequence *c-kitG4*. Moreover, titration experiments of both *c-kitG1* and *c-kitG4* with potassium revealed the formation of parallel G-quadruplex structures in Milli-Q to which no additional ions were added except for anything that came along with the dissolved DNA. There was much evidence to believe the formation of G-quadruplex without any added cation. Apart from the signature bands for parallel G-quadruplex, gradual changes [figure 4-11, inset] in the nature of the CD spectra were also observed during thermal denaturation of *c-kitG4* in Milli-Q. The observed CD spectrum in Milli-Q at 45°C showed a slight shift in the positive intensity to lower wavelength, the appearance of very small positive signals at 298 nm and negative signal at 270 nm. These behaviours are further demonstrated with arrows [figure 4-13].

Multivariate data analysis of merged CD and absorbance spectra for melting of *c-kitG4* in Milli-Q, showed acceptable distribution diagrams [cf. figure 4-13] with two distinct transitions. This agreed well with previously obtained distribution diagram using molecular absorption spectroscopy. Once again, the resolved pure spectral profile showed an interesting CD behavior for the predicted intermediate species. As shown in figure 4-12c, CD measurements identified two positive bands at 260 nm and 298 nm and two negative bands at 240 nm and 270 nm.

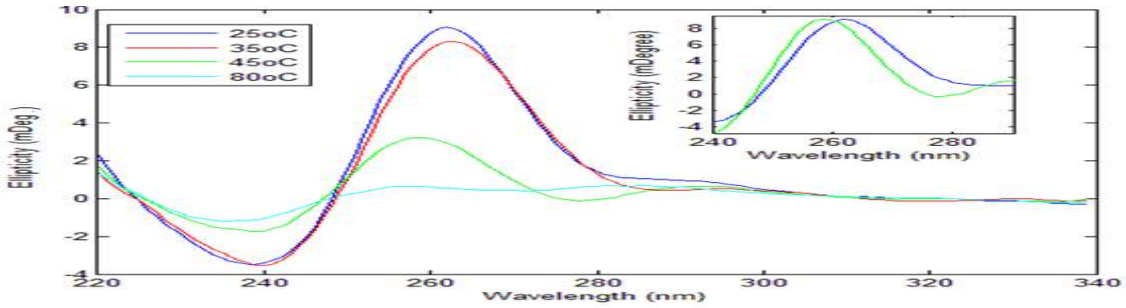


Figure 4-11 Experimental CD spectra for melting of *c-kit* G4 in Milli-Q water. Inset shows slight shift in the positive ellipticity from 262 nm to 258 nm during melting. The blue line corresponds to the fully folded quadruplex form. Spectrum obtained at 80°C is indicated by the cyan line corresponds to the unfolded form.

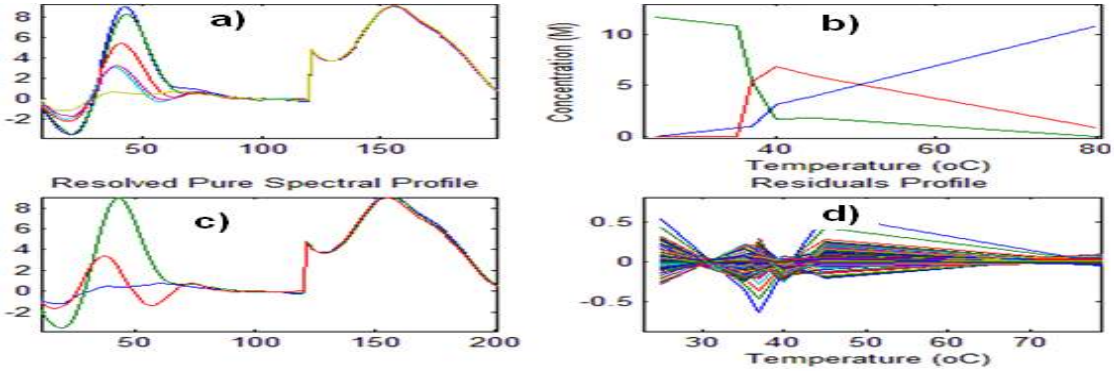


Figure 4-12 Multivariate quantification of various species based on simultaneous analysis of absorbance and CD spectra during thermal denaturation of *c-kitG4* in Milli-Q water. a) Merged experimental CD and absorbance data. b) Resolved concentration profiles for the three optically active species present as mixture. c) Resolved pure spectral profiles d) Unexplained variance.

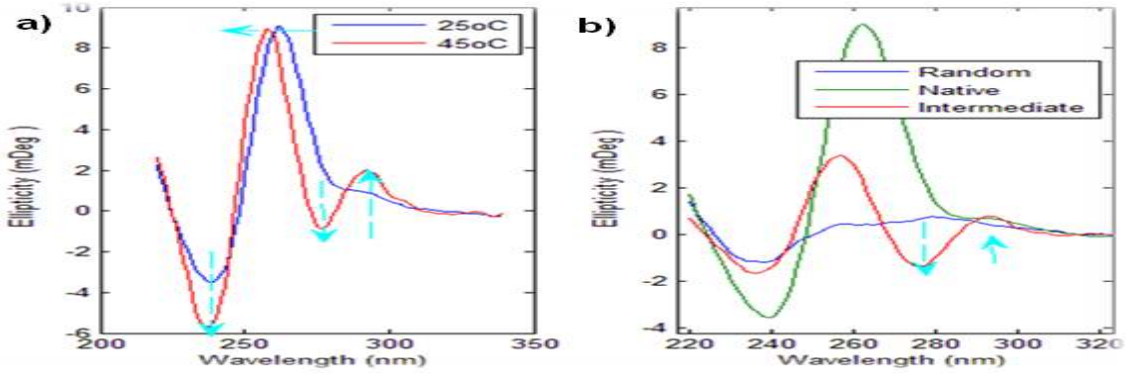


Figure 4-13 a) Experimental CD spectra for melting of *c-kitG1* in Milli-Q (b) MCR-ALS resolved pure spectral profile for proposed three species.

Antiparallel quadruplexes have a maximum between 285 and 295 nm and a minimum at 265 nm. The appearance of new signals at 298 nm and 270 nm suggested the possibility of conformational switch

involving a mixture of parallel and antiparallel G-quadruplex species. This could be interpreted in terms of an intermediate species with glycosidic bond showing some orientations which modify rotation angle between G-tetrads. The formation of potassium ion dependent antiparallel structure by native *c-kitG1* sequence was postulated based on NMR studies. Multivariate quantification showed different distribution diagrams for melting of *c-kitG4* in water and QSB. As previously described, the two structures that coexist in QSB at low temperature were considered to be the blunt end conformer and dangling end conformer. In the absence of potassium, there were two distinct transitions and a dominant population at room temperature suggesting different unfolding mechanism of the same sequence in the presence and absence of potassium ions.

Both circular dichroism and UV melting experiments indicated melting of *c-kitG4* in Milli-Q water involved an intermediate species in its evolutionary history. Furthermore, unfolding kinetics and thermodynamics of *c-kitG4* was different in QSB as compared to just in Milli-Q water. In QSB, *c-kitG4* showed conformational equilibrium between presumably equimolar levels of native (blunt end) and intermediate (dangling end) structure at room temperature; increase in temperature favoring the latter. In Milli-Q, *c-kitG4* showed completely different behavior with two distinct temperature induced transitions. Multivariate quantification of the different species during melting of *c-kitG1* and *c-kitG4* in different buffers was consistent with the hypothesis that the longest loop in *c-kitG1* has some form of ordered structure in potassium solution.

What drives the folding of G-quadruplex in Milli-Q water free of added central cations?

NH_4^+ ion is ubiquitous in DNA preparations as it is part and parcel of the standard deprotection procedure during oligonucleotide synthesis unless special effort is made to avoid it. Ammonium ion has been reported as an effective template to promote G-quadruplex formation. Janez et al. (2005) demonstrated localized binding of two ammonium ions between three G-quartet planes by solution NMR. Ammonium is also known to have very high affinity to DNA among monovalent cations showing non-electrostatic binding mechanism. Based on this, the results could be interpreted as the formation of a low stability G-quadruplex in which the monovalent internal ion that binds to specific site onto G-tetrads is NH_4^+. Obviously, the melting temperature was found to be highly reduced (ca. 48°C) due to lack of non-specific external ions which contribute to partial charge screening and the difference in dehydration energy of K^+ versus NH_4^+(Sket, Crnugelj et al. 2005).

The discussion above demonstrated both *c-kitG1* and *c-kitG4* have already been folded into G-quadruplex structure without any added cation presumably due to ubiquitous ammonium ions. The effect of gradual titration of K^+ ions into the solution of folded oligonucleotides on G-quadruplex structure was further

examined by circular dichroism spectroscopy. Striking resemblances in the appearance of CD spectra were observed with and without addition of potassium ions. Except for small difference in magnitude, no significant difference was observed in the positions of characteristic bands [figure 4-14].

It has been shown that G-quadruplex adopted by *c-kitG1* exhibited two cation binding sites between three of its G-quartets. Monovalent cations in general stabilized G-quadruplexes in the following order: K^+ > NH^+4 > Na^+. The introduction of tighter binding cation into solution (e.g. K^+) was generally expected to results in the replacement of weaker binding cation (e.g. NH^+_4). An intermediate in this type of monovalent cation interconversion process is expected to be mixed mono-K^+–mono-NH^+_4 form(Sket, Crnugelj et al. 2005).

As shown in figure 4-14, progressive addition of tighter binding K^+ ions resulted in very small increase in intensity of both positive and negative bands. The circular dichroism and absorbance data recorded in these titration experiments were well fitted when two components were considered for both *c-kitG1* and *c-kitG4*; which probably corresponded to di-ammonium and di-potassium forms without any distinct CD signal for the hybrid species.

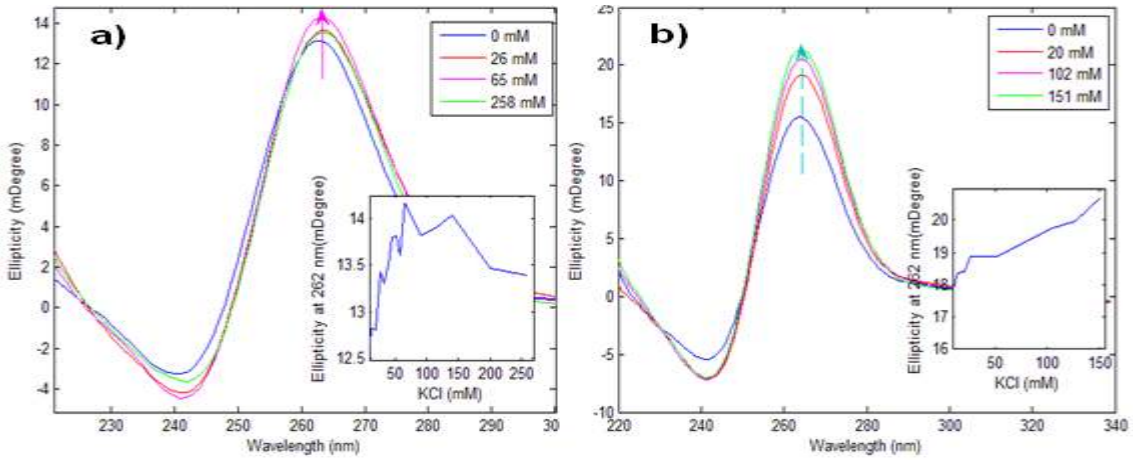


Figure 4-14 Experimental CD spectra for titration of *c-kitG1* (a) and *c-kitG4* (b) with KCl. Inset and arrows show change in positive ellipticity at 262 nm during titration.

Multivariate quantification of di-ammonium and di-potassium forms along the titration experiments confirmed the formation of a more stable G-quadruplex up on addition of potassium ions for both sequences. Satisfactory spectral differences amongst di-cation forms enabled their identification by CD spectroscopy and subsequent quantitative evaluation by multivariate data analysis. However, un-equivocal CD spectral assignment for an intermediate wasn't possible.

4.2. Conclusion

Melting curves for both sequences ($dCG_3CG_3CGCGAG_3AG_4$ and 5'-$dTG_3TG_3TGTGTG_3TG_4$) showed significant hyperchromic shift at 295 nm and CD spectral profiles exhibited characteristic signature spectra for authentic parallel G-quadruplexes. This is an evident for the formation of intramolecular G-quadruplex.

At pH 7.2 in QSB melting temperatures of *c-kitG1* and *c-kitG4* were 75 and 82°C respectively. Cytosine and adenine residues were found to exacerbate unfavorable electrostatic interaction in the backbone. The free-energy change at 310K varied from stable (4.2 kcal mol^{-1}) to very stable (7.4 kcal mol^{-1}) for the wild type and mutant sequence respectively in potassium solution.

Both *c-kitG1* and *c-kitG4* revealed the formation of parallel G-quadruplex in Milli-Q. Two ammonium ions from the DNA preparation were considered to bind specifically between three G-quartet planes. Addition of tighter binding K^+ ions replaced localized ammonium ions forming a more stable structure. Ethanol was also found to stabilize G-quadruplex structure.
Unfolding kinetics and thermodynamics of *c-kitG4* and *c-kitG1* were different in potassium solution but similar in ammonium solution. When *c-kitG1* was heated to 45°C in potassium solution, it produced an intermediate that had the same optical properties as *c-kitG4*. Melting of *c-kitG1* and *c-kitG4* in ammonium solution produced new CD signals at 298 nm and 270 nm suggesting the possibility of conformational switch involving antiparallel G-quadruplex species.

Multivariate quantification of different species showed that the longest loop in *c-kitG1* has some form of ordered structure in potassium solution but not in ammonium solution. Indeed, there is a propensity to form an internal stacking arrangement with in loop 2. G-quadruplex loops aren´t just like unstructured single strands but can have more ordered loop structure with elaborated ion-condensing property.

5. Determination of DNA-Drug Equilibrium Constants using UV-Vis and Circular Dichroism Spectroscopies

5.1. Results and Discussion

Stepwise addition of G-quadruplex solutions to cells containing NMM and TMPyP4 changed the wavelength and intensity of absorption band of the drugs. The Soret band of TMPyP4 at 422 nm changed after binding to G-quadruplex. As can be seen in the figure 5-1, G-quadruplex DNA addition caused red

shift and hyphochromicity of Soret band of both TMPyP4 and NMM but not Soret band of TPPS. Clearly, no molecular interaction was detected between TPPS and G-quadruplex forming sequences.

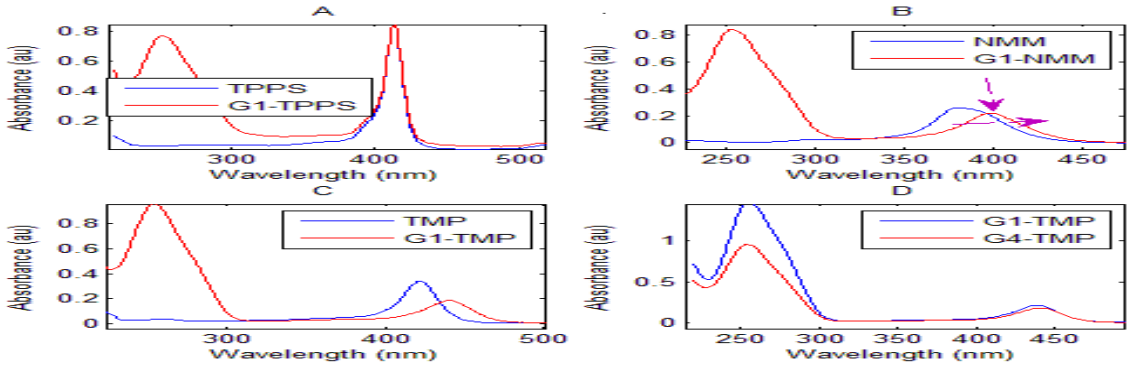


Figure 5-1 Absorption band of TPPS with and without DNA (A), Change in UV-Vis absorption band of TPPS (A), NMM (B) and TMPyP4 (C) upon addition of *C-kitG1*. Shift caused by adding *c-kitG1* and *c-kitG4* (D) on TMPyP4 absorption band.

5.1.1. Chemical Intuition

Inverse titration of TMPyP4 with c-*kitG1* revealed [cf. figure 5-2, a] continuous red shift in the absorbance spectrum of TMPyP4 upon addition of *c-kitG1* until drug to DNA mole ratio was one to one. Once the mole ratio reached 1:1, gradual addition of DNA didn't produce any more red shift. On the other hand, inverse titration TMPyP4 with *c-kitG4* [cf. figure 5-2, a] showed, once the mole ratio reached 2:1, gradual addition of DNA didn't produce any more red shift. In both cases after the specified mole ratios, TMPyP4 was considered to be the limiting reagent. A limiting reagent is a reactant of which there are fewer moles than required by stoichiometry. It determines the extent of chemical reaction. The mole ratio for TMPyP4 interaction was intuitively inferred to be 2:1 for *c-kitG4* and 1:1 for c-*kitG1*.

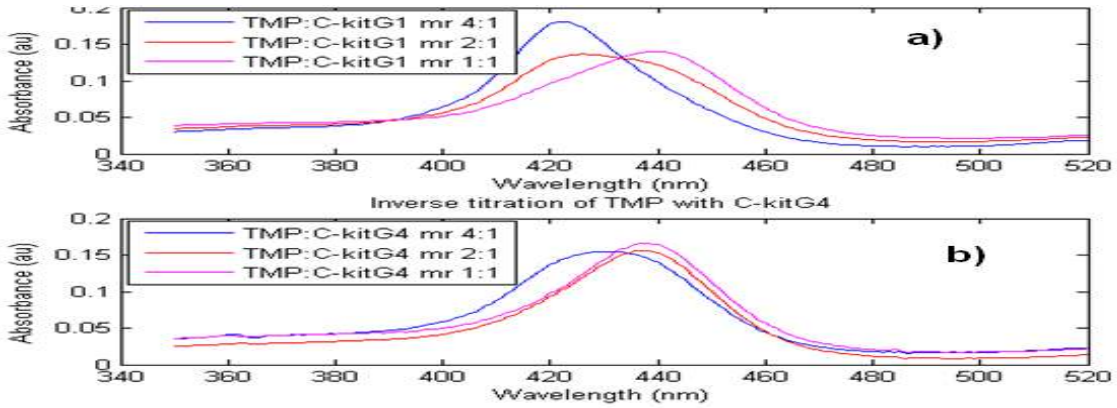


Figure 5-2 Inverse titrations of TMPyP4 with *c-kitG1* (a) and *c-kitG4* (b) in QSB pH 7.2 at 25°C. The corresponding molar ratios are given in the legend.

5.1.2. Joe-Jones Method

In Joe-Jones plot, the absorbance of each solution was measured and plotted against the mole ratio of the reactants. The linear portions of the resulting graph were extrapolated. The points of intersection of extrapolated lines were determined both graphically and mathematically. Graphically, the points at which the extrapolated lines intersect are shown in figure 5-3. Mathematically, the intersection of the lines was calculated by finding the point on x-axis where the two lines have the same y-value. This value was found to be 0.49 for *c-kitG4* and 0.99 for *c-kitG1*. The reciprocal of these intersection points suggested a possible combining ratio of 2:1 for *c-kitG4* and 1:1 for *c-kitG1*.

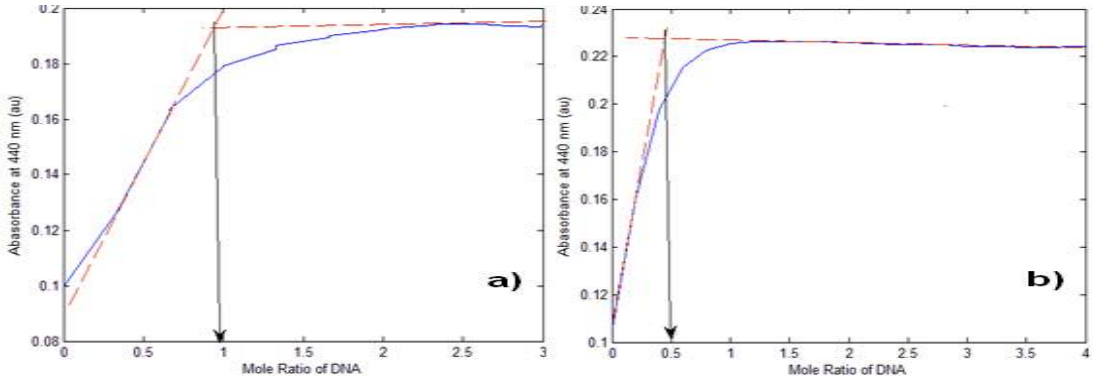


Figure 5-3 Determination of stoichiometry of interaction of *c-kitG1* (a) and *c-kitG4* (b) with TMPyP4 by Joe-Jones method at 25°C in QSB.

5.1.3. Computational Methods

Multivariate hard modeling computational program, EQUISPEC, was used to determine equilibrium constants from mole ratio experiments. Several different chemical models were tested and compared to find the model that best fits the experimental data within experimental error(Bosque-Sendra, Almansa et al. 2003). There was no single criterion with which to make the judgment. Generally the combination of standard errors, quality of fit, physical meaning of constants and distribution of residuals were considered. The combinations of these criteria for the interaction between *c-kitG1* and TMPyP4 are illustrated in figure 5-4. Experimental data fitted well when only three species were considered. Higher interaction models were found to be unstable having very high uncertainties. This could be due to minor species whose concentration was too low to accurately determine by the analytical methods used. The constant for a minor species may prove impossible to determine(Gans, Sabatini et al. 1996). The best model described the interaction of one mole of drug to one mole of *c-kitG1*. Similarly, the combinations of these criteria were used to decide between competing models for the interaction between *c-kitG4* and TMPyP4 [5-5] and *c-kitG4* and NMM [5-6] .Unlike the case of *c-kitG1*, experimental data from titration experiment for *c-kitG4* fitted well when four chemical species were considered. The best model described the interaction of

two moles of either drug to one mole of *c-kitG4* corresponding to four major species in the mixture. These findings are consistent with previous results based on chemical intuition and Joe-Jones plot.

It can be concluded that the combining ratio of porphyrin drugs to *c-kitG4* is 2: 1 and the combining ratio of porphyrin drugs to *c-kitG1* is 1:1 at the observed concentration ranges. This meant there were two distinct and strong binding sites for porphyrins on *c-kitG4* and only one strong binding site on *c-kitG1*. There could still be other binding sites with relatively weaker affinities than the identified ones. Weak affinity binding sites could be revealed at higher drug molar ratios than specified in this experiment or using more sensitive instruments. These binding sites are likely to manifest themselves in high concentration experiments such as NMR.

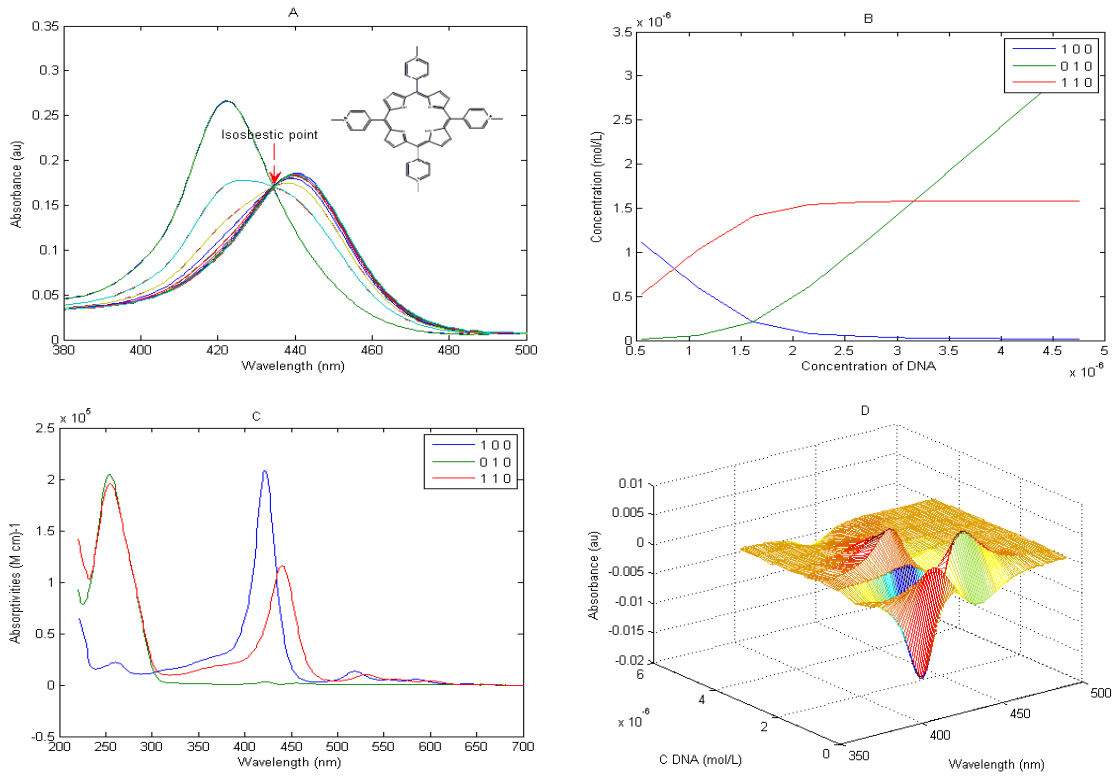


Figure 5-4 Mole-ratio study of the interaction between *c-kitG1* and TMPyP4 in QSB at 25°C. Drug to DNA ratio was progressively changed from 3.0 to 0.3. A) Experimental molecular absorption data. B) EQUISPEC resolved distribution diagram. C) EQUISPEC resolved pure molecular absorption spectra for each spectroscopically active species proposed in the model. D) Residuals profile not explained by the model.

Unsaturated complex (1:1) will be formed by *c-kitG4* when there is no enough ligand to form a saturated complex or when there is a large excess of DNA as in direct titration(Wang, Kumar et al. 1992). The predicted absorptivities of the saturated and unsaturated complex have the same profile.

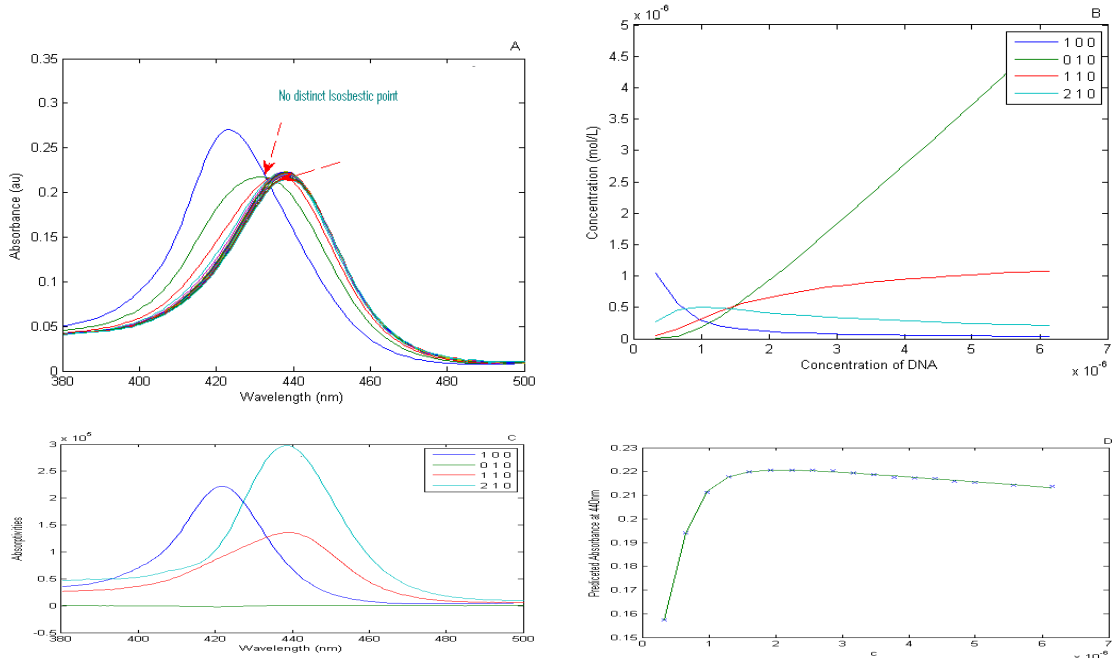


Figure 5-5 Mole-ratio study of the interaction between *c-kitG4* and TMPyP4 in QSB at 25°C. Drug to DNA ratio was progressively changed from 3.0 to 0.3. A) Experimental molecular absorption data. B) EQUISPEC resolved distribution diagram. C) EQUISPEC resolved pure molecular absorption spectra for each spectroscopically active species proposed in the model. D) Fit predicted by the model.

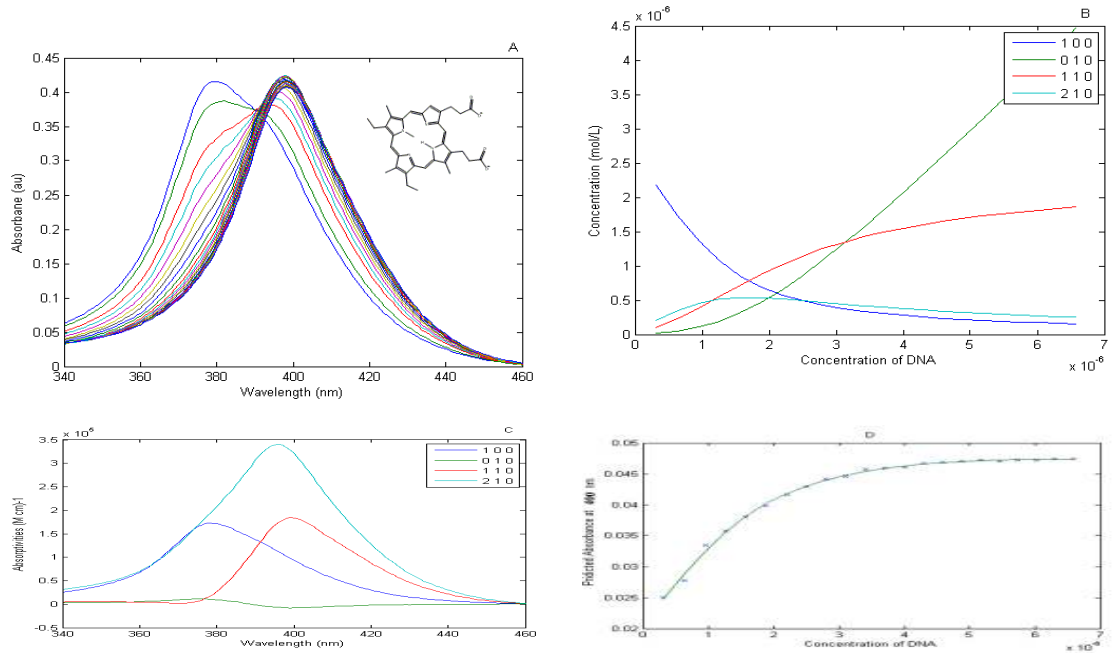


Figure 5-6 Mole-ratio study of the interaction between *c-kitG4* and NMM in QSB at 25°C. Drug to DNA ratio was progressively changed from 3.0 to 0.3. A) Experimental molecular absorption data. B) EQUISPEC resolved distribution diagram. C) EQUISPEC resolved pure molecular absorption spectra for each spectroscopically active species proposed in the model. D) Fit predicted by the model.

The distribution diagram shows the concentration of the saturated complex was very small. It is possible that the saturated and unsaturated complexes are spectroscopically indistinguishable. Equilibrium constants obtained from the best models for each DNA drug combinations are given in the table 5-1. In this work, the logarithm of the equilibrium constant (log K) calculated by the application of EQUISPEC computational program referred to the following equilibria:

$$C - KitG1 + TMPyP4 \leftrightharpoons C - kitG1 - TMPyP4\ complex$$

$$C - KitG4 + 2TMPyP4 \leftrightharpoons C - kitG4 - TMPyP4_2\ complex$$

Sequence	Drug	Log Beta1 (M^{-1})	Log Beta2 (M^{-2})	Residual error	Soret shift(nm)	Hypo. %
c-kitG1	NMM	6.65 ± 0.69	-	0.0069	20	1
c-kitG4	NMM	6.43 ± 0.17	12.37 ± 0.38	0.0089	20	2
c-kitG1	TMPyP4	7.31 ± 0.54	-	0.0526	18	46
c-kitG4	TMPyP4	6.79 ± 0.08	13.53 ± 0.10	0.0015	17	40

Table 5-1 Equilibrium constants for the interactions between intramolecular G-quadruplex forming sequences and water soluble porphyrins in pH 7.2 QSB at 25°C obtained from UV-Vis titrations.

The effect of G-quadruplex interacting agents on the G-tetrad was studied by circular dichroism spectroscopy. The purpose was to determine if any unusual induced CD signature was produced by the action of the drugs. TMPyP4 and NMM are not chiral molecules and therefore they don't have CD spectra of their own.

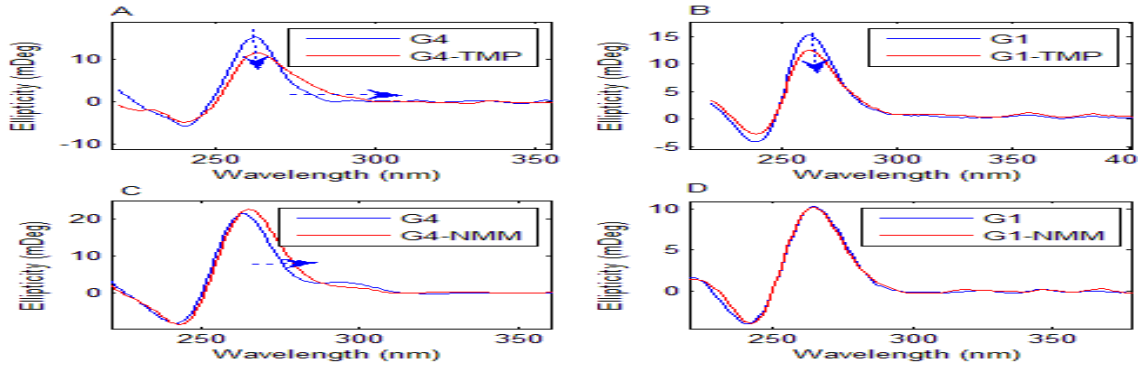


Figure 5-7 The effect of adding TMPyP4 and NMM on CD spectra of *c-kitG1* and *c-kitG4* quadruplexes in QSB at room temperature. Drug to DNA ratios were 3:1 in all cases

Both compounds didn´t produce significant induced signal as seen in figure 5-7. The CD spectra of both *c-kitG1* and *c-kitG4* DNA sequences showed significant reduction upon addition of TMPyP4. Such reduction in molar ellipticity could be either due to unfolding of the optically active structure (as seen during thermal denaturation) or the formation of a complex with altered chirality. Some G-quadruplex

interacting agents were shown to alter chirality of G-quadruplex by reducing existing CD signal and inducing new ones. TMPyP4 has been shown to bind to single strand, double strand and G-quadruplex DNA. The observed decrease in intensity could also be interpreted as the binding of TMPyP4 to unfolded structure more strongly than the folded G-quadruplex.

On the other hand, there was slight increase in the CD spectra of *c-kitG4* DNA and virtually no change in CD spectra of *c-kitG1* upon addition of NMM [figure 5-7, C and D]. This indicated NMM binding didn't alter molar ellipticity of G-quadruplex structure. The results are consistent with reports of NMM binding very weakly but selectively to G-quadruplex structures. CD titrations with *c-kitG1* and *c-kitG4* DNA were also performed with both porphyrins. Circular dichroism studies with NMM and TMPyP4 showed that the binding modes, affinity and selectivity of these compounds towards G-quadruplex structures were very different. This implied that binding modes and selectivity could be highly sensitive to the relative replacement of positive charges by negative charges on the parent porphyrin ring.

Equilibrium constants were estimated by simultaneous analysis of CD and absorbance data. The same quality control measures used in absorbance data analysis were applied in deciding among competing models based on the merged data. As an example, figure 5-8 shows different properties of a proposed one to one interaction model for *c-kitG1* and TMPyP4.

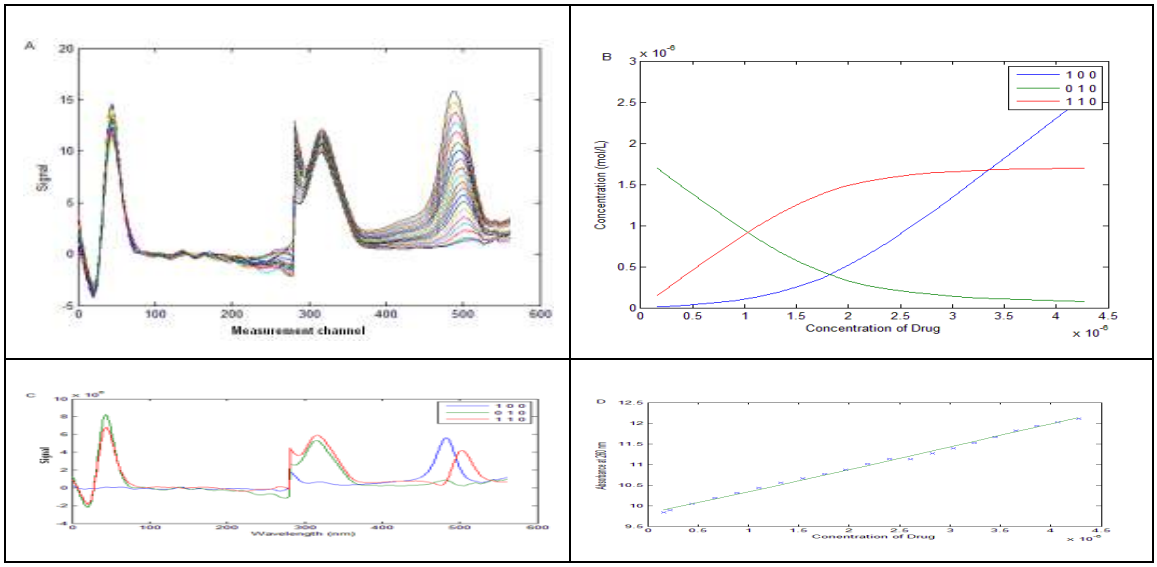


Figure 5-8 Titration of *c-kitG1* with TMPyP4. Drug to DNA ratio was progressively changed from 0.3 to 3.0 at room temperature. (A) Merged experimental CD and scaled absorbance data. (B) EQUISPEC-resolved concentration profiles of three optically active species present along the titration. (C) Resolved pure spectra. (D) Fit at 260 nm

Residual errors associated with CD measurements were generally higher than errors associated with absorbance measurement. This stems from inherently high noise level in CD signals. Direct titration in

circular dichroism experiment was mainly qualitative. It can be used to show changes in molar ellipticity and possible induced signals upon drug binding. This information was valuable to determine a possible binding mechanism.

Quantitative estimation of equilibrium constants from CD direct titration was not as informative as direct titration using molecular absorption spectroscopy. In CD experiments, most of the spectra obtained from direct titration were at low drug DNA mole ratios. As a result, for each drug DNA combination all the best models were with low combining stoichiometry. Satisfactory higher order models were obtained by selectively taking spectra from regions of higher mole ratios. Equilibrium constants determined by this procedure are given in table 5-2. Within the stated experimental errors, the values of equilibrium constants estimated by simultaneous analysis of merged data were essentially equivalent to values obtained molecular absorption studies. These values were further reproduced by Surface Plasmon Resonance studies.

Sequence	Drug	Log Beta1 (M^{-1})	Residual error
c-kitG1	NMM	6.46 ± 0.35	0.4185
c-kitG4	NMM	6.01 ± 0.38	1.105
c-kitG1	TMPyP4	6.99 ± 0.26	2.2663
c-kitG4	TMPyP4	7.25 ± 0.46	1.6254

Table 5-2 Equilibrium constants for the interactions between intramolecular G-quadruplex forming sequences and water soluble porphyrins in pH 7.2 QSB at 25°C obtained from simultaneous analysis of CD and absorption data.

The binding affinity of NMM didn't show significant difference between *c-kitG1* and *c-kitG4*. However, TMPyP4 binds to *c-kitG1* with slightly stronger affinity than *c-kitG4*. To obtain further insight, these interactions were monitored in SPR experiment. SPR experiments provided additional confirmatory information that TMPyP4 was bound more strongly to *c-kitG1* than *c-kitG4*. These and other results could be explained best by considering conformational isomerism of *c-kit G* quadruplex structures and effect of loop substitutions. *C-kitG1* quadruplex has more cytosine residues in the loop which has strongly electronegative oxygen. TMPyP4 is highly positive porphyrin. The interaction between these two oppositely charged partners could involve non-specific electrostatic mechanisms. In addition, both *c-kitG1* and *c-kitG4* can have additional end stacking mechanism of binding.

Why does *c-kitG4* show higher stoichiometry than *c-kitG1*?

C-kit G-quadruplex is structurally heterogeneous due to the four consecutive guanines at the 3' end, each of which could participate in G-tetrad formation. The presence of clear conformational isomers makes this

system extremely complex. A terminal unpaired single nucleotide is called dangling end. Two energetically favorable conformational isomers were proposed for *c-kitG1* based on the results of NMR. These are blunt-end G-tetrad conformer in which the last guanine forms G-tetrad and dangling-end G tetrad conformer in which the last guanine isn´t base paired. Two assumptions are required in order to explain higher stoichiometry of *c-kitG4* than *c-kitG1* in potassium solutions. a) *The dangling-end conformer has higher affinity to porphyrins than the blunt-end conformer*. b) *Loop substitution in c-kitG4 shifted the equilibrium between two conformers favoring the dangling-end conformer*. Several experimental observations of the present study fit with, and therefore support, these two assumptions. Some experimental evidences for these assumptions are given below. Of course, it remains to be seen how detailed molecular dynamic simulations, high resolution structural experiments and *ab intio* quantum calculations agree or disagree with the two assumptions.

The main evidence originated from the recently published NMR structure of *c-kitG1* and its variants. Solution state NMR structure suggested even though, dangling-end conformer of *c-kitG1* was kinetically more favourable, after prolonged incubation the dominant population was eventually the blunt-end conformer(Hsu, Varnai et al. 2009). On the top of that, dynamic coupling was confirmed between the G-tetrads and the long loop 2. After prolonged incubation, cross peaks observed in *c-kitG1* were absent in L3 substituted variants (similar to *c-kitG4*) suggesting blunt-end G-tetrad conformation wasn´t significantly populated in these loop variants. It can be concluded that *c-kitG1* was predominantly in the blunt-end conformer whereas *c-kitG4* was predominantly in dangling-end conformer after prolonged incubation.

Dangling ends are known to greatly increase the stability of RNA and DNA duplexes by staking interaction (Freier, Burger et al. 1983). The 3' terminal base of *c-kitG1* was shown to stack against G20-quadruplex core capping the terminal tetrad plane. Very recently flanking and capping nucleotides are shown to modulate conformational dynamics of human telomeric G-quadruplex structure(Bommarito, Peyret et al. 2000). There are reports in which a *c-myc* sequence forms an unknown mixture of conformers with at least two dominant populations(Freyer, Buscaglia et al. 2007). In this mixture the lower-melting conformer was postulated to have higher affinity for TMPyP4. It is not unlikely for the DNA-drug complex in the dangling-end conformer to be stabilized by interacting with 3' dangling base. Additional binding energies supplied by dangling end account for higher affinity of this end of G-quadruplex. Consequently, *c-kitG4* has more binding sites than *c-kitG1*.

The second X-ray structure of parallel-stranded intermolecular quadruplex in complex with self-assembled daunomycin trimer showed binding by stacking with the terminal G-tetrad(Clark, Pytel et al. 2003). In such an arrangement the trimer is able to optimize stacking interactions with the G-tetrad allowing the daunosamine sugars to dangle in the grooves and form stabilizing hydrogen-bonding interactions.

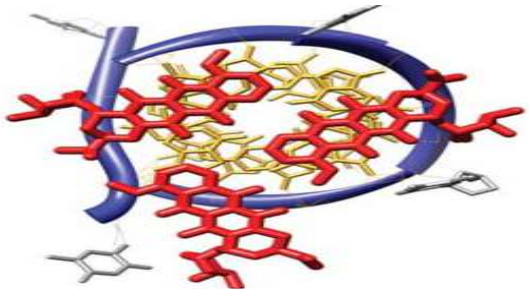


Figure 5-9 Top view of the X-ray structure of the daunomycin trimer with tetramolecular quadruplex-DNA (d[TG4T])4 (PDB entry: 1O0 K). Daunomycin dangles to form hydrogen bonding interaction.

An interesting implication of this binding mechanism is that it can be applied as a probe for drug binding. The reasoning behind this application is that a drug will exhibit a differential binding affinity for dangling-end due to possible additional interactions with the dangling nucleotide. For intramolecular G-quadruplex structures which might be formed in small segments of genomic DNA, even the full-end conformer will practically have dangling nucleotide. This suggests stronger binding affinity *in vivo* than observed *in vitro*. The dangling-end conformer has been shown to be kinetically favored, in-line with continuous assembly and disassembly at different stages of cell cycle.

5.2. Conclusion

The combining ratio of porphyrin drugs to *c-kitG4* was found to be 2:1 and the combining ratio of porphyrin drugs to *c-kitG1* was 1:1 over the ranges of concentration studied (0.3 to 3.0 drug to DNA mole ratios). Both NMM and TMPyP4 didn´t produce significant induced CD signal. The affinity of NMM was not significantly different for *c-kitG1* and *c-kitG4*. TMPyP4 binds to *c-kitG1* with slightly stronger affinity (log beta=7.31$\pm$0.5) than *c-kitG4* (log beta=6.79$\pm$0.1). The CD spectra of both *c-kitG1* and *c-kitG4* DNA sequences decreased upon addition of TMPyP4. The binding modes, affinity and selectivity of these compounds towards G-quadruplex structures were very different. It is proposed that the dangling-end conformer has higher affinity to porphyrins than the blunt-end conformer. Loop substitution in *c-kitG4* shifted the equilibrium between two conformers favouring the dangling-end conformer.

6. Probing affinity and kinetics of binding interactions between porphyrins and *c-kit* G-quadruplex with Surface Plasmon Resonance

The recent development of biosensors, based on Surface Plasmon Resonance (SPR) technology, enables monitoring of a variety of biospecific interactions in real-time. SPR based optical sensor is one of the most promising analytical tools for rapid and real-time analysis of chemical and biological substances in various

analytical matrices with enhanced sensitivity and selectivity. SPR is a high-throughput, label-free spectroscopic technique to monitor biomolecular interactions in real time (Halder and Chowdhury 2005). It has several advantages such as high sensitivity and selectivity in analyzing complex samples, simple procedure, versatility and compatibility with various surface modifications, portable instrumentation, real-time and high-throughput screening.

Getting reliable results from SPR experiments requires careful design and optimization of different variables. Identifying proper regeneration conditions i.e., entirely removing bound analyte molecules is a trial and error process that requires caution. The activity of the ligand should be preserved via out the experimental runs. Unfortunately, no general rule can be drawn in that endeavour. Heterogeneous ligand populations (resulting from partial refolding during purification and possible aggregation) are more likely to lead to heterogeneous binding kinetics for which interpretation will be difficult. As a prerequisite, the quality of the ligand must be high or at least evaluated prior to any SPR experiments. The concentration of analytes should be determined very accurately(Rebecca and David; Homola 2008).

In this study TMPyP4, NMM and TPPS were evaluated against a 21-mer wild type *c-kit* promoter sequence *c-kitG1* and a mutated sequence *c-kitG4* by using SPR. The purpose of the procedure was to determine binding stoichiometries, affinities and kinetics in a specified range of analyte concentration.

6.1. Result and Discussion

There was high immobilized ligand density on the sensor surface as inferred from high observed responses during immobilization. Immobilization efficiency was very high. The amount of *c-kitG4* (1600 RU) immobilized on the sensor surface was higher than that of *c-kitG1* (1200 RU). At least five different concentrations of each analytes were used for steady-state and kinetic measurements before fitting response curves to respective equations to extract thermodynamic and kinetic parameters. The instrument response (RU) in the steady-state region is proportional to the amount of bound drugs and was typically determined by linear averaging over 10-20 seconds depending on the length of the steady-state plateau. The predicted maximum response per bound compound in the steady-state region (RU_{max}) was determined theoretically from the DNA molecular weight, the amount of DNA on the flow cell, the drug molecular weight, and the refractive index gradient ratio of the drug and DNA. This was empirically tested by injecting high drug concentrations.

6.1.1. Steady-State Analysis

BIAcore T100 evalutation software supplied by the manufacturer was used to analyze the binding data. All plots of sensorgrams, report point and evaluation of kinetic and affinity measurements were done using this software. The region 15-20 s before injection stops was taken as the steady- state region and a window width of 5 s was considered. The RU_{Max} value required to convert the observed response (RU) to the standard binding parameter r (moles of drug bound per moles of DNA immobilized) was determined empirically using high drug concentrations. The number of binding sites was estimated fitting plots of RU versus Concentration [figure 6-1].

$$r = \frac{RU}{RU_{Max}}$$

RU is relative response and RU_{Max} is empirically determined maximum response. In SPR experiments, binding stoichiometry arises naturally from the response as saturation of binding sites is approached.

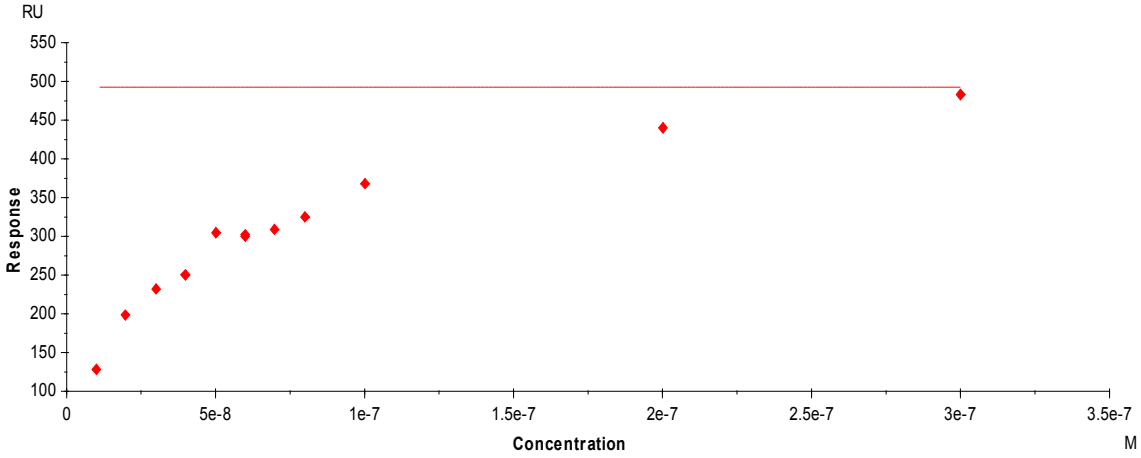


Figure 6-1 Maximum response for a single binding site on *c-kitG1* and observed responses at different concentrations of TMPyP4 in HBS-EP buffer containing 147 mM KCl at 25°C.

The maximum response per bound compound in this steady-state region (RUmax) was determined to be 493.4 RU for the interaction between *c-kitG1* and TMPyP4. In all the cases, the observed RU values at each concentration were less than the predicted RUmax, pointing binding to only one binding site for the drug concentrations considered. Higher combining ratios were expected to give higher response than predicted. It can be concluded both *c-kitG1* and *c-kitG4* interact with the drugs (NMM and TMPyP4) in one to one stoichiometry whereas there was no detectable interaction with TPPS in the concentration ranges considered and experimental conditions discussed above. The binding plots for the concentration dependent binding of the NMM on the *c-kitG1* and *c-kitG4* sequences are shown in Figure 6-2a and 6-2b,

respectively. Direct steady-state binding plots were constructed by averaging the observed SPR responses in the steady-state region. These responses were plotted against the free NMM concentration in the flow solution [figure 6-3]. To obtain the affinity constants, the steady-state data were then fitted to the following single dominant binding site interaction model using the BIAcore T100 evalutation software:

$$R = Conc * Rmax/(Conc + KD) + offset$$

Where KD is equilibrium constant and Conc is the concentration of the drugs in the flow solution. From the model, equilibrium affinity constant and its standard error were extracted. Errors in fitting KD values within the experiments were ±10%.

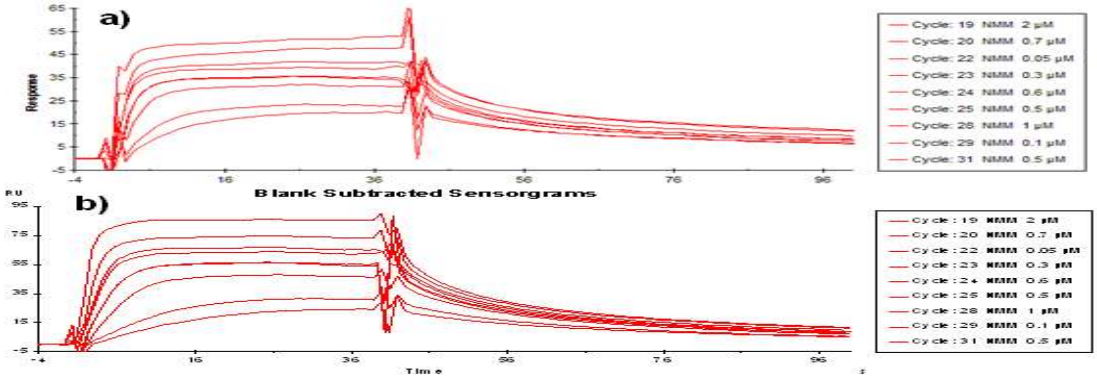
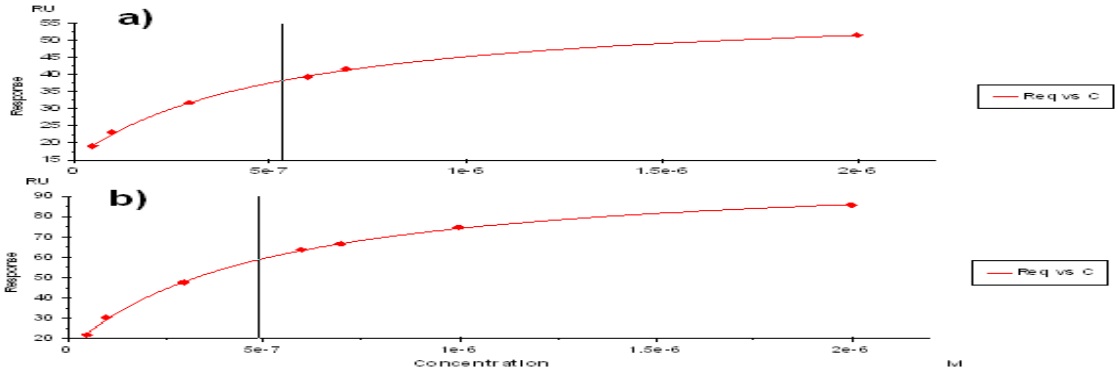


Figure 6-2 SPR sensorgrams for steady-state affinity measurement of the interaction of NMM to immobilized G-quadruplex formed by a 21 mer *c-kitG1* (a) *c-kitG4* (b) sequences in HBS-EP buffer containing 147 mM KCl at 25°C. One concentration was duplicated as shown in the legend. Different concentrations were injected randomly to minimize systematic error.



6-3 Plot of steady-state binding level as a function of concentration for affinity measurement of the interaction between (a) *c-kitG1* (b) *c-kitG4* with NMM at 25°C in HBS-EP buffer containing 147 mM KCl. Fitting. The response points were fit with a one to one binding model. The blank line marks an acceptable affinity constant.

On the other hand, the plot of RU versus the free concentration of TMPyP4 didn´t fit well to the one to one steady-state model [table 6-1]. Higher models were generally more unacceptable (high chi squared value) than dominant binding site model. TMPyP4 was poorly described by the proposed one to one binding model. Global fit on the basis of the analysis of residuals and quality of fit values for the interaction between TMPyP4 and both *c-kitG1* and *c-kitG4* over the concentration range considered was unsatisfactory [table 6-1]. The observed 'terminal depression' before injection stops was also more pronounced in the case of TMPyP4 than NMM [figure 6-4].

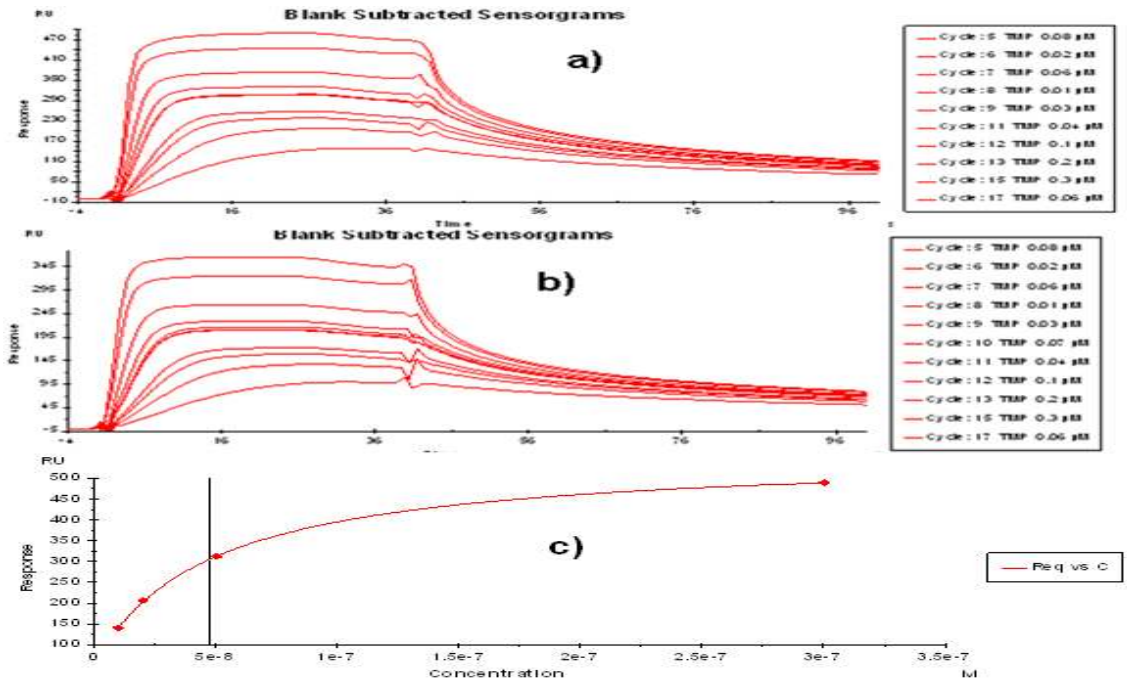


Figure 6-4 SPR sensorgrams for steady-state affinity measurement of the interaction of TMPyP4 to immobilized G-quadruplex formed by a 21 mer a) *c-kitG1* b) *c-kitG4* sequences. The lowest curve corresponds to the lowest concentration and one concentration was duplicated. Different concentrations were injected randomly. c) Plot of steady- state binding level as a function of concentration for steady-state affinity measurement of the interaction between ckitG1 and TMPyP4 in HBS-EP buffer containing 147 mM KCl at 25°C.

TMPyP4 tends to give unstable fitting behaviour .Offsetting the stability of fitting might be due to the presence of binding sites that vary widely in affinity and selectivity for TMPyP4. There was no clear SPR response due to TPPS binding. Blank subtraction gave first positive signal followed by negative signal indicating the movement of the analyte away from the sensor surface as flow progressed. Reference cell subtraction gave completely negative response showing non-specific binding at the reference flow cell was higher than any form of binding at the active flow cell.

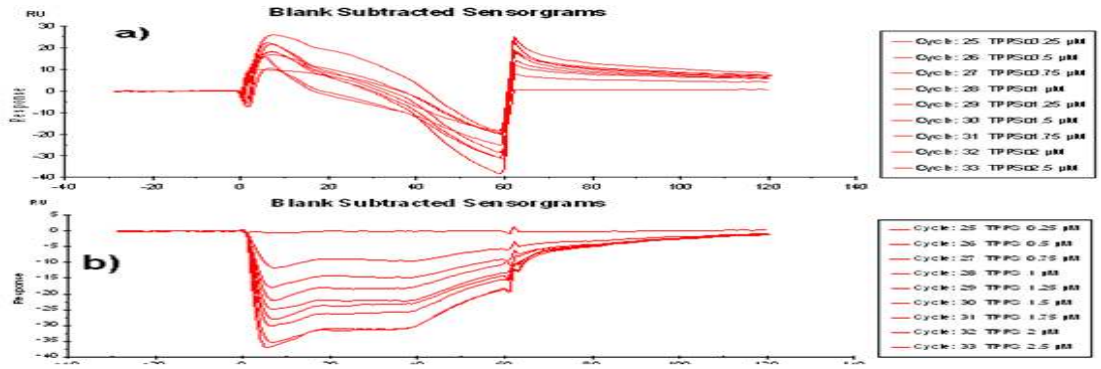


Figure 6-5 Blank subtracted sensorgrams for steady- state affinity measurement of the interaction between *c-kitG1* and TPPS at 25°C. a) The response from the reference cell hasn´t been subtracted. b) The response from the reference cell has been subtracted.

The equilibrium binding constants (Log beta values) are shown for each drug in table 6-1. The binding affinity of NMM to *c-kitG1* was not significantly different from that of *c-kitG4*. Sequence variation on the loops of G-quadruplex structure has no significant effect on the binding affinity of NMM. Rough estimates of binding parameters were obtained for TMPyP4. It exhibited strong binding affinity with each sequence. *C-kitG4* approximated one to one binding with TMPyP4 more closely than c-*kitG1* having chi-squared values 3.7 and 9.4 respectively. There was no significant difference between binding affinities of *c-kitG1* and *c-kitG4* with TMPyP4 according to SPR results. NMM exhibited weaker binding with KD values nearly ten times lower than that of TMPyP4. These results are in agreement with the previous results obtained by using UV-vis spectroscopy and CD-spectroscopy.

Sequence	Drug	Log Beta (M^{-1})	Standard Error	Chi² (RU^2)	df	α	Chi² Crit
c-kitG1	NMM	6.27	0.029	0.15	4	0.05	0.71
c-kitG4	NMM	6.31	0.028	0.51	4	0.05	0.71
c-kitG1	TMPyP4	7.32	0.026	9.39	1	0.05	0.004
c-kitG4	TMPyP4	7.19	0.022	3.74	1	0.05	0.004

Table 6-1: Parameters of steady-state affinity measurements for the binding of porphyrins to *c-kitG1* and *c-kitG4* in HBS-EP buffer containing 147 mM KCl at 25°C

6.1.2. Kinetics Analysis

In this analysis, heterogeneous ligand binding and two state conformational change models gave the best fit to the observed sensorgrams. To reduce heterogeneity due to heterogeneous immobilization, additional experimental effort was done by immobilizing one of the DNA sequence with a five thymine linker

sequence. However, even with a flexible attachment using a linker arm, heterogeneous binding and two state conformational change models described the experimental sensorgrams very well. Fits obtained by heterogeneous binding kinetics are given in figure 6-7.

Moreover, kinetic experiments were conducted at different injection time and flow rate combinations. In all the combinations studied, there was a small decrease in the response plateau near the end of injection time. This 'terminal depression' was found to be concentration dependent. Based on the knowledge of the system from characterization studies(Rankin, Reszka et al. 2005), conformational heterogeneity is known to be inherent property of the immobilized ligand. As described earlier G-quadruplex DNA structures formed by *c-kitG1* showed conformational heterogeneity with the ability to interconvert between these structural states. Consequently, the observed kinetics was not mainly due to heterogeneous immobilization. Particularly, the effect of dangling end and blunt end conformational isomers (figure 4.1) is considered to play important role in the kinetics.

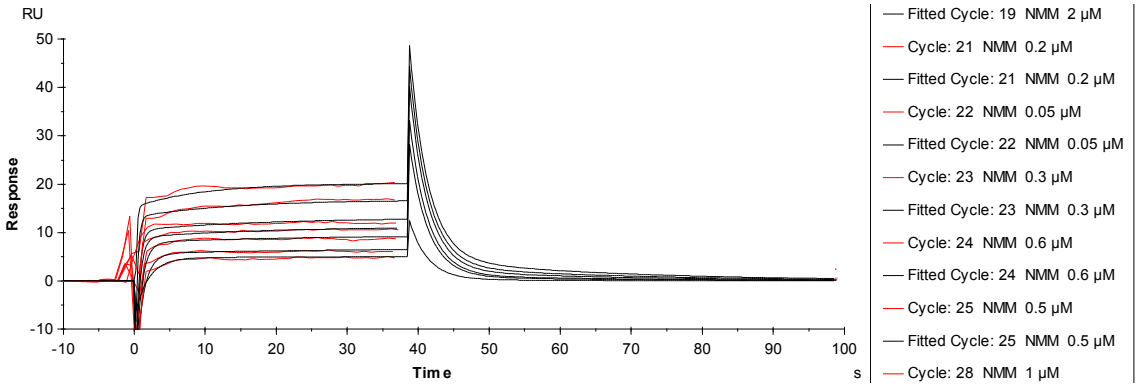


Figure 6-6 Heterogeneous binding kinetic measurement for the interaction between *c-kitG1* with 5T linker sequence and NMM in HBS-EP buffer containing 147 mM KCl at 25°C.

Kinetic parameters extracted from two states conformational change model showed better agreement to steady-state affinity analysis and previous findings than heterogeneous binding model. Binding constants obtained from heterogeneous binding model lacked biological significance and physical reality. Due to the fact that, G-quadruplex structures are polymorphic and show conformational isomerism, two conformational isomers are thought to be the cause of the observed heterogeneity. It is believed that describing the sensorgrams using two states conformational model is very reasonable.

The plot of RU versus real time after injection of NMM was fitted to a two state conformational change model shown below. Two sensorgrams were selected for curve fitting. Kinetic plots of observed (red) and fitted (black) curves for the association and dissociation of NMM on the *c-kitG1* sequence are depicted in figure 6-7. Kinetic parameters (ka and kd values) are given in table 6-2 for NMM.

$$A + B \rightleftharpoons AB \rightleftharpoons AB^*$$

A represents the analyte drug and B denotes the immobilized G-quadruplex DNA ligand.

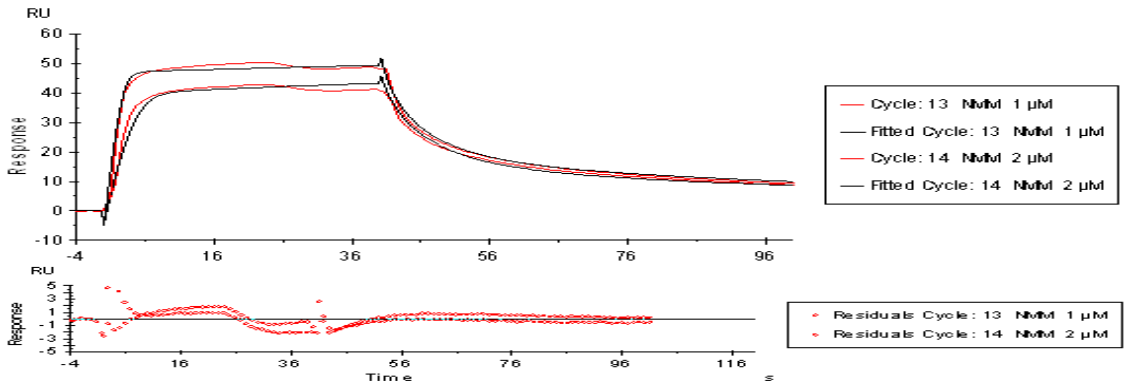


Figure 6-7 Two state conformational change kinetic measurement for the interaction between *c-kitG1* and NMM in HBS-EP buffer containing 147 mM KCl at 25°C. Fitting errors due to random point scatter in any experiment were less than ±25%.

The association rate constant ka1 [= $2.15x10^{+06}$ $M^{-1}s^{-1}$] obtained from the kinetics analysis of the sensorgrams showed fast specific binding of NMM with the *c-kit* G1 and *c-kitG4* sequences. The dissociation rate constant kd1 [=0.98 s^{-1}] further showed specific binding of NMM. The second association rate constant was sufficiently small that the possible conformational change of the bound complex can be considered kinetically slow (80-90 seconds). This could be a possible reason for the observed drop of response at the end of contact times. This agreed with reported very slow interconversion of these conformational subsets in NMR timescale. The steady-state affinity value calculated for the binding of NMM with G-quadruplex is on the same order as that of kinetic values.

Sequence	Drug	ka1 (1/Ms)	kd1 (1/s)	ka2 (1/s)	kd2 (1/s)	Log Beta (M)	Chi^2 (RU^2)	df	α	Chi^2 Crit
c-kitG1	NMM	$2.15\ x10^{+06}$	0.9769	$9.01\ x10^{-3}$	0.01025	6.62	1.23	20	0.05	10.85
c-kitG4	NMM	$1.54\ x10^{+05}$	0.1045	$6.65\ x10^{-3}$	0.009609	6.40	7.12	20	0.05	10.85

Table 6-2 Parameters of kinetic measurement for the binding of NMM to *c-kitG1* and *c-kitG4* in HBS-EP buffer containing 147 mM KCl at 25°C

For binding of TMPyP4 to G-quadruplex forming sequences, the kinetic constants of association and dissociation aren't reproducible for accurate analysis by SPR methods. The binding of TMPyP4 with the

G-quadruplex DNA sequences showed much faster kinetics than binding with NMM. The very high response of TMPyP4 indicated multisite binding at higher concentrations. The very fast dissociation, the very high observed response and difficulty of both steady-state and kinetic modelling of TMPyP4 interactions implied a nonspecific, heterogeneous binding of TMPyP4 to the G-quadruplex unlike that of NMM. SPR demonstrated that there could be heterogeneous binding mechanisms of TMPyP4 that can't be accurately described by the aforementioned steady-state and kinetic models.

TMPyP4 may bind non-specifically more to the *c-kit* G1 structures than *c-kitG4* as the former has very electronegative oxygen on cytosine residues. It isn't a structure selective ligand. Binding sites heterogeneity should be taken into account in designing a model to describe the interaction of non-selective analytes. There should be some measure of heterogeneity in the form of binding site affinity distributions and the heterogeneity index to yield more accurate measures of the binding properties of non-specific analytes. Heterogeneous binding models should be considered for better understanding of the interaction, biological effect and sensor behavior of TMPyP4. NMM is more structurally selective than TMPyP4. Heterogeneous binding (random) binding can't be predicted by simple models. The effect of heterogeneous binding is more pronounced at higher concentrations than at lower concentrations. Consequently, lower concentrations of TMPyP4 give relatively stable fit.

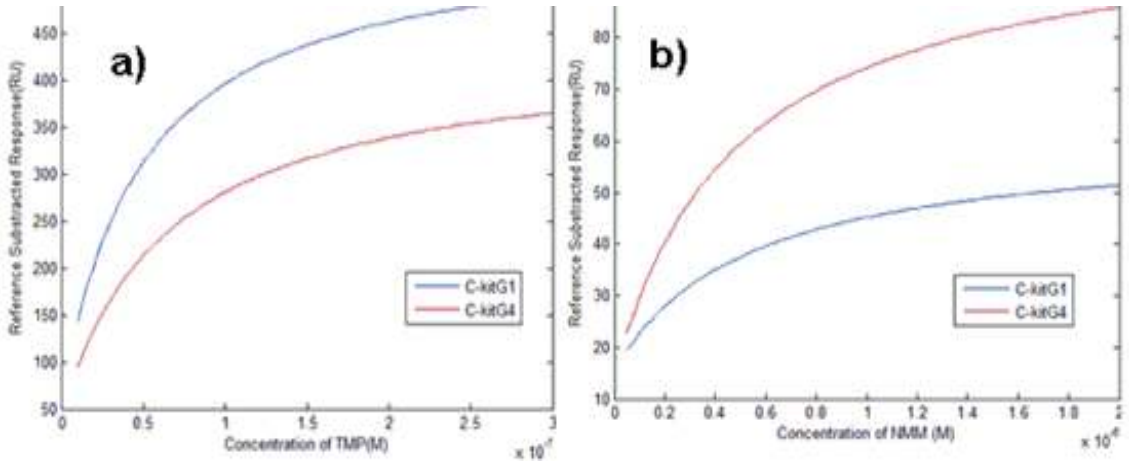


Figure 6-8: Plot of steady-state binding level as a function of concentration for steady- state affinity measurement of the interaction between G-quadruplex and NMM (a) and G-quadruplex and TMPyP4 (b) in HBS-EP buffer containing 147 mM KCl at 25°C.

The differences in the dissociation of NMM from that of the TMPyP4 can be visualized from their corresponding sensorgrams [figures 6-4 and 6-2]. NMM showed relatively stable steady- state region than TMPyP4. The magnitudes of observed responses were also shown to be different for the two different drugs. Generally, it was observed that the response produced by NMM and *c-kitG4* was higher than that of

NMM and *c-kitG1*. On the contrary, the response given by TMPyP4 and *c-kitG1* was higher than that of TMPyP4 and *c-kitG4*. *C-kitG4* gave higher response than c-*kitG1* during ligand immobilization. This could explain the higher response of NMM interaction with *c-kitG4*. On the other hand, TMPyP4 gave higher response with c-*kitG1*. The amount of TMPyP4 bound to the two G-quadruplex forming sequences is significantly different. Sequence mutation has significant effect on the nature of binding of TMPyP4 to G-quadruplex structures.

6.2. Conclusion

Surface Plasmon Resonance was used to determine the binding mechanism, binding affinity and kinetics of three different porphyrins to human c-*kitG1* and mutated *c-kitG4* quadruplex forming sequences. TMPyP4 is cationic porphyrin, whereas TPPS and NMM are anionic porphyrins. TMPyP4, was identified as a drug with the strongest binding (log beta=7.3) whereas TPPS has no detectable binding to both sequences.

The results from SPR experiments showed that NMM binding to both G-quadruplex structures is homogeneous and can be easily described by one to one interaction model whereas TMPyP4 binding is heterogeneous and difficult to model. In homogeneous binding, all binding sites have the same selectivity and sensitivity as opposed to heterogeneous binding in which different binding sites exhibit varying affinity and selectivity. Homogeneous binding of NMM could be the result of binding to a specific structural cavity with defined geometric, physical and electronic property i.e. G-tetrads. On the contrary, TMPyP4 could bind to the G-tetrads as well as other binding cavities in non-structure selective manner. SPR not only confirmed a relatively strong binding (log beta=6.3) of NMM to *c-kit* G-quadruplex structures but also provided solid quantitative information regarding structure specific binding of NMM over heterogeneous binding of TMPyP4. Kinetic modeling using SPR showed the binding of a drug with the G-quadruplex structure is accompanied by a slow conformational change. The interaction kinetics is inferred to be

$$A + B \rightleftharpoons AB \rightleftharpoons AB^* \rightleftharpoons A + B^*$$

Where A is analyte, B is the immobilized ligand and B* is the induced conformer upon drug binding. Such type of conformational change could be one reason for the observed slight response declines at the end of contact times in all sensorgrams. The altered conformer has reduced affinity to both drugs. This indicates that conformational changes upon drug binding should be investigated further using FRET or NMR spectroscopy.

Subsequent replacement of A and C residues in the wild type *c-kitG1* with T residues in the mutan *c-kitG4* sequence has no significant effect on the affinity of NMM. The same base substitution decreased the affinity of TMPyP4. One possible reason is non-specific binding of TMPyP4 at the loops of *c-kitG1* but not at the loops of *c-kitG4*. Heterogeneous binding is believed to be the outcome of electrostatic interaction.

On the basis of these results, NMM is recommended for further study using biomedical techniques. Additional experiments are required to compare the selectivity of NMM with the *c-kit* G-quadruplex versus other established G-quadruplex structures.

7. Spectroscopic Characterization of the Interaction Complex formed by G -quadruplex DNA and Water Soluble Porphyrins

Melting experiments are crucial to understand potential thermodynamic contributions from binding-induced changes in solvation and conformation. This information is a prerequisite for subsequent dissection of binding profiles into contributions from general solvent effects or from specific drug-DNA interactions. Characterizing enthalpy and entropy changes avoids enthalpy-entropy compensations on binding free energies(Breslauer, Remeta et al. 1987). It helps in making the distinction between entropy driven and enthalpy driven binding processes. Binding profiles due to conformational change of G-quadruplex are best detected by their binding enthalpies.

Structural studies on DNA-drug interactions should be complemented by thermodynamic studies. Multivariate data analysis has allowed detection of very weak absorbance and small intrinsic ellipticity changes (macroscopic properties) during melting. The associated structural rearrangements (microscopic properties) during melting can be verified using structural pictures derived from NMR and x-ray studies. Taken together, parallel structural and thermodynamic studies can yield insights into the origins of drug binding, affinities and specificities that neither study alone could provide.

The stability of duplex DNA is known to decrease at very high and very low pH. At low pH, bases in the single strand bind more protons than bases in the duplex favoring denaturation. At high pH, both guanine and thymine can be deprotanted precluding base pairing and increasing charge repulsion. On the contrary, some sequences are known to form double helix only at low pH. This results from hydrogen bonding scheme in which a proton is shared between two bases such as $A.A^+$ and $C.C^+$. Protonated bases don't have to be adjacent as poly (dCT) was shown to form duplex with $C.C^+$ bonds. Little is known about the pH dependence of the interaction of porphyrins with G-quadruplex structures suggesting an interesting

area for further study. G-quadruplex DNA structures themselves were found to be slightly more stable at low pH(Pavel, Joaquim et al. 2009). This was interpreted to be due to the formation of additional hydrogen bonds by protonated bases and reduction of repulsive interactions in the phosphate backbone.

A combination of molecular absorption and circular dichroism spectroscopic techniques were used to detect the influence of DNA binding porphyrins on G-quadruplex properties. The purpose of acid-base titration was to investigate if G-quadruplex structures interact with water soluble porphyrins at very low pH.

7.1. Results and Discussion

In this study, the effect of systematic loop variation of a G-quadruplex forming DNA on its interaction with porphyrin drugs was assessed. Thermal stabilities of different complexes were monitored using CD and UV-Vis spectroscopy by following changes in the spectrum with increasing temperature. Figure 7-1 illustrates a qualitative analysis of the interaction between *c-kitG1* and NMM at two extreme temperatures used in this experiment. At 25°C, the Soret band of NMM was shifted to higher wavelength upon addition of DNA. This effect was observed but diminished at 90°C. As described earlier, NMM can undergo dimerization even at micro molar concentrations. At low temperature DNA drug interaction equilibrium competes with that of dimerization. At high temperature the monomer is practically the most prominent population.

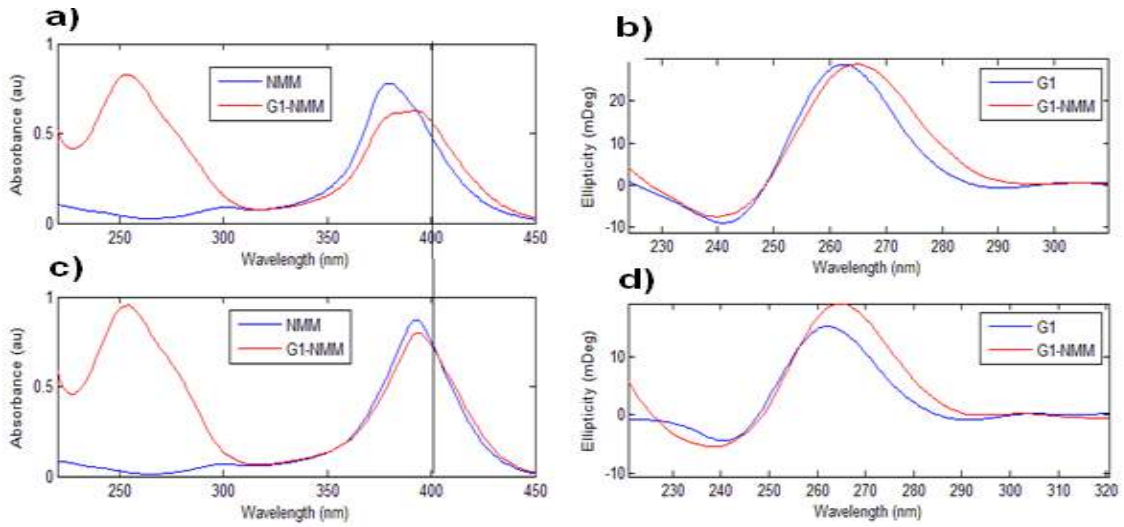


Figure 7-1 a) UV-Vis spectra of free NMM and NMM *c-kitG1* mixture in QSB at 25°C. (b) CD spectra of free c-*kitG1* and NMM *c-kitG1* mixture in QSB at 25°C. (c) UV-Vis spectra of free NMM and NMM c-*kitG1* mixture in QSB at 90°C. d) CD spectra of free *c-kitG1* and NMM *c-kitG1* mixture in QSB at 90°C. Mole ratios in the mixtures are 0.6 and 3.4 for UV-vis and CD respectively.

The corresponding CD spectra for free *c-kitG1* and the complex are also shown in the same figure 7-1 [b and d]. G-quadruplex structures formed by *c-kitG1* and its variant *c-kitG4* were found to be relatively more stable than other G-quadruplex forming sequences. As shown in CD spectrum figure 7-1 (d), there was substantial amount of folded G-quadruplex at 90°C with or without porphyrin drugs. UV-Vis spectrum of NMM was shifted to longer wavelength at high temperature.

One possible explanation for the observed slight shift in the spectrum of NMM at 90°C is interaction of NMM with the unfolded strand. The magnitude of both hypochromic shift and red shift at elevated temperature was sufficiently small that the effect can be explained solely by the presence of stabilized folded DNA. This was shown to be true, as the fraction of folded conformation was found to be significantly enhanced by the presence of NMM in CD measurements (Fig. 1, d). It can be concluded that NMM binds preferentially to G-quadruplex structures. Consequently, the observed elevated melting temperature of the *c-kitG1* NMM complex could be explained due to selective stabilization of G-quadruplex.

Likewise, melting data was collected for the mixture of *c-kitG1* and TMPyP4. TMPyP4 is a pivotal example of G-quadruplex interacting agents. Several literatures accumulated over years of research showed TMPyP4 stabilizes G-quadruplex structures. Nevertheless, this is probably not true for *c-kitG1* quadruplexes. The argument is as follows.

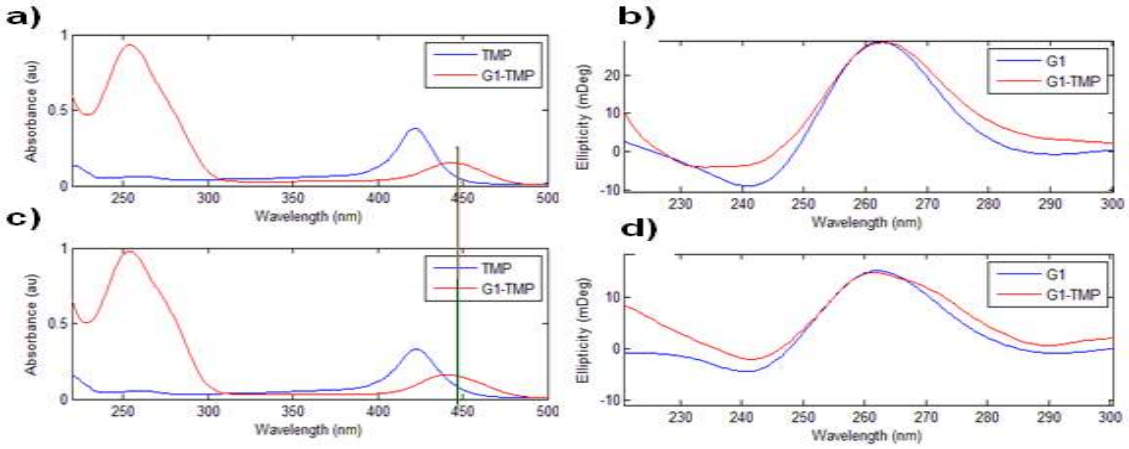


Figure 7-2 a) UV-Vis spectra of free TMP and TMP *c-kitG1* mixture in QSB at 25°C. b) CD spectra of free *c-kitG1* and TMP *c-kitG1* mixture in QSB at 25°C. (c) UV-Vis spectra of TMP and TMP *c-kitG1* mixture in QSB at 90°C. d) CD spectra of free *c-kitG1* and TMP *c-kitG1* mixture in QSB at 90°C. Mole ratios in the mixtures are 0.3 and 2.5 for UV-vis and CD respectively.

Figure 7-2 shows the effect of *c-kitG1* on the Soret band of TMPyP4 and the effect of TMPyP4 on the CD spectra of G-quadruplex. A glance at the UV-Vis spectra obtained at 25°C and 90°C could reveal

temperature had little effect on the interaction between TMPyP4 and *c-kitG1*. Molecular absorption results showed both high hypochomic shift and red shift which characterized binding interaction at the highest temperature used (90°C).

As explained earlier, this could mean two things. a) It could be due to very stable complex in which the porphyrin is still bound to the folded structure. b) It could be due to a stable complex in which the porphyrin is bound to the unfolded structure. Both folded and unfolded structures could potentially bind to TMPyP4. Subplots b and d in figure 7-2 show, CD signals at elevated temperature weren't improved by the presence of TMPyP4. It was actually observed, CD spectral signals diminished upon progressive addition of TMPyP4 during mole ratio experiments. Again, this could mean two things. a) The formation of a stable complex in which the porphyrin is still bound to the regular G-tetrad structure the asymmetry of which has been altered in the complex b) The formation of a stable complex in which the porphyrin is bound to irregular structure which has no structural asymmetry. Additional CD experiments in which the mole ratio of TMPyP4 to DNA was 4 to 1 showed no sign of induced CD signal. TMPyP4 has absorbance bands in the visible region of the spectrum where new CD signals could be induced; any changes in the CD spectrum of the G-quadruplex could be masked by overlapping absorbance spectrum of the drug. As a result, the data collected from UV-Vis and CD spectroscopy were inconclusive whether TMPyP4 binds more strongly to the single strand than the G-quadruplex or vice versa. Thermal denaturation of *c-kitG1* with NMM monitored at 295 nm showed moderately sharp hyperchromicty around 70°C [Figure 7-3, a] which could be due to melting of dimeric NMM promoting G-quadruplex assembly. This mixture showed hypochromicity beyond 80°C. Melting of *c-kitG4* with high concentration of TMPyP4 showed no sign of sharp hypochromicity suggesting a possible non-cooperative melting or melting temperature well beyond studied maximal. Regions of very small hypochromictiy at 40°C, 65°C and one at elevated temperature were observed [Figure-3, b]. Step by step melting of unbound DNA, unsaturated complex and saturated complex could result in complex system with multiple distinct hypochromicity regions.

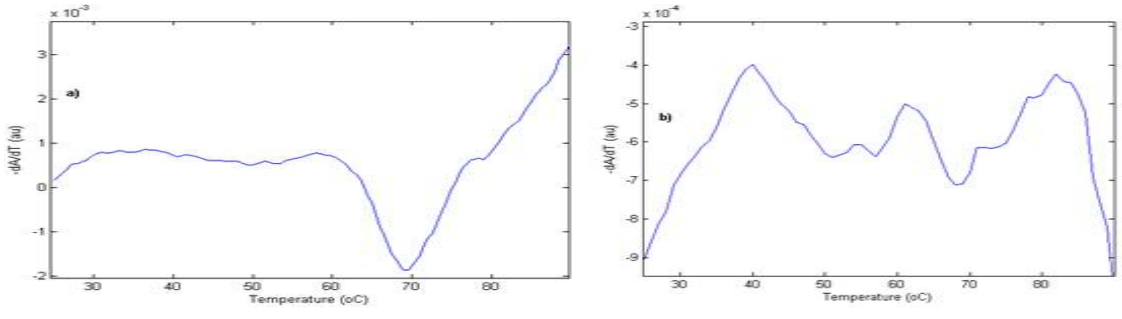


Figure 7-3 a) Thermal denaturation of the mixture of *c-kitG1*and NMM (mole ratio 0.6) in QSB. b) Thermal denaturation of the mixture of *c-kitG4* and TMPyP4 (mole ratio 2.0) in QSB.

MCR-ALS decomposition was performed [figure 7-4] on the various drug complexes of *c-kitG1* melting data based on proposed three components. Only three components, fully native, an intermediate and fully unfolded form were assumed to contribute to the molecular absorbance data. MCR-ALS resolved concentration profiles were utilized to yield thermodynamic parameters of the various complexes. Changes in stability imposed by the different drugs at that particular mole ratio are shown in table 7-1. For *c-kitG4*-drug complexes, MCR decomposed distribution diagram was overlapping. This was like expected as pure *c-kitG4* showed very high melting temperature.

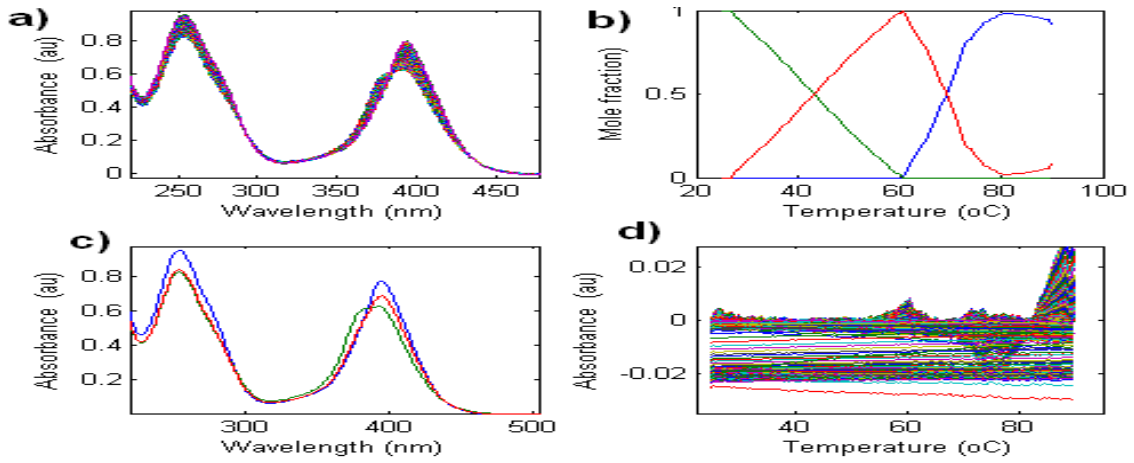


Figure 7-4 a) Experimental absorbance spectra of the mixture of *c-kitG1* and NMM. b) MCR-ALS decomposed distribution diagram for three spectroscopically active species. c) MCR-ALS decomposed pure spectral profile. D) Residuals plot

Ser No.	Complex	Mole Ratio	ΔH^o (kcal/mol)	ΔS^o (cal/Kmol)	ΔG^o (kcal/mol)	T_m (^oC)
1	*c-kitG1* free	-	53.1 ± 0.5	152.4 ± 1.4	5.8 ± 0.6	75 ± 1
2	*c-kitG1*-TPPS	1.0	49.4 ± 1.2	141.1 ± 3.6	5.6 ± 1.7	60 ± 1
3	*c-kitG1*-NMM	0.6	76.3 ± 0.6	223.1 ± 1.8	7.1 ± 0.9	> 88
4	*c-kitG1*-TMPyP4	0.3	31.5 ± 1.2	90.3 ± 0.6	3.5 ± 0.3	67 ± 1

Table 7-1 Thermodynamic parameters obtained from thermal denaturation of *c-kitG1* in QSB with and without different porphyrin compounds.

A two state approximation was considered in the narrow segmented temperature range where the intermediate and the fully unfolded form were the dominants populations. Van't Hoff equation is best suited for two components equilibrium system. During thermal denaturation of the mixture there are multiple possible components. When MCR resolved distribution diagram showed overlapping regions,

Van't Hoff equation wasn't applied. Thermodynamic parameters of *c-kitG4*-drug complexes weren´t determined in this temperature range for this reason.

Based on the thermodynamic parameters estimated and observed CD spectra upon addition of TMPyP4, this cationic porphyrin might bind to single strand DNA with slightly stronger affinity than to *c-kit* G-quadruplex structure. As both CD and molecular absorbance spectroscopy weren't conclusive; additional FRET experiments are required to confirm these observations.

C-kitG1 and *c-kitG4* sequences exhibited different behaviors along pH titrations. Generally, the level of the complex decreased upon progressive titration of the mixture of NMM and *c-kitG1*. This could be due to the formation of increasing level of dimer as low pH is known to facilitate dimerization of NMM. Protons shield electrostatic repulsion between two negatively charged NMM monomers. From this result, the dimeric form was considered to be unable to bind. When the pH of the mixture was lowered much below the pKa of NMM, molecular absorption spectrum of the mixture overlapped with molecular absorption of the pure NMM [figure 7-5, c] in the Soret region. This certainly indicated lack of molecular interaction between NMM and *c-kitG1* at pH below 1.2. Loss of molecular recognition in this region could be due to a) structural changes in *c-kitG1* which affected binding sites, b) inability of protonated NMM to bind c) the combined effect of change in the host and guest molecules. *C-kitG1* quadruplex structure has been found to be very stable at very low pH values. The release of NMM from the complex was particularly observed in pH region around 2 suggesting the release could be primarily due to protonation of NMM. Besides, the product of dissociation of the complex was the protonated form of the drug. Similarly, the level of the *c-kitG4* NMM complex decreased gradually during titration [figure 7-7, a]. NMM was seen to be less basic in the complex as compared to free unbound NMM in a solution [figure 7-7, a].

MCR-ALS decomposition of the experimental data fitted well when three components were considered for the complex of both *c-kitG1* and *c-kitG4* with NMM. *C-kitG1* showed three distinct pH transitions [figure 7-7, a] with mid-point values of 4.6 (indicating dimerization of NMM), 2.3 (indicating protonation of cytosine) and 1.6 (indicating protonation of NMM). The pKa values of free cytosine and NMM are 4.6 and 2.2 respectively(Pavel, Joaquim et al. 2009). The formation of the complex evidently decreased proton affinity of both components. Figure 7-5, b shows there was more deprotonated form of NMM in the mixture than expected for pure NMM.

On the other hand, *c-kitG4* showed only two distinct pH transitions [figure 7-7, b] with mid-point values of 6.1 (indicating dimerization of NMM) and 1.6 (indicating protonation of NMM). Here again, as can be seen from figure 7-5, b NMM in the complex was less basic than the free NMM. The release of NMM

from the complex was particularly observed in pH region around 2 suggesting the release could be primarily due to protonation of NMM.

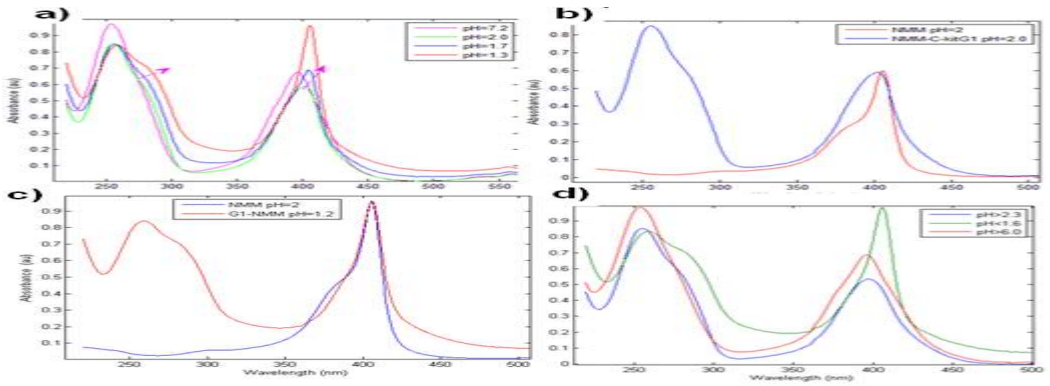


Figure 7-5 Acid-base titration of the mixture of *c-kitG1* and NMM at 25.°C (A) Experimental absorbance spectra. (B, C) Overlaid absorbance spectra of free NMM and the mixture of NMM and *c-kitG1*. (D) Resolved pure molecular absorption spectra for three optically active species present along the titration.

TMPyP4 showed different behaviour from NMM upon pH titrations [figure 7-9]. MCR-ALS decomposition of the experimental data fitted well with three components for its complex with *c-kitG1*. Generally, the level of free TMPyP4 increased upon progressive titration of the mixture with acid. *C-kitG1* showed two pH transitions with mid-point values of 4.1 (indicating protonation of cytosine) and 1.4 (indicating protonation of TMPyP4). The release of TMPyP4 from the complex was particularly observed in pH region around 4.5 suggesting the release could be primarily due to protonation of cytosine. The immediate product of dissociation was observed to be the free base form. Once the free base form was completely released from the complex, it was subsequently protonated. On the contrary, the complex with *c-kitG4* was fitted well with only two components.

This was like expected as neither dimerization of TMPyP4 nor protonation of cytosine was possible in this mixture. Besides, this mixture showed only one pH transition with mid-point value of 1.35 (indicating protonation of TMPyP4). Surprisingly, the immediate product of dissociation was observed to be the protonated form of TMPyP4.

The ongoing discussion indicated two important findings. First, the affinity of porphyrins to protons was significantly lowered upon association with either of G-quadruplex forming sequences. Protonated forms of porphyrins didn´t show any binding event. This is particularly true for associations which involve the central delocalized electrons of porphyrin rings implying the most supported stacking association. Second, NMM and TMPyP4 showed very different binding mechanisms to *c-kitG1*. Protonation of cytosine residues in *c-kitG1* didn´t release bound NMM whereas the same protonation released free base TMPyP4

from the complex. It seems TMPyP4 binds to loop 2 of *c-kitG1* and gets released when this loop is protonated. NMM binds to a different site and gets released only when it is protonated and not when the loop is protonated.

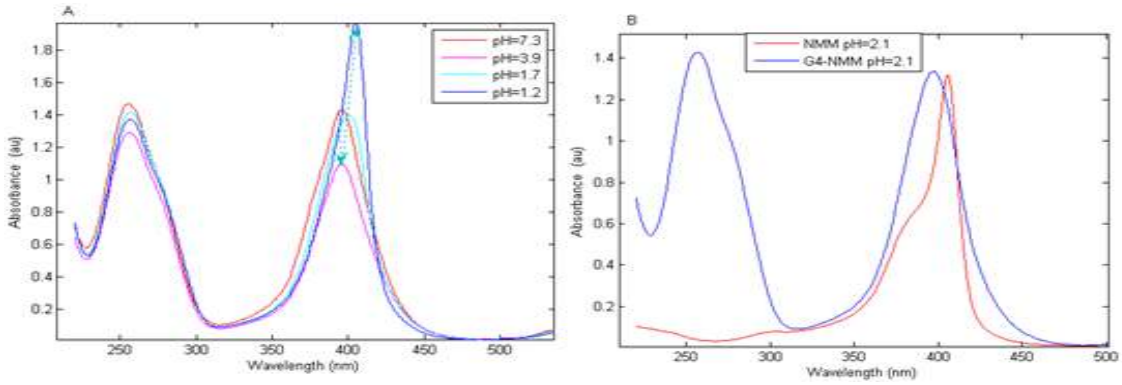


Figure 7-6 Acid-base titration of the mixture of *c-kitG4* and NMM at 25°C. (A) Experimental absorbance spectra. (B) Overlaid absorbance spectra of free NMM and the mixture of NMM and *c-kitG4*.

NMM dimer has Soret band at shorter wavelength than NMM monomer. Change in pH was found to alter the Soret band of NMM due to dimerization.

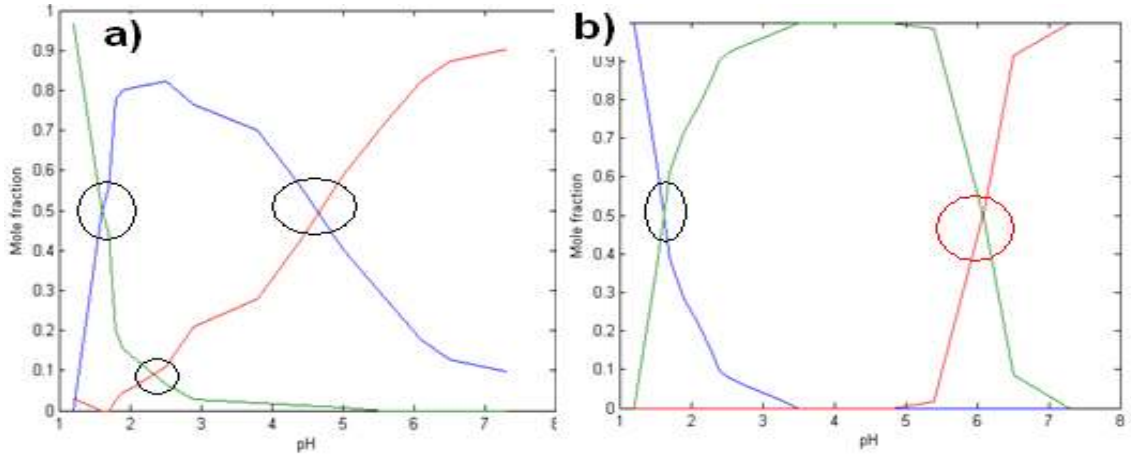


Figure 7-7 MCR-ALS resolved distribution diagram shows three transitions for *c-kitG1*and NMM mixture at 25°C. b) MCR-ALS resolved distribution diagram for acid base titration of *c-kitG4*-NMM mixture shows two transitions. Circles indicate observed pH transitions.

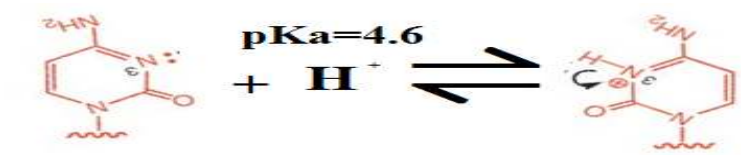


Figure 7-8 Protonation of cytosine decreases the formal charge on sugar phosphate backbone

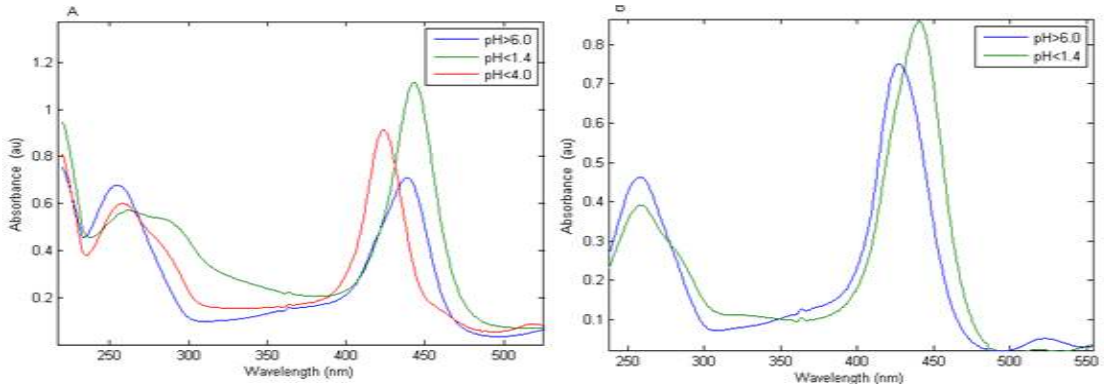


Figure 7-9 A) Resolved pure molecular absorption spectra for three optically active species present along the titration of the mixture of *c-kitG1* and TMPyP4 (B) Resolved pure molecular absorption spectra for two optically active species present along the titration of the mixture of *c-kitG4* and TMPyP4

7.2. Binding Mechanism

Structural features for drug binding on G-quadruplex DNA include the planar end surfaces of the G-tetrad stacked columns, the four grooves, the loops as well as the sugar phosphate backbone(Monchaud and Teulade-Fichou 2008). The different binding mechanisms of drugs to G-quadruplex DNA can be categorized into five general classes namely intercalation, end-stacking, loop binding, groove binding, and peripheral (outside random) binding. In the literature, three of these five mechanisms are proposed for binding of porphyrins to G-quadruplex DNA. These are end-stacking(Park, Shin et al. 2004), loop binding(Martino, Pagano et al. 2009) and peripheral binding. Detailed structural analyses of G-quadruplex-ligand complexes by NMR and X-ray crystallography have demonstrated at least two of these binding sites.

Isothermal titration calorimetry (ITC) and spectrophotometry indicated binding of porphyrins by threading intercalation at each closely similar GpG site, possibly without invoking neighbor exclusion for adjacent sites(Marky, Blumenfeld et al. 1983). This mechanism was supported by dynamic molecular modeling simulations with two DNA G-quadruplexes which show that stable intercalated complexes can be realized. No additional data supported porphyrin insertion within the stack of G-tetrads.

Co-facial end-stacking, which does not require neither the separation of two quartets nor the release of cation is the most frequently observed mode of interaction. This mechanism requires end pasting of ligand onto one or both of the terminal G-tetrads. The main forces involved in end-stacking molecular interactions are π-π stacking interactions between the aromatic ring of drugs and DNA bases. The ΔG° for

the end-stacking binding mode was found by molecular dynamics calculations to be driven by ΔH° with a small unfavorable TΔS° term.

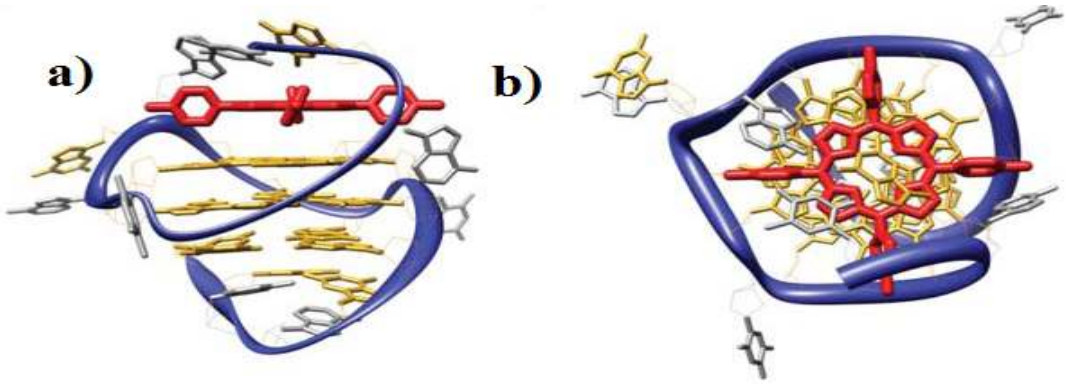


Figure 7-10 Side- (a) and top-views (b) of the NMR structure of human *c-myc* derived quadruplex-DNA in complex with TMPyP4 (PDB entry: 2A5R). TMPyP4 is stacked at the end of G-tetrad

Other binding sites are defined by the surface features of the grooves and/or loop regions. Porphyrin molecules were shown to bind to human telomeric G-quadruplex structure by stacking onto the TTA nucleotides in the external (diagonal) loop structure(Martino, Pagano et al. 2009). This mode of interaction facilitated TMPyP4 molecules to effectively stack between the bases without steric hindrance. Furthermore, it enables ligand-base π-π interactions to be maximized, at an optimal separation of 3.4 Å, and electrostatic interactions to be optimized by interactions between the cationic substituent N-methylpyridinium groups and the phosphate ions. All four pyrrole rings of TMPyP4 are sandwiched between the two T-T base pairs. This binding shows many features of intercalative binding; and is possibly the source of binding controversy.

In outside random binding modes, porphyrins interact with the phosphate group of DNA through electrostatic interaction. Water-soluble cationic ligands exhibit greater affinities for various DNA structures than the anionic ligands. Anionic ligands, in turn have greater selectivity for G-quadruplex DNA over double-stranded DNA or single stranded DNA. Generally, from drug designing point of view, outside random binding is undesirable. These binding modes are in continuous competition with each other depending on medium properties. This indicates the importance of charge on substitute groups on selectivity and affinity(Park, Shin et al. 2004).

The binding mode could be modulated by the chemical features of the porphyrin, the structure and composition of the host DNA, and the solution conditions in the experiment. Medium properties include ionic strength and presence of molecular crowding agent which affect self-association of the porphyrin.

The nature of the porphyrin ring includes the nature of central metal, size, charge and location of the substituent groups on the periphery of the porphyrin(Martino, Pagano et al. 2009).

TMPyP2 is a positional isomer of TMPyP4 in which is the N-methyl group on the pyridyl ring to the porphine core is located at ortho positon. Both stabilize quadruplex DNA to the same extent, but they do so by structurally distinct modes. TMPyP4 binds to the intramolecular human telomere quadruplex DNA predominantly by stacking externally to the guanine tetrads. TMPyP2 binds predominantly to the same quadruplex DNA structure via external binding to the TTA loop. TMPyP2 and TMPyP4 have very different effects on telomerase inhibition(Han, Wheelhouse et al. 1999). Similarly, meta-positional isomers were found to stabilize G-quadruplex structures more favorably than para isomers(Yamashita, Uno et al. 2005).

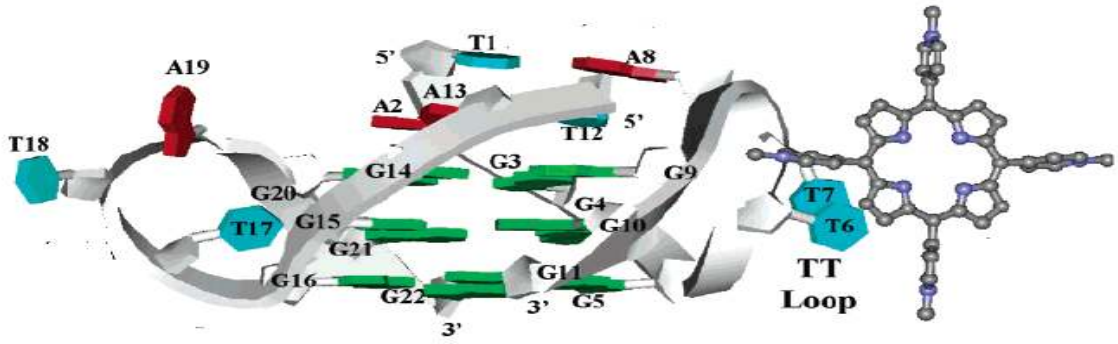


Figure 7-11 Structure of the human telomeric bimolecular G-quadruplex DNA with TMPyP4 bound to the extended TTA loop

It is likely that the cationic side arms of the meta- and orhto- positional isomers direct upward from the porphyrin plane, and interact more effectively with the anionic phosphates than those of the para-isomers. This conformation is favorable for the stabilization of the quadruplex structure, allowing all four cationic groups to interact with the anionic phosphates in the loop of G- quadruplex.

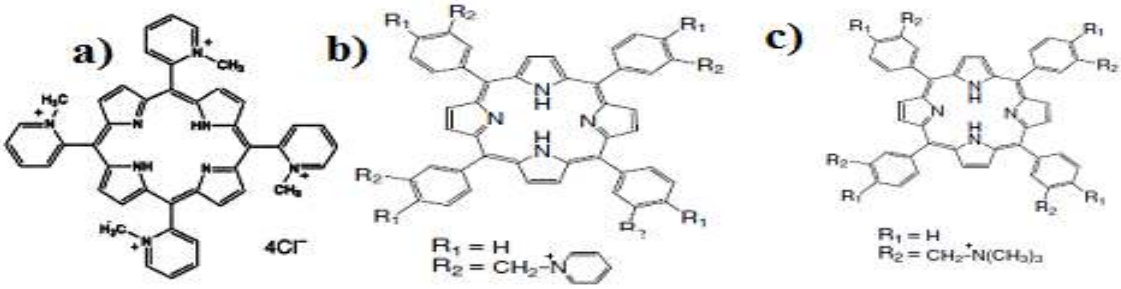


Figure 7-12 Chemical structures of ortho isomer of TMPyP4 (a), meta isomers of TMPyP4 (b) and trimethylammonium derivative of the meta isomer (c).

Because the pyridinium side arms were preferred to the trimethylammonium ones in the stabilization, the aromatic and less bulky pyridinium groups might allow themselves both to interact with the anionic phosphates and to stack with bases in the DNA loop. The cationic side arms of the porphyrins probably stabilize the quadruplex structure by reducing the electrostatic repulsion between the anionic phosphates of the backbone of the lateral or loops. This indicates the importance of the steric hindrance and electrostatic interaction for ligand binding to G-quadruplex loops(Yamashita, Uno et al. 2005).

All available experimental data suggest the presence of three ligand binding sites in the *c-kitG4* quadruplex structure(Martino, Pagano et al. 2009). These are the 5' end of the G-tetrad, the 3' end of the G-tetrad and the longest CGCGA loop. Porphyrin ligands can bind by end-stacking at either end of the G-tetrad or bind by a combination of stacking and electrostatic interaction to loop 2. The favorable free energy change in end-stacking is opposed by unfavorable entropy change. On the other hand the favorable free energy change in loop binding is partly due to a favorable entropy change. These two types of independent binding modes are nonequivalent in which loop binding has higher binding affinities than end stacking.

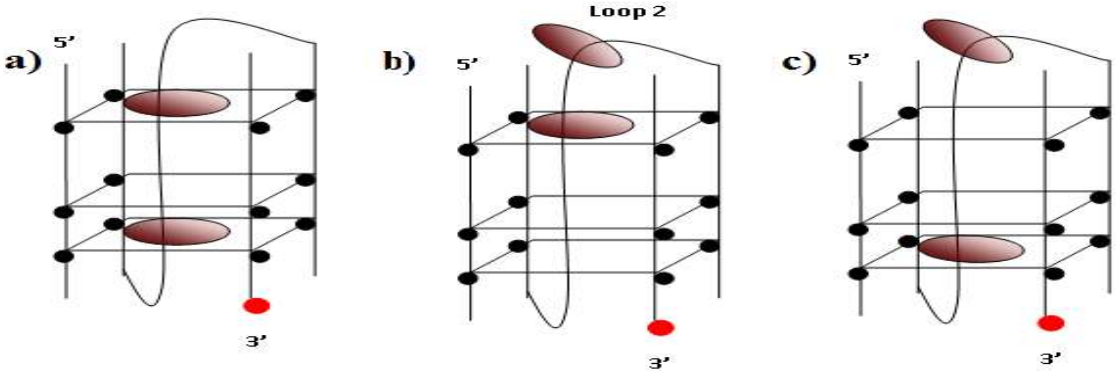


Figure 7-13 Supposed binding modes of cationic porphyrin TMPyP4 to the predominantly dangling end conformer of *c-kit*G4 (a) cap and dangling end-staking; (b) mixed end stacking at the cap and loop binding (c) mixed end-stacking at the dangling end and loop binding

TMPyP4 can bind at three distinct sites in *c-kitG4* as shown in figure 7-13 above. The most probable mechanism is that shown in figure 7-13 a. Likewise, TMPyP4 can bind at either of these sites in *c-kitG1*. Due to the fact that the loops in C-kiG1 are composed of cytosines which contain very electronegative oxygen, cationic porphyrins (TMPyP4) are expected to bind strongly. NMM being anionic porphyrin, it doesn't interact with the sugar phosphate back bone. Hence, NMM is supposed to bind at the 5' end of the G-quartet structure with *c-kitG1* and as shown in figure 7-14 (a) for TMPyP4 with *c-kitG4*. The two TMPyP4-quadruplex molecular structures available to date each show just one of the two possible binding modes. Even though, the end stacking mode of interaction showed less affinity, it is considered to be

more structure specific than its stronger competitor. The heterogeneous binding observed in SPR experiments alerts binding at the loops could be the characteristic feature of promiscuous binders(Coan and Shoichet 2008). Future chemical optimization to enhance the affinity of porphyrin-type molecules for quadruplex structures are advised to take this into consideration.

7.3. Conclusion

Multivariate data analysis allowed detection of very weak absorbance changes and small intrinsic ellipticity changes manifested in response to subtle structural rearrangements during melting, mole ratio and acid-base titration experiments. This was coupled with molecular structures determined with other structural techniques.

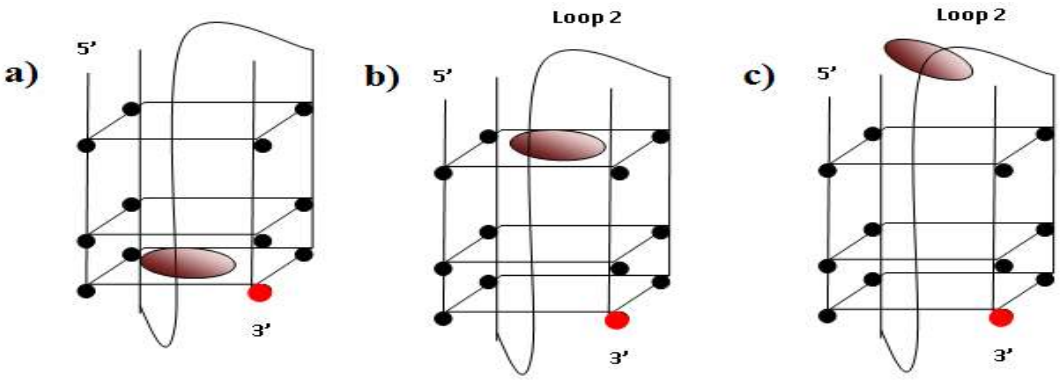


Figure 7-14 Proposed binding modes of cationic porphyrin TMPyP4 to the predominantly blunt end conformer of *c-kitG1* (a) blunt end-staking; (b) cap end -stacking (c) loop binding

By combining these tools, the affinity, specificity and thermodynamic parameters of DNA drug interactions were determined. It should be emphasized curve resolution and modeling techniques helped to extract new structural variants that would otherwise be indistinguishable in these spectroscopic techniques. The resulting information can be further refined based on additional structural information. Spectroscopic techniques along with multivariate data analysis techniques are ideal tools to generate plausible mechanisms. The study showed one such case by producing a possible drug binding mechanism to a complex system of G-quadruplex DNA. Extracting mechanistic conclusion is always associated with some uncertainties and by no means definitive. Competing mechanisms will be postulated; the decision between them is open for rigorous theoretical and experimental test.

8

Materials and Methods

8. Materials and Methods

8.1. Biophysical Characterization of Drugs

Meso-5, 10, 15, 20-tetrakis (p-tri-N-methyl-4pyridyl) porphine ($H_2TMPyP4^{4+}$, here after abbreviated as TMPyP4) and Meso-5, 10, 15, 20-tetrakis (4-sulfotophenyl) porphirine (H_2TPPS^{4-}, hereafter abbreviated as TPPS) were analytical grade reagents purchased from Sigma Aldrich whereas N-methyl mesoporphyrin IX ($HNMM^{2-}$, hereafter abbreviated as NMM) was excellent grade reagent purchased from Frontier Scientific Inc. All compounds were used without further purification. Stock solutions containing 1mM of TPPS, 4 mM NMM and 100 μM of TMPyP4 were prepared in MilliQ water, stored at -20°C and diluted to working concentrations in MilliQ water immediately before use. QSB (QSB) was prepared by adding potassium monobasic phosphate (0.19 g), sodium dihydrogen phosphate (0.23 g), potassium chloride (2.19 g) and magnesium chloride (0.04 g) into MilliQ water (200 ml). The resulting pH was adjusted by adding small volumes of NaOH using Cyberscan 2500 pH meter. The experimental conditions were pH 7.2 and 150 mM ionic strength (147 mM K^+, as KCl, and 1 mM $Mg2^+$, as MgCl2.6H2O). The drugs were characterized as follows.

8.1.1. Test for dimerization of porphyrin drugs

All measurements were made on an Agilent HP8453 photo diode array spectrophotometer. Hellma quartz cuvette (path length of 1.0 cm, 4.50 ml volume) was used. Very diluted drug concentration, absorbance below 0.05 at Soret band in 1 cm path length cell, was prepared and its spectrum was recorded. The solution was successively concentrated until absorbance at Soret band is above 1. This was done by progressive additions of small volumes of the aforementioned stock solutions into 3 ml of QSB. Homogeneity of the resulting solutions was maintained in-line (taking advantage of the stirrer incorporated at the cell holder of the Agilent instrument). Successive additions were done at time intervals of10 minutes. All absorbance spectra were recorded. The spectra obtained from diluted concentrations were multiplied by a factor to give the same absorbance as that of concentrated solutions. The resulting spectra were overlayed and tested if superimposable. Lack of any shift in Soret band (superimposable spectra) indicated either the porphyrin is in monomeric or 100% dimeric form in the concentration ranges covered. The latter is very unlikely as more and more concentration ranges are covered.

8.1.2. Determination of molar absorption coefficients

Samples of working drug solutions (100 μM NMM, 100 μM TMPyP4 and 200 μM TPPS) were measured at a minimum of six concentration levels. The ranges covered were 0.1 to 5 μM for NMM, 0.1 to 3.5 μM

for TMPyP4 and 0.1 to 4 μM for TPPS. Linear regression was used to estimate regression parameters for the relationship between concentration and absorbance. The value of correlation coefficient was determined for characterizing linearity of analytical response in the concentration ranges investigated.

8.1.3. Acid-base properties of porphyrin drugs

Acid-base titrations were monitored in-line (taking advantage of the stirrer incorporated at the cell holder of the Agilent instrument). The pH values were determined with an Orion SA 720 pH/ISE meter and microcombination pH electrode (Thermo). Other experimental conditions were maintained using QSB mentioned above. Samples were prepared at 3.23μM TMPyP4, 3.23μM NMM and 3.28 μM TPPS concentrations in QSB. The pH of the solutions containing each of the above drugs was adjusted by adding small volumes of HCl or NaOH stock solutions. Following this, absorbance spectra were recorded pH stepwise.

8.1.4. CD spectra of porphyrin drugs

CD spectra were recorded on a Jasco J-810 spectropolarimeter equipped with Julabo F-25/HD temperature control unit. Hellma quartz cuvette (path length of 1.0 cm, 4.50 ml volume) was used. CD and absorbance spectra were recorded after adding 70 μL of TMPyP4 and 100 μL of NMM to 3 ml buffer.

8.1.5. Fluorescence spectra of NMM

Fluorescence spectra were obtained using an Aminico-Bowman 8202 luminescence spectrometer with a rotating excitation shutter assembly. Excitation was at 390 nm and peak emission was monitored at 615 nm. All measurements were carried out in QSB at 25°C. The concentration range covered was from 1.6 μM to 35 μM characterizing the shift from a predominantly monomeric solution to predominantly dimeric solution. The excitation in the region of Soret band was provided by monochromatic light with the use of monochromator. Fluorescence spectra were completely calibrated using the known spectral sensitivity of the equipment.

8.1.6. Thermal denaturation of porphyrin drugs

Thermal denaturation experiments were done throughout the temperature range from 25°C to 90°C with a linear temperature ramp of 0.5°C /min and 1 minute hold time. Partial evaporation can be limited by sealing the cuvettes with paraffin. Absorbance spectra were recorded every 0.5°C. Three independent melting experiments were done by adding 50 μL of 100 μM TMPyP4, 100 μL of 100 μM NMM and 50 μL

of 200 μM TPPS working drug solutions into 3 ml of QSB respectively. Each sample was allowed to equilibrate at the initial temperature for 30 min before the melting experiment was started.

8.2. Biophysical Characterization of target DNA

8.2.1. DNA Synthesis and Purification

The oligonucleotides 5'-d[$CG_3CG_3CGCGAG_3AG_4$] -3' and 5' -d[$TG_3TG_3TGTGTG_3TG_4$] 3' corresponding to the G-quadruplex forming element in the promoter region of the human *c-kit* proto-oncogene wild type and mutant *c-kitG4* respectively were prepared on a 1 mM scale using standard 2-cyanoethyl phosphoramidites (Cruachem Ltd.). Syntheses were performed using an automatic DNA synthesizer (Applied Biosystems Mod. 392). Oligomers were removed from solid support and base blocking groups were removed by treatment using standard protocols (concentrated ammonia, 55°C, and overnight). After deprotection, oligonucleotides were purified using oligonucleotide purification cartridges following the protocol suggested by the manufacturer. Finally, purified oligonucleotides were desalted using Sephadex G-25 columns. Each sample was evaporated to dryness and re-dissolved in 1 ml MiliQ water. Oligonucleotide concentration was determined by UV absorbance measurements at 260 nm and 90°C for 10 minutes using a molar absorptivity listed below. These values were obtained by adjusting the nearest-neighbour calculation method.

8.2.2. Thermal denaturation experiment using UV-vis spectroscopy

Melting experiment was carried out by dissolving 4.83 μM *c-kitG1* and 3.23 μM *c-kitG4* solutions in QSB. Melting curves were collected on an Agilent 8453 spectrophotometer using Hellma cells (1cm path length, 4.5 ml volume). Temperature was monitored with Agilent temperature-controlling Peltier unit mounted in the spindle of the Agilent thermoelectric cuvette holder. Oligonucleotides were "annealed" and degassed by raising the temperature to 90°C for 10 min and then cooling to 25°C prior to the melting experiment. Samples were then heated at a linear temperature ramp of 0.5°C/min with data collection beginning at 25°C and ending at 90°C. Absorbance at 295 nm was traced at each temperature value to give melting curves. The spectral bandwidth was 1 nm and the hold time was 1 min. Absorbance wasn´t not corrected for thermal expansion and heat capacity changes (ΔCp) that may accompany quadruplex denaturation. The melting temperature (T_m) values are the average value of at least a pair of T_m values recorded during repeated melting experiments.

8.2.3. Thermal denaturation experiment using circular dichroism spectroscopy

Melting experiment was performed by dissolving 0.87 μM *c-kitG1* and 2.31 μM *c-kitG4* solutions in QSB described elsewhere. CD spectra were recorded on a Jasco J-810 spectropolarimeter equipped with a Julabo F-25/HD temperature-control unit. Hellma quartz cells (1.0 cm path length, 4.5 mL volume) were used. Each sample was allowed to equilibrate at the initial temperature for 30 min before the melting experiment was started. Melting experiments were done in the temperature range from 25°C to 90°C with data pitch 0.5°C/min. CD spectra were recorded every 5°C. Acquisition parameters used were wavelength scan of 339.5 nm to 220 nm, pitch 0.5 nm, continuous mode, 100 nm/min scanning speed, 4 seconds response, 1 second accumulation time and 1 min delay time at each temperature value. A blank dataset taken from the buffer solution alone was also recorded and subtracted from the results.

8.2.4. Effect of environmental factors on G-quadruplex structure

Thermal melting experiment of *c-kitG4* was performed by preparing 1.64 μM solution in 3 ml Milli-Q water. In the case of *c-kitG1*, thermal denaturation was conducted by dissolving 0.66 μM in QSB to which 10% ethanol was added. Melting data were monitored by uv-vis spectrophotometer. The resulting melting temperatures were compared with that of the buffer. To study the effect of ionic strength, 1.8 μM *c-kitG1* and 3.29 μM *c-kitG4* were prepared in QSB buffer. Salt concentrations were increased gradually by adding increasing volumes 4M KCl solution (0-200μL) directly to the *c-kitG1* and *c-kitG4* DNA samples in the CD cells. CD spectra were measured after each salt addition. The effect of specific ion binding on thermal stability and conformational switching of G-quadruplex has been investigated likewise. The CD spectra were recorded for melting of 1.67μM of *c-kitG4* with linear heating rate of 5°C min−1 in Milli-Q water.

8.3. Determination of Equilibrium Constant

Spectroscopic monitoring of DNA-Drug titration experiments was done in-line; taking advantage of the stirrers incorporated at the cell holders of Agilent and Jasco J-810 instruments as mentioned earlier. Other experimental conditions were 25°C in QSB. To aid intuitive judgment of stoichiometric values, absorption spectra were collected when mole ratios of TMPyP4 to *c-kitG1* and *c-kitG4* were close to 4, 3, 2 and 1. The effect of adding small volumes of each DNA targets to the aforementioned mole ratios was qualitatively examined. The resulting effect of shift in equilibrium according to Le Chatelier's principle was weighted against concurrent effect of binding to un-occupied binding site.

UV-Vis procedure was inverse titration i.e. the concentration of the drugs (TMPyP4, NMM and TPPS) was held constant and DNA concentration was varied. First the spectrum of pure drug in QSB was

recorded. This sample was then titrated by increasing volumes of DNA solution (*c-kitG1* and *c-kitG4*) in such a way that the resulting C_{TMPyP4} /C_{DNA} ratios were systematically lowered from 3.00 to 0.33 at least in ten steps. Each addition was done at time intervals of 10 min. Spectroscopic signals from the resulting mixture were recorded after each addition of the DNA solutions. The completion of titration was confirmed when the absorption band of drugs didn´t change significantly. Finally, the spectrum of a pure porphyrin drug was recorded.

On the other hand, CD procedure was direct titration i.e. the concentration of the G-quadruplex forming DNA (*c-kitG1* and *c-kitG4*) was fixed and the concentration of drugs was varied progressively. First the CD spectrum of pure DNA solution in QSB was recorded. Next, progressively increasing concentration of drugs (TMPyP4, NMM and TPPS) was added to the first sample to raise C_{TMPyP4} /C_{DNA} ratios systematically from 0.10 to 3.00 at least in ten steps. Each addition was done at time intervals of 10 min. The CD spectrum of the mixed sample was recorded stepwise. Finally, the spectrum of a pure porphyrin drug was recorded.

8.4. Probing Molecular Recognition

8.4.1. Sample Preparation

The oligonucleotides 5' -d[CG3CG3CGCGAG3AG4] -3'biotin, 5' -d[TG3TG3TGTGTG3TG4] 3'biotin corresponding to the G-quadruplex forming element of promoter region of the human *c-kit* proto-oncogene were prepared on a 1 mM scale using standard 2-cyanoethyl phosphoramidites (Cruachem Ltd.). Syntheses were performed using an automatic DNA synthesizer (Applied Biosystems Mod. 392). Sequences were deprotected using standard protocols (concentrated ammonia, 55 ^{o}C, and overnight). After deprotection, oligonucleotides were purified using oligonucleotide purification cartridges following the protocol suggested by the manufacturer. Finally, purified oligonucleotides were desalted using Sephadex G-25 columns. DNA strand concentration was determined by UV absorbance measurements at 260 nm and 90^{0}C for 10 minutes using a molar absorptivity listed above. These values were obtained by adjusting the nearest-neighbour calculation method. Thermal denaturation experiment and circular dichroism spectroscopy confirmed the formation of G-quadruplex structures by biotin labelled oligonucloeotides.

TMPyP4 and TPPS were purchased from Sigma Aldrich whereas NMM was purchased from Frontier Scientific Inc. All compounds were used without further purification. Stock solutions containing 1 mM of each drug were prepared in MilliQ water, stored at -20 ^{o}C and diluted to working concentrations immediately before use in MilliQ water. The values for the extinction coefficient were as determined

during characterization of drugs. KCl, MgCl2.6H2O, KH2PO4, K2HPO4, HAcO, HCl and NaOH (a.r.) were purchased from Panreac (Spain).

8.4.2. Immobilization of DNA

Biosensor SPR experiments were performed with a four serial channel T100 (BIAcor, Inc.) optical biosensor system and streptavidin-coated sensor chips (Sensor chip SA; BIAcore, Inc.). All DNA samples for quadruplex-binding experiments were used as single strands. The concentration in all cases refers to the strand concentration. The streptavidin sensor chip surface was prepared for DNA immobilization with three consecutive 1 minute injection of regeneration buffer (also called hydration buffer, 1M NaCl in 50 mM NaOH) followed by extensive washing with a running buffer. The running buffer was sterile filtered and degassed HBS-EP buffer (0.01 M HEPES, (pH 7.4), 3 mM EDTA, and 0.005% surfactant Tween 20 with 0.1M NaCl) fortified with 150 mM of KCl at pH 7.2 Quadruplex formation of the oligomers was effected by heating the solution at 90°C for 10 minutes followed by slow cooling at room temperature before non covalent immobilization on chip surface. Characteristic CD peaks with negative intensitites at 240 nm and positive intensities 260 nm confirmed the formation of quadruplex by biotin labelled sequences. Thermal denaturation experiment on biotinlabelled sequences showed there is no change in the melting temperature due to biotinylation.

The 21mer 3'-biotinylated sequences *c-kitG1* d(C GGG C GGG CGCGA GGG A GGG G) and *c-kitG4* d (T GGG T GGG TGTGT GGG T GGG G) in running buffer were immobilized on flowcell 2 and 4 respectively to attain a binding of 1200 RU (ca. 10 μg/ml) as previously described. 3' biotynylated *c-kitG1* with poly A linker sequence (ca. 10 μg/ml corresponding to 1200 RU) was immobilized on flow cell 3. These three oligonucleotides were immobilized in different flow cells on the same sensor chip. Flow cell 1 was coupled with biocytin (ca. 10 μg/ml corresponding to 200 RU) and left blank as a control to characterize non specific binding. The signal from the control flow cell was subtracted from the signals obtained in flow cell 2,3 and 4 to convert absolute response to relative response. After immobilization, the structure was equilibrated by passing running buffer over the surface for 1 hour at 60 ul/min flow rate to attain base line stability. No significant change on base line was detected in this time. All experiments were performed at 25°C using this running buffer. Analysis of interaction was performed by using both kinetics and steady-state methods with multiple injections of different compound concentrations over the immobilized DNA surface.

8.4.3. Biosensor SPR Experiments

Analyte solutions at different concentrations were prepared in the running buffer by serial dilutions from stock solution and injected from 7 mm plastic vials (at 60 μL/min for 40 seconds) with pierceable plastic crimp caps (BIAcore, Inc.). Plastic vaials were randomly positioned in the rack to minimize systematic error. Following this desociation from the surface was monitored for 60 seconds in running buffer. Regeneration was accomplished by dobule injection of 1M NaCl in 50 mM NaOH at 30 μl/min and 15 seconds as neither the running buffer alone nor single injection of the regenration buffer could completely desociate the complex from the surface. The baseline was then reestablished, and the next compound concentration sample was injected. The concentration ranges for TMPyP4 were 0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.1 and 0.2 μM. Two samples (0, 0.06 μM) were replicated. The concentration ranges for NMM were 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 1 and 2 μM. Two samples (0, 0.5 μM) were replicated. The concentration ranges for TPPS were 0, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 μM. Two samples (0, 50 μM) were replicated. Suitable blank control injections (zero concentrations) with running buffer were performed, and the resulting sensorgrams were subtracted from non-zero concentrations to obtain the final concentration-dependent graphs.

8.5. Characterization of Interaction Complex

UV Melting profiles of the mixtures of DNAs and water soluble porphyrin drugs were determined for solutions obtained from each UV-Vis mole ratio experiments. Only in one case *c-kitG4* and NMM mixture solution was derived from CD mole ratio experiment. Subsequently, the mixtures were believed to contain large excess of un-bound DNA. Samples used in this study include *c-kitG1*-TMPyP4 (C_{TMPyP4} /C_{DNA}=0.3), *c-kitG4*-TMPyP4 (C_{TMPyP4} /C_{DNA}=0.2), *c-kitG1*-NMM (C_{NMM} /C_{DNA}=0.6), *G1*-TPPS (C_{TPPS} /C_{DNA}=1.0) and *G4*-NMM mixture (C_{NMM} /C_{DNA}=3.6). Thermal denaturation experiments of the mixtures were conducted as previously described.

CD melting curves were obtained for previously prepared solutions from each CD mole ratio experiments. In this case the mixtures were believed to contain large excess of un-bound drugs. CD melting curves were determined for mixtures of *G1*-TMPyP4 (C_{TMPyP4} /C_{DNA}=2.5), *c-kitG1*-NMM (C_{NMM} /C_{DNA}=3.4) and *c-kitG4*-NMM (C_{NMM} /C_{DNA}=3.5). Melting experiments were preformed as described before.

Acid–base titrations were performed for the mixtures *c-kitG1*-TMPyP4 (C_{TMPyP4} /C_{DNA}=1.9), *c-kitG4*-TMPyP4 (C_{TMPyP4} /C_{DNA}=2.4), *c-kitG1*-NMM (C_{NMM} /C_{DNA}=2) and *c-kitG4*-NMM mixture (C_{NMM} /C_{DNA}=2.0) as described for porphyrin drugs.

9. References

1. (http://oak.ucc.nau.edu/el57/Research.htm)
2. http://www.phy.cam.ac.uk/research/bss/molbiophysics.php
3. Amigo, J. M., A. de Juan, et al. (2006). "A mixed hard- and soft-modelling approach to study and monitor enzymatic systems in biological fluids." Analytica Chimica Acta **567**(2): 245-254.
4. Bakk, A. and R. Metzler (2004). "Two states do not necessarily correspond to a two-state transition: van't Hoff enthalpy in the case of a small entropy." Chemical Physics Letters **398**(1): 190-193.
5. Beale, S. I. and N. C. Chen (1983). "N-Methyl Mesoporphyrin IX Inhibits Phycocyanin, but Not Chlorophyll Synthesis in Cyanidium caldarium." PLANT PHYSIOLOGY **71**(2): 263-268.
6. Besmer, P., J. E. Murphy, et al. (1986). "A new acute transforming feline retrovirus and relationship of its oncogene v-kit with the protein kinase gene family." Nature **320**(6061): 415-421.
7. Biesaga, M., K. Pyrzynska, et al. (2000). "Porphyrins in analytical chemistry. A review." Talanta **51**(2): 209-224.
8. Black, R. J., F. Bray, et al. (1997). "Cancer incidence and mortality in the European Union: Cancer registry data and estimates of national incidence for 1990." European journal of cancer (Oxford, England : 1990) **33**(7): 1075-1107.
9. Blackburn, E. H. (1997). "The telomere and telomerase: nucleic acid-protein complexes acting in a telomere homeostasis system. A review." Biochemistry. Biokhimiiaa **62**(11): 1196-1201.
10. Bommarito, S., N. Peyret, et al. (2000). "Thermodynamic parameters for DNA sequences with dangling ends." Nucleic Acids Research **28**(9): 1929-1934.
11. Bosque-Sendra, J. M., L. Almansa, et al. (2003). "Data Analysis in the Determination of Stoichiometries and Stability Constants of Complexes." Analytical Sciences **19**(10): 1431-1439.
12. Breslauer, K. J., D. P. Remeta, et al. (1987). "Enthalpy-entropy compensations in drug-DNA binding studies." Proceedings of the National Academy of Sciences of the United States of America **84**(24): 8922-8926.
13. Bucek, P., J. Jaumot, et al. (2009). "pH-Modulated Watson-crick duplex-quadruplex equilibria of guanine-rich and cytosine-rich DNA sequences 140 base pairs upstream of the *c-kit* transcription initiation site." Chemistry (Weinheim an der Bergstrasse, Germany) **15**(46): 12663-12671.
14. Clark, G. R., P. D. Pytel, et al. (2003). "Structure of the first parallel DNA quadruplex-drug complex." Journal of the American Chemical Society **125**(14): 4066-4067.
15. Coan, K. E. D. and B. K. Shoichet (2008). "Stoichiometry and physical chemistry of promiscuous aggregate-based inhibitors." Journal of the American Chemical Society **130**(29): 9606-9612.
16. Cogoi, S. and L. E. Xodo (2006). "G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription." Nucleic Acids Research **34**(9): 2536-2549.
17. Daz-Cruz, M. S., J. M. Daz-Cruz, et al. (2000). "Soft- and Hard-Modeling Approaches for the Determination of Stability Constants of Metal-Peptide Systems by Voltammetry." Analytical Biochemistry **279**(2): 189-201.
18. del Toro, M., P. Bucek, et al. (2009). "Targeting the G-quadruplex-forming region near the P1 promoter in the human BCL-2 gene with the cationic porphyrin." Biochimie **91**(7): 894-902.
19. Dixon, D. W. and V. Steullet (1998). "Dimerization of tetracationic porphyrins: ionic strength dependence." Journal of Inorganic Biochemistry **69**(1): 25-32.
20. Duquette, M. L., M. D. Huber, et al. (2007). "G-Rich Proto-Oncogenes Are Targeted for Genomic Instability in B-Cell Lymphomas." Cancer Research **67**(6): 2586-2594.
21. Dyson, R. M., S. Kaderli, et al. (1997). "Second order global analysis: the evaluation of series of spectrophotometric titrations for improved determination of." Analytica Chimica Acta **353**(2): 381-393.
22. Esteban, M., C. Arino, et al. (2000). "Multivariate curve resolution with alternating least squares optimisation: a soft-modelling approach to metal." Trends in Analytical Chemistry **19**(1): 49-61.

23. Freier, S. M., B. J. Burger, et al. (1983). "Effects of 3' dangling end stacking on the stability of GGCC and CCGG double helixes." Biochemistry **22**(26): 6198-6206.
24. Freyer, M. W., R. Buscaglia, et al. (2007). "Biophysical studies of the *c-myc* NHE III1 promoter: model quadruplex interactions with a cationic porphyrin." Biophysical journal **92**(6): 2007-2015.
25. Gampp, H., M. Maeder, et al. (1985). "Calculation of equilibrium constants from multiwavelength spectroscopic data-II Specfit: two user-friendly programs in basic and standard fortran 77." Talanta **32**(4): 251-264.
26. Gans, P., A. Sabatini, et al. (1996). "Investigation of equilibria in solution. Determination of equilibrium constants with the HYPERQUAD suite of programs." Talanta **43**(10): 1739-1753.
27. Gargallo, R., R. Tauler, et al. (1996). "Influence of selectivity and polyelectrolyte effects on the performance of soft-modelling and hard-modelling approaches." Analytica Chimica Acta **331**(3): 195-205.
28. Gaynutdinov, T. I., R. D. Neumann, et al. (2008). "Iodine-125 radioprobing of intramolecular quadruplex conformation of human telomeric DNA in the presence of cationic porphyrin TMPyP4." International journal of radiation biology **84**(12): 984-990.
29. Gazdar, A. F. (1994). "The molecular and cellular basis of human lung cancer." Anticancer research **14**(1B): 261-267.
30. Gemperline, P. J. and E. Cash (2003). "Advantages of Soft versus Hard Constraints in Self-Modeling Curve Resolution Problems. Alternating Least Squares with Penalty Functions." Analytical Chemistry **75**(16): 4236-4243.
31. Goldstein, G. and C. H. Shelly (1973). "Univariate vs. multivariate analysis in neuropsychological test assessment of lateralized brain damage." Cortex; a journal devoted to the study of the nervous system and behavior **9**(2): 204-216.
32. Guo, K., A. Pourpak, et al. (2007). "Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene." Journal of the American Chemical Society **129**(33): 10220-10228.
33. Halder, K. and S. Chowdhury (2005). "Kinetic resolution of bimolecular hybridization versus intramolecular folding in nucleic acids by surface plasmon resonance: application to G-quadruplex/duplex competition in human *c-myc* promoter." Nucleic Acids Research **33**(14): 4466-4474.
34. Han, F. X., R. T. Wheelhouse, et al. (1999). "Interactions of TMPyP4 and TMPyP2 with Quadruplex DNA. Structural Basis for the Differential Effects on Telomerase Inhibition." Journal of the American Chemical Society **121**(15): 3561-3570.
35. Hemesath, T. J., E. R. Price, et al. (1998). "MAP kinase links the transcription factor Microphthalmia to *c-kit* signalling in melanocytes." Nature **391**(6664): 298-301.
36. Hirota, S., K. Isozaki, et al. (1998). "Gain-of-Function Mutations of *c-kit* in Human Gastrointestinal Stromal Tumors." Science **279**(5350): 577-580.
37. Homola, J. (2008). "Surface Plasmon Resonance Sensors for Detection of Chemical and Biological Species." Chemical Reviews **108**(2): 462-493.
38. Hsu, S.-T. D., P. Varnai, et al. (2009). "A G-Rich Sequence within the *c-kit* Oncogene Promoter Forms a Parallel G-Quadruplex Having Asymmetric G-Tetrad Dynamics." Journal of the American Chemical Society **131**(37): 13399-13409.
39. Jaumot, J., N. Escaja, et al. (2002). Multivariate curve resolution: a powerful tool for the analysis of conformational transitions in nucleic acids. Nucl. Acids Res. **30:** e92-.
40. Jeffrey, G. A. and Y. Kinoshita (1963). "The crystal structure of cytosine monohydrate." Acta Crystallographica.
41. Kano, K., H. Minamizono, et al. (1997). "Self-Aggregation of Cationic Porphyrins in Water. Can Ï€â^'Ï€ Stacking Interaction Overcome Electrostatic Repulsive Force?" The Journal of Physical Chemistry A **101**(34): 6118-6124.
42. Kojima, Y., S. Koda, et al. (2001). "Effect of ultrasonic frequency on polymerization of styrene under sonication." Ultrasonics sonochemistry **8**(2): 75-79.

43. Lambert, L., T. Armstrong, et al. (2009). "Incidence, Risk Factors, and Impact of Severe Neutropenia After Hyperthermic Intraperitoneal Mitomycin C." Annals of Surgical Oncology **16**(8): 2181-2187.
44. Lane, A. N., J. B. Chaires, et al. (2008). "Stability and kinetics of G-quadruplex structures." Nucleic Acids Research **36**(17): 5482-5515.
45. Levy, M. Z., R. C. Allsopp, et al. (1992). "Telomere end-replication problem and cell aging." Journal of molecular biology **225**(4): 951-960.
46. Li, N. and S. Y. Tong (1994). "Spectrophotometric study of the interaction of tetraphenylporphyrin tetrasulfonate (TPPS4) with proteins." Talanta **41**(10): 1657-1662.
47. Lipps, H. J. and D. Rhodes (2009). "G-quadruplex structures: in vivo evidence and function." Trends in cell biology **19**(8): 414-422.
48. Luedtke, N. W. (2009). "Targeting G-Quadruplex DNA with Small Molecules." CHIMIA International Journal for Chemistry **63**: 134-139.
49. Margalit, R. and M. Rotenberg (1984). "Thermodynamics of porphyrin dimerization in aqueous solutions." The Biochemical journal **219**(2): 445-450.
50. Margalit, R., N. Shaklai, et al. (1983). "Fluorimetric studies on the dimerization equilibrium of protoporphyrin IX and its haemato derivative." The Biochemical journal **209**(2): 547-552.
51. Marky, L. A., K. S. Blumenfeld, et al. (1983). "Calorimetric and spectroscoptc Investigation of drug - DNA Interactions. I. The binding of netropsin to poly d(AT)." Nucleic Acids Research **11**(9): 2857-2870.
52. Martino, L., B. Pagano, et al. (2009). "Shedding Light on the Interaction between TMPyP4 and Human Telomeric Quadruplexes." The Journal of Physical Chemistry B **113**(44): 14779-14786.
53. Mergny, J. L. and J. C. Maurizot (2001). "Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat." Chembiochem : a European journal of chemical biology **2**(2): 124-132.
54. Michaela, V., ccaron, et al. (2006). "Ethanol is a better inducer of DNA guanine tetraplexes than potassium cations." Biopolymers **82**(3): 253-260.
55. Mikami-Terao, Y., M. Akiyama, et al. (2009). "Antitumor activity of TMPyP4 interacting G-quadruplex in retinoblastoma cell lines." Experimental Eye Research **89**(2): 200-208.
56. Monchaud, D. and M.-P. Teulade-Fichou (2008). "A hitchhiker's guide to G-quadruplex ligands." Organic & biomolecular chemistry **6**(4): 627-636.
57. Monsu Scolaro, L., M. Castriciano, et al. (2002). "Nucleation effects in the aggregation of water-soluble porphyrin aqueous solutions." Physica A **304**(1): 158-169.
58. Onyshchenko, M. I., T. I. Gaynutdinov, et al. (2009). "Stabilization of G-quadruplex in the BCL2 promoter region in double-stranded DNA by invading short PNAs." Nucleic Acids Research **37**(22): 7570-7580.
59. Park, T., J. S. Shin, et al. (2004). "Stacking of meso-Tetrakis(3-N-methylpyridiniumyl)porphyrin on Poly[d(A-T)2]: Importance of the Distance between Porphyrin's Positive Charges." The Journal of Physical Chemistry B **108**(44): 17106-17111.
60. Pavel, B., J. Joaquim, et al. (2009). "pH-Modulated Watson-Crick Duplex-Quadruplex Equilibria of Guanine-Rich and Cytosine-Rich DNA Sequences 140 Base Pairs Upstream of the <I> *c-kit*</I> Transcription Initiation Site." Chemistry - A European Journal **15**(46): 12663-12671.
61. Peter, H. (2003). "Chemistry of Water-Soluble Porphyrins." ChemInform **34**(37).
62. Phillip, A. R., B. Tom, et al. (2007). "Sequence effects of single base loops in intramolecular quadruplex DNA." FEBS letters **581**(8): 1657-1660.
63. Rankin, S., A. P. Reszka, et al. (2005). "Putative DNA Quadruplex Formation within the Human *c-kit* Oncogene." Journal of the American Chemical Society **127**(30): 10584-10589.
64. Rebecca, L. R. and G. M. David "Grading the commercial optical biosensor literature - Class of 2008: lsquoThe Mighty Bindersrsquo." Journal of Molecular Recognition **23**(1): 1-64.
65. Ren, J. and J. B. Chaires (1999). "Sequence and structural selectivity of nucleic acid binding ligands." Biochemistry **38**(49): 16067-16075.

66. Ru-Gang, Z., Z. Li-Jiao, et al. (2007). Theoretical researches on the molecular structures of DNA crosslinks induced by chloroethylnitrosoureas and QSAR analysis. Proceedings of the 1st WSEAS International Conference on Computational Chemistry. Cairo, Egypt, World Scientific and Engineering Academy and Society (WSEAS).
67. Sket, P., M. Crnugelj, et al. (2005). "Identification of mixed di-cation forms of G-quadruplex in solution." Nucleic Acids Research **33**(11): 3691-3697.
68. Wang, Z. X., N. R. Kumar, et al. (1992). "A novel spectroscopic titration method for determining the dissociation constant and stoichiometry of protein-ligand complex." Analytical biochemistry **206**(2): 376-381.
69. Weinraub, D., P. Peretz, et al. (1982). "Chemical properties of water-soluble porphyrins. 1. Equilibriums between some ligands and iron(III) tetrakis(4-N-methylpyridyl)porphyrin." The Journal of Physical Chemistry **86**(10): 1839-1842.
70. Yamashita, T., T. Uno, et al. (2005). "Stabilization of guanine quadruplex DNA by the binding of porphyrins with cationic side arms." Bioorganic & Medicinal Chemistry **13**(7): 2423-2430.
71. Yu, G. L., J. D. Bradley, et al. (1990). "In vivo alteration of telomere sequences and senescence caused by mutated Tetrahymena telomerase RNAs." Nature **344**(6262): 126-132.