The Investigation of DNA-Methyltransferase Interactions in the Adenine Methyltransferases using the Time-resolved Fluorescence of 2-Aminopurine

Eleanor Y.M. Bonnist

Degree of Doctor of Philosophy

The University of Edinburgh

2008

Art, method, way

Abstract

The time-resolved fluorescence of 2-aminopurine (2AP) has been used to investigate DNA base flipping by the adenine methyltransferases and to study aspects of the DNA-enzyme interaction.

2AP is an excellent fluorophore to probe base flipping in the adenine methyltransferases because, as demonstrated in the present work on M.TaqI, the 2AP is delivered into the same position inside the enzyme as the natural target adenine and with the same orientation that prepares the adenine for enzyme catalysis.

2AP emits two types of fluorescence when in DNA. The first is the well-known 370-nm emission, which emanates from 2AP as a monomer species. The second is 450-nm emission and comes from a 2AP which is π-stacked with a neighbouring DNA base, a heterodimer species. Additionally, 450-nm emission is produced by a 2AP-tyrosine or 2AP-phenylalanine heterodimer when a flipped 2AP is π-stacked inside a DNA-methyltransferase complex. Steady state fluorescence of the 2AP-heterodimer has been used to complement the time-resolved investigations.

Combined crystal- and solution-phase studies on M.TaqI have shown that when 2AP is flipped into the active site of M.TaqI it is significantly quenched by face-to-face π-stacking with the tyrosine from the NPPY catalytic motif. Not all of the flipped bases are held inside NPPY; in a minority of complexes, the flipped 2AP experiences very little quenching within the interior of the enzyme. In the sequence of bases recognised by M.TaqI, the thymine opposite the target adenine does not actively cause base flipping, as previously suggested, however, its presence aids the successful delivery of the target base into NPPY.

For the DNA-M.TaqI-cofactor ternary complex, the effect of varying the cofactor has been investigated. The use of 5'-[2(amino)ethylthio]-5'-deoxyadenosine (AETA) or sinefungin as cofactor analogue causes M.TaqI to show different base flipping behaviour compared with the natural cofactor S-adenosyl-L-methionine (SAM) and with the cofactor product S-adenosyl homocysteine (SAH). In the ternary complex containing SAM the flipped base is held the most tightly within the catalytic motif.

M.TaqI mutants have been studied in which the tyrosine (Y) in the NPPY motif is mutated to alanine (A) or phenylalanine (F). Stabilisation of the flipped base inside these mutants is more reliant on edge-to-face π-stacking with phenylalanine 196 and the available hydrogen-bonding in the adenine binding pocket. The NPPF-phenylalanine does not π-stack with the flipped base as NPPY-tyrosine does.

Solution-phase time-resolved fluorescence studies have confirmed that M.EcoRI and M.EcoRV use a base-flipping mechanism to extrude their target bases. For M.EcoRI, with sinefungin cofactor, the majority of the flipped 2APs are not held in the NPPF catalytic motif. When the natural SAM cofactor is used, however, the flipped 2AP strongly associates with NPPF inside M.EcoRI. Non-cognate sequence binding has been investigated, in which M.EcoRI encounters a base that is in almost the same sequence context as the methylation target. M.EcoRI forms some direct contacts with the pseudo-target adenine but does not extrude the base that is in a highly stacked position inside the duplex. The H235N mutant of M.EcoRI, measured under the same conditions as the wild-type enzyme, shows different behaviour to the wild-type enzyme in a small proportion of complexes, when bound to the cognate recognition sequence, and is far more discriminating than the wild-type when bound to the non-cognate sequence.

The M.EcoRV methyltransferase was found to be less efficient at flipping its target base than with M.TaqI or M.EcoRI. When M.EcoRV binds to its GATATC recognition sequence, the base-enzyme interactions of the target (GAT) and non-target (TAT) adenine position are shown to be quite different.

Declaration

I declare that the work presented in this thesis is my own unless otherwise stated by reference.

Signed

Date

Acknowledgements

I would like to express my sincerest gratitude to my supervisors Anita Jones and David Dryden for all their support and guidance throughout my PhD. I would also like to thank Robert Neely for taking the time to show me the ropes and then for his continued help and patience afterwards, for always answering my questions and for using his magic touch to fix the laser.

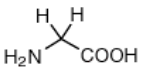
This work has involved a great deal of collaboration and its success has only been possible due to the dedication of my co-workers. I thank Thomas Lenz and Elmar Weinhold in the study on M.TaqI; Ben Youngblood and Norbert Reich in the study of M.EcoRI; and Kirsten Liebert and Albert Jeltsch in the study of M.EcoRV. Also thank you Ben and Mandy for putting me up and showing me around during my visit. I thank Katy Brimble who collected much of the data for the work on long-wavelength emission of 2AP in DNA duplexes.

Thanks to all in the Jones and the Dryden groups, to Patricia Richardson and Mark Tock for technical guidance, to Steven Magennis for assistance in COSMIC, and to Bansi, Nathan and Gemma for their friendship.

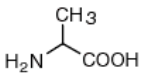
Finally I would like to thank the folk from Edinburgh TAGB and TAGB Scotland for making me feel so welcome, regularly giving me a good kicking and for reminding me of the things that really matter. Thanks to Jim Love, Angus Budge, Luke McGowan, Katie Elliott, James Watson, Paul Knell, Kerry Laird, Brian & Donna Laird, and everyone else, from white belts to world champions.

Most Common Amino Acids

Small

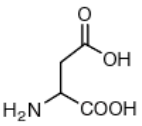


Glycine (Gly, G)

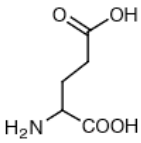


Alanine (Ala, A)

Acidic

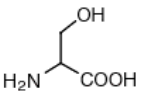


Aspartic Acid (Asp, D)
$pK_a = 3.9$

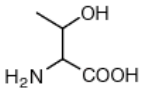


Glutamic Acid (Glu, E)
$pK_a = 4.07$

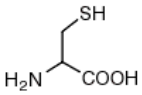
Nucleophilic



Serine (Ser, S)
$pK_a \sim 16$

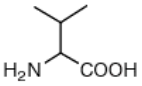


Threonine (Thr, T)
$pK_a \sim 16$

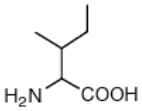


Cysteine (Cys, C)
$pK_a = 8.35$

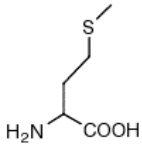
Hydrophobic



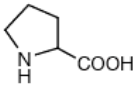
Valine (Val, V)



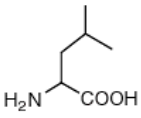
Isoleucine (Ile, I)



Methionine (Met, M)

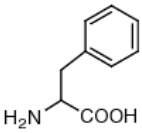


Proline (Pro, P)

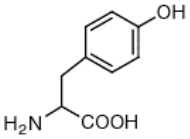


Leucine (Leu, L)

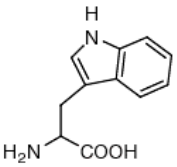
Aromatic



Phenylalanine (Phe, F)

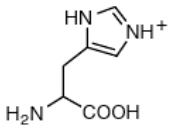


Tyrosine (Tyr, Y)

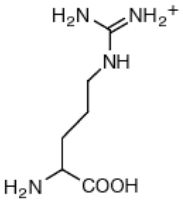


Tryptophan (Trp, W)

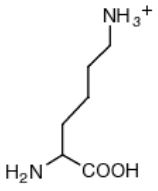
Basic



Histidine (His, H)
$pK_a = 6.04$

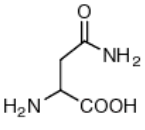


Arginine (Arg, R)
$pK_a = 12.48$

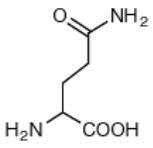


Lysine (Lys, K)
$pK_a = 10.79$

Amide



Asparagine (Asn, N)



Glutamine (Gln, Q)

Contents

Chapter 1

Introduction

Given the title of this thesis, this chapter introduces DNA methyltransferases, the biological system under investigation. DNA methyltransferases are proteins that bind to and modify DNA. Much progress has been made into understanding how these important enzymes work and the way in which they interact with their DNA substrate. Although many different DNA methyltransferases have been discovered, most lack a crystal structure. This has made other studies difficult to interpret, particularly with regards to the principal interaction between methyltransferases and DNA, called *base flipping*. Time-resolved fluorescence spectroscopy has the potential to uncover the nature of the DNA-methyltransferase interface at specific positions along the DNA base sequence. This chapter introduces this fluorescence-based strategy and discusses its advantages over other solution-phase methods. It describes an initial study where time-resolved fluorescence has provided clear and useful information about a particular DNA-methyltransferase complex with much success. Finally this chapter explains how this thesis will advance the current understanding of time-resolved fluorescence applied to DNA methyltransferase investigation.

1.1. DNA methyltransferases and base flipping

DNA methyltransferase enzymes catalyse the transfer of a methyl group from the cofactor S-adenosyl-L-methionine (SAM) to either adenine in the N6 position or cytosine in the C5 or the N4 position. Figure 1.1 shows the structure of SAM and Figure 1.2 shows the structures of the modified bases.

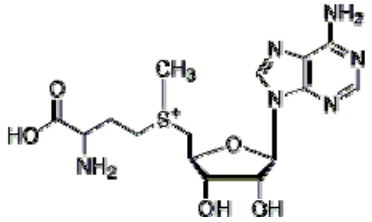


Figure 1.1 S-adenosyl-L-methionine (SAM).

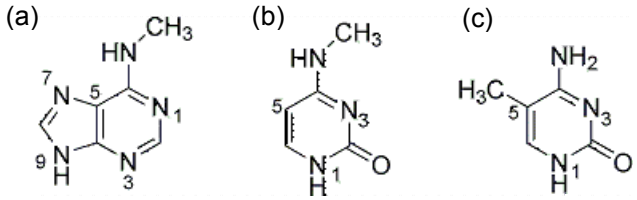


Figure 1.2 (a) N6-methyl adenine, (b) N4-methyl cytosine, (c) C5-methyl cytosine.

Methylation of the natural DNA bases adds extra information to the DNA sequence, without disrupting the structure of the double helix, and affects how the DNA is read and acted on by other proteins inside the nucleus. The target base for methylation lies within a specific base sequence that is recognised by the methyltransferase. This recognition sequence may exist in three states: unmethylated; hemimethylated, where the target is methylated in one strand of the DNA duplex; or fully methylated where the target is methylated in both strands of the duplex, and patterns of these states are built up along the DNA. Although this work focuses on DNA methyltransferases other types exist that methylate RNA, lipids, small molecules and proteins.

DNA methyltransferases are highly abundant in bacteria and mammals and DNA methylation plays important biological roles (see Jeltsch[1] for an excellent review). Methylated DNA is found to be crucial in the control of DNA transcription and

replication, base-mismatch repair, mutagenesis, gene expression, genomic imprinting (the choice of maternal or paternal inheritance), gene silencing and X-chromosome inactivation.

DNA methyltransferases are potential drug targets for a variety of diseases, including cancer, and there is great interest in developing strategies for methyltransferase inhibition. Mammals carry cytosine methyltransferases. Methylated cytosines are found in abundance in the DNA from cancerous cells (hypermethylation)[1;2]. Also a methylated cytosine is more readily converted to thymine than an unmethylated cytosine and so base-mismatch lesions are more likely to occur in methylated DNA[1;2]. The adenine methyltransferases are only found in bacteria. Inhibition of this type of methyltransferase might help to suppress pathogenic bacteria like *Salmonella thyphimurium* and *Neisseria meningitides*[1].

The majority of the methyltransferases identified in bacteria form part of restriction-modification systems (RM systems), which serve to protect bacteria from infection by invading bacteriophages. RM systems consist of a methyltransferase and corresponding endonuclease that act on the same recognition sequence. DNA methylation provides the means of distinction between native and foreign DNA. A methylated sequence will be left intact but an unmethylated one will cause the endonuclease to cleave the DNA, destroying this DNA so that it cannot be replicated inside the cell. Different bacteria carry many RM systems with different recognition sequences to reduce the likelihood that phage DNA escapes excision. RM systems are classified into four different groups according to subunit composition (recognition, methylation and excision subunits), recognition site, cofactor requirement and cleavage position. The methyltransferases presented in this work are all part of Type II RM systems, in which methyltransferase and endonuclease functions are performed by separate enzymes (see Wilson and Murray[3], Roberts *et al.*[4] and Sistla and Rao[5] for more extensive articles on RM systems and the REBASE database[6] for information concerning all aspects of RM systems).

Figure 1.3 shows a generalised scheme for the catalysed DNA methylation and the state of the DNA during this process.

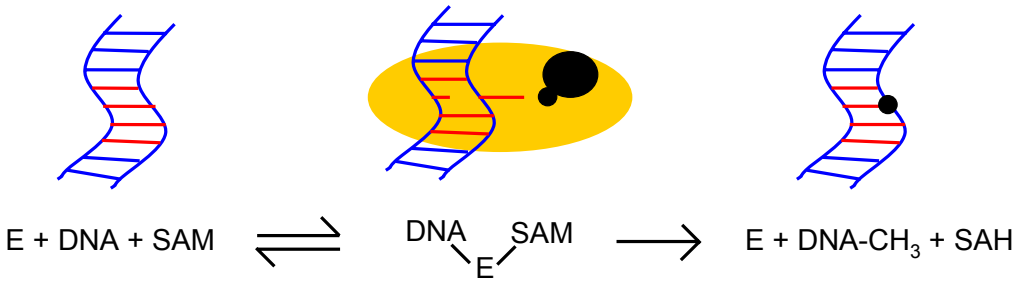


Figure 1.3 The process of methyl transfer to DNA by a DNA methyltransferase enzyme. The DNA is shown in blue and the base sequence recognised by the enzyme is in red. The methyltransferase enzyme is in orange and is denoted E in the scheme. The cofactor S-adenosyl-L-methionine (SAM) is coloured black. SAH is the product formed from SAM after methyl transfer.

The enzyme binds to DNA and to the cofactor SAM to form a ternary complex, catalysis occurs and then methylated DNA is released along with the demethylated version of the SAM cofactor S-adenosyl-L-homocysteine (SAH).

The first stage in Figure 1.3 may occur in an ordered Bi-Bi mechanism with either the enzyme binding to the DNA first and the cofactor second, as with the C5-cytosine methyltransferases[7], or the cofactor first and the DNA second, such as the adenine methyltransferases M.EcoRI[8] and M.EcoRV[9]. Some methyltransferases simply bind their substrates in a random order, such as the adenine methyltransferase M.TaqI for example[7]. Next the enzyme must locate its recognition sequence (shown in red in Figure 1.3) within the huge amount of bases carried by plasmid DNA. This occurs by *facilitated diffusion*, a one-dimensional sliding of the enzyme along the DNA[10]. If the methyltransferase acts in a *processive* mode, as T4 Dam does, the enzyme will continue sliding along the DNA searching for a second recognition sequence after methylating the first one. Alternatively, methylation occurs by a *distributive* mechanism, in which the enzyme dissociates from the DNA after the first methylation event, and then re-associates for the second methylation. This mechanism requires short-range diffusive hopping motions and is employed by all methyltransferases that are part of RM systems.

When the enzyme arrives at its specific target sequence, and in order for methyl transfer, the target base needs to be brought towards the methyl group of the cofactor and into intimate contact with the enzyme. The methyltransferase does this by rotating the target base by up to 180^{o} around the DNA sugar-phosphate backbone so that the base ends up in an extrahelical position[7;11]. This physical mechanism is known as *base flipping* (although *nucleotide flipping* is the correct term since the pentose sugar group is extruded too). Base flipping is illustrated in Figure 1.4, which shows the first example found by Klimasauskas *et al.* for the cytosine methyltransferase M.HhaI[11].

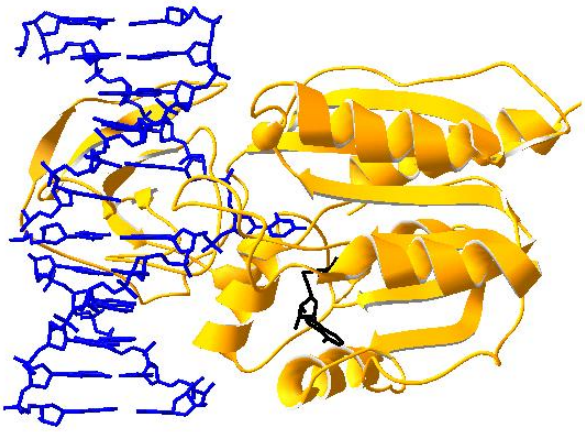


Figure 1.4 The X-ray crystal structure of the base-flipped DNA-M.HhaI-SAH ternary complex[11]. The DNA is blue, the enzyme is orange and the SAH cofactor product is black (structure resolution is 2.60 Å, protein database accession code 1mht).

As can be seen, base flipping imposes a dramatic structural distortion in the duplex in the vicinity of the target base. Besides DNA methyltransferases, base flipping is also performed by some DNA glycosylases[12;13]. Interestingly, when the target cytosine for M.HhaI flipping is removed, leaving an abasic site at that position in the duplex, M.HhaI still moves the sugar-phosphate backbone to the flipped out position[14]. Also, M.HhaI flips an adenine or a uracil base substituted at the target cytosine position. Therefore base flipping appears to be a consequence of DNA backbone manipulation and has no preference for the specific base that is extruded[12].

Characteristic kinetic parameters associated with Figure 1.3 are DNA biding affinity (K_m), the rate of catalysis (k_{cat}), and the specificity of methylation of the target sequence (k_{cat}/K_m). Methyltransferases may have a different affinity for unmethylated and hemimethylated sequences, and generally have higher affinities for longer DNA substrates (e.g. λ-DNA vs. short 10-20 base-pair oligonucleotide duplexes)[7]. Methylated substrate release is the rate-limiting step.

The methyltransferases that have been characterised have very little amino-acid sequence similarity but high similarity with regards to structure[15;16]. With one exception, all share a common core structure referred to as the *SAM-dependent methyltransferase fold* comprising a mixed seven-stranded β-sheet[12]. Strand 7 lies antiparallel to the other six strands and is inserted into the sheet between 5 and 6 (6↓7↑5↓4↓1↓2↓3↓). This structure forms the cofactor binding site and the catalytic centre of the enzyme. As may be realised from Figure 1.2 the DNA methyltransferases can be divided into endocyclic and exocyclic methyltransferases, or into carbon-carbon and nitrogen-carbon bond-forming enzymes (the same groups occur for both). The endocyclic methyltransferases share ten conserved amino acid sequence motifs (labelled I–X), nine of these are conserved across the exocyclic methyltransferases (I-VIII and X). The exocyclic methyltransferases are further subdivided into six groups (α, β, γ, ζ, δ, and ε) according to the possible linear order of three of these conserved motifs[7]; motif I, the SAM binding domain (FXGXG); the target recognition domain, responsible for sequence specific binding (abbreviated as TRD); and motif IV, the catalytic domain (S/N/D PP Y/F/W, sometimes generally referred to as *the DPPY motif*). Most of the exocyclic methyltransferases fall into the α, β and γ subgroups. All the methyltransferases presented in this study are N6-adenine methyltransferases but each is in a different subgroup; M.TaqI is γ, M.EcoRI is β and M.EcoRV is α.

Figure 1.5 illustrates the proposed mechanism by which the N6-adenine amine group is methylated[7]. It is an S_N2 reaction and proceeds with inversion of symmetry at the N6-nitrogen.

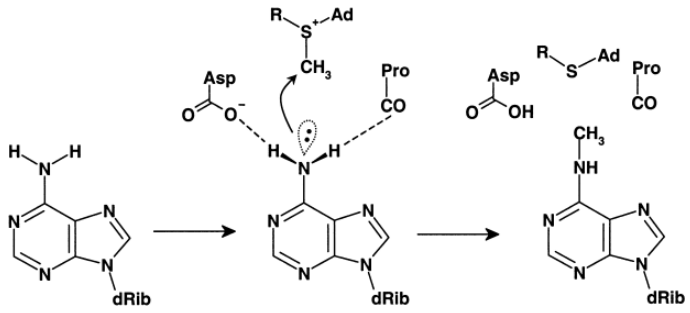


Figure 1.5 The catalytic mechanism of N6-adenine methylation[1]. The figure is based on the structure of the M.TaqI-DNA complex, the aspartate (Asp) and proline (Pro) residues in the figure are features of the D/N/S PP Y/F/W catalytic motif, Ad denotes the adenine moiety of the SAM cofactor.

Figure 1.6 shows the flipped adenine and the NPPY catalytic motif of M.TaqI from a crystal structure solved by Goedecke *et al.*[17]

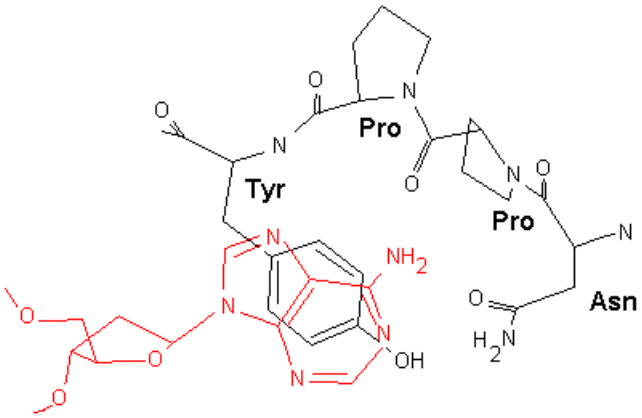


Figure 1.6 The flipped adenine (red) from DNA surrounded by the NPPY catalytic motif amino acids of M.TaqI (black) from the DNA-M.TaqI-cofactor ternary crystal structure[17].

The exocyclic amino group on adenine requires significant activation for nucleophilic attack. As shown in Figure 1.5, it has been proposed that the N6 amino group hydrogens of adenine form a hydrogen bond to the asparagine and to one of the prolines, and that this hydrogen bonding polarises and activates the nitrogen. The methyl group is transferred from SAM before proton release and so a positively charged intermediate forms which is stabilised. The aromatic ring of the adenine is in contact with the conjugated π-system of the tyrosine (and of a nearby

phenylalanine, see Chapter 3). After -$N(CH_3)H_2^+$ deprotonation, the conjugation between the lone pair of N6 and the adenine ring is restored, and then repulsion between the adenine and the tyrosine causes the modified base to move back into the helix[7].

1.2. The Identification of base flipping in DNA methyltransferases

Although all DNA methyltransferases are thought to use a base flipping mechanism it has only been proven in the handful of cases where a base-flipped DNA-enzyme crystal structure is available: M.HhaI[11], M.HaeIII[18], M.TaqI[17], T4 Dam[19], and EcoDam[20]. So few crystal structures have materialised because the co-crystallisation of protein, DNA, and cofactor is a difficult and time-consuming task. Even on production of crystals, reliable X-ray diffraction and reasonable data resolution is not guaranteed. Crystal structures give the most definitive evidence of base flipping and provide an excellent basis for investigating the DNA-methyltransferase interaction; however, the structures reveal little about system dynamics. Also it is assumed that the conformation that crystallises occurs in vivo and has not been influenced by crystal packing constraints. There is, therefore, a need to have solution-based techniques available to confirm the occurrence of the base flipping mechanism and to allow complementary studies on base flipping behaviour.

In an attempt to identify the precise region of DNA enveloped by M.EcoR124I, Mernagh and Kneale carried out hydroxyl radical footprinting[21]. In this technique the hydroxyl radical is generated by the reduction of hydrogen peroxide by iron(II) and cleaves the DNA backbone indiscriminately at the deoxyribose sugars. Backbone deoxyriboses that are protected by bound protein are cut with a lower efficiency[22]. Mernagh and Kneale found that the target adenines on each strand of the enzyme-bound DNA are highly susceptible to hydroxyl radical cleavage compared with the unbound duplex. Hyper-accessibility of these adenines support that these bases might be flipped[21]. Mernagh and Kneale present maps of the DNA substrate base sequence as a line scan of the cleavage pattern (i.e. intensity of each band on the gel) and as a plot of ratio of hydroxyl radical cutting frequencies

between the bound and unbound DNA. There are peaks at the target adenine position for the enzyme complex map but each peak extends over 2-3 bases, therefore the method reveals that a set of bases encompassing the target adenine is likely to be extra helical and exposed but has not isolated a single site of base flipping. It is possible that several bases are extruded by M.EcoR124I, however in the DNA-M.HhaI structure in Figure 1.4 it is clear that the target cytosine is flipped and that the flanking bases are not.

Another chemical footprinting method employs $KMnO_4$ oxidation. Thymine oxidation yields 5,6-dihydroxy-5,6-dihydrothymine (thymine glycol) and then selective cleavage of the glycosidic bond with piperidine allows DNA strand breakage[23]. In principle, an exposed thymine should show enhanced reactivity with the oxidiser. This is indeed what was observed by Serva *et al.* for DNA-M.HhaI and DNA-M.TaqI complexes where the target base in the recognition sequence of each enzyme was replaced by a thymine and the complexes subjected to $KMnO_4$ oxidation[24]. For the M.HhaI-DNA complex there is a clear single gel band corresponding to the target cytosine position. For the DNA-M.TaqI complex footprint, however, the bands from base sites either side of the methylation target position are also significantly intense. This footprinting method has failed to identify that only one base is flipped by M.TaqI[17] (see Chapter 3). An inherent problem with this method is that oxidation only occurs on thymine base and not on the natural cytosine or adenine targets for these methyltransferases. Although methyltransferases are known to flip their target regardless of which specific base is there[12] Serva *et al*. admit that the substitution of thymine is problematic[24]. Introducing a thymine into the recognition sequence changes the recognition sequence and creates a base pair mismatch, so that a T-T (pyrimidine) pair is present in the M.TaqI sequence and a T-G pair in the M.HhaI sequence. Serva *et al.* note that mismatch lesions may have additional conformational flexibility and so the background signal of the unbound DNA footprint will be affected. Also, protein binding might induce conformational changes to the duplex which are unique to the mismatched base pair. They admit the possibility that the environment of a flipped thymine inside M.HhaI may be markedly different from a flipped cytosine in ternary

complex. Serva *et al*. further propose that there might be open and closed ternary complex conformations to explain why a flipped thymine is exposed and susceptible to permanganate oxidation inside M.HhaI when a flipped cytosine is tightly surrounded and shielded by the cofactor and active site loop[24].

To exploit the possibility that a flipped adenine is in close contact with an aromatic amino acid side chain inside the adenine binding pocket of M.EcoRV, Beck and Jeltsch[25] used cross-linking of 5-iododeoxyuracil (^{5I}U). When photoactivated, ^{5I}U becomes a uracil radical which will cross-link to phenylalanine, tyrosine, tryptophan, methionine and histidine. Cross-linking is enhanced if a π-stacking interaction exists between the ^{5I}U base position and an aromatic amino acid side chain[26;27]. Beck and Jeltsch substituted ^{5I}U for each adenine in the GATATC recognition sequence of M.EcoRV in turn. In this sequence the first adenine is methylated by M.EcoRV but the second is not. ^{5I}U at the target adenine position showed the highest conversion to cross-linked product, however, there was a substantial yield for the non-target adenine position too. Beck and Jeltsch do not expect M.EcoRV to flip both adenines so reason that the non-target adenine position has cross-linked with an intercalating aromatic side chain from the enzyme[25]. In practice, ^{5I}U cross- linking enhancement is not unique to the flipped base.

Klimasauskas *et al.* devised an NMR-based method to monitor M.HhaI base flipping and prepared a DNA substrate in which the target cytosine bases were replaced with a fluorine-19 labelled cytosine analogue, 5-fluorocytosine[28]. An additional 5-fluorocytosine was positioned three nucleotides away to provide an internal reference. On addition of enzyme and cofactor to the DNA the ^{19}F signal shifted upfield, from approximately -86 to -97 ppm, and this was taken to indicate that 5-fluorocytosine at the target position had been flipped[28]. Methyltransferases greatly accelerate base flipping compared with spontaneous DNA breathing, however, when Klimasauskas *et al*. measured spin relaxation times to monitor base flipping dynamics their results reported that M.HhaI did not increase the rate of base pair opening much beyond the intrinsic rate of base pair opening[28] (~10 ms). The conditions required for this experiment are problematic. Signal intensity is highly

dependent on ionic strength and some measurements could not be run when ionic strength was reduced because the signal became too weak to distinguish from background. Also, the data acquisition times are greatly impractical. DNA-enzyme complexes were measured continuously for 11, 48 and even 120 hours and it is not surprising that some samples precipitated out. Since many enzymes degrade after a few hours this technique is deemed unsuitable for studying other methyltransferases.

The attempts to identify base flipping by methyltransferases in solution described above have had limited success. An improved approach needs to encompass certain criteria. It needs to be non-destructive to the sample, to work under mild conditions, to be non-disruptive to enzyme behaviour, to provide unambiguous results and to have a reasonable data collection time. This discussion now turns to fluorescence methods since fluorescence potentially adheres to most of these requirements.

1.3. Detecting base flipping with fluorescence methods

1.3.1. The principles of fluorescence and fluorescence decay

The energies associated with radiation from the ultra-violet and visible region of the electromagnetic spectrum match the spacing between successive electronic energy levels in atoms and molecules. Consequently the absorbance of this incoming light by atoms and molecules causes the outermost electrons within them to be promoted to higher energy states, as is illustrated schematically in Figure 1.6. The process is also referred to as *electronic excitation* since absorbance leaves these molecules in an energetically elevated state compared with the ground state electronic configuration.

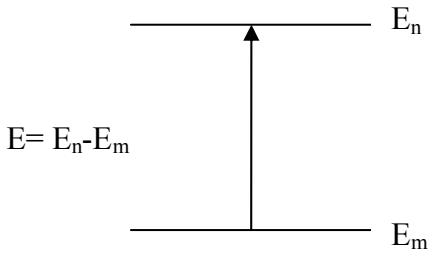


Figure 1.6 Absorption of radiation with energy, E, resulting in an electronic transition from the ground state, m, having a potential energy E_m, to an excited state, n with potential energy E_n.

The energy (E) associated with an electronic transition is discrete and corresponds to a particular frequency (ν) or wavelength (λ) characteristic for that transition. The relationship between these is given in Equation 1.1, h is Planck's constant and c is the speed of light.

Equation 1.1
$$E = h\nu = \frac{hc}{\lambda}$$

For molecules, using molecular orbital terminology, the promoted electron comes from the highest occupied molecular orbital (HOMO) and goes to the lowest unoccupied molecular orbital (LUMO). The excited state can be one of two spin types; a singlet state, in which the electrons in the HOMO and LUMO have opposite spins, or a triplet state, in which these electrons have parallel spins. This is shown pictorially in Figure 1.7.

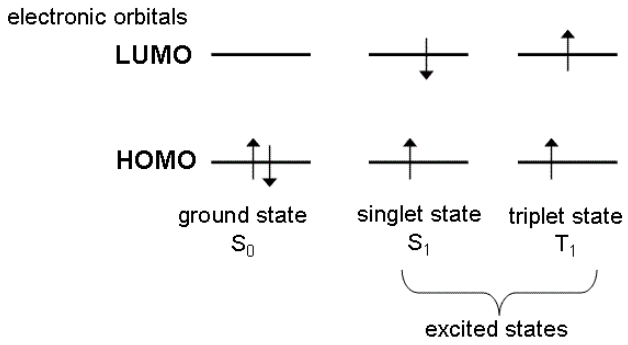


Figure 1.7 An illustration of ground and excited state spin states in molecules. HOMO is the highest occupied molecular orbital, LUMO is the lowest unoccupied molecular orbital, electrons are represented by arrows and their orientation represents the spin of the electron.

In accordance with quantum-mechanical selection rules, the spin of an electron cannot change but in real systems *spin flipping* is possible. This is because states of different multiplicities can mix, termed *spin-orbit coupling*, and so the existing spin states in real systems are not pure. Thus, there is a small but finite probability that a spin-forbidden singlet to triplet transition can occur. This concept is illustrated next in the discussion of excited state relaxation.

Once excited, a molecule can relax back down to its electronic ground state by a number of different pathways. Figure 1.8 presents a Jablonski diagram of electronic (S,T) and vibrational (ν) energy levels to help describe the excitation and de-excitation pathways. This diagram is a simplified scheme that neglects molecular shape and size. In the S_n^{ν} terminology, n refers to the n^{th} singlet electronic level and ν to a higher-than-ground state vibrational level. This thesis is concerned with photophysical relaxation processes, photochemical processes also exist but will not be discussed here.

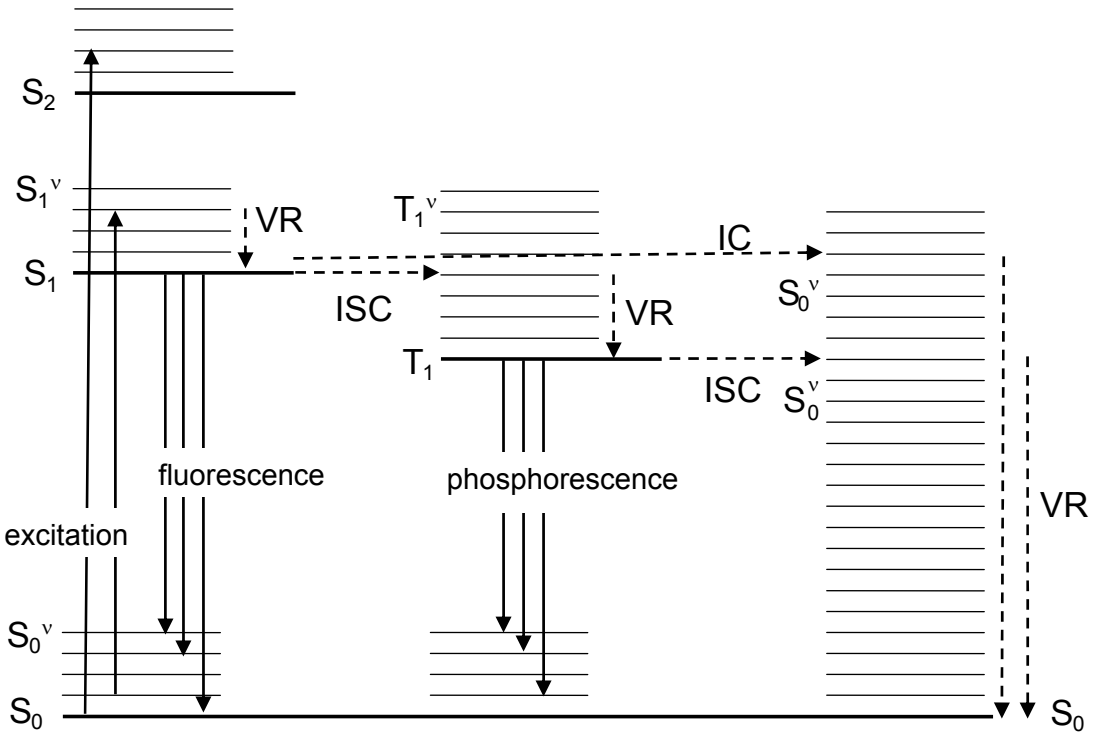


Figure 1.8 A Jablonski diagram illustrating electronic excitation and excited state relaxation pathways. S_n and T_n denote the n^{th} electronic singlet and triplet state levels respectively, ν denotes vibrational sub-levels within each electronic state, solid-line arrows represent radiative transitions, broken-line arrows represent non-radiative transitions, VR is vibrational relaxation, ISC is inter-system crossing and IC is internal conversion.

An electron can be directly promoted to S_1 or a higher electronic level during excitation. Usually, higher vibrational energy states within the S_1 level become populated. After excitation, the excited molecule rapidly moves to the lowest vibrational state of S_1 by vibrational relaxation (VR). VR defines energy loss through vibrational modes, the excess energy is transferred to heat and dissipated into the surrounding environment. Processes that return the molecule to its ground state with the emission of a photon are termed *radiative* processes. Fluorescence describes the radiative transition between states of the same spin (e.g. singlet to singlet) and phosphorescence defines a radiative transition between states of differing spin. The triplet state is reached through inter-system crossing (ISC). ISC describes the transition between states of different spin with no loss in electronic and vibrational energy. Internal conversion (IC), describes another isoenergetic transition but between states of the same spin. In Figure 1.8, IC transfers the energy

from the electronically excited S_1 state, to higher vibrational levels in the electronic ground state S_0^v. As can be seen, combinations of IC, ISC and VR can result in *non-radiative* relaxation.

Each pathway described above has an approximate rate associated with it, giving the state a corresponding lifetime (see later for a formal definition). Their timescales are compared in Figure 1.9.

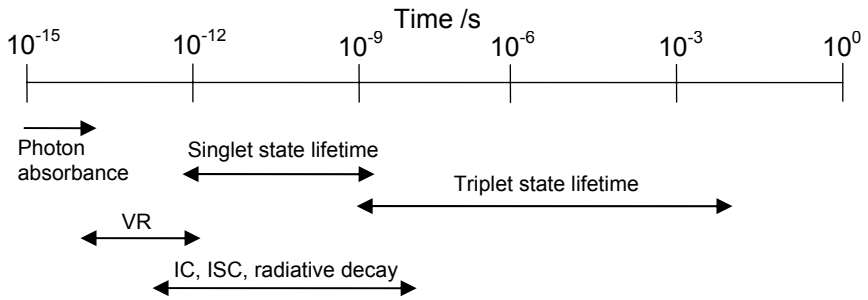


Figure 1.9 Approximate timescales for photon absorbance, photon emission, vibrationl relaxation (VR), inter-system crossing (ISC), internal conversion (IC), and excited state lifetimes.

VR is the fastest relaxation process and is complete before the occurrence of any radiative emission. The spin-allowed fluorescence persists for around 10^{-8} s whereas, by comparison, phosphorescence lasts longer and occurs later. This is because the rates of ISC and phosphorescence are relatively low given that these are both spin-forbidden process. The probability of ISC, and thus its rate, can be enhanced by increasing the singlet-triplet state mixing (spin-orbit coupling). This can be brought about by the presence of heavy atoms such as iodine. IC is spin allowed and thus is a rapid transition. IC relies on the activation of additional vibrational modes and this non-radiative process can be suppressed by immobilising the molecules in a rigid matrix.

Radiative and non-radiative relaxation mechanisms compete. Figure 1.8 only illustrates the intramolecular processes governing the fate of the excited state but intermolecular interaction opens up additional pathways for excited state relaxation. *Quenching* describes the process in which an external molecule causes ground state

return, further suppressing radiative emission. There are static and dynamic modes of quenching. In dynamic quenching the excited molecule loses its excess energy through collisions with the quencher. In static quenching the excited molecule and the quencher combine to form a non-emissive complex or a complex that emits photons of significantly reduced energy. Quenching can also occur by energy transfer when the HOMO-LUMO electronic energy gap of the nearby quencher molecule matches the energy needing to be dissipated by the donor excited molecule.

The measurable properties of fluorescence are wavelength (λ, see above) intensity (I) and lifetime (τ). Fluorescence intensity is defined as the number of photons emitted per unit time. I is directly related to the population of excited state molecules M* and the collective rate constant for radiative processes, k_r. Typically the rate of phosphorescence is three orders of magnitude smaller than the rate of fluorescence and so k_r is dominated by fluorescence.

Equation 1.2 $$I = k_r[M^*]$$

The total fluorescence decay rate of the excited state population is first order and given in Equation 1.3. k_F is the combined rate constant for the radiative (r) and non-radiative (nr) processes of de-excitation, t is time and [M*] is the concentration of excited state molecules.

Equation 1.3 $$\frac{-d[M^*]}{dt} = k_F[M^*] = (k_r + k_{nr})[M^*]$$

The instantaneous population of M* at a given time is obtained by the integration of Equation 1.3 to give Equation 1.4, $[M^*]_0$ is the concentration of excited state molecules at t = 0.

Equation 1.4 $$[M^*](t) = [M^*]_0 \exp(-k_F t)$$

By combining Equations 1.4 and 1.2 an expression can be obtained for the decay of fluorescence intensity with time:

Equation 1.5 $$I(t) = I_0 \exp(-k_F t)$$

Given that fluorescence lifetime is the inverse of the fluorescence decay rate:

Equation 1.6 $$\tau = \frac{1}{k_F}$$

τ can be calculated from experimental measurement of the decay of fluorescence intensity:

Equation 1.7 $$I(t) = I_0 \exp\left(\frac{-t}{\tau}\right)$$

For simple systems fluorescence lifetime is further defined as the time taken for the fluorescence intensity to fall to 1/e of its initial magnitude I_0.

Another useful characteristic quantity is the fluorescence quantum yield Φ, this is a measure of the efficiency of a fluorophore:

Equation 1.8 $$\Phi = \frac{\text{number of photons emitted}}{\text{number of photons absorbed}} = \frac{k_r}{k_r + k_{nr}}$$

1.3.2. The use of 2-Aminopurine

The solution-phase and fluorescence-based strategy adopted in this thesis was first proposed by Allan and Reich[29] for the investigation of the adenine methyltransferase M.EcoRI. Since adenine is effectively non-fluorescent, the target adenine in the DNA recognition sequence of the enzyme is replaced with a closely related and fluorescent analogue, 2-aminopurine (2AP).

The structures of 2AP and adenine and their fluorescence quantum yields are given in Figure 1.9. As can be seen 2AP differs from adenine (6-aminopurine) only by the location of the exocyclic amine group, but this repositioning results in a dramatic increase in fluorescence quantum yield, of almost 3 orders of magnitude.

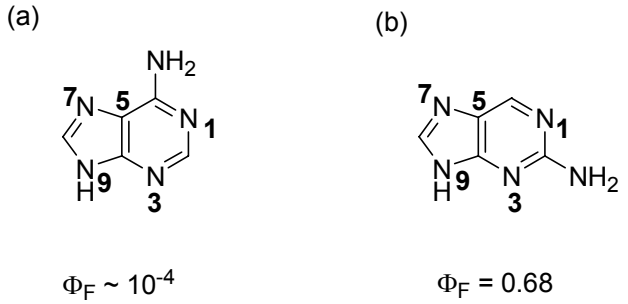


Figure 1.9 (a) 2-aminopurine (2AP) and (b) adenine, and their fluorescence quantum yields Φ_F.

2AP is highly photostable under physiological conditions and will hydrogen-bond and stack with the natural DNA bases. Thus, 2AP can be substituted into DNA at a specific sequence position and maintains the structure of the double helix. 2AP in DNA is usually excited with a 310- or 320-nm wavelength and maximum emission occurs at approximately 370 nm. These irradiation wavelengths do not excite the natural bases (λ_{max} ~280 nm) or the fluorescent amino acids tyrosine and tryptophan when the DNA is bound to protein, hence 2AP emission can be selectively excited and monitored. Consequently 2AP has proved popular for the study of local DNA conformation and dynamics[30-33], and of DNA- and RNA-protein interactions[29;34-36].

When 2AP is placed inside DNA its fluorescence decreases dramatically. The main cause of the fluorescence quenching is charge transfer to the neighbouring natural DNA bases and it is mediated by base-stacking interactions and base dynamics[37]. In accordance with the oxidation potentials of the individual DNA bases in solution, the most efficient charge transfer partner is guanine[38].

Since a 2AP molecule is strongly fluorescent when free in solution, but very weakly fluorescent when inside DNA, the fluorescence intensity of an extrahelical 2AP is high whereas the intensity of an intrahelical 2AP is low. Based on this premise, Allan and Reich rationalised that a stark increase in 2AP emission should be observed if the 2AP is brought extrahelical by methyltransferase-induced base flipping. Indeed, binding of M.EcoRI to 2AP-labelled DNA, with 2AP at the target position for methylation, induced a 14-fold increase in the fluorescence intensity of the 2AP[29].

Subsequently, 2AP intensity measurements were carried out on other methyltransferases in solution. The intensity increases for each (compared with the free DNA) are reported in Table 1.1.

Enzyme	Reported 2AP intensity increase(s) compared with unbound DNA (x-fold)	Ref.
M.HhaI	54	39
T4 Dam	50	40
M.EcoRI	14, 6	29;41
M.TaqI	13	39
M.KpnI	10-12	42
M.EcoPI	8-10	43
M.EcoKI	10	44
EcoDam	2.8	45
M.RsrI	1.7	41
M.EcoRV	1.4	9

Table 1.1 The x-fold increase reported for 2AP emission on methyltransferase binding to 2AP-labelled DNA, with 2AP substituted at the target site for methylation.

The extents of the observed intensity increases vary. Most enzymes cause around a 10-fold increase in 2AP intensity on DNA-enzyme binding, however, the signal

increases from EcoDam, M.RsrI and M.EcoRV binding are negligible and the rises observed for M.HhaI and T4 Dam binding are especially high. For the M.EcoRV and M.EcoPI enzymes a detectable signal increases also arises when 2AP is substituted at non-target sequence positions in the DNA, i.e. non-methylated positions. In the M.EcoPI-DNA complexes 2AP at a non-target position was equally fluorescent as 2AP at the target position, and in the M.EcoRV-DNA complexes emission from a non-target 2AP was higher than the target position 2AP.

These collective observations have brought about some interesting questions. One is whether the different sizes of signal increase, effected by different methyltransferase enzymes, are a result of different extents of flipping by each enzyme or by differently quenching environments for the 2AP once flipped inside that particular enzyme. Another is whether the increases observed for non-methylated base positions mean that these non-target bases are also flipped, or are even flipped instead of the target base. There is now a general consensus that a 2AP intensity increase is merely consistent with the unstacking of this base from the duplex. It is clear that the methyltransferases induce significant conformational change to the DNA at the site(s) probed by the 2AP, but the experiment cannot go much further in uncovering the nature of that conformation.

Recently, Neely *et al.* have adapted the fluorescence method developed by Allan and Reich and have chosen to monitor fluorescence lifetime rather than the fluorescence intensity of the 2AP, since lifetimes provide a more sensitive measure of DNA conformation and conformational change. Neely *et al.* replaced the cytosine target for M.HhaI base flipping (see Figure 1.4) with 2AP and compare the lifetimes of the 2AP-labelled DNA when the duplex is unbound and then bound by M.HhaI[36].

The fluorescence decay of 2AP inside DNA is complex and is best described by four lifetimes[30;31;36;37;46;47]. Multiple lifetimes identify multiple quenching environments for the 2AP. These different quenching environments relate to different conformations of the duplex, since there is only one 2AP in each duplex and in the same sequence position along the DNA, and the 2AP remains covalently attached to

the sugar-phosphate backbone. The first of the four lifetimes is tens of picoseconds in magnitude. This short lifetime, commonly referred to as τ_1, relates to a conformation of the duplex where the 2AP is highly base-stacked with its neighbouring DNA bases, allowing the most efficient charge transfer. Then there are two intermediate lifetimes, one around hundreds of picoseconds and the other a few nanoseconds. These correspond to duplex conformations where the 2AP experiences imperfect stacking conditions with the neighbouring bases, on account of the freedom of bases to twist or roll for example. The final lifetime, nearing 10 ns, is comparable to the lifetime of a solvent-exposed 2AP. In this duplex conformation the 2AP is envisaged to be extrahelical and in a solvent-like environment. Figure 1.10 provides a simplistic picture of the variation of stacking arrangements that would give rise to different lifetimes of the 2AP.

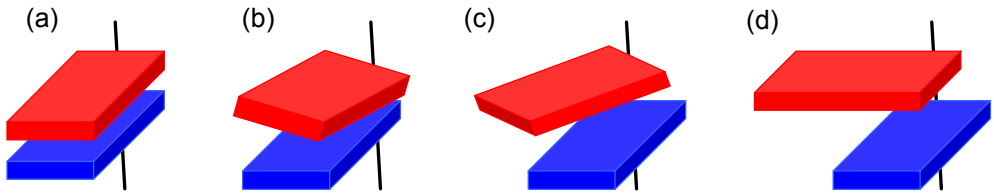


Figure 1.10 An illustration of various stacking conditions between bases (red and blue) tethered to the DNA backbone (black), with base-base overlap decreasing from (a) to (d).

In reality, each lifetime characterises a distribution of duplex conformations that provide similar quenching environments for the 2AP. For the purpose of this technique the duplex conformations present and, more importantly, the changes to these conformations can reliably be monitored through discrete lifetime components and the distribution of the emitting population amongst them. Typically, the shortest lifetime component constitutes well over half of the emitting population while the longest lifetime is usually found in just 10 percent abundance, thus most of the 2AP in DNA is tightly stacked and intrahelical.

For the investigation of M.HhaI, Neely, Daujotyte *et al.* produced a set of DNA-enzyme crystal complexes in which a variety of different base sequence positions had been replaced by 2AP[36]. Four duplex positions were labelled in turn; APtarget,

in which 2AP replaces the flipped cytosine; APopp, in which 2AP is the base opposite the flipped cytosine; APadj, located immediately adjacent to the flipped cytosine and in the same strand of the duplex; and APout, at a sequence position outside of the enzyme-bound region. From the crystal structures of these complexes (which all resemble Figure 1.4), the 2AP in the APtarget-M.HhaI complex was flipped, whereas 2APadj, 2APopp and 2APout all remain stacked within the DNA interior when in the enzyme complex[36]. Neely *et al.* then compared the fluorescence decay parameters of the 2AP in each position inside the enzyme complex. The APout lifetime parameters are characteristic of 2AP in unbound DNA, having the short τ_1 lifetime in 64% of the complexes. This is consistent with the fact that the 2AP in this complex is intrahelical and remains well stacked inside the duplex. The APtarget lifetime signature, the flipped base position, shows no τ_1 component since the 2AP does not stack with its neighbouring DNA bases in the enzyme complex. This is testament to base flipping. The decay parameters for APopp and APadj also reported that the majority of the 2APs in these sequence positions are intrahelical.

Across the four sets of lifetime data (one set for each duplex) it was clear that the lifetime signature of APtarget was unique and definitively identifies base flipping for this sequence position. Additionally, lifetime measurements report subtle differences in the intrahelical stacking of 2AP in APopp and APadj compared with APout for the enzyme bound duplexes[36]. Thus Neely *et al.* have shown that a time-resolved fluorescence approach uncovers the intricacies of even minor DNA conformational change, providing that the transient species occur within the timescale of excited state lifetime. Finally, Neely and co-workers repeated these measurements on the same DNA-enzyme combinations in solution and on the unbound DNA duplexes. The lifetime signatures of each solution-phase complex were nearly identical to their crystal-phase counterpart, furthermore comparison with the unbound DNA decay parameters verified that the differences between each duplex in the enzyme complex originate from enzyme binding. Through their work, Neely *et al.* provide a foundation for the use of 2AP fluorescence lifetimes as a viable solution-based method of probing base flipping. They suggest that the scope of investigation of base flipping and of DNA-methyltransferase interactions can be greatly expanded.

1.4. Project Overview

This thesis seeks to further establish the use of time-resolved fluorescence of 2AP in the study of DNA methyltransferases and aims to demonstrate that the current understanding of DNA-methyltransferase interactions has been developed by its application.

The systems chosen are all adenine methyltransferases. Chapter 3 presents a crystal- and solution-phase lifetime study of the M.TaqI methyltransferase. After fully characterising the solution-phase behaviour with the aid of a crystal structure, this chapter investigates how aspects of each part of the DNA-enzyme-cofactor ternary complex specifically affect base flipping; the role of the base which is paired with the flipping target in DNA; the mutation of the catalytic motif in the enzyme; and the effect of using molecules that are analogous to SAM as the cofactor.

Chapters 4 and 5 present studies on methyltransferases for which no crystal structure exists, M.EcoRI and M.EcoRV, and so are solely reliant on solution-based methods of investigation. Chapter 4 looks at the effect on base flipping behaviour of M.EcoRI when a single base variation is introduced into the recognition sequence bound by this enzyme (a *non-cognate* sequence). Also, a mutant of M.EcoRI is studied which has been identified as deficient at bending the DNA substrate compared with the wild-type enzyme and is more discriminating against the methylation of the adenine in a non-cognate DNA base sequence. Chapter 5 uses fluorescence lifetimes to distinguish differences in the DNA base-M.EcoRV interface for two bases within the DNA recognition sequence, one base is the target for methylation and the other is not.

Chapter 6 returns to a fundamental study of the fluorescence of the 2AP molecule inside a DNA duplex. In this chapter a novel emission band from the 2AP-labelled DNA is identified and characterised, and Chapter 7 demonstrates that this emitting species can be highly informative of the methyltransferase-DNA interaction for adenine methyltransferases.

1.5. References

(1) Jeltsch, A. *Chembiochem* **2002**, *3*, 275-293.

(2) Zingg, J. M.; Jones, P. A. *Carcinogenesis* **1997**, *18*, 869-882.

(3) Wilson, G. G.; Murray, N. E. *Annual Review of Genetics* **1991**, *25*, 585-627.

(4) Roberts, R. J.; Belfort, M.; Bestor, T.; Bhagwat, A. S.; Bickle, T. A.; Bitinaite, J.; Blumenthal, R. M.; Degtyarev, S. K.; Dryden, D. T. F.; Dybvig, K.; Firman, K.; Gromova, E. S.; Gumport, R. I.; Halford, S. E.; Hattman, S.; Heitman, J.; Hornby, D. P.; Janulaitis, A.; Jeltsch, A.; Josephsen, J.; Kiss, A.; Klaenhammer, T. R.; Kobayashi, I.; Kong, H. M.; Kruger, D. H.; Lacks, S.; Marinus, M. G.; Miyahara, M.; Morgan, R. D.; Murray, N. E.; Nagaraja, V.; Piekarowicz, A.; Pingoud, A.; Raleigh, E.; Rao, D. N.; Reich, N.; Repin, V. E.; Selker, E. U.; Shaw, P. C.; Stein, D. C.; Stoddard, B. L.; Szybalski, W.; Trautner, T. A.; Van Etten, J. L.; Vitor, J. M. B.; Wilson, G. G.; Xu, S. Y. *Nucleic Acids Research* **2003**, *31*, 1805-1812.

(5) Sistla, S.; Rao, D. N. *Critical Reviews in Biochemistry and Molecular Biology* **2004**, *39*, 1-19.

(6) The Restriction Enzyme Database (REBASE). 2007. http://rebase.neb.com

(7) Bheemanaik, S.; Reddy, Y. V. R.; Rao, D. N. *Biochemical Journal* **2006**, *399*, 177-190.

(8) Reich, N. O.; Mashhoon, N. *The Journal of Biological Chemistry* **1990**, *265*, 8966-8970.

(9) Gowher, H.; Jeltsch, A. *Journal of Molecular Biology* **2000**, *303*, 93-110.

(10) Halford, S. E.; Marko, J. F. *Nucleic Acids Research* **2004**, *32*, 3040-3052.

(11) Klimasauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. *Cell* **1994**, *76*, 357-369.

(12) Cheng, X.; Roberts, R. J. *Nucleic Acids Research* **2001**, *29*, 3784-3795.

(13) Stivers, J. T.; Jiang, Y. L. *Chemical Reviews* **2003**, *103*, 2729-2759.

(14) O'Gara, M.; Horton, J. R.; Roberts, R. J.; Cheng, X. D. *Nature Structural Biology* **1998**, *5*, 872-877.

(15) Dryden, D. T. F. Bacterial methyltransferases; In *S-adenosylmethionine-dependent methyltransferases : structures and functions*; Cheng, X. D., Blumenthal, R. M., eds. World Scientific: 1999.

(16) Martin, J. L.; McMillan, F. M. *Current Opinion in Structural Biology* **2002**, *12*, 783-793.

(17) Goedecke, K.; Pignot, M.; Goody, R. S.; Scheidig, A. J.; Weinhold, E. *Nature Structural Biology* **2001**, *8*, 121-125.

(18) Reinisch, K. M.; Chen, L.; Verdine, G. L.; Lipscomb, W. N. *Cell* **1995**, *82*, 143-153.

(19) Horton, J. R.; Liebert, K.; Hattman, S.; Jeltsch, A.; Cheng, X. D. *Cell* **2005**, *121*, 349-361.

(20) Horton, J. R.; Liebert, K.; Bekes, M.; Jeltsch, A.; Cheng, X. D. *Journal of Molecular Biology* **2006**, *358*, 559-570.

(21) Mernagh, D. R.; Kneale, G. G. *Nucleic Acids Research* **1996**, *24*, 4853-4858.

(22) Tullius, T. D.; Dombroski, B. A. *PNAS* **1986**, *83*, 5469-5473.

(23) Rubin, C. M.; Schmid, C. W. *Nucleic Acids Research* **1980**, *8*, 4613-4619.

(24) Serva, S.; Weinhold, E.; Roberts, R. J.; Klimasauskas, S. *Nucleic Acids Research* **1998**, *26*, 3473-3479.

(25) Beck, C.; Jeltsch, A. *Biochemistry* **2002**, *41*, 14103-14110.

(26) Willis, M. C.; Hicke, B. J.; Uhlenbeck, O. C.; Cech, T. R.; Koch, T. H. *Science* **1993**, *262*, 1255-1257.

(27) Wong, D. L.; Pavlovich, J. G.; Reich, N. O. *Nucleic Acids Research* **1998**, *26*, 645-649.

(28) Klimasauskas, S.; Szyperski, T.; Serva, S.; Wuthrich, K. *EMBO J.* **1998**, *17*, 317-324.

(29) Reich, N. O.; Allan, B. W. *Biochemistry* **1996**, *35*, 14757-14762.

(30) Guest, C. R.; Hochstrasser, R. A.; Sowers, L. C.; Millar, D. P. *Biochemistry* **1991**, *30*, 3271-3279.

(31) Nordlund, T. M.; Andersson, S.; Nilsson, L.; Rigler, R.; Graslund, A.; Mclaughlin, L. W. *Biochemistry* **1989**, *28*, 9095-9103.

(32) Johnson, N. P.; Baase, W. A.; von Hippel, P. H. *PNAS* **2004**, *101*, 3426-3431.

(33) Somsen, O. J. G.; Keukens, L. B.; de Keijzer, M. N.; van Hoeck, A.; van Amerongen, H. *Chemphyschem* **2005**, *6*, 1622-1627.

(34) Frey, M. W.; Sowers, L. C.; Millar, D. P.; Benkovic, S. J. *Biochemistry* **1995**, *34*, 9185-9192.

(35) Krosky, D. J.; Song, F. H.; Stivers, J. T. *Biochemistry* **2005**, *44*, 5949-5959.

(36) Neely, R. K.; Daujotyte, D.; Grazulis, S.; Magennis, S. W.; Dryden, D. T. F.; Klimasauskas, S.; Jones, A. C. *Nucleic Acids Research* **2005**, *33*, 6953-6960.

(37) Ross, J. B. A.; Osman, R.; Rachofsky, E. L. *Biochemistry* **2001**, *40*, 946-956.

(38) Barton, J. K.; O'Neill, M. A. *Journal of the American Chemical Society* **2004**, *126*, 11471-11483.

(39) Weinhold, E.; Serva, S.; Klimasauskas, S.; Holz, B. *Nucleic Acids Research* **1998**, *26*, 1076-1083.

(40) Malygin, E. G.; Evdokimov, A. A.; Zinoviev, V. V.; Ovechkina, L. G.; Lindstrom, W. M.; Reich, N. O.; Schlagman, S. L.; Hattman, S. *Nucleic Acids Research* **2001**, *29*, 2361-2369.

(41) Szegedi, S. S.; Reich, N. O.; Gumport, R. I. *Nucleic Acids Research* **2000**, *28*, 3962-3971.

(42) Bheemanaik, S.; Bujnicki, J. M.; Nagaraja, V.; Rao, D. N. *Biological Chemistry* **2006**, *387*, 515-523.

(43) Reddy, Y. V. R.; Rao, D. N. *Journal of Molecular Biology* **2000**, *298*, 597-610.

(44) Su, T. J.; Connolly, B. A.; Darlington, C.; Mallin, R.; Dryden, D. T. F. *Nucleic Acids Research* **2004**, *32*, 2223-2230.

(45) Liebert, K.; Hermann, A.; Schlickenrieder, M.; Jeltsch, A. *Journal of Molecular Biology* **2004**, *341*, 443-454.

(46) Neely, R. K. The Photophysical Properties of 2-Aminopurine and its Application as a Probe of DNA-Protein Interactions. Thesis 2005.

(47) Hochstrasser, R. A.; Carver, T. E.; Sowers, L. C.; Millar, D. P. *Biochemistry* **1994**, *33*, 11971-11979.

Chapter 2

Experimental Methods

2.1. Preparation of materials

2-Aminopurine ribonucleoside (2APr) was purchased from Sigma-Aldrich. It was dissolved in distilled water and used at a 2 μM concentration for steady state and time-resolved fluorescence measurement. The 2AP-labelled DNA duplexes and stocks of the single strand oligonucleotides were kindly provided by the following people:

2AP-T and 2AP-D
Elmar Weinhold, Institut für Organische Chemie der RWTH, Aachen

CBM or GPA(T) and A4 or APT(T)
Norbert Reich, University of California Santa Barbara

AP1 or GPT(T) and AP2 or TPT(T)
Albert Jeltsch, Jacobs University Bremen

GPC(T), APA(C) and GPT(C)
Geoff Kneale, University of Portsmouth

GPG(T), GPG(G) and APC(T)
David Dryden, University of Edinburgh

Duplexes were prepared by combining the 2AP-labelled strand with a 50% excess of the unlabelled strand to avoid the possibility of having any unpaired 2AP-labelled single strands in the sample. The DNA was buffered as advised by the donating party, and in accordance with stability requirements of the particular methyltransferase enzyme with which the duplex would be used. Buffer solution composition is detailed in the relevant results chapter. The DNA was annealed by heating to 95 ^{o}, holding the sample at this temperature for approximately 5 minutes, and then leaving it to cool slowly back to room temperature, typically over a period of several hours.

All water stocks, buffer solutions, enzyme samples and cofactors were checked for background fluorescence prior to their use in fluorescence experiments.

2.2. Steady-state fluorescence measurements

Steady state fluorescence spectra were measured on a Horiba Jobin Yvon Fluoromax® fluorimeter, equipped with a 150 W xenon lamp and separate excitation and emission monochromators.

To measure an emission spectrum, the excitation monochromator was fixed at one wavelength, the emission monochromator scanned across an appropriate wavelength range, and the emission signal intensity monitored at each different emission wavelength. To measure an excitation spectrum, the emission monochromator was fixed at one wavelength, the excitation monochromator scanned across an appropriate wavelength range, and emission signal intensity monitored at each different excitation wavelength.

The slit width set for both monochromaters corresponded to a 5 nm-wavelength resolution for the chosen excitation and emission wavelength. Integration times were set to between 0.1 and 0.4 s per wavelength increment, with the longer integration time used for lower signal intensities, so as to obtain a reasonably smooth spectral trace. Finally, measured signal intensities were internally corrected for wavelength-

dependent lamp performance and for background by the Instruments SA, Inc. DataMax software.

Solution-phase samples were measured in a Starna fused silica fluorescence micro cuvette with a capacity of 100 μl. Crystals of the DNA-M.TaqI methyltransferase-AETA complexes (see Chapter 3) were provided by Thomas Lenz. The crystals were sealed inside fused silica capillaries, 1 mm in diameter. The capillary was carefully mounted in a piece of foam that was held by a clamp and stand (see Section 2.3 below). Steady state fluorescence measurements of these crystals were carried out on a Horiba Jobin Yvon Fluoromax-P® fluorimeter. The excitation beam and fluorescence emission were directed between the spectrometer and the sample through Glen Spectra fused silica fibre optic bundles.

2.2.1. Measurement of samples at 77 K

The selected pre-annealed duplexes were dissolved in 10 M LiCl such that the concentration of the 2AP-labelled strand was approximately 3.3 μM. Between 600 μl and 1 ml of the DNA/LiCl solution was transferred to an NMR-style tube made of fused silica (prepared in house) and set aside. Separately, a Horiba Jobin Yvon fused silica liquid nitrogen Dewar assembly was filled with liquid nitrogen. The NMR-style tube containing the sample was attached to the apparatus lid and slowly lowered into the Dewar. On cooling, the tube was lifted out momentarily to check that the sample had been transformed a uniform glass which was free from cracks. The re-assembled apparatus, shown in Figure 2.1, was mounted inside the Fluoromax® and emission spectra were recorded as described in Chapter 6. During the experiment it was ensured that the outside of the Dewar stayed free from condensation.

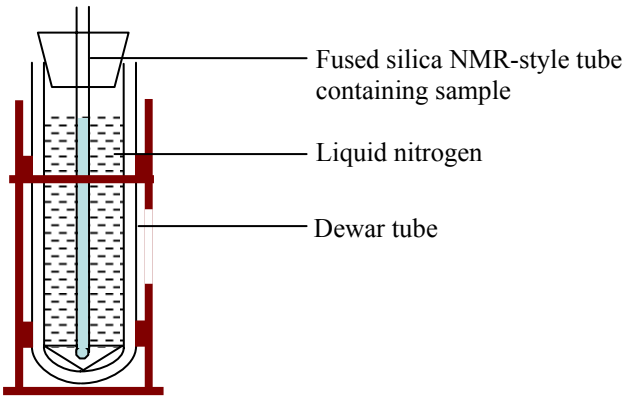


Figure 2.1 Apparatus used for the measurement of the fluorescence from 2AP-labelled DNA at 77 K.

Corresponding room-temperature measurements of the DNA/LiCl samples were carried out using the same assembly but without the liquid nitrogen. Background scans were performed with only the LiCl solution inside the fused silica tube for both temperature regimes. These background scans were subtracted from the measured spectra using the DataMax software to generate background-corrected spectra for each DNA duplex.

2.3. Time-resolved fluorescence measurements

Fluorescence decays were measured using the time-correlated single photon counting technique (TCSPC)[1] on an experimental set up provided in the Collaborative Optical Spectroscopy Micromanipulation and Imaging Centre (COSMIC, University of Edinburgh). The arrangement is described in Figure 2.2.

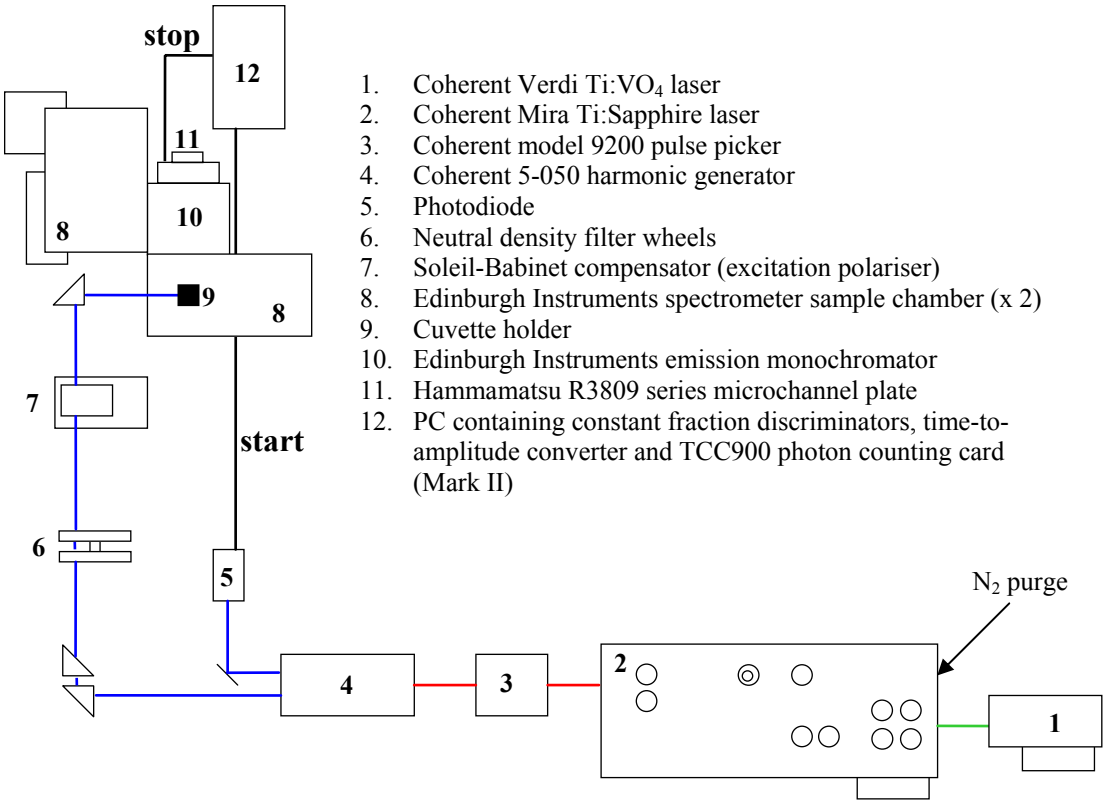


Figure 2.2 Arrangement for time-correlated single photon counting.

The excitation source consists of a Coherent Mira Ti:Sapphire femtosecond laser, which is tuneable over a 700-1000 nm wavelength range, pumped by a Coherent Verdi Ti:VO_4 laser operating at 10 W. The Mira output is mode-locked and produces ~200 fs pulses at a 76 MHz repetition rate. These are sent through a Coherent 9200 pulse-picker, which reduces the repetition rate to 4.75 MHz, such that a pulse is emitted every 210 ns. The beam is directed into a Coherent 5-050 harmonic generator where the desired excitation wavelength is obtained.

For the DNA-enzyme experiments (Chapters 3-5), the Mira was operated at approximately 960 nm and purged with a gentle stream of nitrogen, and the beam was frequency-tripled to generate 318-nm wavelength pulses. For the DNA experiments (Chapter 6), the Mira was operated at 720 nm and frequency-doubled to generate 360-nm wavelength pulses. Excitation wavelength resolution was approximately 10 nm. The excitation beam was attenuated with neutral density filters, if necessary, vertically polarised with a Soleil-Babinet compensator, directed into the Edinburgh Instruments spectrometer and loosely focussed into the centre of the cuvette holder.

Fluorescence decays were measured with a card-based system, using TCC900 photon counting electronics in conjunction with an Edinburgh Instruments spectrometer and F900 software. The *start* signal for the time-to-amplitude converter (TAC) was provided by a photodiode, triggered with the laser output from the harmonic generator. For experiments requiring 318-nm excitation (the third harmonic), the residual second harmonic beam was used to trigger the photodiode. For experiments requiring 360-nm excitation, either the excitation beam was split and sent to the photodiode or the attenuated fundamental beam was used as the trigger.

Fluorescence emission from the sample provided the *stop* signal. Sample emission was directed through a polariser set at the magic angle (54.7 ^{o}) and detected by a Hammamatsu R3809 microchannel plate photomultiplier tube. The emission wavelength was selected with an Edinburgh Instruments monochromator with the slits set for a wavelength resolution of 10-15 nm, depending on the intensity of emission. The 15-nm bandpass was used for 2AP-labelled DNA excited at 381 nm because emission intensity was so low. A 10-nm bandpass was used for the DNA-enzyme samples and for the DNA excited at 360 nm. Any scattered light from the excitation beam was blocked with a Schott longpass filter selected for the appropriate wavelength.

Prior to any experimental measurement the system was optimised by scattering the attenuated excitation beam into the spectrometer using a Ludox solution and

maximising the *stop* count rate detected by the F900 software. The wavelength of the excitation beam was checked by performing an emission scan on the unpolarised beam and with an Edinburgh Instruments IR-sensitive photomultiplier tube (connected to the spectrometer but not shown in Figure 2.2). The system was further optimised, and its performance monitored, by detecting the count rate produced by the fluorescence of a 2 μM aqueous 2AP solution. When excited at 318 nm and fluorescence emission detected at 370 nm, a 5000-Hz count rate from this standard works well for the DNA-enzyme experiments. Finally an instrument response function (IRF) was recorded using the Ludox solution. The IRF for this system had a 50-ps FWHM.

The Ludox scattering solution was held in a Starna 5-ml capacity fused silica fluorescence cuvette for the IRF measurement and system optimisation. Solution-phase samples were measured inside a Starna fused silica micro-cuvette, with 100-μl capacity, a 2 x 1 mm excitation window and a 2 x 10 mm emission window. In general, a 2-μM concentration of the 2AP-labelled DNA is used for the DNA-enzyme experiments. This proves a good compromise between getting a reasonable count rate from the DNA when unbound, using the enzyme at a practical concentration and keeping sample consumption to a minimum. The specifics of DNA-enzyme sample composition are given in the relevant results chapters. Figure 2.3 shows how crystal-phase DNA-enzyme samples were mounted inside the spectrometer.

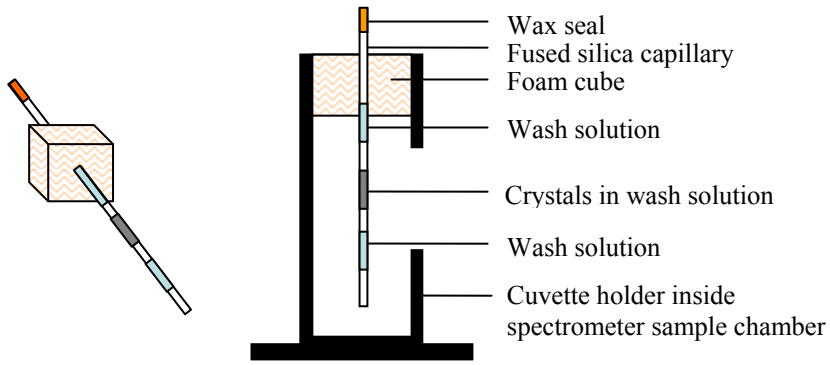


Figure 2.3 The mounting of crystal samples for fluorescence measurement.

The crystals were housed inside 1-mm diameter, fused silica capillaries. These sample capillaries were prepared by Thomas Lenz. The capillary was pushed through a hole in the centre of a foam cube, cut slightly larger than the dimensions of the cuvette holder. The cube was placed in the top of the cuvette holder and the height of the capillary was adjusted through the foam so that the crystals met with the excitation beam.

Initially an emission scan was measured to check the state of the sample and its wavelength of maximum emission. Fluorescence decays were measured over a 50 ns time window divided into 4095 channels, giving each channel a 12 ps width. Data was acquired until 10,000 counts had collected in the channel of peak intensity. The point of sample excitation was displaced by ~ 200 channels along the 4095-channel collection window to allow the background count level to be determined. Decays were collected for several different emission wavelengths from each sample.

2.3.1. Fluorescence decay analysis

Typical decay data are shown in Figure 2.4. Each individual decay was analysed with the Edinburgh Instruments F900 software, which uses a *non-linear least squares* curve fitting approach.

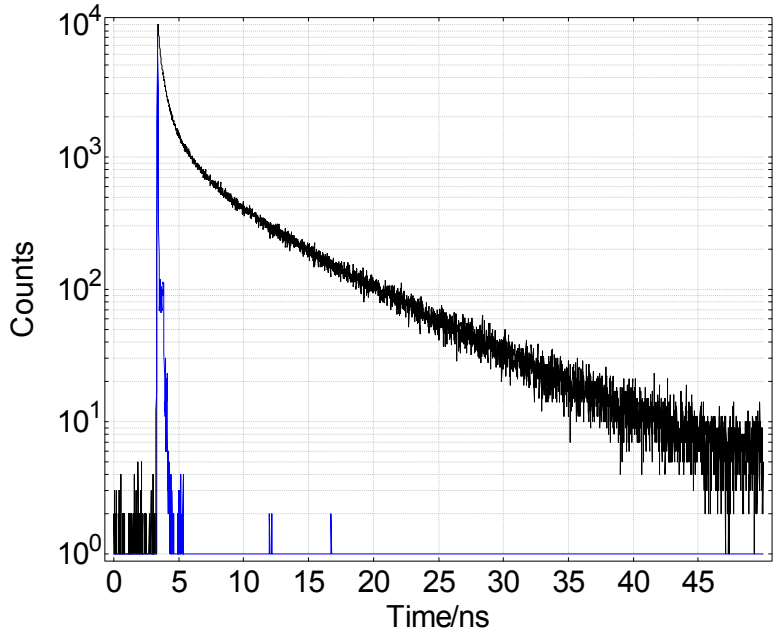


Figure 2.4 Fluorescence decay data for 2AP-labelled DNA, collected with the time- correlated single photon counting set up (black trace). The blue trace is a separately measured instrument response function.

The observed decay, I(t), is a convolution of the instrument response function, P(t), and the experimental decay of the 2AP, G(t),:

Equation 2.1 $$I(t) = P(t) \otimes G(t)$$

G(t) is obtained by iterative reconvolution and uses a previously recorded instrument response function (IRF) as P(t).

It is assumed that G(t) takes an exponential form and thus is fitted to Equation 2.2, where B is an integer corresponding to the background level, t is time, τ_i is the lifetime of the i^{th} component and A_i is the fractional amplitude of the i^{th} component.

Equation 2.2

$$G(t) = B + \sum_i A_i \exp\left(\frac{-t}{\tau_i}\right)$$

F900 creates a test function G(t)' with A_i and τ_i as variable parameters. G(t)' is convoluted with the IRF (Equation 2.1) to generate a calculated fitting function Y(t). Data points for the calculated function, $Y(t_i)$, are compared with the experimentally observed data points, $y(t_i)$, through the *chi squared* parameter, χ^2:

Equation 2.3

$$\chi^2 = \sum_i W_i [y(t_i) - Y(t_i)]^2$$

This is the weighted sum, with weighting factor W_i, of the squares of the deviations of corresponding points from the experimental test function. Equation 2.4 expresses χ^2 with the commonly used weighting factor:

Equation 2.4

$$\chi^2 = \sum_i \left\{ \frac{[I_0(t_i) - Y(t_i)]^2}{I(t_i)} \right\}$$

Where $I_0(t)$ is the background-corrected intensity decay and I(t) is the uncorrected function:

Equation 2.5

$$I_0(t_i) = I(t_i) - B$$

During the fitting procedure, χ^2 is minimised and is further evaluated through the *reduced chi squared* parameter χ_v^2:

Equation 2.6

$$\chi_v^2 = \frac{\chi^2}{n_2 - n_1 + 1 + p}$$

where n_1 and n_2 are the first and last data point, respectively, of the data fitting range, and p is the number of variable parameters. For a good fit, χ_v^2 is close to unity.

The suitability of the calculated function is further assessed by visual inspection and by the inspection of a plot of the weighted residual $r(t_i)$ for each point.

Equation 2.7

$$r(t_i) = \sqrt{W_i[I_0(t_i) - Y(t_i)]} = \frac{I_0(t_i) - Y(t_i)}{\sqrt{I(t_i)}}$$

Plotted residuals from successful fits should be randomly distributed about zero.

Decay curves are fitted to Equation 2.2 using the smallest number of exponential terms as possible that will give a reasonable value for χ_v^2, a good visual fit and randomly varying residuals. Further exponential terms are only included to Equation 2.2 if they significantly reduce the value of χ_v^2. A decay is fitted from the rising edge of the decay curve, a few channels from the peak, to the end of the decay curve or a point where the photon counts for that channel were approximately 100, as appropriate. The background values used in fitting are estimated from the number of counts arriving before the rising edge of the fluorescence decay and were fixed for the fitting procedure.

For example, Figure 2.5 presents the decay from Figure 2.4 fitted with 3 exponential terms in Equation 2.2. Figure 2.6 shows the same decay fitted with four exponentials. The result of adding a fourth exponential term has resulted in a significantly lower χ_v^2, a more uniform variation of the residuals and, visually, a better fit of the calculated function to the experimental data. Thus the decay in Figure 2.4 is best described by a four-exponential function.

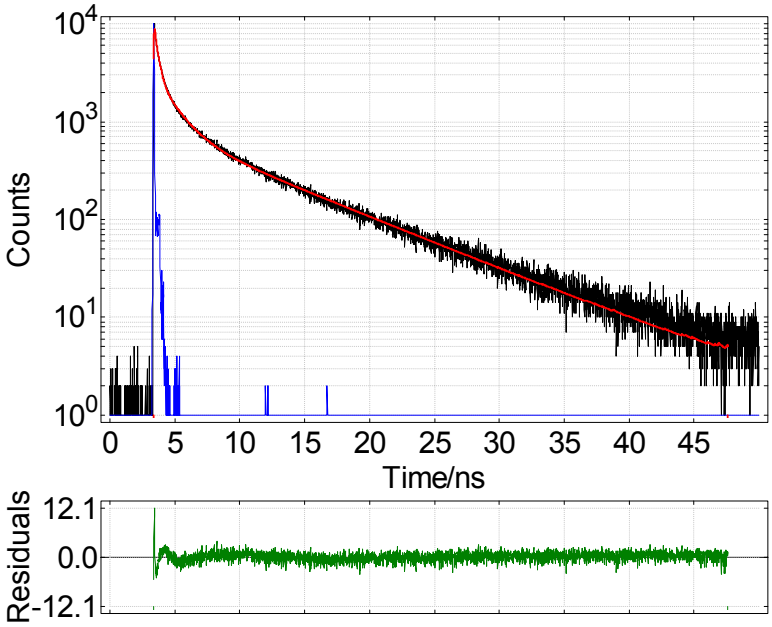


Figure 2.5 Fluorescence decay data for 2AP-labelled DNA (black), the calculated function generated by F900 (red) and a plot of the residuals between the decay data and the calculated function (green). The decay was fit with three exponentials, by iterative reconvolution with a separately measured instrument response function (blue), between the 276th and the 3900th time channel, and using a background level as 1 count. The resulting lifetimes calculated for the red function are 0.25, 1.49 and 8.07 ns and the reduced chi squared = 1.349.

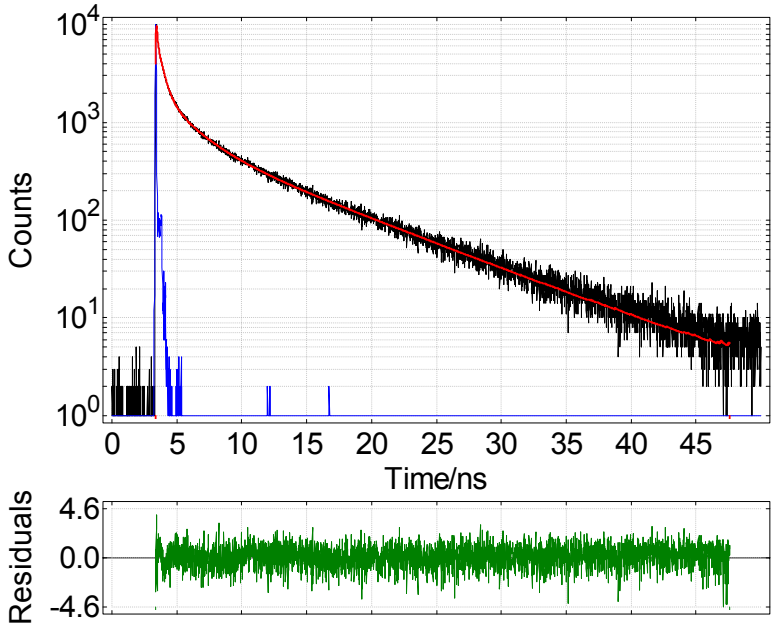


Figure 2.6 Fluorescence decay data for 2AP-labelled DNA (black), the calculated function generated by F900 (red) and a plot of the residuals between the decay data and the calculated function (green). The decay was fit with four exponentials, by iterative reconvolution with a separately measured instrument response function (blue), between the 276th and the 3900th time channel, and using a background level as 1 count. The resulting lifetimes calculated for the red function are 0.08. 0.44, 2.13, and 8.39ns, and the reduced chi squared = 1.071.

For a given sample, decays are collected at several emission wavelengths (typically 370, 390 and 410 nm for the DNA-enzyme experiments). Each decay is fitted individually with F900 and then the set is fitted globally to common values of τ_i. Global fitting is carried out by Edinburgh Instruments FAST software.

The lifetimes reported in the caption of Figure 2.6 encompass the range of values calculated for the decays in this thesis. It should be noted that each lifetime is characteristic of a distribution. The standard deviation associated with each lifetime is given in parenthesis and reflects what is typical of such values: 0.08 (0.00629) ns, 0.44 (0.01500) ns, 2.13 (0.05455) ns and 8.39(0.03554) ns.

The decay parameters of interest from Equation 2.2 are τ_i, the fluorescence lifetime of the i^{th} emitting species, and A_i, the fractional amplitude of the i^{th} emitting species. Useful quantities that can be calculated from these decay parameters, and that are referred to in the results chapters to follow, are the quantum yield of the sample relative to free 2AP riboside in solution, Φ_{rel}, (τ_{2APr} = 10.6 ns):

Equation 2.8

$$\Phi_{rel} = \frac{\sum_i A_i \tau_i}{\tau_{2APr}}$$

the contribution of the n^{th} emitting species to the steady state intensity (for a multiexponential decay):

Equation 2.9

$$\frac{A_n \tau_n}{\sum_i A_i \tau_i}$$

and the average lifetime of the sample:

Equation 2.10

$$\frac{\sum_i A_i \tau_i}{\sum_i Ai}$$

2.4. References

(1) Phillips, D.; O'Connor, D. V. *Time-correlated Single Photon Counting*; Academic Press: 1984.

Chapter 3

The M.TaqI Methyltransferase

3.1 Introduction

M.TaqI was the very first adenine methyltransferase to be co-crystallised with its substrate DNA. The resulting crystal structure demonstrated that, like the cytosine methyltransferases, an adenine modifier also performs base flipping; therefore M.TaqI is an excellent subject for investigating base flipping in the adenine methyltransferases through 2AP time-resolved fluorescence.

The M.TaqI methyltransferase comes from *Thermus Aquaticus* and is part of a restriction-modification system. This protein recognises the palindromic TCGA sequence, and the adenine in both DNA strands must be methylated to avoid cutting of the DNA by the restriction enzyme R.TaqI[1].

The crystal structure of the ternary complex of M.TaqI with DNA and an analogue of the cofactor, SAM, was solved by Goedecke *et. al.*[2] and is shown in Figure 3.1. In this structure, the enzyme is bound to a 10 base pair duplex and has extruded the target adenine (shown in red) from the duplex interior. This 47, 856 Da, 421 amino acid protein folds into a structure approximately 80 x 55 x 40 Å in dimension. The amino acid chain is divided into two equal units, the N-terminal domain and the C-terminal domain, which are connected through a short loop thus forming a C-shaped cleft structure.

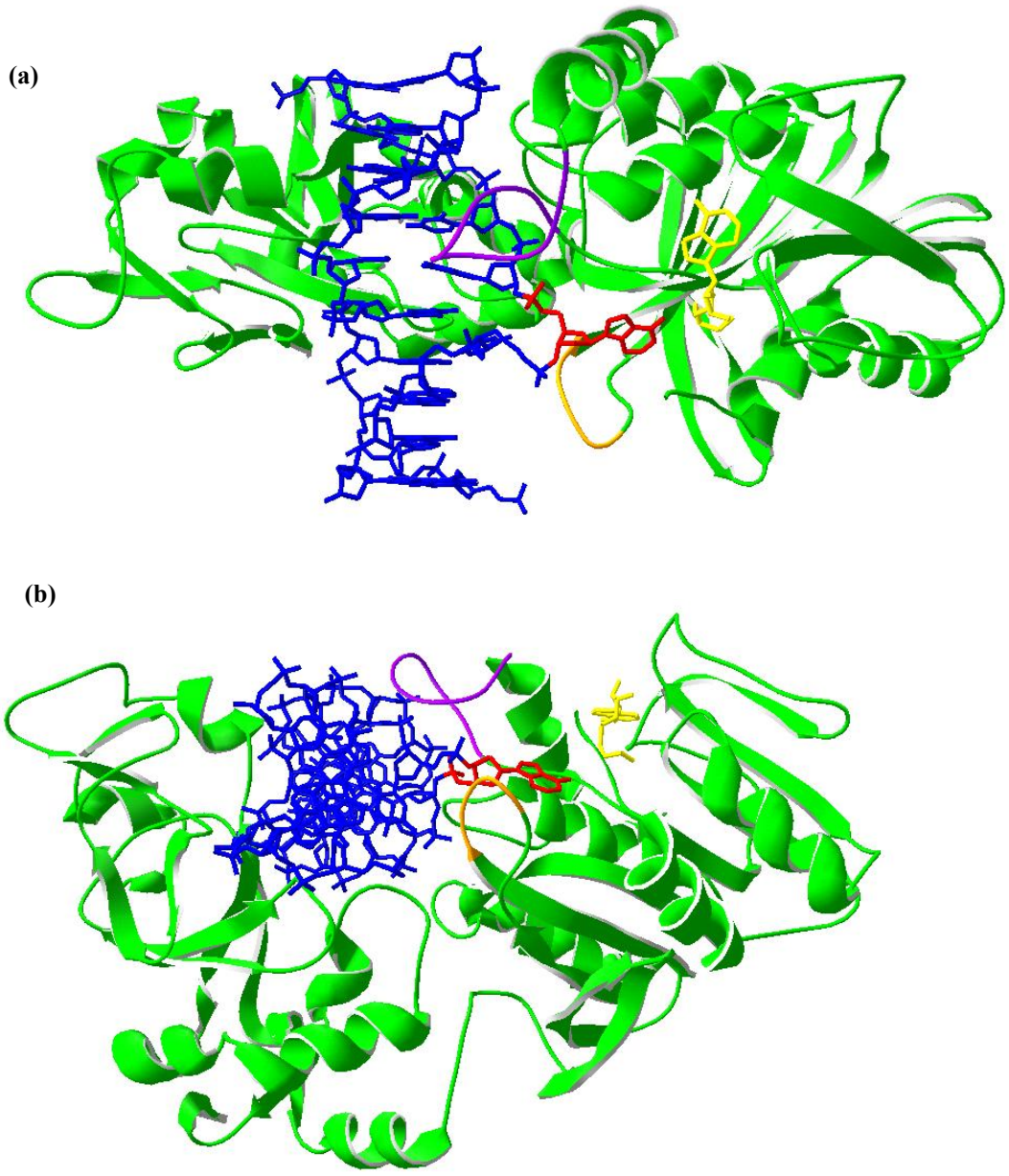


Figure 3.1 The three dimensional crystal structure of the ternary DNA/M.TaqI/AETA complex with 2.0 Å resolution reported by Goedecke *et al.* (PDB accession code 1G38)[2]. The main chain of M.TaqI is green, loop I is purple and loop II is orange. The cofactor analogue AETA is yellow, the DNA is blue and the flipped target adenine is red. (a) is a side view, (b) is a bottom view resulting from a 90° rotation of view (a). In both views the N-terminal domain of the enzyme is on the right and the C-terminal domain is on the left of the bound DNA.

The N-terminal domain, comprising amino acids 1-243 and situated to the right of the DNA duplex in Figure 3.1, is the catalytic domain and contains the cofactor binding motif PAxAxGP and the catalytic motif NPPY. This domain is nearly spherical and contains 9 β- sheets flanked by several α-helices. The C-terminal domain, formed by the remaining amino acids and situated to the left of the duplex in Figure 3.1, is slightly elongated and contains a 3-helix bundle. This is the recognition domain responsible for sequence-specific DNA substrate binding.

The structure of a binary complex of M.TaqI with the cofactor SAM, published by Labahn 7 years previously[3], enables comparison of the structure of the enzyme with and without the bound DNA. Evidently, the changes in the structure of the enzyme on DNA- binding are subtle. In the absence of DNA the C shaped binding cleft is 21 Å wide, easily accommodating of a DNA duplex and when DNA is bound the two domains close around the duplex effecting a 5 degree change in the connecting hinge loop. Two other looped segments of the enzyme become highly ordered on DNA-enzyme binding. Loop I (amino acids 110-121, coloured purple in Figure 3.1) presents a barrier to back-flipping and loop II (amino acids 197-201, coloured orange in Figure 3.1) forms part of the target adenine binding pocket.

Figure 3.2 maps the M.TaqI-DNA interface[2]. As can be seen, 12 direct contacts form between the enzyme and the DNA phosphodiester backbone. M.TaqI hydrogen-bonds with all of the bases in the DNA recognition sequence with one exception, the T15 thymine which is paired to the flipped target adenine. This thymine, however, is in van der Waals contact with proline 393.

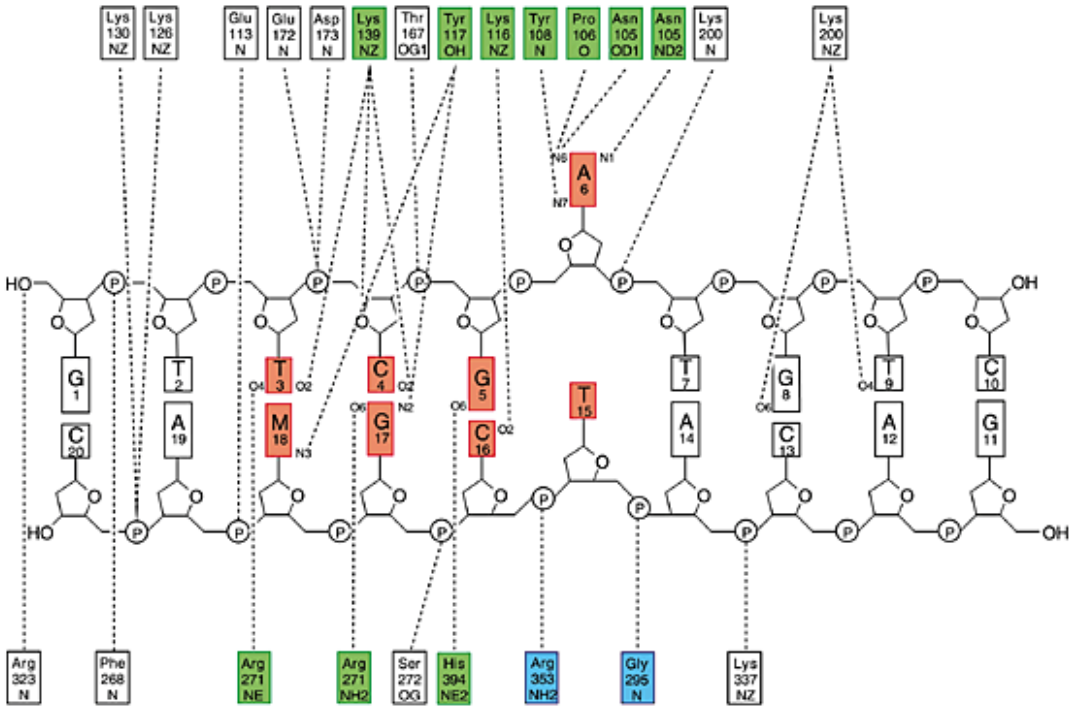


Figure 3.2 Schematic representation of hydrogen bonds and salt bridges between M.TaqI and the DNA substrate compiled by Goedecke *et al.*[2] Green backgrounds designate the specific M.TaqI residues in contact with the bases (red) of the DNA recognition sequence. M.TaqI residues with non-specific contacts with backbone phosphodiester groups are in white and the amino acids involved in DNA compression are in blue.

M.TaqI binding distorts the DNA causing a 3 Å widening of the minor groove and a 2 Å compression of the duplex in the region of the AT target base pair. The target adenine and its pentose sugar are rotated around the DNA backbone, such that the amine N6 of the target base lies 4.7 Å from the sulphur group of the cofactor. This is a suitable distance for methyl transfer, whereas the corresponding distance for an unflipped intrahelical adenine, an estimated[3] 15 Å, is not. Figure 3.3 expands the region of the M.TaqI-bound duplex around the target base position.

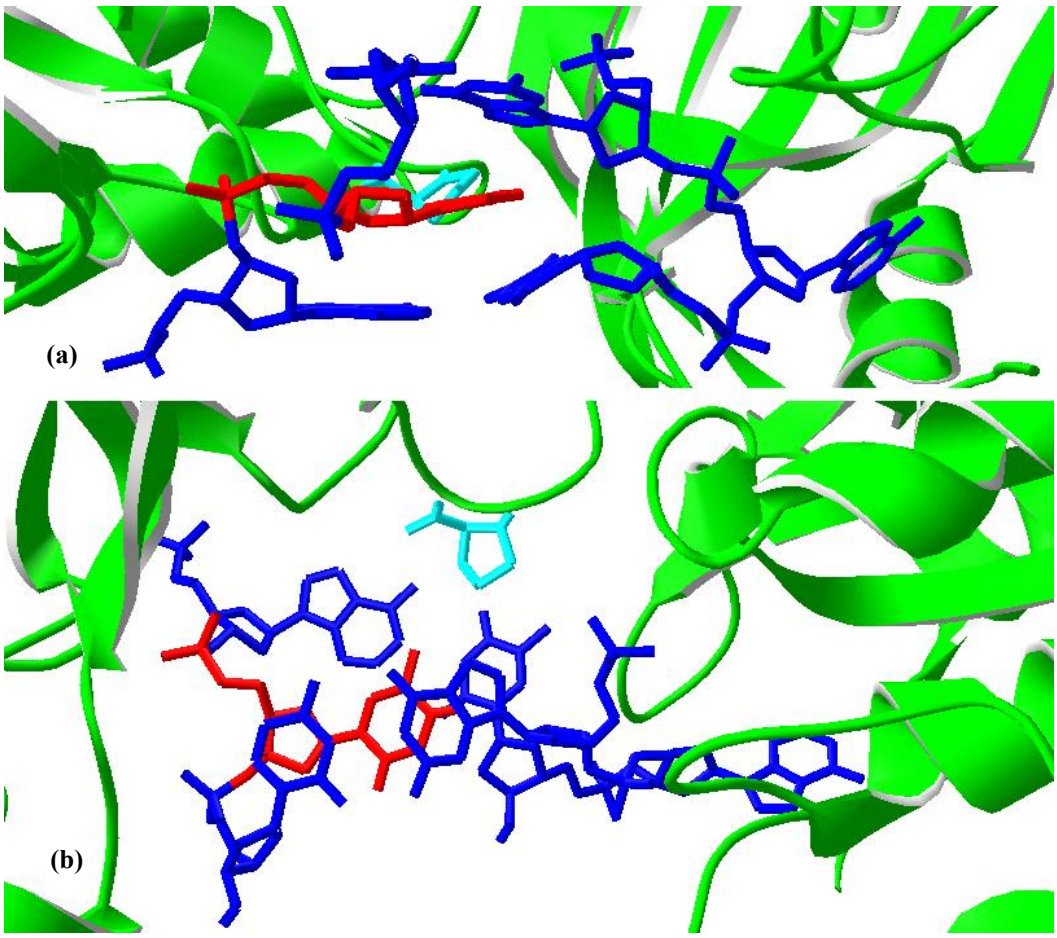


Figure 3.3 Enlargement of the Goedecke DNA/M.TaqI/AETA structure[2] in the vicinity of the flipped target adenine. The DNA duplex is shown in blue, with the thymine partner of the flipped base coloured red. The M.TaqI structure is displayed as the green ribbon with the proline 393 amino acid coloured turquoise (2.0 Å resolution). (a) is a side view , (b) is a top view looking down the long axis of the DNA helix.

The orphaned thymine, formerly paired with the target adenine, sits in the gap left by the flipped adenine and π-stacks with bases in the opposite oligonucleotide strand. As mentioned before, this thymine is not in strong contacts with the enzyme. The position of the partner thymine in the M.TaqI-DNA-cofactor complex is unusual. Its role is two-fold; maintenance of the base stacking in the duplex interior and prevention of the return of the flipped base to the duplex interior. In all other methyltransferase structures the enzyme is observed to perform these functions through amino acid intercalation[2;4].

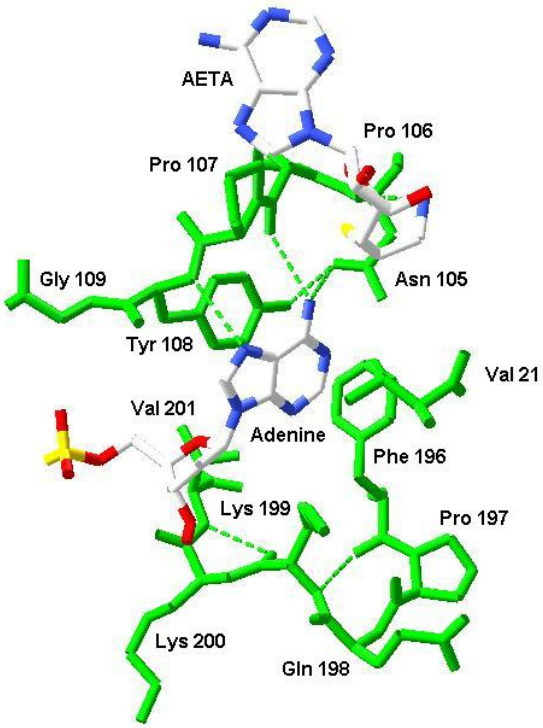


Figure 3.4 An image of the M.TaqI amino acids (green) surrounding the flipped adenine, taken from the 2.0 Å resolution structure reported by Goedecke *et al.*[2]. The cofactor analogue AETA is also displayed and the broken green lines represent hydrogen bonds calculated by Swiss PDB viewer.

The highly ordered amino acid environment of the flipped adenine, as it sits inside M.TaqI, shown in Figure 3.4, features three sequence motifs that are conserved among all adenine methyltransferases, denoted IV, VIII and X. In the catalytic NPPY motif (motif IV, amino acids 105-108) hydrogen bonds form between asparagine 105 and N6 of adenine; the carbonyl oxygen of proline 106 and N6 of adenine; and the amide nitrogen of tyrosine 108 and N7 of adenine. Hydrogen bonding to the N6 of adenine polarises the nitrogen lone pair thus activating it for methyl transfer. An additional NPPY interaction is face-to-face π-stacking between the top face of the aromatic ring system of adenine and the aromatic side chain of tyrosine 108. On the bottom face of the adenine, C6 and N1 are in hydrophobic contact with valine 21 (motif X). Phenylalanine 196, in motif VIII, is positioned for edge-to-face π-stacking to the 1-3 edge of the adenine heterocycle. Also of significance in this motif is the lysine 199 in hydrophobic contact with the N3 of adenine. This lysine is believed to act as a hinged gate that is open before base flipping and closed afterwards[2].

Prior to the Goedecke ternary structure much had been inferred from the Labahn binary structure which reveals how the amino acid sequence folds and shows the binding site of the natural cofactor SAM. Labahn's M.TaqI structure was mapped onto the structures of the cytosine methyltransferase HhaI and the oxygen catechol methyltransferase[5]. These three methyltransferases have almost identical cofactor binding sites. The M.TaqI cofactor binding site was investigated crystallographically and Schluckebier *et al.* observed differences in the way in which different biologically relevant SAM analogues are incorporated inside the cofactor binding pocket of the enzyme[6]. Similarities in the structure and spatial position of the catalytic domain of M.TaqI to the other methyltransferases suggested that M.TaqI would employ base flipping too. Holz *et al.* subsequently tested this notion by replacing the target base for M.TaqI methylation with 2AP[7]. The 13-fold increase in 2AP emission was taken as strong evidence that M.TaqI-induced base flipping does occur, although this level of signal enhancement is smaller than the 54-fold increase produced for base flipping by the cytosine methyltransferase M.HhaI. Furthermore, a binding pocket for the flipped adenine was proposed. Within this pocket, Pues *et al.* made mutations to the aromatic residues, tyrosine 108 and phenylalanine 196, to test the idea that these amino acids were responsible for proper orientation of the extrahelical adenine[1].

This chapter presents an investigation of M.TaqI-induced DNA base flipping through time-resolved fluorescence. As was done by Holz *et al.*[7], the fluorescent probe 2AP is substituted into the substrate DNA of M.TaqI so as to be the target for base flipping. Lenz has produced M.TaqI crystals analogous to the Goedecke ternary complex but containing the 2AP-labelled DNA[8]. Time-resolved measurements of the 2AP fluorescence in these crystals have allowed the resulting decay parameters to be related to the refined crystal structure and indicate the response of the lifetimes of 2AP to specific M.TaqI interactions. Solution-phase measurements were also undertaken to establish whether the same lifetime signature is obtained from a solution-phase ternary complex. In addition, the decay parameters are compared with those of the unbound DNA to asses the conformational changes that the DNA

has been subjected to on enzyme binding. Thus it will be demonstrated that M.TaqI-induced base flipping can be observed unambiguously in solution, without the uncertainty encountered by Holz *et al.* in steady state fluorescence detection[7].

A second crystal and solution phase study is also presented, which investigates the role of the thymine partnering the flipped target. Recall that the partner thymine is believed to stabilise the hole (abasic site) formed after target base flipping[2] (Figure 3.3). Blumenthal *et al.* postulated that this thymine might actively force the target adenine out to get to its interstrand-stacking position[9]. Here, lifetime measurements have been carried out using a duplex with 2AP at the flipping target site, but where the partner thymine has been removed, creating an abasic site opposite the 2AP.

In light of the cofactor binding study by Schluckebier *et al.*[6] 2AP lifetime data are used to investigate the effect of different cofactors on base flipping. Finally, time-resolved fluorescence is employed to add to the work by Pues *et al.*[1] in learning how the flipped base is stabilised when the catalytic NPPY motif of M.TaqI is altered.

3.2 Materials and Methods

2AP ribonucleoside was purchased from Sigma-Aldrich, dissolved in water and diluted to a micromolar concentration for lifetime measurements. DNA-M.TaqI co-crystals, 2AP labelled duplexes, M.TaqI enzyme stocks (wild type and mutants Y108A and Y108F), cofactors and buffers were provided by Thomas Lenz and Elmar Weinhold (RWTH Aachen). Details on crystal growth and structure determination are documented in Lenz *et al.*[8], included in appendix B.

The DNA substrates used for this study are shown in Table 3.1. The 10-mer substrates were used in the enzyme crystal complexes and the 14-mer substrates were used for all solution phase measurements. All have 2AP positioned at the target methylation site inside the M.TaqI recognition sequence (i.e. the flipping site) and the fluorophore is either paired with thymine (2AP-T) or opposite an abasic site (2AP-D). Since there are two target base positions per duplex, the unlabelled deoxyoligonucleotide strand contains N6-methyl adenine to direct enzyme binding orientation to the 2AP-labelled side of the substrate.

Duplex	Base Sequence
2AP-T 10-mer	5'- GT **T CGP** TGTC -3' 3'- CA **MGCT** ACAG -5'
2AP-D 10-mer	5'- GT **T CGP** TGTC -3' 3'- CA **MGCD** ACAG -5'
2AP-T 14-mer	5'-GCCGC **T CGP** TGCCG -3' 3' CGGCG **MGCT** ACGGC-5'
2AP-D 14-mer	5'-GCCGC **T CGP** TGCCG -3' 3' CGGCG **MGCD** ACGGC-5'

Table 3.1 The base sequences of M.TaqI substrate duplexes. The M.TaqI recognition sequence is shown in bold, P designates the fluorophore 2-aminopurine, M is N6-methyl adenine and D is a 1,2 dideoxy-D-ribose abasic site analogue.

Ternary DNA-enzyme co-crystals consisted of 2AP labelled 10-mer DNA in complex with M.TaqI and the cofactor analogue 5'-[2-(amino)ethylthio]-5'-deoxyadenosine (AETA, Table 3.2 (b)). Crystals, provided by T. Lenz, were mounted in fused silica capillaries for measurement (see Chapter 2). Co-crystals were too small to allow fluorescence measurements to be made on a single crystal, so fluorescence was collected from many crystals of the same type (Figure 3.5).

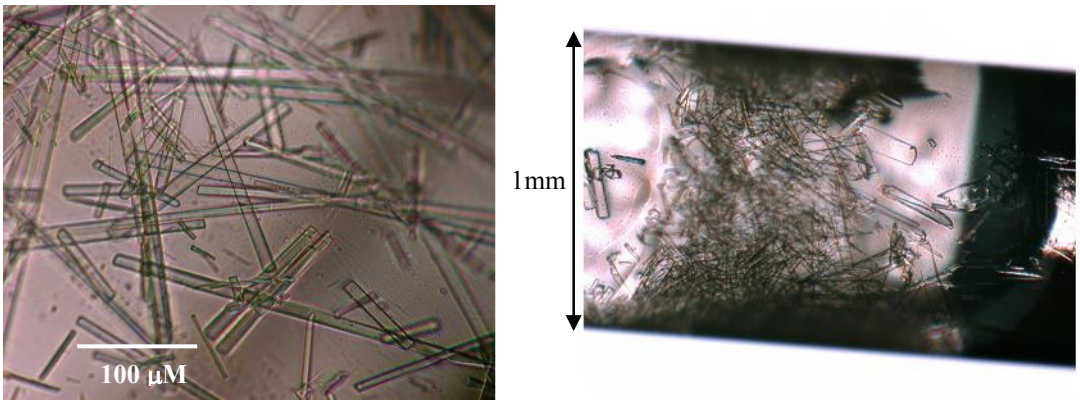


Figure 3.5 Images of the 2AP-T/M.TaqI/AETA crystals inside a fused silica capillary

Solution phase samples consist of (1) free DNA, 2 μM in the 2AP labelled strand; (2) 2 μM DNA, 8 μM M.TaqI and (3) 2 μM DNA, 8 μM M.TaqI and 1 mM cofactor (analogue). Since K_d for this 2AP-labelled DNA with M.TaqI[7] is 120 nM all the DNA in these mixtures will be bound by enzyme. All solutions were buffered with 50 mM KOAc, 20 mM Tris/HOAc, 10 mM $Mg(AcO)_2$ and 1mM DTT (pH 7.9). 0.01% Triton X-100 was included in solutions containing enzyme, both wild type and mutant.

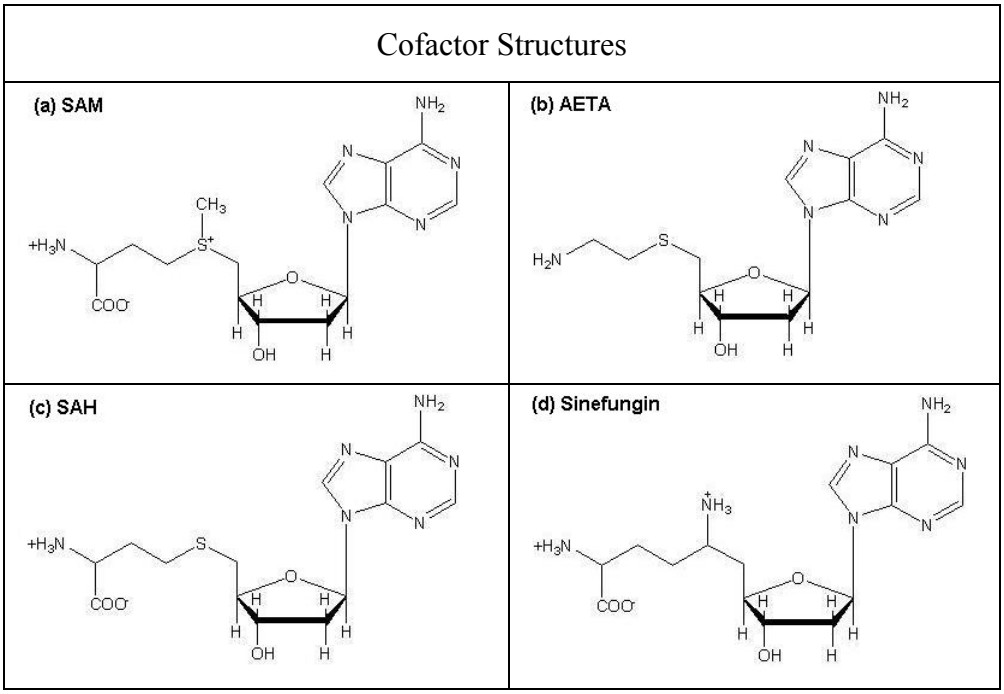


Table 3.2 Structure of the natural cofactor S-adenosyl-L-methionine (a), and analogues 5'-[2-(amino)ethylthio]-5'-deoxyadenosine (b), S- adenosyl homocysteine (c) and sinefungin (d).

Table 3.2 displays structures of the natural cofactor S-adenosyl-L-methionine (SAM) and synthetic cofactor analogues S-adenosyl homocysteine (SAH), 5'-[2-(amino)ethylthio]-5'-deoxyadenosine (AETA) and sinefungin (Sf). Dissociation constants determined for these cofactors with M.TaqI are as follows[6]: $K_{d\ SAM}$ 2.0 μM; $K_{d\ SAH}$ 2.4 μM; $K_{d\ Sf}$ 0.34 μM. The M.TaqI methyltransferase will bind to DNA equally as well whether cofactor is already bound to the enzyme or not (i.e. there is no specific order of binding in this system). For the cofactors experiment, however, the DNA and enzyme were mixed first, allowed to incubate for a few minutes and then cofactor added. Mutant enzyme complexes were also prepared in this order but only the natural cofactor SAM was used.

Steady state and time-resolved fluorescence measurements on crystals and solution phase samples were carried out as described in Chapter 2. The full set of calculated decay parameters for the duplexes and DNA-enzyme complexes in this chapter are given in Tables A1 to A9 in appendix A.

3.3 Results and Discussion

3.3.1 2AP-T/M.TaqI/AETA crystals: structure and fluorescence lifetimes

Figure 3.6 shows the refined crystal structure of M.TaqI in complex with the AETA cofactor molecule and the 2AP-labelled DNA duplex, with 2AP substituted at the target site for methylation[8].

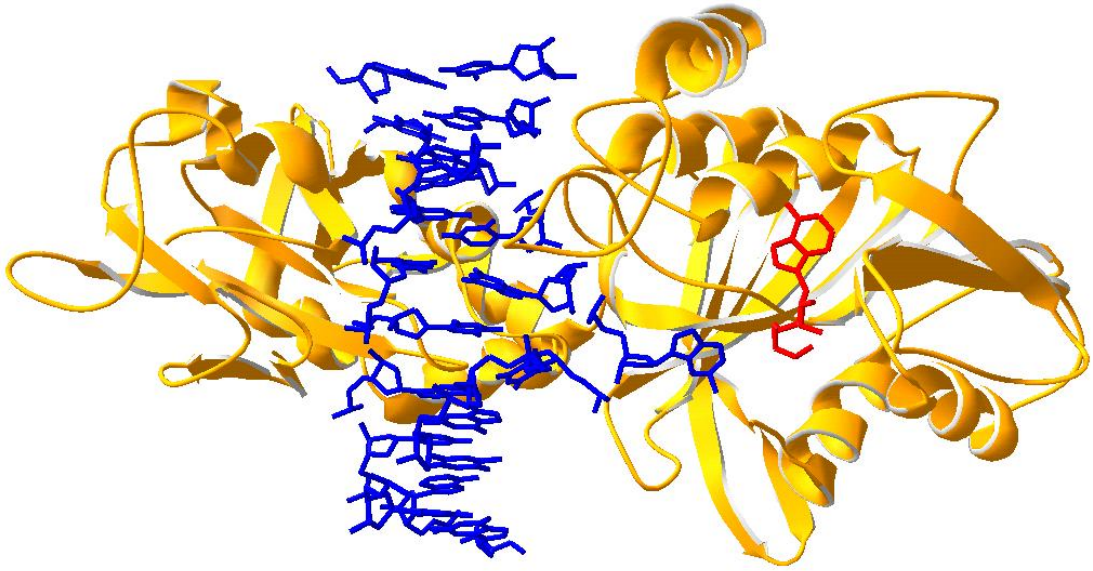


Figure 3.6 The three dimensional crystal structure of the ternary 2AP-T/M.TaqI/AETA complex with 2.4 Å resolution reported by Lenz *et al.* (PDB accession code 2IBS)[8]. M.TaqI is orange, AETA is red and the DNA is blue.

By comparison with Figure 3.1, Figure 3.6 establishes that when 2AP replaces the target adenine in substrate DNA, M.TaqI binds the labelled duplex in a similar manner to the native DNA and flips out the 2AP. Clearly, the presence of this extrinsic probe in the M.TaqI recognition sequence does not appreciably obstruct the specificity and action of the enzyme (prior to catalysis). A closer examination of the flipped 2AP inside M.TaqI and a superposition of the corresponding flipped adenine environment are provided in Figure 3.7. The cofactor analogue AETA has been excluded for clarity but it sits in a very similar conformation in each complex.

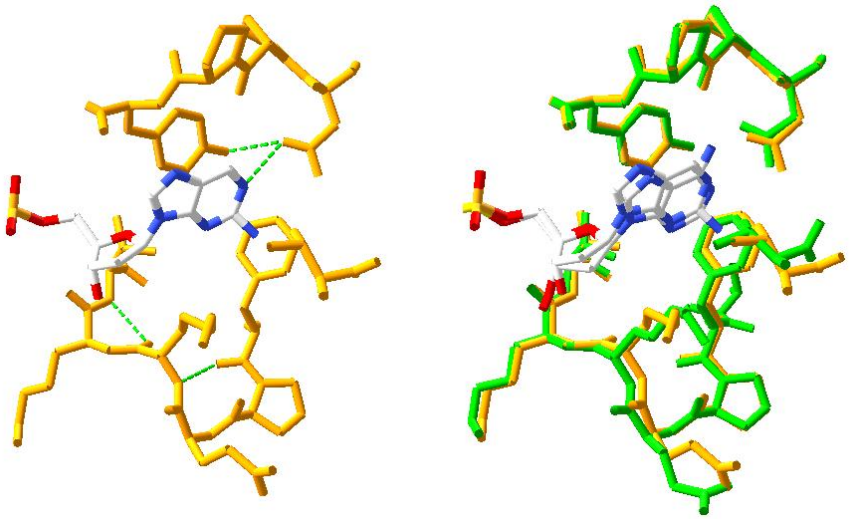


Figure 3.7 Crystal structures of the flipped 2-aminopurine (in the 2AP-T/M.TaqI/AETA complex) showing, on the left, the surrounding amino acids (orange) and, on the right, a superposition of the lefthand structure with the amino acids surrounding the flipped adenine of the Goedecke structure (green). Hydrogen bonds calculated by Swiss PDB viewer are displayed as dashed green lines.

In contrast, when 2AP replaces the target base in substrate DNA for the cytosine methyltransferase M.HhaI the flipped base analogue is held in a different location to a flipped cytosine or uracil[10]. Whereas the M.HhaI catalytic site is designed to accommodate a pyrimidine base, it does not, evidently, permit stabilisation of the larger purine base. Figure 3.7 shows that the M.TaqI catalytic site is not challenged in this way, since the adenine- stabilising pocket also accepts the very adenine-like 2AP molecule.

With regards to specific 2AP-M.TaqI interaction, the position of the flipped 2AP allows for face-to-face contact with tyrosine 108, edge-to-face contact with phenylalanine 196 and hydrogen bonding with asparagine 105. Unlike the adenine structure, Swiss PDB Viewer has not calculated a hydrogen bond between the N7 of 2AP and the tyrosine amide nitrogen but since the separation is 2.9 Å such a hydrogen bond will likely form. The differing position of the exocyclic amine group on 2AP causes asparagine 105 to hydrogen bond with 2AP N1, with loss of the proline 106–flipped base hydrogen bond. The flipped base-phenylalanine 196 face-to-edge contact is altered such that a protruding N2 amine group points towards the

centre of the benzene ring face. Compared with the adenine structure, the conformations of the amino acids around 2AP are highly conserved. Although specific interactions between these amino acids and each flipped base differ, the types of interaction are essentially similar. Consequently, from the perspective of the flipped target base, 2AP reliably represents the adenine-M.TaqI interaction.

The emission spectrum of 2AP fluorescence in the 2AP-T/M.TaqI/AETA crystals is presented in Figure 3.8. Maximal emission of 2AP in this crystalline complex, λ_{max}, comes at 360 nm. Table 3.3 presents the calculated global fluorescence decay parameters (lifetime (τ_i) and fractional amplitude (A_i)) for decays measured from the crystals at three different wavelengths that encompass λ_{max} across the emission profile. It is reasonable to assume that fluorescence comes from 2AP that is flipped and resides within the M.TaqI catalytic site since the crystals used for fluorescence decay measurements were from the same batch as those sent for X-ray diffraction.

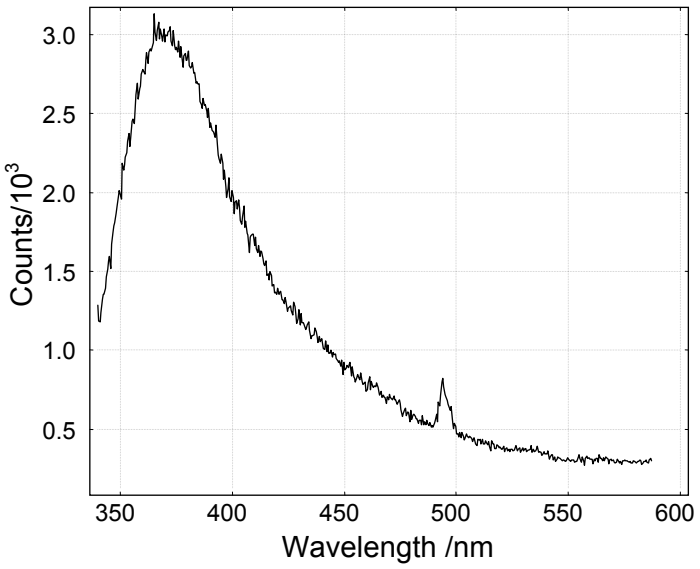


Figure 3.8 Fluorescence emission spectrum of the 2AP-T/M.TaqI/AETA crystals under 320-nm excitation

Global lifetimes /ns

$\tau_2 = 0.93$ $\tau_3 = 2.30$ $\tau_4 = 8.59$

Emission λ /nm	A_2	A_3	A_4	Local χ^2
360	0.49	0.51	0.00	1.043
370	0.48	0.51	0.01	1.052
390	0.46	0.51	0.03	1.106

Global $\chi^2 = 1.067$

Table 3.3 The global fluorescence decay parameters for the 2AP-T/M.TaqI/AETA crystals.

Three lifetimes are needed to describe the decay from the flipped 2AP in the crystalline complex. The decay components in Table 3.3 have been numbered 2, 3 and 4. The number 1 has been reserved for a species with a lifetime shorter than those found in this complex, one of the order of ~10 picoseconds. This will become apparent later when the decay parameters of 2AP in the unbound DNA are presented. A multi-exponential decay indicates that the ternary complex exists in a number of conformational states in which 2AP experiences different interactions and hence different quenching rates. In a simple interpretation, each lifetime can be considered to represent a distinct conformational state and the fractional amplitude indicates the occupancy of this state. Realistically, each lifetime is likely to represent a distribution of conformations in which 2AP experiences similar quenching rates. It can be seen that although three lifetimes are needed to accurately describe the emitting species, the two shorter 0.93 and 2.3 ns lifetimes (τ_2 and τ_3) with fractional amplitudes of 0.46 and 0.51, respectively, account for the vast majority of the population whereas the minor 8.9 ns lifetime component (τ_4) only constitutes 3% (these A factors correspond to the decay measured at 390 nm).

Like the Goedecke ternary complex (native DNA), the 2AP-T/MTaqI/AETA crystals are found to have two symmetry-independent, yet similar, complexes in the asymmetric unit[8]. These complexes, termed A and B, have different positions of Lys 199 and slightly different positions of Phe 196 in close proximity to the 2AP (see Figure 3.4 for amino acid labelling). A point of interest is that while two complex conformations arise from the X-ray data the lifetime data reports three. The heterogeneity apparent in the lifetime data may be related to complexes A and B however, it cannot be ruled out that the lifetimes of 2AP reported in Table 3.3 are

affected by conformational heterogeneity that is unresolved in the crystal structure. The long lifetime component τ_4 has such a small fractional amplitude (0.03) that its corresponding conformations are unlikely to be observed in the crystal structure due to the small contribution to electron density. Evidently, fluorescence lifetimes report that the DNA-M.TaqI-cofactor complex is conformationally heterogeneous in the crystal phase as was also observed for DNA-M.HhaI crystals[10].

The lifetime of 2AP ribonucleoside (2APr) in the enzyme-stabilising buffer for M.TaqI is 10.3 ns. The dominant lifetimes for 2AP in the M.TaqI crystals (0.93 and 2.3 ns) are much shorter than this, indicating that the 2AP experiences far greater quenching in the enzyme complex than in free solution. So what is the nature of this quenching? When 2AP is inside the M.TaqI active site, the crystal structure shows that the 2AP π-stacks face-to-face with tyrosine 108 and is also in close proximity to phenylalanine 196, allowing a potential edge-to-face π-stacking interaction (Figure 3.7). These π-stacking interactions are believed to be responsible for the quenching of the enzyme-stabilised base. This is consistent with the generally accepted view that it is π-stacking that governs the quenching of 2AP in DNA and not hydrogen bonding[11-17]. Since tyrosine 108 forms the strongest type of π-stacking association with the 2AP, it is thought to be the most influential in 2AP quenching.

Further insight into the impact of amino acid π-stacking on 2AP lifetime is provided by lifetimes measured previously for the DNA/M.HhaI crystal[10]. Figure 3.9 shows the crystal structure of a flipped 2AP inside the cytosine methyltransferase M.HhaI and Figure 3.10 compares the lifetime parameters of 2AP flipped into M.HhaI with those of 2AP flipped into M.TaqI. As can be seen, the decay parameters representing 3 distinguishable M.HhaI complex conformations are quite different from those for M.TaqI. In the major conformation of the M.HhaI/DNA complex (76%) 2AP has a much longer, 10.9 ns, lifetime. This conformation is the one likely to be represented in Figure 3.9. In the amino acid environment of 2AP inside M.HhaI, specific 2AP-enzyme interactions are limited to hydrogen bonding. Thus, in the absence of a π-stacking interaction, the hydrogen bonding evokes only a low quenching rate for the 2AP.

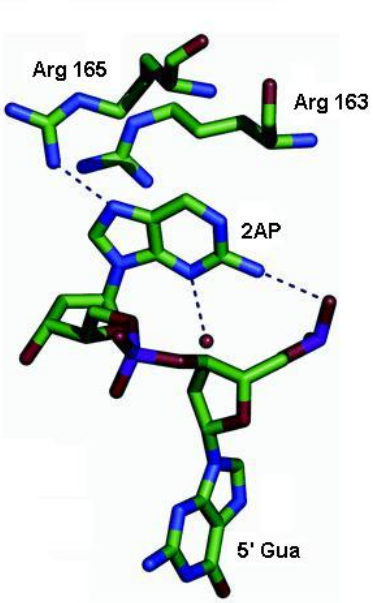


Figure 3.9 The crystal structure of 2-aminopurine (2AP) flipped into the M.HhaI methyltransferase, reported with 2 Å resolution by Neely *et al.*[10] Hydrogen bonds (dashed lines) form with two arginines (Arg) of M.HhaI, the phosphate group of the DNA backbone towards the 5' guanine (5'Gua), and a water molecule (red sphere).

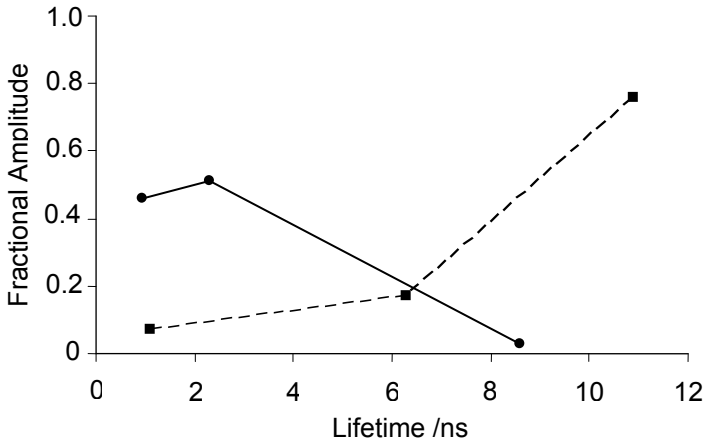


Figure 3.10 A plot of the decay parameters for 2AP flipped inside M.TaqI (circles and solid line) and flipped inside M.HhaI (squares and dashed line).

It is interesting that a flipped 2AP can display such different lifetimes. Evidently the quenching environment provided for the flipped 2AP is unique to the enzyme. A

useful quantification of the extent of 2AP quenching, calculable from the decay parameters, is provided by Φ_{rel}, the quantum yield of the probe relative to free 2AP ribonucleoside (Equation 2.8, Chapter 2). The Φ_{rel} value for 2AP in M.TaqI is 0.2 while that for 2AP in M.HhaI is 0.9. Were these DNA-enzyme crystal complexes to be studied by steady state fluorescence spectroscopy, the 2AP-M.HhaI emission intensity would be roughly 4 times as high as the 2AP-M.TaqI signal although both signify base-flipped complexes. The M.TaqI NPPY motif is a permutation of the D/N PP Y/F/W sequence combination conserved in all the N6-adenine methyltransferases and it is thought that these methyltransferases stabilise a flipped adenine in the same way, i.e. hydrogen bonding through the D/N PP part and π-stacking with the aromatic Y/F/W part. It is envisaged that other adenine modifiers will also quench 2AP flipped into their active site.

Now that the lifetime signature of flipped 2AP within the M.TaqI active site has been characterised and related to the molecular environment this study will proceed to solution phase measurements.

3.3.2 2AP-T/M.TaqI/AETA solution-phase fluorescence lifetimes

Table 3.4 presents the solution-phase decay parameters for 2AP-labelled DNA as the free duplex, in binary complex with M.TaqI and in ternary complex with M.TaqI and cofactor analogue, AETA, and Figure 3.11 presents these data plotted as A vs. τ. The crystal-phase decay parameters have been included in the same formats for further comparison.

Composition		τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	Φ_{rel}
2AP-T/M.TaqI/AETA	C	-	0.93	2.3	8.6	-	0.46	0.51	0.03	0.18
2AP-T/M.TaqI/AETA	S	-	0.55	1.9	7.3	-	0.54	0.38	0.08	0.15
2AP-T/M.TaqI	S	-	0.41	1.9	7.3	-	0.21	0.66	0.13	0.21
2AP-T	S	0.03	0.57	2.4	7.8	0.81	0.06	0.09	0.04	<0.055

Table 3.4 The global decay times for the 2AP-T duplex, its binary and ternary complexes with M.TaqI/AETA in solution (S) and for the crystalline (C) ternary complex, A factors are reported for decays observed at a 390-nm emission wavelength. Also reported are the calculated quantum yields relative to free 2AP riboside Φ_{rel}.

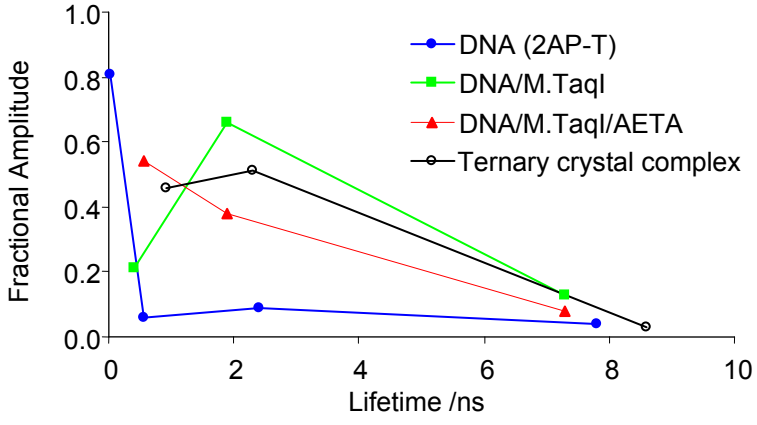


Figure 3.11 A plot of the decay parameters reported in Table 3.4 for duplex 2AP-T and its complexes.

For a solution based technique to be effective, it is essential that a clear indication of the occurrence of base flipping is apparent in the lifetime data. This must include a realisation of how the native conformational state of DNA is altered on encountering enzyme and so it is necessary to also characterise the state of the 2AP in an unbound duplex.

The fluorescence decay of 2AP in DNA (Table 3.4, 2AP-T) is highly complex and needs a four-exponential function to describe it. In the first of four resolved conformations, 2AP has a very short 30 ps lifetime indicating that it is quenched with extremely high efficiency. This corresponds to a conformation in which 2AP is π-stacking very effectively with the neighbouring natural bases. The longest lifetime, 7.8 ns, indicates much less quenching and is representative of a conformation where 2AP is not stacking particularly well, perhaps protruding out of duplex and subjected to some solvent exposure. The intermediate lifetimes, τ_2 and τ_3, relate to conformations of the duplex between those of 1 and 4, where 2AP experiences imperfect stacking conditions. As can be seen from the accompanying A factor distribution, 81% of the duplexes are very highly stacked and the 2AP is soundly buried inside the DNA double helix. In consequence a highly reduced quantum yield ensues (<0.05). The decay parameters for 2AP in the free duplex that constitutes the M.TaqI substrate are typical of this fluorophore in DNA, where four separate species are generally observed[11;18;19]. The particularly short value for τ_1 here is attributed to the presence of an immediate guanine neighbour, since guanine is the most efficient charge transfer quencher of 2AP[20;21] (further discussion on the behaviour of 2AP in DNA is reserved for Chapter 6).

For the ternary complex in solution, three sets of conformations are resolved, giving lifetimes of 0.55, 1.9, and 7.3 ns for the 2AP. The A factors indicate that the proportions of these species are 54, 38 and 8% respectively. These decay parameters inform that 92% ($A_2 + A_3$) of the conformations of the ternary complex substantially quench 2AP (relative to free 2APr) and the small proportion of the remaining conformations give a 2AP lifetime approaching τ_{2APr}. Essentially, the crystal phase parameters also give this picture, reporting a similar partitioning of complex

conformations which each provide similar extents of quenching for the 2AP. It follows then, that the crystal and solution phase complexes are equivalent and that the structure of flipped 2AP inside the catalytic site of M.TaqI, shown in Figure 3.7, describes the major DNA-enzyme interaction in solution.

The solution-phase lifetimes are shorter than the crystal-phase lifetimes; however, the differences between corresponding crystal- and solution-phase lifetimes are small and A factor distributions are very similar. The same range of lifetimes is observed such that τ_2 remains sub-nanosecond, τ_3 is approximately 2 ns and τ_4 is nearing 10 ns. The lifetimes in solution clearly indicate that 2AP experiences similar extents of quenching to those in the crystal, which suggests that the conformations adopted by the complex in solution then prevail when the enzyme is crystallised. A plausible reason for lifetime shortening is that conformations in solution are less constrained:an increased level of mobility in solution leads to an improved quenching rate is in line with the observation that the excited state decay of 2AP in duplexes is influenced by base dynamics[16;20;21].

An obvious difference between the decay of the free duplex (2AP-T) and the ternary complex (2AP-T/M.TaqI/AETA) is the loss of τ_1 (Table 3.4, Figure 3.11). The depopulation of this duplex conformation, where the 2AP is highly stacked, reports that the base has been flipped. Interpretation of the changes to lifetimes 2 to 4 in solution is greatly aided by the crystal phase lifetime parameters. Even though there is very little change in the values of $\tau_{2\text{-}4}$ between unbound DNA and the ternary complex the stark increase in their combined A factors and the knowledge that similar lifetimes are shown by the 2AP inside the active site of M.TaqI in the crystal, indicate that specific 2AP-M.TaqI interactions are responsible. It is likely that the set of conformations represented by τ_2 and τ_3 for 2AP-T/M.TaqI/AETA also includes imperfectly stacked 2AP, i.e. 2AP that is still intrahelical in the DNA-enzyme complex. The proportion of unflipped M.TaqI complexes is thought to be low.

The removal of the short component, τ_1, on base flipping has been observed before for the previously mentioned M.HhaI methyltransferase. The working definition of "short" in this context is a lifetime under 100 picoseconds, descriptive of 2AP in its most efficiently stacked intrahelical duplex conformation[11;13;18;19]. However, in the DNA-M.HhaI-cofactor decays in solution a short component is still present[10]. In these ternary M.HhaI complexes 16% of the bound duplexes appear to still bear an intrahelical 2AP at the target position for base flipping. Evidently, the definition of the 2AP decay response to base flipping should include an appreciable reduction in the population of τ_1 (A_1). Indeed, in the unbound M.HhaI DNA substrate, conformation τ_1 is in 53% abundance. The ability of M.TaqI to depopulate completely the substrate conformation τ_1 in the solution phase, is a mark of the proficiency of M.TaqI in extruding its target base.

So far, the decay data for the binary 2AP-T/M.TaqI complex has been neglected. The lifetimes of the binary complex are remarkably similar to those of the ternary complex, indicating only subtle changes to the enzyme-bound duplex with the addition of AETA. The cofactor analogue appears to redistribute the complex conformations (A_2 and A_3) and causes a slight increase in the decay rate of the probe.

Φ_{rel} values predict an approximate 4-fold increase in 2AP emission intensity on addition of M.TaqI to the DNA, falling to a 3-fold overall increase when cofactor is included in the complex. These changes fall far short of the observed 13-fold increase published by Holz *et. al.*[7]. Such inconsistency arises from the profound difficulty in measuring absolute emission intensities for 2AP in the free DNA. The relative quantum yield for the 2AP-T duplex is very low, less than 5% of that of the free 2AP ribonucleoside. Each set of duplex conformations, characterised by lifetimes 1 to 4, do not contribute equally to the total steady state emission and relative contributions are calculable from the decay parameters (see Equation 2.9, Chapter 2) Although the species τ_1 is 81% abundant, its short lifetime means that it only constitutes 4.2% of the total emission intensity. The major contributors to the steady state intensity are species with the longer lifetimes, τ_3 accounts for 36.8% and τ_4 for 53.2% although together these species only occur in 13% of the duplexes.

Though sensitivities of fluorometers vary it is expected that a typical instruments will only detect the emission from this 13%. Distinguishing signal from noise is further obstructed by low 2AP concentrations thus overestimation of intensity increases on flipping arise from an underestimate of unflipped intensity, because signal is small relative to background. Measurements of the higher emission levels for 2AP in the enzyme binary and ternary complexes are less prone to error. The Φ_{rel} values indicate that on addition of AETA emission is reduced to 71% of the binary complex intensity. Independent steady state experiments find the extent of this reduction to be 87% (unpublished data from Thomas Lenz see section 3.3.4). In this instance, comparing more reliable intensity changes, lifetime and steady state data are in good agreement.

3.3.3 2AP-D/M.TaqI/AETA crystal structure and crystal- and solution-phase fluorescence lifetimes

Figure 3.12 shows the crystal structure of the ternary complex of M.TaqI/AETA with the 2AP-D duplex. 2AP-D denotes DNA with 2AP at the target flipping position but with the pairing thymine removed. Interest in this duplex arose from the desire to know whether the partner thymine is actively involved in the base flipping mechanism, that is, whether this thymine forces the target base out of the duplex and into the enzyme.

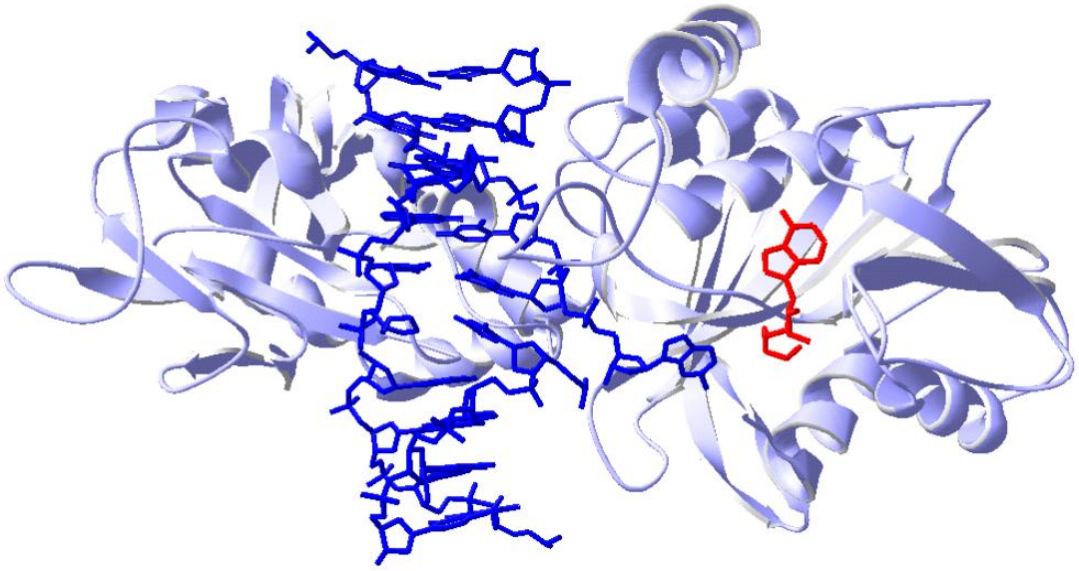


Figure 3.12 The three dimensional crystal structure of the ternary 2AP-D/M.TaqI/AETA complex, with 1.7 Å resolution, reported by Lenz *et al.* (PDB accession code 2IBT)[8]. M.TaqI is lilac, AETA is red and the DNA is blue.

In the above structure, 2AP has been flipped from the 2AP-D duplex into the M.TaqI interior and is held in close proximity to the bound cofactor molecule. Recall that in the native DNA structure (Figure 3.1), the partner thymine has van der Waals contact with the M.TaqI proline 393 and Blumenthal speculated that the base-flipping process might start with this proline pushing on the partner thymine[9]. It is evident from the crystal structure in Figure 3.12 that the removal of this thymine does not prevent base flipping. Furthermore, its former function, in restoring base-stacking in the duplex interior when the target base has been flipped out, is performed by

another base instead; the thymine adjacent to the flipping target (situated below the flipped 2AP in Figure 3.12).

Figure 3.13 presents the environment of the 2AP flipped from 2AP-D and also a superposition of the environments of flipped adenine and 2AP flipped from 2AP-T (once again, the AETA has been omitted for clarity). Evidently, the base flipped from the abasic site duplex is stabilised inside M.TaqI in a highly conserved manner.

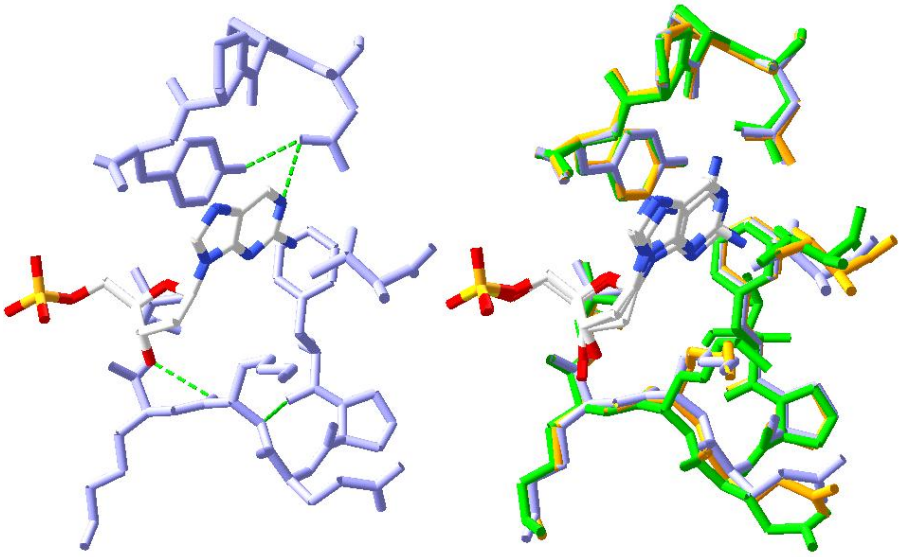


Figure 3.13 Crystal structures of the flipped 2AP (in the 2AP-D/M.TaqI/AETA complex) showing, on the left, the surrounding amino acids (lilac) and, on the right, a superposition of the lefthand structure with the amino acids surrounding the flipped 2AP from 2AP-T (orange) and the flipped adenine of the Goedecke structure (green). Hydrogen bonds calculated by Swiss PDB viewer are displayed as dashed green lines.

Solution Composition		τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	Φ_{rel}
2AP-D/M.TaqI/AETA	C	-	0.92	2.3	7.6	-	0.44	0.48	0.08	0.21
2AP-D /M.TaqI/AETA	S	-	0.28	2.8	9.6	-	0.33	0.30	0.37	0.42
2AP-D/M.TaqI	S	-	0.23	3.3	9.7	-	0.26	0.18	0.56	0.57
2AP-D	S	0.04	0.26	2.5	8.3	0.62	0.18	0.09	0.11	<0.11

Table 3.5 The global decay times for the 2AP-D duplex, its binary and ternary complexes with M.TaqI/AETA in solution (S) and for the crystalline (C) ternary complex, A factors are reported for decays observed at a 390-nm emission wavelength. Also reported are the calculated quantum yields relative to free 2AP riboside $\Phi_{rel.}$

Like its 2AP-T counterpart, the decay signature of the 2AP-D/M.TaqI/AETA crystalline complex is three-exponential (Table 3.5) informing that, as before, a mixture of conformations exists in the crystal phase. The resulting lifetimes and A factors are almost identical to those of the 2AP-T crystalline complex (Table 3.4) and report that 2AP is held within the NPPY motif in 92% (A_2 + A_3) of the flipped complexes. The crystal structure overlays in Figure 3.13 do indeed show that the 2AP in both crystals experiences the same enzyme-interactions. Once again, Lenz finds two independent structures for the 2AP-D-enzyme crystals[8] and τ_2 and τ_3 characterise these complex conformations.

As expected, the free 2AP-D duplex decay requires a four-exponential function to describe it. The lifetimes differ from the free 2AP-T duplex and this is attributed to the slightly different conformations adopted when the DNA contains an abasic site in the vicinity of the probe. The A factors report a lesser proportion of the most highly stacked conformation, 62% compared with 81%, showing that the probe is able to sample more of the imperfectly stacked and extrahelical conformations. Rachofsky *et al*. have also performed time-resolved fluorescence measurements on 2AP opposite an abasic site and similarly observed four-exponential decays with an appreciable proportion of the τ_1 species[22]. As a consequence of the higher abundance of long-lifetime species, the relative quantum yield of 2AP-D is higher than 2AP-T. Previous steady state measurements on 2AP-abasic site duplexes are in agreement that there is still a high degree of stacking of the 2AP in the helix, but a 2AP opposite an abasic site is less well stacked than the corresponding normal duplex[23;24]. Even in the absence of an interstrand pairing partner, much of the 2AP in the 2AP-D duplexes remains tightly intrahelical.

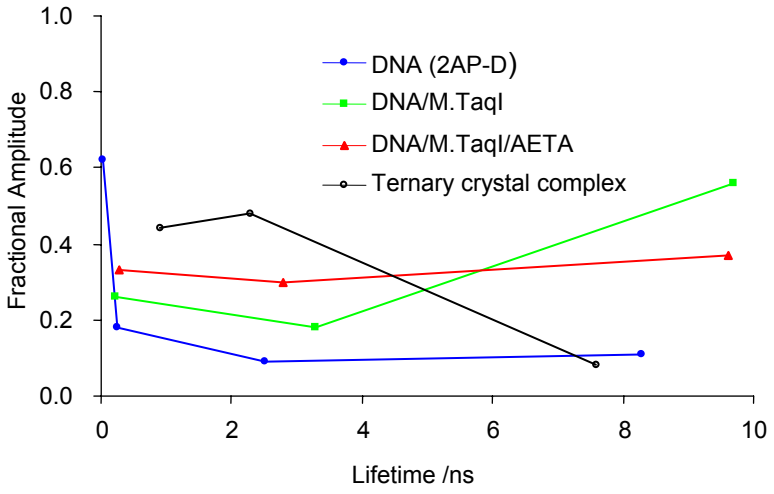


Figure 3.14 A plot of the decay parameters reported in Table 3.5 for duplex 2AP-D and its complexes.

In solution the 2AP-D duplex shows the familiar response of the decay parameters to base flipping (Figure 3.13). Once again, τ_1 disappears on enzyme binding, signifying removal of the 2AP from the most buried intrahelical position in the duplex. Changes in the remaining parameters suggest that the duplex conformations with imperfectly-stacked intrahelical 2AP become conformations with enzyme-stabilised extrahelical 2AP. As before, lifetimes 2 and 3 are likely to come from both enzyme-associated and duplex-associated 2AP. The values of A_2 and A_3 for the binary and ternary complexes report that the amount of 2AP in catalytic conformations rises with the inclusion of cofactor. What is most interesting is that τ_4 and A_4 are quite different between the crystalline and solution-phase complex. In solution a substantial amount of the flipped 2AP experiences very little quenching inside M.TaqI.

The 9.6 and 9.7 ns lifetimes of the solution-phase 2AP-D-M.TaqI binary and ternary complexes, respectively, are appreciably longer than the 7.3-8.6 ns lifetimes determined for the other M.TaqI complexes, both in crystal and solution (Tables 3.4 and 3.5). Additionally, A_4 is especially large for the 2AP-D complexes. Therefore, it is concluded that the 2AP that exhibits this lifetime is flipped but not located within the quenching environment of the M.TaqI NPPY active site. A lifetime of 9.7 ns is close to that of solvent-exposed 2AP (10.3 ns), however, acrylamide quenching

experiments performed by Lenz reveal that 2AP in the 2AP-D/M.TaqI/AETA is not solvent-exposed[8]. This suggests that there exists a conformation in which the flipped 2AP is protected from solvent exposure by enzyme binding, but is not located within the active site. There is a crystal structure available demonstrating that another methyltransferase-flipped target adenine has not entered the active site pocket but has bound to an alternative site on the surface of the enzyme EcoDam[25]. Another example of an alternative binding site (exo site) is provided by the base-flipping hOGG1 DNA glycosylase[26]. A possible functional role of such an exo site in M.TaqI is the preferential binding of a methylated target adenine in order to provide a checkpoint and avoid double methylation of the target base[8]. Similarly Liebert *et al.* think that the EcoDam conformation might resemble an intermediate in the base-flipping pathway but one that could equally occur pre- or post- methyl transfer[25].

Unlike with 2AP-T system, the crystal- and solution-phase decay parameters for 2AP-D/M.TaqI/AETA differ appreciably and, in consequence, the structure of the ternary complex provided by Figure 3.13 is far less informative of what happens in solution. Thus the crystalline system in which 2AP is located in the M.TaqI active site appears to represent only a subset of the conformations present in solution. This raises the question as to whether a change in complex conformation occurs on crystallisation or whether the conformations characterised by τ_4 do not easily crystallise. Since the 2AP-D complex crystals were grown in the same way as the 2AP-T ones (with the use of a seed crystal) and the 2AP-T ternary complexes do not change on crystallisation, the latter reason is more plausible. It would be interesting to know, however, if the seed crystals bias the conformations of the new ones.

On the basis of the crystal structures alone, as shown in, Figure 3.13, it would have been concluded that in the complexes made with an undamaged duplex (2AP-T) and an abasic site duplex (2AP-D), the flipped base experiences the same enzyme-interactions. This conclusion could not have explained why fluorescence emission in solution from the 2AP-D enzyme complexes is higher ($\Phi_{rel} = 0.42$) than in the 2AP-T enzyme complexes ($\Phi_{rel} = 0.15$).

So why should there be differences in flipped base conformation between the natural and the abasic site substrate DNA when the enzyme has little direct contact with the partner thymine base (see Figure 3.2)? Base flipping is thought to be caused by manipulation of the DNA backbone by the methyltransferase enzyme[27]. Clearly, with the partner thymine removed, M.TaqI still performs base flipping but the resulting DNA-enzyme complex adopts conformations in which the flipped target is held outside of the NPPY catalytic motif in a position assumed inaccessible for methyl transfer. Therefore there may be subtle differences in the way in which the duplex backbone responds to the compression induced by M.TaqI when the partner thymine is absent compared with undamaged DNA. Thus the partner base appears to be important (albeit not obligatory) for delivering the target base into the active site and avoiding catalytically incompetent conformations.

3.3.4 2AP-T/MTaqI cofactors study

So far, all the M.TaqI ternary complexes studied have incorporated AETA, an unreactive variant of the natural cofactor molecule S-adenosyl-L-methionine (SAM), that is, one without a labile methyl group available for methyl transfer. The use of AETA enabled successful crystal growth which subsequently allowed lifetimes to be related back to structural conformation, but what is of most interest is the biologically relevant DNA-M.TaqI interaction, which ultimately involves SAM. Time-resolved fluorescence permits conformational study of the DNA/M.TaqI/SAM complex and has been used here to gain an indication of how closely the AETA complex resembles the SAM complex in solution. Two other cofactor analogues of interest are included in this study, S-adenosyl homocysteine (SAH) and sinefungin (Sf or adenosyl ornithine). SAH is the cofactor product created from SAM after the occurrence of methyl transfer and sinefungin is a potent inhibitor of methyltransferases. Sinefungin shares a similar charge distribution to SAM and is found to have an improved affinity towards methyltransferases compared with the natural cofactor, approximately ten times higher in M.TaqI[6]. The structures of these cofactors are presented in Table 3.2.

The decay parameters for the 2AP-T substrate in ternary complex with M.TaqI and each cofactor are reported in Table 3.6. It should be noted that the AETA parameters come from the global fitting of three decays, the SAM decay was also measured at three separate wavelengths but parameters are quoted from the fitting of one decay and the SAH and sinefungin parameters have been calculated from a single decay.

2AP-T/M.TaqI/cofactor	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	Φ_{rel}
Free DNA (no enzyme)	0.03	0.57	2.4	7.8	0.81	0.06	0.09	0.04	<0.05
Binary (no cofactor)	-	0.41	1.9	7.3	-	0.21	0.66	0.13	0.21
AETA	-	0.55	1.9	7.3	-	0.54	0.38	0.08	0.15
SAM	0.12	0.94	3.3	9.9	0.32	0.32	0.20	0.16	0.24
SAH	0.09	0.84	2.5	8.7	0.36	0.35	0.21	0.08	0.14
Sinefungin	0.07	0.74	3.3	9.2	0.55	0.22	0.13	0.10	0.15

Table 3.6 Individual decay parameters for the 2AP-T/M.TaqI/cofactor ternary complex with various cofactors. Parameters are reported from decays measured with a 390-nm emission wavelength. Also reported are the calculated quantum yields relative to free 2AP riboside, Φ_{rel}. Free DNA, binary complex and AETA data are global parameters, SAM decays were measured at 3 wavelengths (see appendix A, Table A8) but individual parameters are reported.

Considering the SAM decay parameters first, an obvious difference between these and the well-characterised AETA parameters is that the 2AP decay when M.TaqI binds SAM is four-exponential. Compared with the decay signature of the free DNA, A_1 has been reduced which indicates that base flipping has occurred. 32% of the DNA-enzyme complexes are unflipped and the intrahelical 2AP is likely to form less-tightly stacked arrangements with the neighbouring DNA bases. The distribution of the flipped complexes with SAM shows a similar pattern to the flipped complexes with the AETA cofactor; most of the flipped conformations provide a sub-nanosecond quenching condition for the stabilised 2AP and the less-quenching conformations constitute progressively lesser proportions ($A_2>A_3>A_4$). Once again, τ_2 and τ_3 report substantial quenching of the 2AP by M.TaqI and τ_4 indicates that 2AP is quenched very little compared with a solvent-exposed 2AP.

It is acknowledged that the τ_2 and τ_3 lifetimes for the SAM complex are longer than their counterparts in the AETA complex, where the 2AP is know to interact with the catalytic NPPY motif. Even so, it is concluded that the flipped 2AP is still stabilised in NPPY in the majority of the ternary complexes made with SAM. Further evidence for this is presented in Chapter 7 and the reason for an increase in the 2AP lifetimes will be explained.

In the SAH and sinefungin complexes 2AP also has a four-component decay with lifetimes remarkably similar to those of the SAM complex. Since these parameters come from analysis of a single decay, the absolute magnitudes of τ_i and A_i are less certain. If the lifetimes were calculated by global analysis, it may be that there are effectively no differences between corresponding lifetime values for different M.TaqI-cofactor complexes (and recall that each lifetime simply characterises a distribution of similarly quenching conformations for the 2AP). Yet again, all the calculated lifetimes are longer compared with those of the complexes made with the AETA cofactor. This will be returned to in Chapter 7. More noteworthy are the differences in A factor distribution of the distinguishable complexes formed. A_1 reports that a much higher proportion of the duplexes are unflipped in the sinefungin complex than when SAM or SAH is the bound cofactor. $A_{2\text{-}4}$ vary across the SAH and sinefungin complexes too but these differences are small and not deemed to be important. Most revealing is that the SAM and SAH partitioning of complexes is quite similar whereas the sinefungin complex partitioning is markedly different. The DNA-M.TaqI complexes made with a sinefungin cofactor favour unflipped conformations.

Even though the order of DNA and cofactor binding is not important for the M.TaqI methyltransferase these samples were prepared by addition of the DNA to the enzyme first. Initial interest in performing fluorescence lifetime measurements with various cofactor analogues originated from the observation that each analogue causes a slightly different change in the steady state fluorescence from DNA-enzyme binary complex to DNA-enzyme-cofactor complex (unpublished data from T. Lenz). Figure 3.15 shows these measured intensity changes (red data bars). When any of the presented cofactors are added to the binary complex 2AP emission intensity decreases. The Φ_{rel} values, calculated from lifetime parameters (Table 3.6, blue data bars in Figure 3.15) are in good agreement with steady state observation with the exception of the SAM complex. Errors on the steady state data will be quite large due to the difficulty in assessing the absolute intensity of these low emission complexes whereas lifetime measurements are independent of concentration and provide a more reliable method of assessing these changes. Though steady state

emission intensities of the ternary complexes are quite similar lifetime data reveals that the choice of cofactor may cause complexes formed to be very different, AETA vs. sinefungin for example.

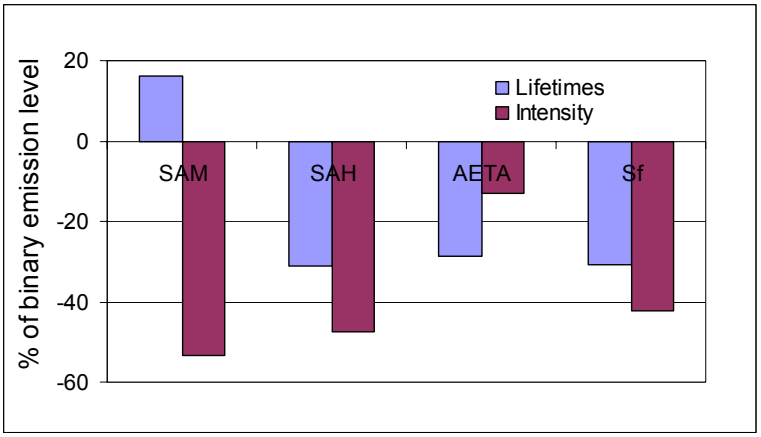


Figure 3.15 Changes in emission level effected when various cofactors are added to the 2AP-T/M.TaqI binary complex. The changes predicted from lifetime data (blue) are compared with experimentally observed changes from steady state measurements (red)

As noted, the addition of SAM to the DNA-enzyme binary complex causes the steady state intensity to decrease however Φ_{rel} values calculated from the decay parameters predict increased emission instead. Further steady state measurements on this ternary complex reveal a distortion to the emission profile under the excitation condition used for time-resolved experiments, where 320-nm excitation has been employed rather than 310 nm. These spectra are shown in Figure 3.16. It appears that the usual 2AP spectrum is broadened by another emission band emerging to the red. Steady state emission has been measured at 360 nm and will have monitored the normal emission expected from 2AP. Fluorescence decays have been measured under conditions in which 2AP gives out additional emission from the new red-shifted band, thus identifying why steady state and time-resolved experiments disagree. An investigation into the origin of the red-shifted emission band is the subject of Chapters 6 and 7. Chapter 7 specifically addresses red-shifted emission from 2AP inside DNA-enzyme complexes and further compares the ternary complexes formed between 2AP-T, M.TaqI and each of the cofactors discussed in this section.

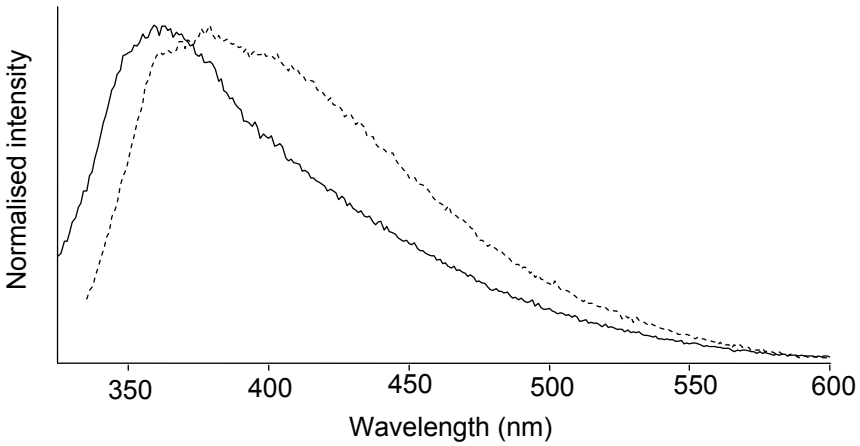


Figure 3.16 Normalised emission spectra for the 2AP-T/M.TaqI/SAM complex in solution with 310-nm excitation (solid line) and 320-nm excitation (dashed line) wavelengths.

The existence of an equilibrium mixture of flipped and unflipped complexes has biological relevance to the methyltransferase system. A flipped base that can un-stack from and re-stack into the duplex while the DNA is bound to enzyme, allows sampling of multiple extrahelical conformations including, but not exclusively comprising, ones favoured for catalysis. That many conformations can be sampled implies a looser association of the flipped base with the enzyme and this is functionally useful when it is considered that tight binding compromises catalysis[27]. Localised substrate release is essential for catalytic turnover. It allows a bound enzyme to move further along the DNA in search of the next recognition sequence (in a processive mechanism) or to detach from the DNA completely and re-bind to the DNA at a sequence much further away (distributive mechanism).

Presenting the decay data as discrete components conceals how broad a distribution of duplex conformations each lifetime component might encompass. Three components for the AETA complex over the four components for the SAM complex implies that a smaller range of duplex conformations are permissible in the AETA complex, so it appears that the flipped 2AP in the AETA complex is more constrained in solution. That AETA is able to lock 2AP into well defined, highly stable conformations perhaps facilitates the subsequent formation of crystalline complexes. Conversely, the reduced ability of the SAM complex to persistently hold

the duplex in a particular conformation might explain why ternary complexes containing SAM have yet to be obtained.

The time-resolved fluorescence of 2AP has allowed investigation of the influences of the cofactor analogues SAH and sinefungin on the flipped base compared with SAM. Previously Schluckebier crystallised each cofactor in a binary complex with M.TaqI and the refined structures show variation in the way that the amino acid tail part of these molecules lies within the enzyme[6]. The tails of SAH and sinefungin are arranged similarly to each other but differently to SAM, such that the SAH and sinefungin tails can form direct contacts with the NPPY motif. Though less informative of cofactor binding in the presence of bound DNA, these findings generate interest into whether SAH and sinefungin prevent adenine from entering NPPY or at least provide a barrier to its delivery into the catalytic motif. Liebert *et al.* have shown that an adenine flipped by the *E. coli* DNA adenine methyltransferase (EcoDam) does not go into the DPPY active site when SAH is bound[25] and have found a binary EcoDam-SAH crystal structure where the tail of SAH is folded into the adenine binding pocket and directly occupies DPPY[28]. Similarly, DPPY from the T4 Dam-SAH complex is not fully ordered around a flipped adenine[29]. However when sinefungin is the cofactor in the T4 Dam enzyme, adenine sits in DPPY in a catalytic conformation (i.e. hydrogen bonding with DPP and π-stacking with Y). Lifetime data reveal that both cofactor analogues permit base flipping by M.TaqI, and that the environments of the flipped base are only subtly different to the permissible conformations adopted with SAM. The flipped 2AP behaves very similarly with all three cofactors but, contrary to the binary crystals study, behaviour of the SAH complex is more similar to SAM than to sinefungin. The substitution of sinefungin for SAM decreases the extent of flipping of the bound DNA. This may be due to the sinefungin tail blocking the NPPY motif but the lifetime data cannot be used as confirmation.

Clearly, AETA and SAM cause M.TaqI to treat a target 2AP quite differently. SAH and sinefungin may be more appropriate substitutes for SAM in reliably reproducing the natural behaviour of the enzyme towards substrate DNA. With the newly

acquired lifetime data further proposals can be made as to what aspect of cofactor structure has the greatest effect on the M.TaqI-target base interaction. The AETA molecule instigates maximal active site trapping of the flipped 2AP over the other cofactors and this might relate to the absence of the zwitterionic $C(NH_2)(CO_2H)$ terminus. Removing charge on the sulfur seems to effect little variation but perhaps redistribution of the positive charge towards the methyl group position, as in sinefungin, effects a more persistent change in the enzyme. These queries provide a basis for new lines of investigation on the DNA-M.TaqI-cofactor complex and for cofactor engineering in the SMILing technique[30] (Sequence specific Methyltransferase Induced Labelling of DNA).

3.3.5 2AP-D/M.TaqI cofactors study

Table 3.7 presents lifetime data for analogous ternary complexes studied in the previous section but with the abasic site duplex 2AP-D.

2AP-D/M.TaqI/cofactor	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	Φ_{rel}
Free DNA (no enzyme)	0.04	0.26	2.5	8.3	0.62	0.18	0.09	0.11	<0.11
Binary (no cofactor)	-	0.23	3.3	9.7	-	0.26	0.18	0.56	0.57
AETA	-	0.28	2.8	9.6	-	0.33	0.30	0.37	0.42
SAM	-	0.99	5.3	9.7	-	0.06	0.28	0.66	0.74
SAH	-	-	4.0	9.8	-	-	0.23	0.77	0.80
Sinefungin	-	0.57	4.8	9.4	-	0.10	0.27	0.63	0.69

Table 3.7 Individual decay parameters for the 2AP-D/M.TaqI/cofactor ternary complex with various cofactors. Parameters are reported from decays measured with a 390-nm emission wavelength. Also reported are the calculated quantum yields relative to free 2AP riboside, Φ_{rel}. Free DNA and AETA data are global parameters, SAM decays measured at 3 wavelengths (see appendix A, Table A9) but individual parameters are reported.

Like the 2AP-T duplex, the new 2AP-D ternary complexes are different from the corresponding binary complex and from the 2AP-D/M.TaqI/AETA complex. As can be seen, τ_1 is absent in the decay signatures for the ternary complexes made with SAM, SAH and sinefungin, informing that complete base flipping has occurred. In contrast, the M.TaqI complexes with 2AP-T and these cofactors included some

conformations where 2AP was intrahelical (Table 3.6). The predominant conformation for the 2AP-D-containing complex with each cofactor is one in which 2AP experiences little quenching (τ_4). The value of τ_4 suggests that this conformation corresponds to 2AP in the exo site environment discussed in section 3.3.3. τ_3, at 4-5 ns, is longer than the τ_3 from the AETA complex, indicative of 2AP that is extrahelical and somewhat quenched, but having a weaker association with enzyme amino acids than in the AETA complex. For the complexes made with SAM and sinefungin there is a small population of a species with sub/near-nanosecond lifetime for the 2AP. Though these τ_2 values differ, it is feasible that both represent conformations where 2AP has some π-stacking interaction with tyrosine 108 within the active site of M.TaqI.

This brief study of the 2AP-D duplex in ternary complex with M.TaqI has revealed that the 2AP opposite an abasic site is flipped but does not become stabilised within the NPPY catalytic site when SAM, SAH or sinefungin is bound to the enzyme. As for the complexes made with the 2AP-T duplex, 2AP is inhibited from forming or maintaining an interaction with NPPY by each of these three cofactors (either directly or indirectly). In the 2AP-D-M.TaqI ternary complex with each of the cofactors in solution, the flipped 2AP is held in an environment of minimal quenching, outside the catalytic site, in a dominant or significant proportion of the population. This confirms that this behaviour is a consequence of the absence of the thymine partner, rather than an effect of the cofactor molecule.

3.3.6 M.TaqI mutants: mutation of the NPPY motif

Before solution of the Goedecke ternary structure, and given that aromaticity in NPPY was thought to be important to the catalytic mechanism, Pues *et al.* generated M.TaqI mutants[1] where the aromatic tyrosine (Y108) was replaced by either a phenylalanine (Y108F) or an alanine (Y108A). These mutations are illustrated in Figure 3.17). When tyrosine is replaced by another aromatic amino acid, catalytic activity was found to be similar to the wild-type activity level, whereas the mutant with no aromatic side chain on the 108 position had virtually no methylating ability. 2AP intensity was employed to characterise DNA substrate binding and base-flipping abilities of the mutants and, interestingly, different emission intensities were observed from the various DNA-M.TaqI complexes (wild type and mutants). Pues concluded that base-flipping occurs in both mutants and hypothesised that when flipped, the 2AP is subjected to differently stabilised and hence differently quenching environments inside these mutants. With confirmation of the face-to-face π-stacking stabilisation of adenine from crystal structure, a further hypothesis about these mutants arose. It is thought that removal of the hydroxyl group might enhance π-stacking with the flipped base causing Y108F to be better that the wild type at stabilising a flipped base. 2AP lifetime data should reveal the intricacies of the variation in steady state emission from these mutants and inform on the nature of the stabilising environments provided by each mutant when NPPY is altered.

Tables 3.9 and 3.10 present the decay parameters for the Y108F and Y108A mutants respectively. Decay parameters for the wild type 2AP-T/M.TaqI/SAM complex are reproduced in Table 3.8 for comparison. The predicted emission changes from the lifetime data agree with the intensity data (unpublished data from Thomas Lenz) in that the Y108F mutant effects an increase then a decrease in the intensity from free DNA to binary to ternary-SAM complex, whereas the Y108A mutant effects subsequent increases. Both types of fluorescence measurement report that 2AP has higher emission in the Y108A ternary complex than in the Y108F ternary complex.

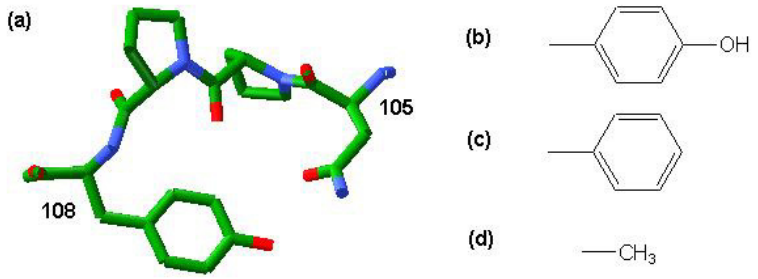


Figure 3.17 (a) The $N^{105}PPY^{108}$ catalytic motif, (b) the structure of the tyrosine side chain in the wild type catalytic motif, (c) the phenylalanine side chain of the Y108F mutant and (d) the alanine side chain of the Y108A mutant.

Solution composition	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	Φ_{rel}
2AP-T/M.TaqI/SAM	0.12	0.94	3.3	9.9	0.32	0.32	0.20	0.16	0.24

Table 3.8 Decay parameters for the 2AP-T duplex with wild type M.TaqI and SAM for the 390-nm decay.

Solution composition	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	Φ_{rel}
2AP-T alone	0.03	0.17	1.6	7.9	0.83	0.14	0.02	0.01	0.015
2AP-T/Y108F M.TaqI	0.08	0.69	5.2	9.2	0.22	0.10	0.12	0.55	0.55
2AP-T/Y108F M.TaqI/ SAM	0.06	0.71	5.3	9.5	0.47	0.19	0.10	0.25	0.28

Table 3.9 Individual decay parameters for the 2AP-T duplex, free and in binary and ternary complex with the Y108F M.TaqI mutant and SAM, for the decays measured at a 390-nm emission wavelength. Also reported are the quantum yields relative to free 2AP riboside, Φ_{rel}.

Solution composition	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	Φ_{rel}
2AP-T alone	0.03	0.17	1.6	7.9	0.83	0.14	0.02	0.01	0.015
2AP-T/Y108A M.TaqI	0.08	0.73	4.0	9.5	0.49	0.16	0.12	0.23	0.26
2AP-T/Y108A M.TaqI/SAM	0.09	0.74	4.2	9.0	0.16	0.39	0.07	0.38	0.38

Table 3.10 Individual decay parameters for the 2AP-T duplex, free and in binary and ternary complex with the Y108A M.TaqI mutant, for the decays measured at a 390-nm emission wavelength. Also reported are the quantum yields relative to free 2AP riboside, Φ_{rel}.

First it must be acknowledged that the free 2AP-T lifetime parameters are different from those reported earlier in Table 3.4 indicating variation in the state of the DNA used for the current experiment. New lifetimes for the intermediate components τ_2 and τ_3 inform of different quenching environments for the imperfectly stacked duplex conformations. Nevertheless the longest and shortest lifetimes have the same value as before, and A factors report that the proportion of the most highly stacked duplex conformation is still approximately 80%.

In the Y108F parameters of Table 3.9 the changes in A_1, informative of the degree of flipping, report that much more of the target 2AP is flipped in the binary than in the ternary Y108F complex. Furthermore the Y108F mutant is less efficient at base-flipping in the ternary complex than the wild-type enzyme (the latter leaves only a third of complexes with the intrahelical (τ_1) conformation). It appears that the addition of SAM to Y108F allows a up to quarter of the 2AP that was flipped in the binary complex to return to the interior of the helix, this quarter being taken from the 2APs residing in an environment of minimal enzyme interaction (τ_4, 9.2 ns). The Y108A mutant binary complex shows flipping in approximately 50% of the duplexes which then increases to 84% with the inclusion of SAM. Interestingly the Y108A mutant flips more of the bound complexes than the wild-type enzyme, providing that SAM is present. For both these mutants in ternary complex, a substantial proportion of the flipped 2AP adopts τ_4 indicating that a preferred position of the flipped base in the enzyme interior is an unquenched one. This is not the case for the analogous wild-type enzyme complex. Thus the mutants provide further evidence of the presence of an exo site inside M.TaqI, as was observed for the 2AP-D-M.TaqI complexes.

The values of the intermediate lifetimes for each mutant, τ_2 and τ_3, relating to 2AP inside the enzyme, are quite similar to each other and are less like the corresponding wild-type enzyme lifetimes. Incidentally, all the lifetimes in Table 3.8 (the wild type parameters) are lengthened by an additional factor (see in Chapter 7) and are actually much shorter than indicated than for the 2AP-mutant interaction. As will be seen in

Chapter 7, the lifetimes for the mutant decays are not artificially lengthened and so truly represent the 2AP quenching conditions being scrutinised.

The τ_2 and τ_3 lifetimes of the mutant decays are common to both mutants. Since the aromaticity at the 108th position is quite different in these mutants, these lifetimes cannot represent a π-stacking interaction between the 2AP and the 108th amino acid in M.TaqI. Species τ_2 must either be a 2AP-phenylalanine 196 interaction, since this is the other aromatic side chain nearby, or an imperfectly-stacked 2AP inside the duplex interior. It is feasible that τ_3 is another 2AP-phenylalanine 196 interaction, but one that is less quenching than a typical face-to-face π-stacking relationship. Thus it is presumed that τ_3 is governed by an edge-to-face π-stacking geometry. The original hypothesis of enhanced π-stacking stability from the tyrosine of wild type to the phenylalanine of Y108F mutant does not hold.

It could be assumed that complex conformations where 2AP is held in the NPPY motif (or NPPF for the Y108F mutant) in a geometry suited for catalysis will be contributing to the populations, A_2 and A_3, of the species 2 and 3, but the combined population, $A_2 + A_3$ (given in parenthesis), does not relate to catalytic competence. Y108F (29%) is half as effective as wild type (52%) at methylating a given proportion of DNA but the catalytic activity of Y108A (46%) is 400 times less and practically obsolete. It follows then that it is impossible to distinguish catalytic conformations of the flipped base within the conformational populations represented by these two decay parameters.

The inability of the lifetime data to rationalise relative catalytic ability indicates that there is, understandably, a higher degree of complexity in these DNA-mutant enzyme intermediates than can be contained in just four separable sets of lifetimes. It has not been possible to distinguish a clear difference between the two mutants in the available stabilising environments for a flipped 2AP within the active site when the 108th amino acid has been altered but there is an observed difference in the partitioning of the target base amongst available conformations.

A further conclusion from the mutant study by Pues was that catalytic activity relies on aromaticity at both position 108 and position 196, since even when NPPY is maintained, loss of aromaticity at the phenylalanine 196 position causes disappearance of methylating ability[1]. For an NPPF motif (Y108F mutant) slight reduction in catalytic activity compared with wild type can sensibly be connected with the greater proportion of unflipped 2AP present in the ternary complex as revealed by the decay parameters. Evidently the 108 and 196 aromatic groups act together so slight alterations in 108 do not dramatically affect enzyme function. An NPPA motif effects removal of catalytic activity but not of phenylalanine 196 stabilisation. Decay parameters find that Y108A has a greater flipping ability than the wild type enzyme and that the flipped complex is substantially stable, even without an aromatic sidechain at the 108 position. In approximately half of the flipped Y108A complexes, the DNA-enzyme interaction is one which does not result in quenching of the 2AP.

3.4 Conclusion

From the refined X-ray crystal structure of M.TaqI bound to a DNA duplex incorporating 2AP at the target site, it is clear that this methyltransferase does base flip the fluorescent adenine analogue. Furthermore the flipped 2AP is held inside the M.TaqI active site in the same position as the natural target adenine, establishing that this fluorescent probe is a good reporter for the adenine-M.TaqI interaction. Multiple lifetimes arising from the decay of 2AP in the crystalline ternary M.TaqI complex show that a range of conformations can be adopted by the base-flipped complex which in turn permits multiple conformations of the flipped base inside the enzyme binding pocket. Evidently the conformational states adopted by the crystalline DNA-enzyme complex are more varied than the refined crystal structure suggests. The characteristic lifetimes of 2AP in the M.TaqI active site are much shorter than that of a solvent-exposed 2AP, demonstrating that, unlike for M.HhaI base flipping, an M.TaqI-flipped base is still significantly quenched. The specific 2AP-M.TaqI interaction responsible for the quenching is the face-to-face π-stacking stabilisation of the 2AP by the aromatic tyrosine within the NPPY catalytic motif.

Remarkably similar lifetimes occur for the same DNA-M.TaqI-cofactor complex in solution which leads to the conclusion that the DNA-enzyme conformations which exist in solution are maintained when the enzyme is crystallised. It is important to a completely solution-based technique that base flipping is clearly identifiable in the lifetime data. For M.TaqI this identifier is the disappearance of the short lifetime species which is characteristic of highly stacked DNA with most of the 2AP intrahelical. Since 80% of the unbound DNA harbours a tightly intrahelical 2AP, M.TaqI is obviously very efficient at extruding its target base.

In order to investigate the role of the thymine which partners the target base in the M.TaqI recognition sequence, this thymine was removed from the duplex leaving 2AP opposite an abasic site. In the crystal phase, lifetimes of the flipped 2AP of this modified duplex (2AP-D) match those of the flipped Watson-Crick duplex (2AP-T), confirming that these lifetimes are indeed characteristic of an NPPY stabilised 2AP.

The solution-phase decay parameters for the abasic-site duplex show that, in this case, the DNA-enzyme conformations that have crystallised are only a selection of the set which exist in solution. When in solution, an equally favoured conformation of the 2AP-D duplex is one where 2AP forms minimal specific contact to M.TaqI despite having been flipped into the interior of the enzyme.

The decays of the target 2AP become more complicated when the natural cofactor SAM is bound to M.TaqI in place of the unreactive AETA cofactor analogue. Parameters are indicative of both enzyme-stabilised and duplex-stacked 2APs showing that, in the presence of the natural cofactor, the target base is not completely locked into the active site and may also reside in the duplex interior. Substitution of the cofactor analogues SAH and sinefungin cause M.TaqI to treat the target base in the same manner as if SAM were bound with only subtle differences in the partitioning of the equilibrium of DNA-enzyme conformations. SAH and sinefungin reproduce the enzyme's native behaviour towards substrate DNA more closely than the AETA analogue.

Finally, variants of M.TaqI, where the tyrosine (Y108) of the catalytic motif NPPY had been mutated, were examined. Either aromaticity was removed (alanine) or altered (phenylalanine) so as to remove or supposedly enhance the π-stacking ability of the 108th residue, respectively. On encountering substrate DNA, both mutants were found to flip a target 2AP but to hold it in quite a different way to the wild-type enzyme, without apparent use of NPPY-like stabilisation of the flipped base. Interestingly the attempt to create two mutants which are starkly different in active site design gives rise to M.TaqI variants that manipulate the DNA in a very similar fashion.

3.5 References

(1) Pues, H.; Bleimling, N.; Holz, B.; Wolcke, J.; Weinhold, E. *Biochemistry* **1999**, *38*, 1426-1434.

(2) Goedecke, K.; Pignot, M.; Goody, R. S.; Scheidig, A. J.; Weinhold, E. *Nature Structural Biology* **2001**, *8*, 121-125.

(3) Labahn, J.; Granzin, J.; Schluckebier, G.; Robinson, D. P.; Jack, W. E.; Schildkraut, I.; Saenger, W. *PNAS* **1994**, *91*, 10957-10961.

(4) Cheng, X.; Roberts, R. J. *Nucleic Acids Research* **2001**, *29*, 3784-3795.

(5) Schluckebier, G.; Ogara, M.; Saenger, W.; Cheng, X. D. *Journal of Molecular Biology* **1995**, *247*, 16-20.

(6) Schluckebier, G.; Kozak, M.; Bleimling, N.; Weinhold, E.; Saenger, W. *Journal of Molecular Biology* **1997**, *265*, 56-67.

(7) Weinhold, E.; Serva, S.; Klimasauskas, S.; Holz, B. *Nucleic Acids Research* **1998**, *26*, 1076-1083.

(8) Lenz, T.; Bonnist, E.; Pljevaljcic, G.; Neely, R. K.; Dryden, D. T. F.; Jones, A. C.; Scheidig, A. J.; Weinhold, E. *Journal of the American Chemical Society* **2007**, *129*, 6240-6248.

(9) Blumenthal, R. M.; Cheng, X. D. *Nature Structural Biology* **2001**, *8*, 101-103.

(10) Neely, R. K.; Daujotyte, D.; Grazulis, S.; Magennis, S. W.; Dryden, D. T. F.; Klimasauskas, S.; Jones, A. C. *Nucleic Acids Research* **2005**, *33*, 6953-6960.

(11) Guest, C. R.; Hochstrasser, R. A.; Sowers, L. C.; Millar, D. P. *Biochemistry* **1991**, *30*, 3271-3279.

(12) Zewail, A. H.; Barton, J. K.; Schiemann, O.; Fiebig, T.; Wan, C. *PNAS* **2000**, *97*, 14052-14055.

(13) Ross, J. B. A.; Osman, R.; Rachofsky, E. L. *Biochemistry* **2001**, *40*, 946-956.

(14) Jean, J. M.; Hall, K. B. *PNAS* **2001**, *98*, 37-41.

(15) Somsen, O. J. G.; Hoek, V. A.; Amerongen, V. H. *Chemical Physics Letters* **2005**, *402*, 61-65.

(16) Barton, J. K.; O'Neill, M. A. *Journal of the American Chemical Society* **2004**, *126*, 11471-11483.

(17) Fiebig, T.; Wan, C. Z.; Zewail, A. H. *Chemphyschem* **2002**, *3*, 781-788.

(18) Neely, R. K. The Photophysical Properties of 2-Aminopurine and its Application as a Probe of DNA-Protein Interactions. Thesis 2005.

(19) Nordlund, T. M.; Andersson, S.; Nilsson, L.; Rigler, R.; Graslund, A.; Mclaughlin, L. W. *Biochemistry* **1989**, *28*, 9095-9103.

(20) O'Neill, M. A.; Becker, H. C.; Wan, C. Z.; Barton, J. K.; Zewail, A. H. *Angewandte Chemie-International Edition* **2003**, *42*, 5896-5900.

(21) O'Neill, M. A.; Barton, J. K. *Journal of the American Chemical Society* **2004**, *126*, 13234-13235.

(22) Rachofsky, E. L.; Seibert, E.; Stivers, J. T.; Osman, R.; Ross, J. B. A. *Biochemistry* **2001**, *40*, 957-967.

(23) Pompizi, I.; Haberli, A.; Leumann, C. J. *Nucleic Acids Research* **2000**, *28*, 2702-2708.

(24) Stivers, J. T. *Nucleic Acids Research* **1998**, *26*, 3837-3844.

(25) Horton, J. R.; Liebert, K.; Bekes, M.; Jeltsch, A.; Cheng, X. D. *Journal of Molecular Biology* **2006**, *358*, 559-570.

(26) Banerjee, A.; Yang, W.; Karplus, M.; Verdine, G. L. *Nature* **2005**, *434*, 612-618.

(27) Bheemanaik, S.; Reddy, Y. V. R.; Rao, D. N. *Biochemical Journal* **2006**, *399*, 177-190.

(28) Liebert, K.; Horton, J. R.; Chahar, S.; Orwick, M.; Cheng, X. D.; Jeltsch, A. *Journal of Biological Chemistry* **2007**, *282*, 22848-22855.

(29) Horton, J. R.; Liebert, K.; Hattman, S.; Jeltsch, A.; Cheng, X. D. *Cell* **2005**, *121*, 349-361.

(30) Pljevaljcic, G.; Pignot, M.; Weinhold, E. *Journal of the American Chemical Society* **2003**, *125*, 3486-3492.

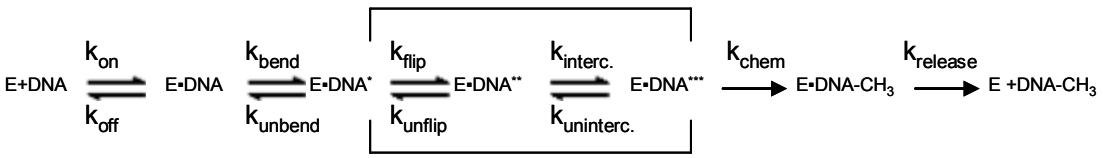
Chapter 4

The M.EcoRI Methyltransferase

4.1. Introduction

The *Escherichia coli* RI methyltransferase (M.EcoRI) is one of the many methyltransferases where no crystal structure has been obtained and Hopkins and Reich, the first people to employ 2AP as an indicator of base flipping, used M.EcoRI as their base flipping enzyme[1]. M.EcoRI (molecular weight 38,050 Da, 326 amino acids[2]) is another type II restriction-modification methyltransferase i.e. DNA cutting is performed by a separate enzyme. This methyltransferase binds DNA non-specifically and adopts a facilitated diffusion mechanism to locate its cognate GAATTC sequence[3;4] before methylating the second adenine along. M.EcoRI is known to encompass two additional base pairs either side of its recognition sequence resulting in enzyme contact with a total of ten base pairs[5] and sequence specific binding results in a 52^{o} bend to the DNA[6]. Unlike M.TaqI, M.EcoRI has an ordered binding mechanism; the cofactor must first be bound before a stable DNA-enzyme complex can form[7]. Sequence and motif alignments reveal that M.EcoRI is not closely related to other adenine methyltransferases since matches for motifs I-III, V, VII, VIII and X are poorly observed[5], nevertheless, a good match is found for the catalytic motif IV, comprising NPPF (amino acids 139-142)[5;8;9].

Scheme 4.1 describes the processes that have been characterised within the mode of action of M.EcoRI[10].



Scheme 4.1 The proposed mechanism of action for the M.EcoRI methyltransferase (E)[10], the square brackets encompass intermediates which have conformations favourable for catalysis.

When M.EcoRI encounters substrate DNA the enzyme binds, bends the DNA, flips out the target base and stabilises the flipped duplex by amino acid intercalation. The flipped base undergoes catalysis and the enzyme then releases the methylated DNA. These processes and their rates have been identified by various different means; binding (and release) by the fluorescence anisotropy of rhodamine-x isothiocyanate-labelled DNA[11]; flipping by the steady state fluorescence of 2AP-labelled DNA[1]; bending by atomic force microscopy and Förster Resonance Energy Transfer (FRET) measurements between fluorophores placed at either end of the substrate duplex[11;12]; and catalysis by the incorporation of ^{3}H radiolabelled methyl groups to DNA after controlled incubation times with varying enzyme concentrations[12].

M.EcoRI binding to the 2AP-labelled DNA causes a 14-fold intensity increase which was specifically attributed to the disruption of base stacking, hence flipping[1]. The FRET measurements revealed an unexpected decrease in the energy transfer between the FRET pairs after the DNA-bending signal[5]. Intercalation was interpreted by the reasoning that an end-to-end expansion of the DNA is caused by amino acid invasion into the duplex. The use of *two-colour DNA* encompassing both 2AP and a FRET pair allowed distinguishable, real-time monitoring of binding, bending and flipping[11]. Measurement at 4 ^{o}C slowed these processes enough to establish their relative order[10] and rate of occurrence: bending happens quickly and is followed by a near-simultaneous flipping and intercalation.

E·DNA, E·DNA*, E·DNA** and E·DNA*** are regarded as *conformational intermediates*, that is, DNA-M.EcoRI complexes which hold the bound duplex in a distinct conformation. Substrate release is the rate-limiting step and, with regards to catalysis, two separate processes are observed[11]. Under pre-steady state conditions, M.EcoRI gives a burst of product formation (burst kinetics) followed by a slower and rate-limiting product release phase. The reason for such behaviour is though to stem from the separate conformational intermediates identified, one conformation forms early and this intermediate allows catalysis and facile substrate release while another conformation, also catalytic, forms more slowly and, when methylated, this intermediate does not detach from the enzyme as easily.

The key to the specificity of M.EcoRI is thought to lie in the partitioning of and transition between these intermediates manifested through differences in conformation[10]. Consider the sequence GAATT**T** i.e. one that contains a single base difference to the cognate sequence. When M.EcoRI encounters this non-cognate sequence the enzyme binds with almost equal affinity. Some methylation of the second adenine occurs but catalytic efficiency is 3500-fold lower[10] than for methylation of the cognate sequence. Youngblood and Reich rationalise that the partitioning between bent, flipped and intercalated complexes with the non-cognate duplexes is different from that of the cognate duplex and that there is a bias away from the intercalated species for this non-cognate DNA[10]. Though M.EcoRI readily binds to a nearly identical DNA sequence, it is thought that destabilisation of the conformations which lead to catalysis prevents methylation of non-target sites.

The M.EcoRI mutant H235N is a less efficient methylator but much more selective of the target sequence. For this mutation, histidine 235 (H) has been replaced with an asparagine (N). H235N binds cognate DNA with a 10-fold lower affinity than the wild-type enzyme and is bending-impaired[12]. Some bending is detected by FRET measurements but not enough to be identified with atomic force microscopy (personal communication with B. Youngblood). Compared to wild-type M.EcoRI, steady state 2AP measurements indicate a 30% higher quantum yield for mutant-induced flipping[12]. Since supporting studies find that H235N does not show burst kinetics (k_{chem} is several orders of magnitude lower than wild-type. Scheme 4.1) and has a lower rate of 2AP unstacking, the higher intensity 2AP signal was interpreted as a different and less quenching environment for the flipped base rather than a higher population of flipped complexes. Under pre-steady state conditions the mutant may be just as good as the wild-type at methyl transfer but there is a slower transition to the catalytic conformation. Furthermore, while wild-type M.EcoRI will readily methylate non-cognate "target" sites under suitably forcing conditions, the H235N mutant does not, showing that the mutant is much more discriminating than the wild-type[12].

Now, over 10 years on from the original 2AP base flipping paper, the time-resolved fluorescence of 2AP allows a more intricate investigation of base flipping by the

M.EcoRI methyltransferase. This chapter presents decay parameters for M.EcoRI with 2AP-labelled cognate DNA in order to study the interaction of this enzyme with its target base. The near-identical GAATTT sequence is also investigated so as to observe how M.EcoRI treats a non-target adenine in a substrate that is as readily bound as the cognate DNA. Finally, equivalent measurements are made with the H235N mutant and the same two DNA substrates to compare 2AP flipping and stabilisation with the wild-type and to help rationalise why a non-target base is ignored by this mutant but not by the wild-type enzyme.

4.2. Materials and Methods

M.EcoRI was purified from MM294 *E. coli* cells with the plasmid pXRI as previously described[5]. 6 litres of culture were induced with 1 mM IPTG at an $OD_{600\ nm}$ 0.4 for 3 hours. Cells were sonicated in extraction buffer (200 mM NaCl, 6.5 mM K_2HPO_4, 3.5 mM KH_2PO_4, 1 mM EDTA and protease inhibitor mixture (Sigma)). The lysate was centrifuged (30 min at 4 °C), the supernatant was loaded onto a 20-ml phosphocellulose (Whatman) column and the protein was eluted with a salt gradient between 200 mM and 800 mM NaCl. Fractions containing M.EcoRI were loaded onto a 10-ml hydroxyapatite (Bio-Rad) column equilibrated with extraction buffer. The protein was eluted with a gradient between 6.5 mM and 1M KH_2PO_4, and dialyzed overnight in extraction buffer containing 10% glycerol. The dialyzed protein was loaded onto a biorex (Bio-Rad) column, and the protein was eluted with a gradient of 200 mM and 800 mM NaCl. Protein from this procedure was typically greater than 95% pure by densitometry. Enzymes were stored at –80 ºC with 10 % (V/V) glycerol, 0.3-0.4 M NaCl, 10 mM potassium phosphate (pH 7.8), 1 mM EDTA, and 7.5 mM β-mercaptoethanol.

The H235N M.EcoRI mutant and 2AP-labelled duplexes were provided by Benjamin Youngblood and Norbert Reich (University of California, Santa Barbara). Sinefungin was purchased from Sigma-Aldrich and dissolved in water.

The sequences of the 2AP-labelled 14-mer oligonucleotides are listed in Table 4.1. All bottom strand target adenines were methylated to force the orientation of the enzyme to target the top strand for flipping.

Duplex	Base Sequence
CBM	5′-AGAC **GAP TTC** CGAA-3’ 3’- TCTG **CTT MAG** GCTT-5’
A4	5′-AGAC **GAP TTT** CGAA-3’ 3’-TCTG **CTT MAA** GCTT-5’

Table 4.1 The base sequences for M.EcoRI substrate duplexes. The recognition sequence for M.EcoRI is emboldened, P designates 2-aminopurine and M designated N6-methyladenine.

Decays were measured for free DNA and for the ternary complex formed from 2 μM DNA, 6 μM M.EcoRI and 14 μM sinefungin. All solutions were buffered with 100 mM NaCl, 50 mM Tris, 5 mM EDTA and 1 mM DTT (pH 8). In preparation of the ternary complex enzyme and sinefungin were combined first and allowed to incubate for 5 minutes before the addition of the DNA. Time-resolved fluorescence was carried out as described in Chapter 2. The full set of calculated decay parameters for the duplexes and DNA-enzyme complexes in this chapter are given in Tables A10 to A15 in appendix A.

4.3. Results and Discussion

4.3.1. Wild-type and mutant enzyme with cognate sequence DNA

Table 4.2 presents the global decay parameters for the cognate sequence DNA substrate (CBM, Cognate Binding Methylated) alone and in complex with wild-type M.EcoRI or the H235N mutant enzyme.

Solution Composition	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	Φ_{rel}
CBM alone	0.14	0.54	2.6	8.7	0.50	0.35	0.09	0.06	0.10
CBM/wild-type M.EcoRI/Sf	0.17	1.27	4.4	8.4	0.24	0.14	0.09	0.53	0.49
CBM/mutant M.EcoRI/Sf	0.16	0.88	3.9	8.0	0.25	0.13	0.08	0.54	0.46

Table 4.2 Global lifetimes (τ_i) and A factors (A_i) for the CBM duplex free, and in complex with wild-type or H235N mutant M.EcoRI and sinefungin. The A factors are reported for the 390-nm decays. Also reported are the calculated quantum yields relative to free 2AP riboside Φ_{rel}.

The *CBM alone* decay parameters show that 2AP at the target position in this particular sequence context is not as strongly quenched as in the M.TaqI sequence (see Chapter 3, Table 3.4). The most highly stacked duplex conformations give 2AP a 140 ps lifetime, an order of magnitude higher than τ_1 for 2AP-T, and this condition occurs in only 50% of the duplexes. That said, these decay parameters are still typical of 2AP in DNA and reflective of the local sequence context of this fluorophore[13] (see Chapter 6). It should be remembered that what is of most interest here is how these parameters change when the DNA is bound by the methyltransferase. For the unbound M.EcoRI substrate 85% of the DNA is regarded as being in highly stacked conformations i.e. ones that give rise to sub-nanosecond lifetimes.

Considering wild-type M.EcoRI complex first (wt), it can be seen that up to 76% of the complexes are flipped as only 24% of the 2APs have a sub-nanosecond lifetime. Although sub-nanosecond lifetimes can also represent a flipped 2AP, as found for

M.TaqI flipping, the resolved τ_1 value for CBM-wt is not very different from τ_1 for CBM alone suggesting that the intrahelical 2AP environment has just been slightly perturbed. The remaining lifetimes, $\tau_{2\text{-}4}$, are considered to be mainly extrahelical conformations. The sets of conformations characterised by τ_2 and τ_3 are expected to contain some duplex-stacked 2AP but since these lifetimes are substantially different from their counterparts in the free duplex, it can be assumed that most conformations in each set are M.EcoRI-flipped complexes. The longest lifetime component provides a quenching environment similar to the one for the free duplex, where 2AP is already removed from the duplex interior. This extrahelical conformation is preferred by the enzyme, being adopted by 53% of the DNA-enzyme complexes. The lifetime data predict a 5-fold increase in the emission of 2AP when cognate DNA is bound by M.EcoRI. This is approximately a third of the signal rise published by Hopkins but is consistent with the 6-fold steady state increase published by Szegedi *et al.*[14] and also observed by B. Youngblood for recently prepared enzyme (personal communication).

The catalytic motif in M.EcoRI comprises NPPF and it is envisaged that a flipped 2AP will be stabilised within it in a similar manner to the NPPY of M.TaqI; hydrogen bonding with the asparagine (N) and face-to-face π-stacking with phenylalanine (F). In addition there is scope for an edge-to-face π-stacking interaction with phenylalanine 232 in motif VIII of M.EcoRI, which is thought to form part of the adenine binding pocket[1]. Phenylalanine 232 is equivalent to the phenylalanine 196 of M.TaqI (see Figure 3.4). Should NPPY-like stabilisation occur inside M.EcoRI, lifetimes of 0.55 and 1-3 ns are expected (see Table 3.4, the 2AP-T/M.TaqI/AETA crystal- and solution-phase complexes). The 1.27 and 4.4 ns lifetimes for the CBM-wt complex are sufficiently short as to be representative of 2AP quenching in the M.EcoRI active site but are also likely to include unflipped and duplex-associated 2APs (particularly τ_2). Presuming that some of the conformations represented by species τ_2 and τ_3 place 2AP in the adenine binding pocket of M.EcoRI and near NPPF, these will be the ones that can undergo catalysis. Therefore only the complexes with these intermediate lifetimes are thought capable of methyl transfer since these might hold the 2AP in the suitable geometry. The

combined A factors for these components suggest an upper limit of 23% for the proportion of DNA-M.EcoRI complexes that have catalytic competency.

It is interesting to compare the DNA-M.EcoRI complex to the M.TaqI Y108F mutant as both have NPPF as their catalytic motif. From the lifetimes of the DNA-Y108F complex (Table 3.9), little can be said with regards to direct NPPF stabilisation. This is because the 2AP held inside Y108F does not strongly associate with the catalytic motif, unlike the wild-type M.TaqI (see Chapter 3). More helpful is the value of τ_3 in both the Y108F and the Y108A M.TaqI mutants. This took a 4.0-5.3 ns lifetime and was interpreted as the quenching condition for 2AP when sitting edge-to-face with the π-system of phenylalanine 196. Thus, the designation of the 4.4 ns lifetime in the DNA-M.EcoRI complex is supported.

Evidently, the M.EcoRI-flipped complexes formed do not favour catalytically competent conformations. Given that such behaviour is paradoxical to the role of an enzyme, it is believed that the use of sinefungin as cofactor is the cause. This is with regards to flipped base geometry and partitioning of conformational intermediates of course, since 2AP cannot be methylated and sinefungin has no labile methyl group. In Chapter 3 the choice of cofactor was shown to influence the extent of base flipping for M.TaqI, 55% of the DNA-M.TaqI complexes with sinefungin are unflipped compared with just 32% of the complexes with the natural cofactor SAM (Table 3.6). Hopkins and Reich specifically report that for M.EcoRI an equivalent substrate release step occurs at a lower rate when the cofactor is sinefungin instead of SAH[5]. Further support for this concept is provided in Chapter 7.

In this case interpretation of M.EcoRI behaviour must be done with caution since the sinefungin complex may give completely different partitioning to the catalytically competent SAM complex. The decay parameters for the CBM/wild-type/Sf complex suggest that a flipped 2AP mostly samples non-catalytic conformations inside M.EcoRI. This means that many of the complexes being observed in the lifetime data are ones that do not feature in the mechanism of Scheme 4.1 since the reaction coordinate only includes precursors that will lead to catalysis.

The multiple-lifetime decay for cognate sequence DNA bound to M.EcoRI shows that this complex exists in a range of conformational states, which are resolvable into four sets. Evidently the target base can sample several different conformations when flipped but it is not known whether these conformations necessarily lie on the mechanistic pathway (though some inevitably will). With regards to the model proposed in Scheme 4.1, the unflipped conformations, τ_1 and those within τ_2, will relate to the bound and bent species (E·DNA and E·DNA*). In the flipped complexes it is unlikely that the lifetime data can distinguish between E·DNA** (not intercalated) and E·DNA*** (intercalated). Both are already formed under the equilibrium conditions of this experiment and probably provide similar quenching environments for the flipped 2AP, thus $\tau_{2\text{-}4}$ will be reflective of both. With regard to conformations that result in two observably different release rates, if it is assumed that the substrate can only be released when unflipped and unbent[11] a scenario is envisaged where the unflipped conformations permit a faster rate of release than the extrahelical conformations that must first re-stack into the DNA.

In the decay of the H235N mutant complex the A factor distribution appears unchanged but all the lifetimes are appreciably shortened compared with the wild-type decay. Table 4.2 shows that for both wild-type and mutant enzymes approximately 25% of the bound duplexes remain unflipped. The τ_1 value for this mutant enzyme complex is comparable with the other τ_1 values in Table 4.2, which is expected if species 1 is not an enzyme-associated 2AP. Similarly, in most of the flipped mutant-complexes the 2AP has a lifetime indicative of very little quenching. Although most of the DNA-mutant complexes are apparently unbent at equilibrium, this has not prevented base flipping. The values taken by the intermediate lifetimes τ_2 and τ_3, however, show that the mutant and wild-type enzyme behave differently.

The simplest interpretation of τ_2 and τ_3 shortening is that there is a difference in quenching of the 2APs that are flipped into the adenine binding pocket of the M.EcoRI mutant. $\tau_{2\text{-}4}$ for the mutant complexes differ appreciably from their free-DNA equivalents but less so from their equivalents in the wild-type complex supporting that these encompass mostly enzyme-associated 2APs. In this instance,

let us consider how the H235N mutation might directly affect the environment of flipped base. The 235th amino acid is within the 228-236 chain that has been mapped to the motif VIII loop[12] and, by analogy with M.TaqI, is thought to form part of the adenine binding pocket of M.EcoRI. In M.TaqI the equivalent amino acid is lysine 199 (positioned within the orange loop in Figure 3.1 and shown in the adenine binding pocket of M.TaqI in Figure 3.3). This lysine is believed to act as a hinged gate at the DNA-M.TaqI interface, open before base flipping and closed afterwards[15]. The structures of lysine, histidine and asparagine are presented in Figure 4.1.

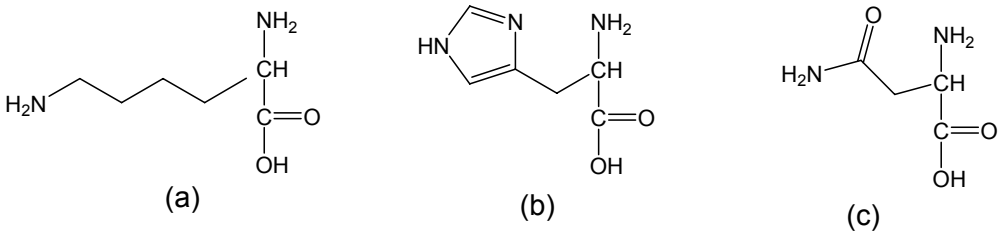


Figure 4.1 Structures of the amino acids (a) lysine [K] (b) histidine [H] and (c) asparagine [N]

Substitution of histidine 235 by an asparagine removes the acid/base properties and hydrogen bonding capabilities of the amino acid at this position but maintains steric bulk in the side chain. In the H235N mutant the asparagine at position 235 might cause the 2AP to sit closer to NPPF or phenylalanine 232 in the complexes where the 2AP is inside the adenine binding pocket. This would lead to improved quenching of the 2APs by the phenylalanine side chains. A tighter binding condition for the flipped base inside the mutant is consistent with the observation that the mutant releases the CBM substrate 4 times more slowly than the wild-type enzyme[12]. It should be noted, however, that the difference in the molecular environment of the flipped base inside H235N is only subtle and present in just 25% of the complexes. This would explain why the difference in DNA release rate is so minor.

An alternative interpretation of the differences in τ_2 and τ_3 relate to Scheme 4.1 and to the idea of a different partitioning between the different conformational intermediates that the enzyme complex can take. In this instance it is more important to recall that the distributions characterised by τ_2 and τ_3 encompass both flipped and unflipped complexes. The lifetimes for the mutant complexes have values that are

closer to the intermediate lifetimes of the free duplex, particularly τ_2, and this may indicate that there is a higher proportion of unflipped complexes in these two sets. Furthermore, it is reasonable to accept that no change in A factor has been detected with the current resolution. Allan *et al.* observe that the rate of base flipping by the mutant is clearly lower compared with the wild-type (depending on the proportion of enzyme:DNA, using a 100s of nanomolar concentration range)[12] so there could well be more unflipped mutant complexes present at equilibrium.

4.3.2. Wild-type and mutant enzyme with non-cognate sequence DNA

Table 4.3 presents the global decay parameters for the wild-type and H235N mutant M.EcoRI with the A4 non-cognate substrate, containing a GAATT**T** sequence and incorporating 2AP in the second adenine position. Recall that this sequence differs from the cognate GAATT**C** sequence by a single base pair modification in the duplex.

Solution Composition	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	Φ_{rel}
A4 alone	0.15	0.46	2.5	9.9	0.51	0.36	0.04	0.09	0.12
A4/wild type M.EcoRI/Sf	0.17	0.74	3.6	9.2	0.49	0.23	0.09	0.19	0.25
A4/mutant M.EcoRI/Sf	0.15	0.49	2.7	9.6	0.51	0.33	0.06	0.10	0.13

Table 4.3 Global lifetimes (τ_i) and A factors (A_i) for the A4 duplex free, and in complex with wild-type or H235N mutant M.EcoRI and sinefungin. The A factors are reported for the 390-nm decays. Also reported are the calculated quantum yields relative to free 2AP riboside Φ_{rel}.

The lifetime signature of the A4-wild-type complex shows that some base flipping of the 2AP has occurred, but that half of the complexes remain unflipped (τ_1, A_1). τ_2 and τ_3 for the enzyme complex differ from τ_2 and τ_3 for the free A4 duplex indicating that the enzyme has induced changes in the molecular environment of 2AP. It appears that for this non-cognate duplex, M.EcoRI has been able to perturb the imperfectly-stacked duplex conformations but not those where 2AP is tightly intrahelical. The extrahelical and solvent exposed 2AP from the unbound A4 DNA

(τ_4, 9.9 ns) is presumed to be in the M.EcoRI interior although the extent of quenching of this 2AP species in complex has not changed. Compared with the cognate sequence (Table 4.2), lifetimes and A factor distribution differ when M.EcoRI binds the non-cognate sequence showing that cognate and non-cognate DNA-enzyme conformations are not the same. This, together with the high proportion of intrahelical 2AP, supports the previous hypothesis of Youngblood and Reich that M.EcoRI partitions A4 away from the catalytic conformations in Scheme 4.1[10].

The decay parameters of A4 change very little when H235N binds this duplex. This is consistent with the study by Allan *et al.* which determined that the H235N mutant is much more discriminating against non-target sites than wild-type M.EcoRI[12]. Thus, it appears important to the H235N mutant that the terminal base in GAATTC is a cytosine. Evidently the asparagine in the 235-position has a role in sequence recognition prior to base flipping.

Instead of concluding that wild-type M.EcoRI is fooled into flipping a non-target base a more feasible interpretation lends itself to the context of the facilitated diffusion mechanism[3] (the way in which methyltransferases are believed to sample stretches of DNA[16]). *In vivo* methyltransferases must locate their recognition sequence amongst vast lengths of non-cognate base combinations. Diffusion alone makes the likelihood of cognate sequence encounter improbable but in the facilitated diffusion mechanism, the methyltransferase stays bound to non-cognate substrate and slides along a portion of the DNA searching for a cognate sequence. Therefore, it is necessary that M.EcoRI forms some persistent interaction with non-specific (non-cognate) DNA sequences in order to stay connected to the DNA whilst searching. Thus what is being observed in the decay of a non-target 2AP could be a non-specific M.EcoRI-duplex association.

4.4. Conclusion

From the original M.EcoRI base flipping study it was assumed that on enzyme binding all the target 2APs in the sample were stabilised extrahelically. Time-resolved fluorescence reveals that, at equilibrium, up to a quarter of complexes contain an unflipped target base. The majority of the flipped target 2APs sit in an environment inside the enzyme which is relatively unquenched and, in conjunction with the M.TaqI mutants study, it is concluded that these preferred conformations do not allow catalysis. The flipped, non-catalytic intermediates observed here do not feature on the reaction coordinate and arise as a result of the choice of sinefungin as cofactor. Further demonstration of this is given in Chapter 7.

When M.EcoRI is in complex with a non-cognate sequence, one that is closely related to the cognate sequence, the decay of 2AP in the analogous target position shows that only the imperfectly stacked duplex conformations have been disturbed. The resulting DNA-M.EcoRI contacts are believed to be non-specific yet stable enough to allow the enzyme to stay bound.

The bending-impaired H235N mutant of M.EcoRI also flips the target base from the cognate sequence. Most of the flipped DNA-enzyme complexes with this mutant hold a similar conformation to the wild-type enzyme, however, a small proportion of mutant complexes give 2AP a different lifetime to the wild-type equivalents. It is unclear as to whether the mutant provides a different environment for the flipped 2AP or whether less of the DNA-mutant complexes are flipped. In this case the decay parameters do not have the sufficient resolution to distinguish between these two possibilities. When bound to the non-cognate sequence, H235N does not change the decay signature of a 2AP at the analogous target site. In this regard the mutant is more discriminating towards this non-cognate sequence than the wild-type since the mutant leaves the non-target base alone.

4.5. References

(1) Reich, N. O.; Allan, B. W. *Biochemistry* **1996**, *35*, 14757-14762.

(2) The Restriction Enzyme Database (REBASE). 2007. http://rebase.neb.com

(3) Surby, M. A.; Reich, N. O. *Biochemistry* **1996**, *35*, 2201-2208.

(4) Surby, M. A.; Reich, N. O. *Biochemistry* **1996**, *35*, 2209-2217.

(5) Hopkins, B. B.; Reich, N. O. *Journal of Biological Chemistry* **2004**, *279*, 37049-37060.

(6) Garcia, R. A.; Bustamante, C. J.; Reich, N. O. *PNAS* **1996**, *93*, 7618-7622.

(7) Reich, N. O.; Mashhoon, N. *The Journal of Biological Chemistry* **1990**, *265*, 8966-8970.

(8) Malone, T.; Blumenthal, R. M.; Cheng, X. D. *Journal of Molecular Biology* **1995**, *253*, 618-632.

(9) Landry, D.; Barsomian, J. M.; Feehery, G. R.; Wilson, G. G. *Methods in Enzymology* **1992**, *216*, 244-259.

(10) Youngblood, B.; Reich, N. O. *Journal of Biological Chemistry* **2006**, *281*, 26821-26831.

(11) Allan, B. W.; Reich, N. O.; Beechem, J. M. *Biochemistry* **1999**, *38*, 5308-5314.

(12) Allan, B. W.; Garcia, R.; Maegley, K.; Mort, J.; Wong, D.; Lindstrom, W.; Beechem, J. M.; Reich, N. O. *Journal of Biological Chemistry* **1999**, *274*, 19269-19275.

(13) Neely, R. K. The Photophysical Properties of 2-Aminopurine and its Application as a Probe of DNA-Protein Interactions. Thesis 2005.

(14) Szegedi, S. S.; Reich, N. O.; Gumport, R. I. *Nucleic Acids Research* **2000**, *28*, 3962-3971.

(15) Goedecke, K.; Pignot, M.; Goody, R. S.; Scheidig, A. J.; Weinhold, E. *Nature Structural Biology* **2001**, *8*, 121-125.

(16) Halford, S. E.; Marko, J. F. *Nucleic Acids Research* **2004**, *32*, 3040-3052.

Chapter 5

The M.EcoRV Methyltransferase

5.1. Introduction

The final methyltransferase to be investigated in this study is M.EcoRV. This 299 amino acid, 34,637 Da molecular weight enzyme[1] recognises the GATATC sequence and methylates the first (5') adenine. The catalytic motif of M.EcoRV comprises DPPY (amino acids 193-196)[2]. This methyltransferase brought the link between 2AP emission enhancement and base flipping into doubt and exposes ambiguity in the 2AP intensity experiment.

2AP substituted at the methylation site was found to show little increase in fluorescence intensity on enzyme binding, 1.4 fold, but when the second adenine position was labelled a much larger signal increase, of the order of 10-fold, was observed[3]. Could both intensity increases here be due to base flipping? More likely is that significant conformational changes have taken place at both these sites. After confirmation that the first adenine is methylated whereas the second adenine is not, and given that the base sequence is bent by M.EcoRV, Gowher and Jeltsch interpreted both signal rises as due to 2AP unstacking[3]. These intensity measurements suggest that the environment of each 2AP inside M.EcoRV is unique but are not informative of the nature of these environments.

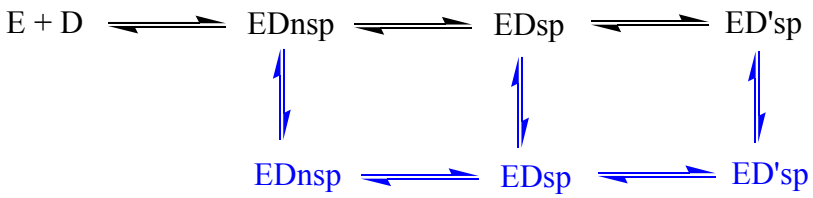
On binding to its DNA recognition sequence, M.EcoRV connects to the major groove[4;5] and effects a 60^{o} bend towards this groove[6]. The enzyme makes contact with less than 12 base pairs and at least 4 phosphate groups (NpNpNpGpA)[4;5] and separate loops of the enzyme bind to the GAT and the ATC parts of the recognition sequence[7]. The target-containing GAT half is the more tightly contacted of the two[8] and, in the opposite DNA strand, the thymine partnering the target adenine is not recognised by M.EcoRV, as concluded from the methylation kinetics of the target opposite an abasic site[9].

M.EcoRV is classed as an α-type N-methyltransferase. It is thought to have evolved from the closely related GATC-specific Dam methyltransferases[7], two of which have recently been structurally characterised with their DNA substrates; EcoDam[10] and T4 Dam[11]. Incidentally, M.EcoRV will methylate a GATC adenine *in vivo*[7]. M.EcoRV and the Dam methyltransferases all share the DPPY catalytic motif and are believed to have a common GAT recognising module. Beck *et al.* identify three M.EcoRV loops which match Dam family sequences and proposed the following functions on the basis of site directed mutagenesis[7]; loop I (I10-L19), which precedes motif X, forms backbone contacts and recognises the TC part of the base sequence; loop II (R128-Q148) is responsible for DNA recognition of the GAT bases; loop III (V210-G208), which follows the catalytic motif VI, bends the DNA and allows specific M.EcoRV-DNA contacts to form.

Besides 2AP fluorescence, the cross-linking of 5-iododeoxyuracil (^{5I}U) was employed by Beck and Jeltsch to see whether a less ambiguous indication of M.EcoRV base flipping could be obtained[12]. When photoactivated, ^{5I}U becomes a uracil radical which will cross-link to phenylalanine, tyrosine, tryptophan, methionine and histidine[13;14]. Cross-linking is enhanced if a π-stacking interaction exists between the ^{5I}U position and an aromatic amino acid side chain[15]. Within the GATATC sequence a high cross-linking yield was found for both adenine positions, with the target base position showing the highest conversion[12]. Taken with the results of the 2AP study, Beck and Jeltsch propose that an aromatic amino acid from loop III, intercalates the DNA helix and π-stacks with the second adenine, thereby causing the DNA bending, and that the non-target adenine sits at the centre of the bend.

Like the previous subject, M.EcoRI (Chapter 4), M.EcoRV methylates its substrate distributively and must be bound to cofactor before a specific DNA-enzyme complex can form[3] (binary DNA-enzyme complexes do form but these are said to be bound in a non-specific manner). M.EcoRV also displays a burst of methylation activity in its

catalytic mechanism. Scheme 5.1 shows the current model for M.EcoRV-DNA binding[3].



Scheme 5.1 The M.EcoRV-DNA binding equilibrium[3], E represents M.EcoRV, D represents DNA, nsp denotes non-specific and sp denotes specific. Black and blue colouring indicates open and closed complexes respectively.

M.EcoRV binds its DNA substrate in two binding modes, specific (sp) and non-specific (nsp), with the non-specific complexes forming quickly followed by their slow conversion to specific complexes[9]. Jeltsch *et al.* find that the non-specific binding is almost as strong as specific binding, and that the non-specific complex is based on an electrostatic interaction whereas the specific complex involves a hydrophobic interaction[9]. Thus, the specific complex is assumed to be a base flipped complex where the target base is held inside a hydrophobic binding pocket within the enzyme. During specific complex formation, contacts to the GAT part of the recognition sequence form first. Then the ATC contacts form which mediate increased specificity (over near-related GATC sites)[12]. Further kinetic studies using 2AP in the non-target adenine position find two separable specific complexes present in slow equilibrium, EDsp and ED'sp, in which the hydrophobic environments of the 2AP differ[3]. Additionally, Gowher and Jeltsch observe two substrate release rates for the DNA-enzyme complexes and propose that both open and closed versions of ED exist, with the open state allowing exchange of the DNA substrate and the closed state tightly binding DNA. The ratio of DNA-enzyme complexes in the open and closed formation[3] is roughly 1:10.

In the present work 2AP is substituted for adenine at each position within GATATC in turn and the decays are measured for unbound and M.EcoRV-bound labelled-substrates. Time-resolved fluorescence is used to see whether M.EcoRV does indeed

use a base flipping mechanism to extract its target base. Lifetime signatures allow comparison of the nature of the M.EcoRV-DNA interaction at each adenine position and help rationalise why a larger intensity increase arises from a 2AP at the non-target adenine site.

5.2. Materials and Methods

M.EcoRV enzyme and 2AP-labelled duplexes were provided by Kirsten Liebert and Albert Jeltsch (Jacobs University Bremen). S-adenosyl homocysteine (SAH) was purchased from Sigma-Aldrich and dissolved in water.

The sequences of the 2AP-labelled 20-mer oligonucleotides are listed in Table 5.1. All bottom strand target adenines were methylated to force the orientation of the enzyme to target the top strand.

Duplex	Base Sequence
AP1	5'-GATC GTA **GPT ATC** GCA TCGA-3' 3'-CTAG CAT **CTA TMG** CGT AGCT-5'
AP2	5'-GATC GTA **GAT PTC** GCA TCGA-3' 3'-CTAG CAT **CTA TMG** CGT AGCT-5'

Table 5.1 The base sequences of M.EcoRV substrates, the M.EcoRV recognition sequence is emboldened, P denotes 2-aminopurine and M denotes N6-methyl adenine.

Decays were measured for free DNA and the ternary complex formed from 2 μM DNA, 6 μM M.EcoRI and 100μM SAH. All solutions were buffered with 50 mM NaCl, 50 mM HEPES, 1 mM EDTA and 0.001% Lubrol (pH 7.5). In preparation of the ternary complex enzyme and SAH were combined first and allowed to incubate for 5 minutes before the addition of the DNA. Time-resolved fluorescence was performed as described in Chapter 2. The full set of calculated decay parameters for the duplexes and DNA-enzyme complexes in this chapter are given in Tables A16 to A19 in appendix A.

5.3. Results and Discussion

5.3.1. Unbound M.EcoRV substrate DNA: 2AP lifetimes for the adenine positions

Table 5.2 presents the decay parameters for 2AP in each adenine position in the recognition sequence of M.EcoRV. These positions are shown in Figure 5.1; AP1 denotes 2AP at the methylation target position and AP2 denotes 2AP two bases away from the methylation target.

Duplex	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	Φ_{rel}
AP1	0.02	0.25	1.6	5.9	0.94	0.03	0.01	0.02	<0.015
AP2	0.11	0.40	4.5	-	0.46	0.53	0.01	-	0.029

Table 5.2 Lifetimes and A factors for unbound M.EcoRV substrates, reported for a 370-nm emission wavelength. Also reported are the calculated quantum yields relative to free 2AP riboside Φ_{rel}.

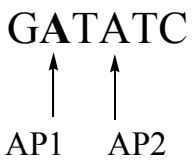


Figure 5.1 The 2-aminopurine labelling of the M.EcoRV recognition sequence, the methylation target (the AP1 position) is emboldened.

Like many DNA duplexes, the decay for AP1 is 4-exponential but two particular features are worth noting. τ_1 is exceptionally short indicating that, in these duplex conformations, 2AP is quenched with high efficiency and so is very tightly stacked with the neighbouring guanine. These duplex conformations are highly abundant, constituting 94% of the sample. Also in the τ_4 conformations, 2AP is significantly quenched compared with a solvent-exposed 2AP (10.6 ns) and A_4 is very low. Evidently, few 2APs are extrahelical in the AP1 duplex.

The decay for AP2 is 3-exponential. Whereas most decays from 2AP in DNA are 4-exponential[16-18], Rachofsky *et al.* also assign a 3-exponential function to the decays of 2AP in duplexes with the local sequence contexts A(2AP)A and T(2AP)T and with the probe opposite an abasic site[19]. As for AP1, AP2 adopts very highly stacked conformations with 99% of the 2APs having a sub-nanosecond lifetime. Again, there is no extrahelical τ_4 component. The conformational behaviour of AP2 appears to be a consequence of specific base sequence since Rai *et al.* find that 2AP inside the ATAT duplex sequence is highly constrained and immobile, and inaccessible to acrylamide quenching[20]. For both these duplexes calculated Φ_{rel} values are especially low and report that the emission intensity of AP2 is twice that of AP1.

5.3.2. Bound M.EcoRV substrates

Table 5.3 presents the decay parameters for each duplex bound to M.EcoRV. As can be seen, the differing lifetime signatures show that different enzyme-interactions occur with the target and non-target adenine on recognition sequence binding.

Solution Composition	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	Φ_{rel}
AP1/M.EcoRV/SAH	0.09	0.62	2.9	7.6	0.45	0.23	0.19	0.13	0.16
AP2/M.EcoRV/SAH	0.07	0.86	3.2	9.3	0.20	0.18	0.30	0.32	0.39

Table 5.3 Global lifetimes for DNA-M.EcoRV complexes with the AP1 and AP2 duplexes and A factors reported for decays observed at a 380-nm emission wavelength. Also reported are the calculated quantum yields relative to free 2AP riboside Φ_{rel}.

On first inspection, the decay parameters of AP1 bound to M.EcoRV look to be those of a typical unbound duplex, four lifetimes spanning a familiar range and an A factor distribution weighted towards the shorter lifetime components. When compared with the lifetime signature of the free duplex (Table 5.2), however, it is clear that M.EcoRV has effected a large structural distortion in AP1.

The value of τ_1 is characteristic of an intrahelical 2AP and A_1 reports that 45% of 2APs at the target position remain unflipped. Incidentally Gowher and Jeltsch

expect a mixture of non-specific (i.e. unflipped) and specific complexes at equilibrium[3] (*vide infra*). The combined A factors of components 2-4 have risen from 6% to 55%, indicating that a substantial degree of flipping has taken place. All these lifetimes have undergone a marked increase in size compared with their free duplex counterparts (Table 5.3 vs. Table 5.2), which shows that the quenching environment of 2AP has changed when the duplex is bound by M.EcoRV. The conformations that result in the sub-nanosecond τ_2 are likely to contain intrahelical, imperfectly stacked 2APs. Given that a tyrosine-associated 2AP inside M.TaqI has a 0.55 ns lifetime however (Table 3.4), it is plausible that this 0.62 lifetime results from a 2AP that has π-stacked with the tyrosine from DPPY inside M.EcoRV. Similarly, the 2.9 ns τ_3 of the DNA-M.EcoRV complex can be envisaged as 2AP inside the adenine binding pocket of M.EcoRV but less strongly associated with the tyrosine. Now that the AP1 duplex is bound, an extrahelical τ_4 component has emerged in its decay signature. The 7.6 ns lifetime is assigned to a flipped complex conformation with 2AP stabilised in the M.EcoRV interior but in a weakly quenching environment.

As mentioned earlier, Beck *et al.* believe that M.EcoRV shares the same GAT-interacting module as the Dam methyltransferases[7]. Recently crystal structures for T4 Dam and EcoDam show a flipped adenine in close proximity to the DPPY motif[10;11] and imply that DPPY stabilisation of a flipped 2AP by M.EcoRV is feasible. However, these Dam enzymes differ from M.EcoRV in that they do not bend their substrate. Also the Dam methyltransferases carry out methylation processively (rather than distributively) hence the catalytic conformation in the Dam methyltransferases will only form if the natural cofactor SAM is present, rather than SAH or sinefungin. Interestingly, the T4 Dam structure shows the flipped adenine stabilised in a non-catalytic conformation on the surface of the enzyme[11]. Similar behaviour to this may be being observed in the AP1/M.EcoRV/SAH complexes characterised by τ_4 since these are flipped but the 2AP is not strongly associated with DPPY.

Changes to the AP2 lifetime signature on M.EcoRV complex formation, specifically τ_4, support the notion of base unstacking during DNA bending. Compared to the free AP2 duplex, τ_4 is unique to the DNA-enzyme complex and in relatively high proportion. The long 9.3 ns lifetime reports that the 2AP of this species is experiences very little quenching, neither by the M.EcoRV amino acids nor by neighbouring partner bases in the duplex. Although τ_4 might signify an exo site environment, as was found for the target base position inside M.TaqI (Chapter 3), these data support the likelier model that the 2AP at the AP2 sequence position is at the apex of the bent DNA duplex. This interpretation is made in conjunction with the work by Cal and Connolly[6], and by Gowher and Jeltsch[3]. Overall, the population of sub-nanosecond species has fallen from 99% to 38% and indicates a reduction in the stacking of the non-target 2AP in about 60% of the DNA-enzyme complexes.

Since enzyme binding to AP2 gave the larger signal increase, Gowher and Jeltsch employed AP2 to determine the rates of formation of the species in Scheme 5.1[3]. The reported rate constants for interconversion of these complexes in the presence of SAH indicate that the relative concentrations of the three complexes in the equilibrium mixture are 0.23 EDnsp, 0.34 EDsp and 0.43 ED'sp (though it should be noted that the errors associated with the rate constants are high). In the lifetime data of AP2/M.EcoRV/SAH (Table 5.3), species 1 and 2 are assigned to non-specific DNA-enzyme complexes. In these complexes, 2AP has a sub-nanosecond lifetime and so has not been successfully unstacked from the neighbouring DNA bases by M.EcoRV (as would occur if the base sequence were bent by the enzyme). With this assignment, lifetime data predict a 38% proportion for EDnsp. Though the relative population of EDnsp in this set of complexes is found to be higher than in the complexes measured by Gowher and Jeltsch, τ_1 and τ_2 probably encompass some highly stacked but specific complex conformations too, such that 38% indicates an upper limit for EDnsp.

At this point, some justification must be given for the above interpretation of the AP2-M.EcoRV lifetime data since, if the context were unknown, the changes to the decay parameters for 2AP in this sequence position might also be perceived as a flipping response. Gowher and Jeltsch have confirmed that the second adenine in the recognition sequence is not methylated and so must not enter the catalytic DPPY motif[3]. In addition, the target adenine in the AP2 substrate is methylated in the presence of natural cofactor, following the observation of a large increase in fluorescence intensity from 2AP at the non-target position. Comparing τ_4 for the two duplexes in their enzyme complex, the 9.3 ns lifetime of AP2 is appreciably longer than the 7.6 ns lifetime of AP1. This is consistent with the AP2 probe sitting in an extrahelical position at the apex of the bent duplex while the AP1 probe has a higher level of enzyme interaction. Assuming that the duplex is both bent and flipped, and considering that the A factor distribution in the AP1-M.EcoRV decay signature favours sub-nanosecond lifetimes, it appears that the target adenine position is more protected from the effects of DNA bending than the non-target position. Since these adenines are separated by only one base it is difficult to envisage how the AP1 probe could be so removed from the effect of bending without being strongly associated inside the active site of the enzyme.

Further reasoning comes from the consideration of how the decay parameters of 2AP at each adenine position change when taken up in the enzyme complex. Recall that, when unbound, each duplex has a high proportion of sub-nanosecond lifetime 2AP species (99% for AP2 and 97% for AP1). Sub-nanosecond lifetimes indicate highly quenched conformations. On binding of AP2 the population of highly quenched conformations is 38%, a large fraction of these are assigned to non-specific binding. From the remaining complexes, species 2 and 3, which are attributed to specific binding conformations, 32% of the flipped 2APs are virtually unquenched (9.8 ns τ_4). In contrast, for the AP1 duplex, 68% of the population remains highly quenched on M.EcoRV binding and there is no unquenched population. This occurrence for AP1 is entirely consistent with quenching in the active site that is know to be characteristic of the M.TaqI NPPY catalytic motif (see Chapter 3). Also, if components τ_1 and τ_2 from AP1/M.EcoRV/SAH were considered to be non-specific

(i.e. unflipped) complex conformations, as they were for the AP2 complex, then the parameters in Table 5.3 report very different fractions of non-specific complex formation; 68% from AP1 data compared with just 38% for AP2 data predictions. Thus the response of the AP2 decay parameters must monitor stacking alleviation whereas the response of AP1 parameters must demonstrate flipping.

The lifetime data predict that on enzyme binding the AP1 duplex emission intensity should increase approximately 10-fold; however, in the original steady state measurements, Gowher and Jeltsch noted virtually no change in intensity[3]. Recent measurements by Jeltsch confirm that there is indeed an observable intensity increase on AP1-M.EcoRV complex formation (personal communication) but quantification is expected to be highly error-prone since Φ_{rel} for AP1 is so low. The Φ_{rel} values determined from the lifetime data predict that 2AP emission from the AP2-M.EcoRV complex is higher than in the AP1-M.EcoRV complex, in agreement with steady state measurements. The Φ_{rel} values indicate a 13-fold rise in AP2 emission on M.EcoRV binding, compared with published values of 8.8- and 11.2-fold[3;12] increases.

Previously, Beck and Jeltsch proposed that the bent substrate might be intercalated by M.EcoRV and that the non-target adenine lies in proximity to an aromatic amino acid side chain[12]. When M.HhaI flips its cytosine target this methyltransferase intercalates the DNA at the former location of this cytosine, so as to occupy the gap formed in the duplex. For a 2AP adjacent to the target cytosine, this intercalation causes little change to the lifetime signature compared with the free DNA[21]. In the M.HhaI-DNA complex, 2AP adjacent to the flipped cytosine is stabilised by hydrogen bonding with the aliphatic arginine and isoleucine, such that the 2AP remains intrahelical. Unlike the M.HhaI-bound substrate, the M.EcoRV-bound DNA is bent. On this basis, if the intercalation behaviour of M.EcoRV were to completely compensate for the loss of stacking to the 2AP in AP2 then, as for M.HhaI, the unbound and bound duplex decay signatures would be indistinguishable. With this reasoning, the lifetime data for the bound AP2 duplex suggest that 2AP at the non-target adenine position is not stacked with an aromatic M.EcoRV side chain on

duplex bending. Given that the 2AP has 3.2 and 9.3 ns lifetimes in the specifically bound complexes where the duplex is bent (species 3 and 4 in Table 5.3), the 2AP is either only moderately quenched or unquenched in these conformations. If a strong π-stacking interaction were to occur with an aromatic side chain a much higher rate of quenching, and therefore shorter lifetimes, would be expected. The significant population of species 3 and 4 for this enzyme complex, especially the long-lifetime species 4, is informative of the bending event having successfully peeled away the neighbouring bases and broken the 2AP base stacking association.

5.4. Conclusion

Once again, the time-resolved fluorescence of 2AP has allowed a more intricate investigation of the way in which a methyltransferase interacts with its substrate and the results yielded are less ambiguous than previous steady state efforts. The differing lifetime signatures which arise from substituting 2AP in place of the adenines in GATATC show that M.EcoRV manipulates these sequence positions uniquely. The target position is flipped and stabilised in DPPY, as indicated by the lifetimes obtained being comparable with those given by the DNA-M.TaqI complex in a catalytic conformation. The non-target position lifetimes show alleviation of quenching in the enzyme complex and are consistent with the notion of unstacking upon DNA bending.

When M.EcoRV binds its recognition sequence, lifetime data show that there is still significant quenching of 2AP at the target adenine position in the resulting DNA-enzyme complex. However, in approximately a third of the complexes made with 2AP at the non-target adenine position, 2AP is effectively unquenched, hence a larger emission signal arises when 2AP is placed at the non-target adenine site.

The decay parameters of the ternary complexes indicate four conformational sets and reveal that up to 45% of the complexes remain unflipped. This is in keeping with the current binding model, which predicts an appreciable presence of non-specifically bound complexes at equilibrium. Lifetime data for the target adenine position report three distinguishable specifically-bound conformations whereas the decay data for the non-target adenine position show significant populations for only two specifically-bound conformations. The kinetic methods used by Gowher and Jeltsch were only sensitive enough to detect the larger changes produced by labelling the non-target adenine position and so also only resolved two. Hence the ability to measure the changes to the target adenine position has allowed a higher degree resolution for studying DNA-M.EcoRV binding.

5.5. References

(1) The Restriction Enzyme Database (REBASE). 2007. http://rebase.neb.com

(2) Roth, M.; Helm-Kruse, S.; Friedrich, T.; Jeltsch, A. *Journal of Biological Chemistry* **1998**, *273*, 17333-17342.

(3) Gowher, H.; Jeltsch, A. *Journal of Molecular Biology* **2000**, *303*, 93-110.

(4) Szczelkun, M. D.; Connolly, B. A. *Biochemistry* **1995**, *34*, 10724-10733.

(5) Szczelkun, M. D.; Jones, H.; Connolly, B. A. *Biochemistry* **1995**, *34*, 10734-10743.

(6) Cal, S.; Connolly, B. A. *Journal of Biological Chemistry* **1996**, *271*, 1008-1015.

(7) Beck, C.; Cranz, S.; Solmaz, M.; Roth, M.; Jeltsch, A. *Biochemistry* **2001**, *40*, 10956-10965.

(8) Newman, P. C.; Nwosu, V. U.; Williams, D. M.; Cosstick, R.; Seela, F.; Connolly, B. A. *Biochemistry* **1990**, *29*, 9891-9901.

(9) Jeltsch, A.; Friedrich, T.; Roth, M. *Journal of Molecular Biology* **1998**, *275*, 747-758.

(10) Horton, J. R.; Liebert, K.; Bekes, M.; Jeltsch, A.; Cheng, X. D. *Journal of Molecular Biology* **2006**, *358*, 559-570.

(11) Horton, J. R.; Liebert, K.; Hattman, S.; Jeltsch, A.; Cheng, X. D. *Cell* **2005**, *121*, 349-361.

(12) Beck, C.; Jeltsch, A. *Biochemistry* **2002**, *41*, 14103-14110.

(13) Willis, M. C.; Hicke, B. J.; Uhlenbeck, O. C.; Cech, T. R.; Koch, T. H. *Science* **1993**, *262*, 1255-1257.

(14) Wong, D. L.; Pavlovich, J. G.; Reich, N. O. *Nucleic Acids Research* **1998**, *26*, 645-649.

(15) Norris, C. L.; Meisenheimer, P. L.; Koch, T. H. *Journal of the American Chemical Society* **1996**, *118*, 5796-5803.

(16) Guest, C. R.; Hochstrasser, R. A.; Sowers, L. C.; Millar, D. P. *Biochemistry* **1991**, *30*, 3271-3279.

(17) Nordlund, T. M.; Andersson, S.; Nilsson, L.; Rigler, R.; Graslund, A.; Mclaughlin, L. W. *Biochemistry* **1989**, *28*, 9095-9103.

(18) Neely, R. K. The Photophysical Properties of 2-Aminopurine and its Application as a Probe of DNA-Protein Interactions. Thesis 2005.

(19) Rachofsky, E. L.; Seibert, E.; Stivers, J. T.; Osman, R.; Ross, J. B. A. *Biochemistry* **2001**, *40*, 957-967.

(20) Rai, P.; Cole, T. D.; Thompson, E.; Millar, D. P.; Linn, S. *Nucleic Acids Research* **2003**, *31*, 2323-2332.

(21) Neely, R. K.; Daujotyte, D.; Grazulis, S.; Magennis, S. W.; Dryden, D. T. F.; Klimasauskas, S.; Jones, A. C. *Nucleic Acids Research* **2005**, *33*, 6953-6960.

Chapter 6

Long wavelength emission of 2AP in DNA

6.1. Introduction

This study now turns to the further characterisation of the fluorescence properties of 2-aminopurine in DNA. In order to follow the concepts that will be presented, it is necessary to recollect some general aspects of DNA structure.

DNA is the acronym for *deoxyribonucleic acid* and defines a polymer of deoxyribonucleotide monomers. The deoxyribonucleotide for adenine, formally 2'-deoxyadenosine 5'-phospate, is shown in Figure 6.1(a). It comprises the base, a deoxyribose sugar and a phosphate group. Monomers are joined together through a phosphodiester link, forming a chain of sugar-phosphate groups (the DNA backbone) with the base branching off, Figure 6.1(b).

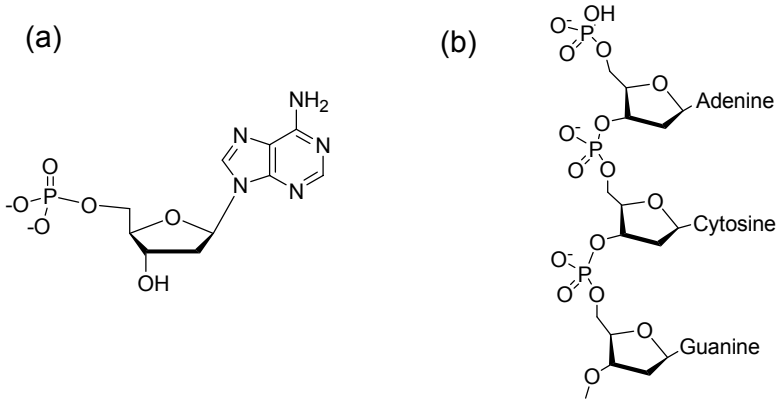


Figure 6.1 (a) the deoxyribonucleotide 2'-deoxyadenosine 5'-phospate and (b) a deoxyribonucleic acid.

Two of these chains (single strands) pair up to form double-stranded DNA and hold together with hydrogen bonds, which form between bases on opposite single strands. Adenine pairs with thymine and cytosine pairs with guanine and, usually, the bases associate through Watson-Crick type base pairing, Figure 6.2.

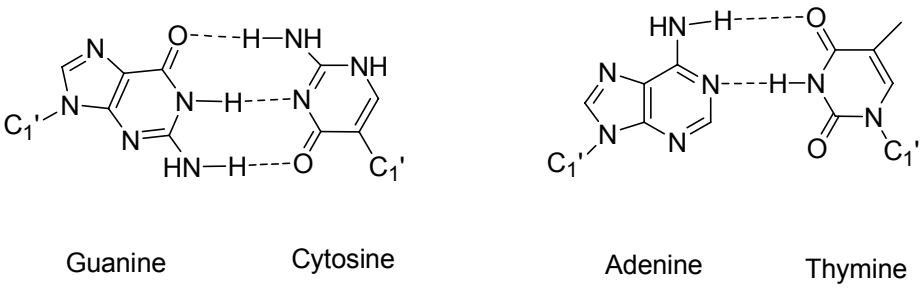


Figure 6.2 Watson-Crick base paring between the DNA bases.

The structure formed by double-stranded DNA is a double helix, this is shown in Figure 6.3 along with the dimensions of B form DNA, the most common form.

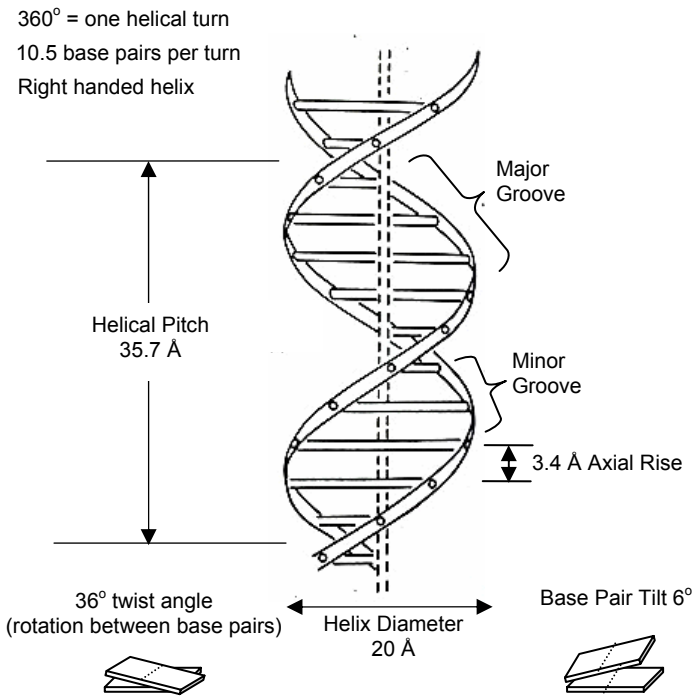


Figure 6.3 The structure of the DNA double helix. The DNA backbone is shown in ribbon-form, the base pairs are the horizontal shapes linking the two ribbons and the dashed vertical line is the long axis of the helix. The dimensions shown define B-form DNA.

Although Figure 6.3 presents a well-defined, rigid structure, the DNA molecule is far from static. Base pairs continually break and re-form on the millisecond-timescale (DNA breathing) and the DNA backbone is flexible.

DNA structure is strongly governed by its base sequence, by the non-covalent π-stacking interactions between adjacent bases and base pairs[1;2]. Each base has characteristic charge distribution across it, reliant on the electronegativity of each atom in its σ-framework. Additionally there is π-electron density above and below the plane of the aromatic ring. As can be seen, the pairing of successive bases along each single strand brings each base close to its neighbour and base pair geometries must be optimised to minimise electrostatic repulsion and to avoid steric clashes between exocyclic groups.

Figure 6.4 shows the excitation and emission spectra for 2AP in DNA. 2AP in DNA is usually excited with a 310- or 320-nm wavelength and its emission collected at 370 nm. In this regard, the emission from 2AP in DNA closely resembles the emission from a free 2AP molecule in solution.

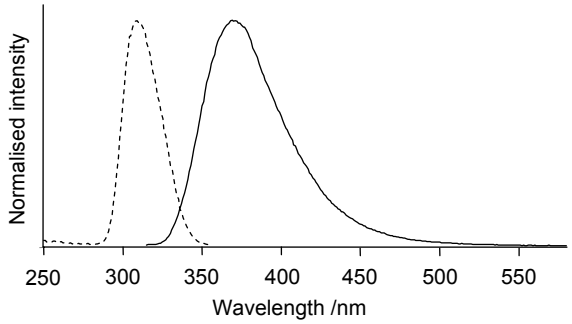


Figure 6.4 The excitation (dashed line) and emission spectrum (solid line) of 2-aminopurine in DNA

Rist, Fiebeig *et al.*[3] observed that the excitation of 2AP with 343 nm (a longer wavelength than is ordinarily used) yields an additional emission band with maximum intensity at 440 nm. This emission was observed for two different duplexes, one containing pairs of adjacent 2AP molecules and the other containing several single 2AP molecules situated between natural DNA bases. The long-wavelength emission band was more intense for the adjacent 2APs and when Rist irradiated this duplex with 350 nm, the 440-nm band became more pronounced and the 370-nm band almost vanished. At this stage 2AP excimers were being sought so red-shifted emission from the 2AP was anticipated, however, subsequent excitation spectra of the 440-nm emission yielded two peaks, 310 nm (the usual for 2AP) and

370 nm. Once again, the longer-wavelength excitation band was more intense for adjacent 2APs. Consequently, Rist *et al.* identified that there was a ground-state interaction between adjacent 2APs and this created a partially delocalised species which can be directly excited.

More recently, Wang, Fiebig *et al.* carried out femtosecond broadband pump-probe spectroscopy on 2AP in the trimers (2AP)(2AP)C and A(2AP)C[4]. In this study they recreated the local environments of the 2AP in the duplexes discussed above, and sought to selectively excite the 2AP and observe electron transfer from the excited state. Wang, Fiebig *et al.* proposed a separate decay scheme for each trimer given that different relaxation times were observed. However their interpretation was based on the premise that the species being excited was a single 2AP, initially detached from any interaction with its neighbours, but a 316-nm excitation wavelength was used. The earlier work by Rist, Fiebig *et al*.[3] clearly showed that with this excitation wavelength, ground-state dimer species would have been excited too.

Long wavelength emission has also been observed by Guest *et al.* from 2AP in singly labelled 7-mer oligonucleotide duplexes[5]. In this instance, 320-nm excitation brought about a red-shifted emission band in the tail of the 370-nm emission profile. This was attributed to a 2AP-guanine exciplex. Marti *et al.* reported 450-nm emission from 2AP inside a novel molecular beacon[6]. Here the 2AP is in a single stranded nucleic acid chain that forms a stem-loop structure. Red-shifted emission was observed on 310-nm excitation when the 2AP was unpaired and situated in a looped region of the chain.

Theoretical calculations find that there is significant interaction between 2AP and an adjacent natural base in the ground state[7;8]. Ground state molecular orbital mixing causes the electron density from the HOMO of 2AP to extend to the neighbouring base. Consequently the oscillator strength between ground and excited state is decreased with increased base stacking, to greater effect when 2AP stacks with another purine. There is some disagreement as to whether there is one accessible

excited state or two for the 2AP as a result of 2AP-base stacking. Configuration interaction singles (CIS) finds only one accessible exited state. The oscillator strength between this and the ground state 2AP is purely sequence dependent, reduced by a neighbouring cytosine but increased by a neighbouring guanine[8]. Time-dependent density functional theory (TDDFT) finds two accessible excited states, S_1 and an additional low-lying LUMO S_2, in which 2AP has different geometries and is subjected to differing quenching mechanisms[7]. However, the lower energy $S_1 \rightarrow S_0$ transition has a weak oscillator strength and is assumed to be a non-emitting *dark state*. Therefore both methods conclude that fluorescence from the base-stacked 2AP molecule occurs from a state which is close in character to the 2AP monomer.

Long wavelength emission is a feature of natural DNA and there is debate as to whether an excimer species is responsible for this emission (see Kwok *et al.*[9] and references therein). The photophysical properties of natural DNA are of great interest, particularly the charge transfer mechanisms that the molecule employs. When DNA is exposed to UV light the excess energy is dissipated with high efficiency, which is crucial in order to avoid permanent photodamage to genetic material. Transient absorption spectra of $(dA)_{20}$ after different time delays lead Kwok to agree with the excimer model[9] however Kononov concludes that stacked DNA bases can be directly excited, since excitation spectra of (poly(A)) reveal a separate excitation band for the long wavelength emission not observed in absorption[10].

Neely *et al.* have shown that 360-nm excitation of a 2AP-labelled DNA duplex gives a long wavelength emission band with maximal intensity at 430 nm for which there is a distinct excitation profile[11]. In conjunction with strong 420-nm emission from columns of π-stacked molecules in crystalline 2AP, they propose that the red-shifted emission in DNA comes from a 2AP which is highly stacked with its neighbouring base(s). This study presents steady-state and time-resolved measurements of 2AP in DNA in a range of sequence contexts and aims to further characterise this long wavelength emission of 2AP. As will be seen in the next chapter, long wavelength

emission also arises in DNA-methyltransferase complexes and is informative of the interactions occurring in the methyltransferases currently being investigated.

6.2. Materials and Methods

The series of duplexes, providing various sequence contexts for the 2-aminopurine, are listed in Table 6.1. In the preparation of these, the labelled strand (between 15 and 25 μM) was combined with a 50% excess of the unlabelled strand to avoid the presence of unpaired 2AP. The DNA was buffered (100 mM NaCl, 10 mM Tris-Cl, 1 mM EDTA) and annealed by heating to 95^{o} and then leaving to cool slowly back to room temperature. Single strands samples were prepared by dilution of the original stocks to 20 μM levels in the same buffer.

Steady state fluorescence was performed as described in Chapter 2. Multiple emission spectra were recorded for excitation wavelengths in the range 290-400 nm at 10 nm intervals. Multiple excitation spectra were recorded with signal intensities monitored in the range 360-500 nm at 10 nm intervals. Both types of spectra were measured with a 5-nm band pass. Spectral intensities were assessed by the integration of a 20-nm wide strip encompassing the maximal emission wavelength and values were converted to the 20 μM-concentration equivalent.

For low temperature measurement, selected duplexes were diluted in 10 M LiCl (~ 3.3 μM in 2AP), transferred to a fused silica Dewar assembly (Horiba Jobin Yvon) and cooled with liquid nitrogen to 77 K, thus forming an optically transparent glass. Emission spectra for 310- and 360-nm excitation wavelengths were measured.

Time-resolved fluorescence spectroscopy was performed as described in Chapter 2. Excitation radiation of 360 nm was generated by frequency doubling the 720-nm output of the Ti:Sapphire laser. Fluorescence decays of the long-wavelength emission were measured at 430, 450 and 470 nm and fitted globally. The full set of calculated decay parameters for the duplexes in this chapter are given in Tables A13, and A20 to A24 in appendix A.

Duplex	Base Sequence
APT(T)	5’-AGACGA**P**TTTCGAA-3’
	3’-TCTGCTTAAAGCTT-5’
APC(T)	5’-CACGGGCCTA**P**CGATATCGTGCGTACGAGC-3’
	3’-GTGCCCGGATTGCTATAGCACGCATGCTCG-5’
TPT(T)	5’-GATCGTAGAT**P**TCGCATCGA-3’
	3’-CTAGCATCTATAGCGTAGCT-5’
GPT(T)	5’-GATCGTAG**P**TATCGCATCGA-3’
	3’-CTAGCATCTATAGCGTAGCT-5’
GPA(T)	5’-TGCAG**P**ATTCGAGGTCGACGG-3’
	3’-ACGTCTTAAGCTCCAGCTGCC-5’
GPC(T)	5’-TGCAGAATTCGAGGTCGACGG-3’
	3’-ACGTCTTAAGCTCC**P**GCTGCC-5’
GPG(T)	5’-GACAGTATCAGGCGCCTCCCCACAA-3’
	3’-CTGTCATAGTCCGCGG**P**GGGGTGTT-5’
APA(C)	5’-TGCA**P**AATTCGAGGTCGACGG-3’
	3’-ACGTCTTAAGCTCCAGCTGCC-5’
GPT(C)	5’-TGCAGPATTCGAG**P**TCGACGG -3’
	3’-ACGTCTTAAGCTCCAGCTGCC-5’
GPG(G)	5’-ACTGGTACAGTATCAGG**P**GCTGACCCACAACATCCG-3’
	3’-TGACCATGTCATAGTCCGCGACTGGGTGTTGTAGGC-5’

Table 6.1 Base sequences of the DNA duplexes. Duplexes are designated XPY(Z) where P is 2AP, X and Y are the bases immediately 5’and 3’, respectively, to 2AP and Z is paired with 2AP.

6.3. Results

6.3.1. Red-shifted emission from 2AP-labelled duplexes at room temperature

Figure 6.5 shows the emission spectra that arise from exciting the GPA(T) duplex (see Table 6.2 for sequence) at 310 and 360 nm. These results are typical of the entire set of duplexes listed in Table 6.1. The longer excitation wavelength produces an emission band with maximum intensity at 450 nm, which is considerably weaker than the 370-nm emission band.

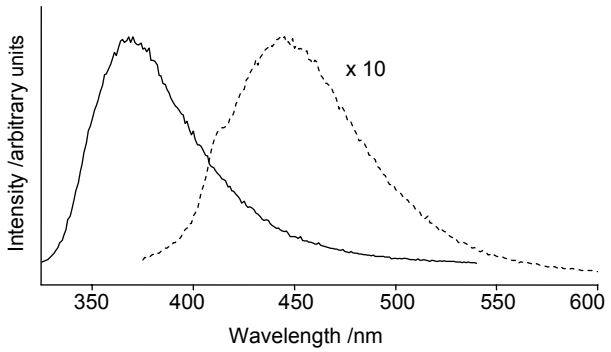


Figure 6.5 Emission spectra of the GPA(T) duplex, excited at 310 nm (solid line) and 360 nm (broken line, intensity multiplied by 10).

Table 6.2 compares the intensity of the 450-nm (long) emission to that of the 370-nm (short) emission for each duplex. The relative intensity of the long-wavelength emission from 2AP varies with sequence context, ranging from approximately 3 to 30% of the short-wavelength emission intensity.

Duplex	Relative intensity
APA(C)	0.03
GPT(C)	0.03
APT(T)	0.04
APC(T)	0.05
GPA(T)	0.11
GPC(T)	0.22
GPG(T)	0.24
GPG(G)	0.27
GPT(T)	0.28
TPT(T)	0.32

Table 6.2 The relative intensity of the long wavelength (450 nm) emission, excited at 360 nm, relative to that of the short wavelength (370 nm) emission, excited at 310 nm, for duplexes in solution at room temperature.

If the relative emission intensities at the two wavelengths are considered separately, as is presented in Table 6.3, it is apparent that 370-nm emission is much more dependent on sequence context, and so the quenching efficiency of this short-wavelength emission governs the variation of relative intensities in Table 6.2.

Duplex	370-nm band	450-nm band
APA(C)	0.9	1.0
GPT(C)	1.0	1.0
APT(T)	0.6	0.7
APC(T)	0.2	0.3
GPA(T)	0.2	0.6
GPC(T)	0.08	0.6
GPG(T)	0.07	0.6
GPG(G)	0.08	0.8
GPT(T)	0.03	0.3
TPT(T)	0.02	0.2

Table 6.3 The intensity of each emission band relative to that of GPT(C) for duplexes in solution at room temperature. The 370 nm emission is excited at 310 nm and the 450 nm emission is excited at 360 nm.

Figure 6.6 plots the excitation spectra for the GPA(T) duplex. It indicates that the long-wavelength emission is maximally excited at 360 nm. The short-wavelength excitation band is still present when emission is monitored at 450 nm since the short-wavelength emission profile still has intensity at this wavelength, as is apparent in Figure 6.5.

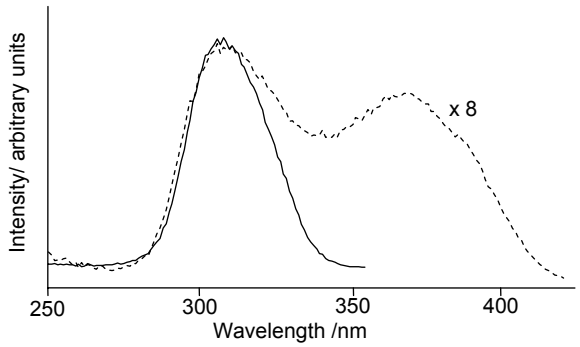


Figure 6.6 Excitation spectra of the GPA(T) duplex, detected at 370 nm (solid line) and 450 nm (broken line, intensity multiplied by a factor of 8).

Figure 6.7, which shows the excitation spectra for emission wavelengths of 440 nm and longer, indicates that the 310-nm peak never completely disappears from the excitation profile, even when a signal observation wavelength as long as 490 nm is employed. This occurs because there is still an appreciable level of short-wavelength emission in the spectral region spanned by the long-wavelength emission band.

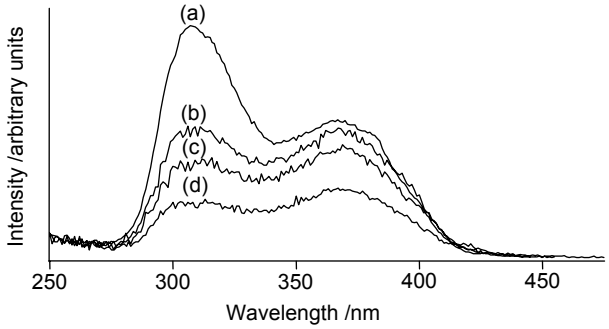


Figure 6.7 Excitation spectra of the GPA(T) duplex detected at (a) 440 nm, (b) 460 nm, (c) 470 nm and (d) 490 nm.

As the excitation band for the long-wavelength emission overlaps with the short-wavelength excitation band, 310-nm excitation can still stimulate long wavelength emission from 2AP. In some sequence contexts, providing the long-wavelength emission is significantly intense, long wavelength emission can visibly distort the short-wavelength emission profile. This is demonstrated in Figure 6.8 which shows the 370-nm emission profile for GPG(G), where the long/short relative emission intensity is 0.27 (Table 6.2). In contrast the GPT(C) duplex appears undistorted

because long-wavelength emission is weak, with the long/short relative intensity being 0.03 for this duplex.

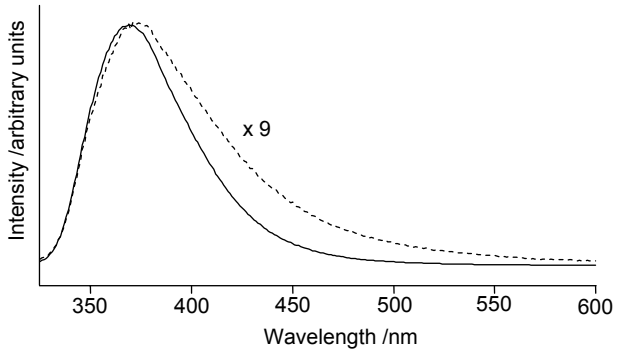


Figure 6.8 Comparison of the emission spectra of GPT(C) (solid line) and GPG (G) (broken line, intensity multiplied by a factor of 9), excited at 310 nm.

Fluorescence decay parameters for the long and short-wavelength emission of 2AP in the APT(T), APC(T) and GPG(T) duplexes are presented in Table 6.4. The 450-nm decay curves follow a 4-exponential function, like the well-characterised 370-nm decays, however lifetimes for the long-wavelength emission are notably longer. Furthermore the 450-nm lifetimes and A factors are similar across these differing sequences whereas the 370-nm equivalents are more variable and dependent on the neighbouring bases of the 2AP. The short component τ_1, for example, is especially short at 370 nm when the 2AP is adjacent to a guanine but the same cannot be said of the τ_1 of the 450-nm emission.

Duplex		τ_1 / ns	τ_2 / ns	τ_3 / ns	τ_4 / ns	A_1	A_2	A_3	A_4
APT(T)	450 nm	0.15	1.0	3.5	11.5	0.49	0.29	0.17	0.05
	370 nm	0.15	0.46	2.5	9.9	0.51	0.36	0.04	0.08
APC(T)	450 nm	0.17	0.74	3.6	12.6	0.44	0.38	0.13	0.05
	370 nm	0.12	0.37	2.7	10.8	0.88	0.08	0.02	0.02
GPG(T)	450 nm	0.19	0.97	3.6	9.2	0.53	0.26	0.16	0.05
	370 nm	0.04	0.44	2.7	10.7	0.69	0.11	0.09	0.11

Table 6.4 The global fluorescence decay parameters for duplexes at emission wavelengths of 450 nm and 370 nm, in solution at room temperature. 370 nm emission is excited at 320 nm and 450 nm emission is excited at 360 nm.

6.3.2. Duplexes at 77 K

Quenching of 2AP in duplex DNA is largely a dynamic process. O'Neill and Barton[12] observed a dramatic increase in the 370-nm emission of 2AP in DNA when cooled to 77 K and Neely and Jones[13] have subsequently shown that the most highly quenching duplex conformation, characterised by τ_1, is no longer present at this low temperature. Table 6.5, which quantifies the changes to the maximal emission intensity of each band on cooling, shows that the intensity of the long wavelength emission also increases. In these three duplexes long and short wavelength emission enhancement is approximately the same.

Duplex	370-nm emission	450-nm emission
APT(T)	6	7
GPC(T)	4	4
GPG(G)	6	7

Table 6.5 The intensities of the long and short wavelength emission bands for duplexes at 77 K, relative to their respective intensities at 298 K (the duplexes are in LiCl for both temperatures).

Figure 6.9 shows the emission spectra of APT(T) at 77 and 298 K. Both low-temperature emission profiles are blue-shifted due to the alleviation of solvent relaxation. Long-wavelength emission is noticeable in the tail of the short-wavelength emission profile under 310-nm excitation and low-temperature conditions.

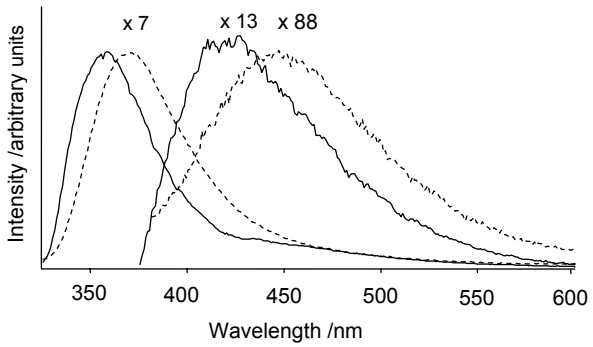


Figure 6.9 Emission spectra (solid lines) of the APT(T) duplex frozen at 77 K, excited at 310 and 360 nm. The corresponding spectra at 298 K are also shown (broken lines). Intensities have been multiplied by the factors shown.

6.3.3. Single-strand oligodeoxynucleotides at room temperature

The long-wavelength emission produced by 2AP on 360-nm excitation also occurs in single strand oligodeoxynucleotides so, evidently, red-shifted emission in the duplex is not a result of hydrogen bonding between 2AP and its base-pairing partner. 2AP-labelled single strands from four of the duplexes have been measured and Table 6.6 presents the relative long/short wavelength emission intensities. Table 6.7 presents the relative emission intensities at each wavelength.

Oligo	Relative intensity
APA	0.01
GPA	0.05
GPC	0.05
TPT	0.01

Table 6.6 The intensity of the 450 nm emission, excited at 360 nm, relative to that of the short wavelength (370 nm) emission, excited at 310 nm, for single-strand oligodeoxynucleotides in solution at room temperature.

Oligo	370-nm band	450-nm band
APA	1.0	1.0
GPA	0.4	1.5
GPC	0.2	0.8
TPT	0.8	0.5

Table 6.7 The intensity of each emission band relative to that of APA for single-strand oligodeoxynucleotides in solution at room temperature. 450 nm emission is excited at 360 nm, 370 nm emission is excited at 310 nm.

Long-wavelength emission from single strands is only a few percent of the 370-nm emission, not because the long-wavelength emission is weaker than inside duplexes, but because the 370-nm emission is not as quenched in a single strand environment.

Ramreddy *et al.*[14;15] have observed that in the 370-nm fluorescence decay parameters of 2AP in single strands there is no short component nearing 10 ps. The shortest measured lifetime for their single strands is around 500 ps. The very short lifetime emerges only when the single strands are annealed with their complimentary partners into duplexes.

6.4. Discussion

It is proposed that long wavelength emission from 2AP in DNA characterises conformational states of the duplex in which 2AP forms a highly overlapped, π-stacked heterodimer with one of its adjacent bases. This dimer species has a lower lying excited state compared with the 2AP monomer species so results in emission that is red-shifted from the usual 370-nm emission band. The long wavelength emission has a distinctive excitation band, verifying that this species exists in the ground state, and is therefore not an excimer (which will only form when 2AP is electronically excited). As red-shifted emission also arises from 2AP in single strand oligonucleotides the dimer species must be bound by base stacking interactions and not by hydrogen bonding. The conventional 370-nm emission arises from an isolated 2AP monomer species which is not appreciably affected by such base stacking in that no stable dimer species can form although, clearly, some 2AP-base interactions remain which permit rapid charge transfer away from the excited state 2AP.

It is reasoned that helical twist in the duplex provides the means to dimer formation. Helical twist describes the relative rotation of the bases about an axis perpendicular to the plane of the base pair, shown in Figure 6.10. It governs base overlap and, in turn, the extent of the non-covalent π-stacking interaction between two adjacent bases.

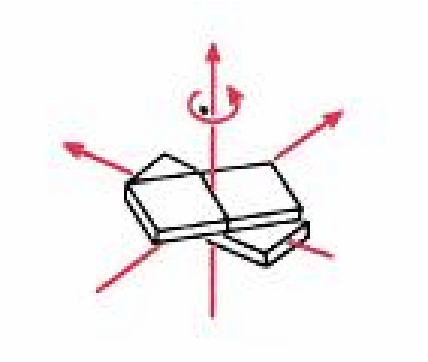


Figure 6.10 A schematic illustration of helical twist between two adjacent base pairs

To illustrate the effect of helical twist on base overlap in DNA, Figure 6.11 shows the crystal structure of 2AP and its neighbouring bases inside a duplex[16]. This DNA is bound to the methyltransferase M.HhaI. The 2AP has been positioned outside of the recognition sequence of the enzyme and it has been shown that enzyme binding has not perturbed the base stacking condition of the duplex in this region[16]. Figure 6.11(a) and 6.11(b) show that 2AP has slightly better overlap with the 3' guanine. If helical twist between these two bases is further reduced, as has been done in the hypothetical structure of Figure 6.11(c), maximal overlap and thus a favourable dimerisation condition is met.

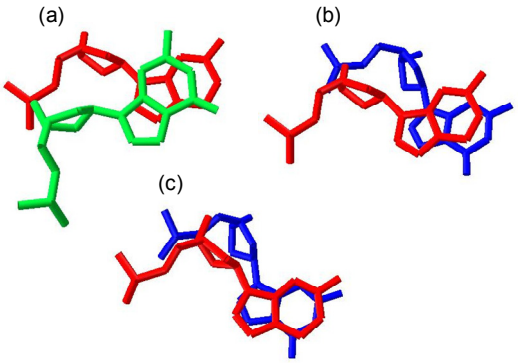


Figure 6.11 Detailed views from the crystal structure of a duplex with 2AP in the context of GPG(T)[16], looking down the long axis of the helix, to illustrate the overlap of (a) 2AP (red) with 5'G (green) and (b) 2AP (red) with 3'G (blue). A hypothetical structure of a 2AP-G nucleodimer, obtained by decreasing the twist between 2AP and 3'G is shown in (c).

It is envisaged that there is a limiting value of helical twist that permits dimer formation. Below this value 2AP dimers will form and labelled duplexes will emit

long-wavelength emission, whereas above the limiting value only short-wavelength emission will arise. Quantum mechanical calculations have shown that for B-form DNA, with an average helical twist value of 36 ^{o} per base pair, a stacked 2AP will fluoresce at the same wavelength as a solvated 2AP[7;8]. Thus a higher degree of stacking is needed for dimer formation than the B-form DNA structure provides. Such rotational offset does not occur in the 2AP crystals investigated by Neely *et al.*[11] and the overlap between adjacent bases is much greater compared with DNA. Consequently, red-shifted emission from 2AP in the crystal-phase is highly intense.

Lifetime parameters for 370-nm emission indicate that the most highly populated duplex conformation of the short wavelength species is the highly-stacked τ_1 arrangement which allows very efficient charge transfer, as indicated by a lifetime of the order of 10s of picoseconds. Neely and Jones have shown that τ_1 disappears when the duplex is frozen indicating that these are thermally accessible conformations[13]. The continued presence of long-wavelength emission at 77 K shows that the stacking condition responsible for the dimers is not a purely dynamic condition and must constitute a minimum energy structure of the duplex.

In decay data for the long wavelength emission of the APT(T), APC(T) and GPG(T) duplexes (Table 6.4), the average lifetime of 2AP in these sequences is 1.44, 1.45 and 1.39 ns respectively (see Equation 2.10). The similarities of these suggest little variation in quantum yield with sequence context. Under steady state conditions the intensity of APT(T) at 450 nm is more than twice APC(T) so relative intensities at 450 nm must relate to the populations of dimers present in these duplexes. Assuming that average lifetimes for the other duplexes show this trend, the data in Table 6.3 can be used to assess the relative populations of dimer species permitted by each sequence.

Some sequences have notably higher dimer populations than others and it appears that dimer population is reliant on the base pairing that the 2AP adopts in these duplexes. The mismatched 2AP-C and 2AP-G make wobble[17] and Hoogsteen pairs respectively and these give higher relative intensities at 450 nm over the 2AP-T

Watson-Crick arrangements. Mismatched base pairs allow greater freedom of base twist so could make dimer formation more facile. For example, consider the 2AP-G stacking in crystal structure data from a GPG(G) duplex shown in Figure 6.12 (once again this is an M.HhaI-bound duplex; the 2AP is outside of the recognition sequence and intrahelical)[16;18].

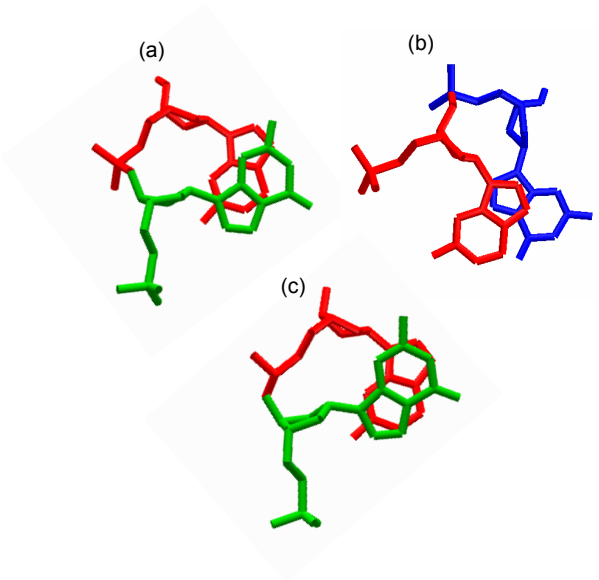


Figure 6.12 Detailed views from the crystal structure of a duplex with 2AP in the context of GPG(G)[18], looking down the long axis of the helix, to illustrate the overlap of (a) 2AP (red) with 5'G (green) and (b) 2AP (red) with 3'G (blue). A hypothetical structure of a 2AP-G nucleodimer, obtained by decreasing the twist between 2AP and 5'G is shown in (c).

Compared with the structure in Figure 6.11, the 2AP is rotated by 180^{o} around the N9-C1' bond such that this causes differing base overlap with the adjacent guanines. It is envisaged that, for this GPG(G) duplex, a heterodimer will most likely form with the 5' guanine (Figure 6.12(c)). In addition, increased base stacking might occur to compensate for the less favourable paring conditions.

The TPT(T) sequence has a low dimer population despite having been shown to exist an unusually highly stacked duplex conformation. Rai *et al.* observe 2AP inside the ATAT sequence to be highly constrained and immobile and resistant to acrylamide quenching[19]. Lifetime data for this duplex, presented as the AP2 duplex in Table 5.2, show that the 2AP does not occupy an extrahelical position, demonstrated by the

absence of τ_4. Furthermore the highly quenched conformations with sub-nanosecond lifetimes constitute 99% of the population. It is apparent then that the TPT(T) stacking geometry permits very efficient charge transfer but is not advantageous for dimer formation.

In comparing the lifetime data for 370- and 450-nm emission it is evident that the short-wavelength emission is much more dependent upon sequence than the long-wavelength emission. This is particularly noticeable in the τ_1 and A_1 parameters. Though the APT(T) and GPG(T) duplexes have similar populations of nucleodimers, for example, these sequences show a factor of 6 difference in their 450:370 nm fluorescence intensity ratio because of a large difference in quantum yield at 370 nm.

The intensity data can provide an estimation of the relative proportions of dimers to monomers in a given sample. The APT(T) duplex has similar average lifetime for its 450- and 370-nm decays 1.18 and 1.55 ns respectively. If the monomer and dimer species have comparable oscillator strengths and, as the average lifetimes indicate, similar quantum yields the ratios given in Table 6.2 reflect the fractional proportion of dimer to monomer in each duplex. This implies that the dimer species accounts for between 1 and 4% across the entire set of duplexes, given the relative intensities of 450-nm emission in Table 3. In this case, approximately 95% of the 2AP in these duplexes would be monomer species and inter duplex variation would be chiefly governed by the sequence-dependent quenching of the 370-nm emission. However, if the oscillator strength of the dimer were smaller than the monomer, the population of dimers in some of these sequences would have to be overly large to produce the higher 450:370 nm intensity ratios reported in Table 6.2.

On occasion, there has been some discrepancy between the relative quantum yields determined from lifetime parameters and those from steady state intensities for 2AP emission in DNA in duplexes measured by this author and by Neely (personal communication). Specifically, globally fitted decays across a 370- to 410-nm envelope tend to yield Φ_{rel} values that are larger than expected. It is now apparent that, as illustrated by Figure 6.4, the short wavelength emission band contains some

contribution from the emitting dimers and this may distort the decays measured at wavelengths longer than the 370-nm emission maximum. All the 450-nm decay lifetimes are longer than their 370-nm counterparts and it is expected that lifetimes calculated for a 390-nm decay, for example, will also be lengthened if emission from the two 2AP species were to mix. Some evidence of this has been observed during the individual decay fitting of sets of 2AP-labelled DNA decays by this author, this will be illustrated in the next chapter. Besides nucleodimer emission, it must be noted that these discrepancies might also arise, in part, from the presence of fast-occurring quenching processes that are too short to be detected by the resolution of the current TCSPC technique.

The multiexponential decay of the long wavelength emission indicates that the nucleodimer exists in a multiplicity of conformational states that are distinguished by different non-radiative decay rates. Conformational heterogeneity of the dimer has a number of possible origins: (i) in each duplex, 2AP may form a nucleodimer with either its 5' or its 3' neighbour; (ii) there will be a range of twist angles (relative orientations of its partner base) over which the nucleodimer remains bound; (iii) the nucleodimer as a whole may adopt a range of conformations relative to its neighbouring bases in the duplex. It is assumed that quenching of the excited dimer occurs through interaction with neighbouring bases and hence variation in the non-radiative decay rate arises primarily from (iii). Thus the heterogeneous decay times of the dimer, like those of the monomer, can be interpreted in terms of the geometric relationship between the dimer as a whole and the local duplex structure (while recognising that each decay time must represent a distribution of conformational states with similar quenching rates). As was the case for 2AP-monomer decay paramters, τ_1 represents conformations of the duplex in which the 2AP dimer is closely associated with its adjacent bases and τ_4 represents the DNA conformation in which the dimer has poor overlap with the adjacent bases. The monomer τ_4 species takes a value close to the solvent exposed monomer and is believed to be extrahelical. It is not clear whether the dimer can adopt a similarly extrahelical position nevertheless τ_4 at 450 nm gives a lower limit for the dimer quenching in duplex DNA.

Magnitudes of the lifetime components τ_1, τ_2 and τ_3 at 450 nm are broadly similar to those at 370 nm showing that each species undergoes similar rates of decay. This implies that the monomer and heterodimer are subjected to similar quenching mechanisms. A notable difference in the GPG(T) decay parameters is the value of τ_1. This takes an especially short value for the monomer species next to a guanine but no τ_1 values as low as 10s of picoseconds are measured at 450 nm in the three duplexes. Neely and Jones have shown that the τ_1 370-nm conformation does not occur at low temperatures, even for a GPG sequence, and that such conformations are attained through thermal motion[13]. It is proposed that rapid dimer quenching is absent because the dimers are geometrically constrained in the DNA conformation rather than because a unique quenching mechanism exists. On freezing these duplexes to 77 K the 450-nm emission intensity increases by the same factor as the 370-nm emission does. Evidently, the same reduction in quenching rate ensues when conformational fluctuations of the duplex are suppressed.

6.5. Conclusion

2AP gives a second emission band with maximum emission at 450 nm when in duplex DNA and in single strand oligonucleotides. This additional emission occurs for all DNA sequence contexts and its intensity is far less sequence dependent than the well-characterised 370-nm emission. The long-wavelength emission originates from a heterodimer species in which the 2AP is π-stacked with one of the neighbouring natural bases. The 2AP heterodimer is present as a ground state molecule and is maximally excited with 360-nm irradiation. When frozen at 77 K 2AP continues to show 450-nm emission which demonstrates that the heterodimer is not exclusively formed by base dynamics. Like the 370-nm 2AP monomer fluorescence, the fluorescence decay of heterodimer emission is 4-exponential and reflects conformational heterogeneity of DNA structure. The lifetimes of 2AP in a heterodimer are all longer than those of the 2AP monomer. Finally, it is clear that dimer emission contributes to the fluorescence of 2AP excited at 310 nm. If the dimer emission is appreciably high then the tail of the 370-nm emission profile becomes visibly distorted, but usually such features are subtle enough as to go unnoticed when the conventional 300- or 320-nm excitation is used.

6.6. References

(1) Hunter, C. A. *Bioessays* **1996**, *18*, 157-162.

(2) Hunter, C. A. *Angewandte Chemie-International Edition in English* **1993**, *32*, 1584-1586.

(3) Rist, M.; Wagenknecht, H.-A.; Fiebig, T. *Chemphyschem* **2002**, *8*, 704-707.

(4) Wang, Q.; Raytchev, M.; Fiebig, T. *Photochemistry and Photobiology* **2007**, *83*, 637-641.

(5) Guest, C. R.; Hochstrasser, R. A.; Sowers, L. C.; Millar, D. P. *Biochemistry* **1991**, *30*, 3271-3279.

(6) Marti, A. A.; Jockusch, S.; Li, Z. M.; Ju, J. Y.; Turro, N. J. *Nucleic Acids Research* **2006**, *34*.

(7) Jean, J. M.; Hall, K. B. *PNAS* **2001**, *98*, 37-41.

(8) Hardman, S. J. O.; Thompson, K. C. *Biochemistry* **2006**, *45*, 9145-9155.

(9) Kwok, W. M.; Ma, C. S.; Phillips, D. L. *Journal of the American Chemical Society* **2006**, *128*, 11894-11905.

(10) Kononov, A. I.; Bukina, M. N. *Journal of Biomolecular Structure & Dynamics* **2002**, *20*, 465-471.

(11) Neely, R. K.; Magennis, S. W.; Parsons, S.; Jones, A. C. *Chemphyschem* **2007**, *8*, 1095-1102.

(12) O'Neill, M. A.; Barton, J. K. *Journal of the American Chemical Society* **2004**, *126*, 13234-13235.

(13) Neely, R. K.; Jones, A. C. *Journal of the American Chemical Society* **2006**, *128*, 15952-15953.

(14) Ramreddy, T.; Sen, S.; Rao, B. J.; Krishnamoorthy, G. *Biochemistry* **2003**, *42*, 12085-12094.

(15) Ramreddy, T.; Rao, B. J.; Krishnamoorthy, G. *Journal of Physical Chemistry B* **2007**, *111*, 5757-5766.

(16) Neely, R. K. The Photophysical Properties of 2-Aminopurine and its Application as a Probe of DNA-Protein Interactions. Thesis 2005.

(17) Sowers, L. C.; Boulard, Y.; Fazakerley, G. V. *Biochemistry* **2000**, *39*, 7613-7620.

(18) Neely, R. K.; Daujotyte, D.; Grazulis, S.; Magennis, S. W.; Dryden, D. T. F.; Klimasauskas, S.; Jones, A. C. *Nucleic Acids Research* **2005**, *33*, 6953-6960.

(19) Rai, P.; Cole, T. D.; Thompson, E.; Millar, D. P.; Linn, S. *Nucleic Acids Research* **2003**, *31*, 2323-2332.

Chapter 7

Long-wavelength emission from 2AP inside DNA-methyltransferase complexes

7.1. Introduction

Chapter 6 established that 2AP in DNA has a 450-nm emission band originating from π-stacked 2AP-base heterodimers. Su *et al.* found that such 2AP long-wavelength emission also arises from labelled DNA bound to the adenine methyltransferase M.EcoKI, which shows a 430-nm emission λ_{max}, maximally excited at 340 nm[1]. Strong 430-nm fluorescence only occurs when 2AP has been placed at the target position for M.EcoKI base flipping and when the DNA is specifically part of a ternary complex made with the natural cofactor SAM. The 430-nm emission evolves over the course of approximately 5 minutes and is accompanied by a corresponding decrease in the 2AP fluorescence at 370 nm. Su *et al.* deduced that a new 2AP ground state species was present, termed 2AP*. 2AP* is not a chemically modified version of the 2AP and is only stable when bound to the methyltransferase. Su *et al.* concluded that 2AP* forms by a specific three-dimensional arrangement of protein, SAM and DNA around the 2AP itself. Mutational studies suggest that the presence of phenylalanine 196 in the NPPF catalytic motif of M.EcoKI is essential for 2AP* generation. Furthermore the rate of appearance of 2AP* is similar to the rate of substrate methylation and so Su *et al.* propose that 2AP* formation is related to a step in the overall reaction prior to methyl group transfer from SAM to DNA.

In this chapter it is shown that the base-flipped DNA-M.TaqI crystal complex (Chapter 3) produces long-wavelength emission when directly excited at 360 nm. This demonstrates that long-wavelength emission from 2AP can be indicative of a specific 2AP-methyltransferase interaction inside the active site. Subsequently, 360-nm excitation of the DNA-enzyme complexes formed by M.TaqI, M.EcoRI and

M.EcoRV in solution is discussed and it will be show that the red-shifted emission from 2AP can be useful in the investigation of DNA-methyltransferase interactions.

7.2. Materials and Methods

The DNA duplexes and methyltransferase complexes investigated in this chapter are:

(1) The 2AP-T/M.TaqI/AETA crystals; 2AP-T/M.TaqI/cofactor solution phase ternary complexes with the AETA, SAM, SAH and sinefungin cofactors; and the 2AP-T/M.TaqI mutant ternary complexes Y108A and Y108F made with the SAM cofactor. Duplex sequences, concentrations and buffers for the M.TaqI complexes are detailed in Section 3.2.

(2) The CBM/M.EcoRI/SAH ternary complex and a CBM/M.EcoRI/SAM ternary complex in which SAM is used at a 100 μM concentration. Duplex sequence, all other concentrations and buffer for the M.EcoRI complexes are detailed in Section 4.2.

(3) The DNA/M.EcoRV/SAH complex with the AP1 and AP2 duplexes. Duplex sequences, concentrations and buffer for the M.EcoRV complexes are detailed in Section 5.2.

Fluorescence emission spectra were recorded using the Fluoromax® spectrofluorometer for excitation wavelengths in the range 310-500 nm with additional 300-nm excitation for the unbound duplexes. The 2AP-T/M.TaqI/AETA crystals were measured using a Fluoromax-P® spectrofluorometer in the manner described in Chapter 2. Emission intensities were calculated by the integration of a 10-nm wide strip under the emission profile encompassing the emission λ_{max}. A 10-nm width was chosen to avoid the strong Raman band that occurs at the emission λ_{max} for these DNA-methyltransferase complexes when 320-nm excitation is employed

7.3. Results and Discussion

7.3.1. Long-wavelength emission from DNA-M.TaqI complexes

Figure 7.1 shows the emission profile that result from 360- and 310-nm excitation of the 2AP-T/M.TaqI/AETA crystals. As for the M.EcoKI system[1], a second emission band arises. λ_{max} emission for the red-shifted band in these M.TaqI crystals is approximately 420 nm. Compared with the unbound duplexes of Chapter 6 the long-wavelength band is relatively strong, possessing an emission intensity which is a third of the 360-nm short-wavelength emission. It is known from the X-ray structure of these crystals (Figure 3.6) that the 2AP is held in a catalytic conformation inside M.TaqI and that the fluorophore π-stacks with tyrosine 108 therefore the 440-nm emission in Figure 7.1 must come from a 2AP-tyrosine heterodimer.

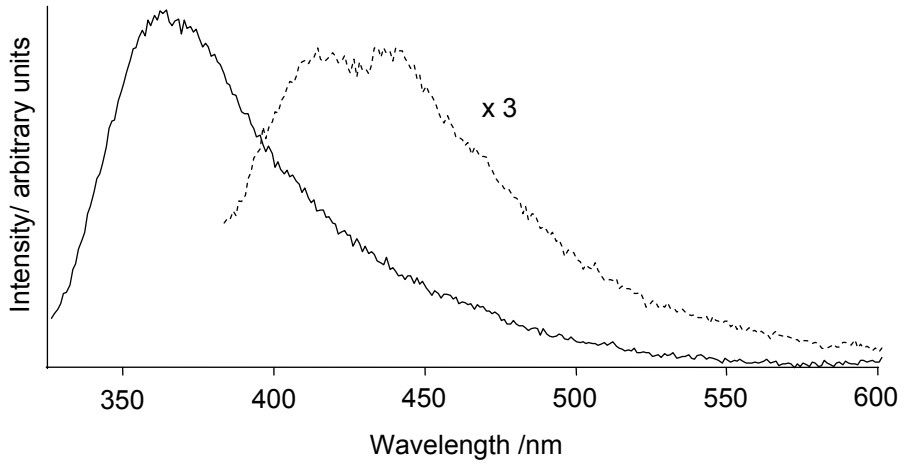


Figure 7.1 Emission spectra, excited at 310 nm (solid line) and 360 nm (broken line, intensity multiplied by a factor of 3), of the 2APT/M.TaqI/AETA crystalline complex.

Since the intensity of heterodimer emission is high in this sample, fluorescence from the long-wavelength emission band will contribute to the 310-nm-excited fluorescence emission profile (i.e. to the solid line spectrum in Figure 7.1)

Decay data collected at separate observation wavelengths over a 360 to 410-nm emission window yields the parameters reported in Table 7.1 when the decays are individually fitted.

Wavelength /nm	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4
360	-	0.90	2.15	6.16	-	0.45	0.52	0.03
370	-	0.90	2.21	6.74	-	0.45	0.52	0.03
390	0.29	1.30	2.82	9.87	0.10	0.59	0.29	0.02
410	0.23	1.31	3.21	11.53	0.15	0.61	0.21	0.03

Table 7.1 Lifetimes and A factors calculated from the individual fitting of decays from the 2AP-T/M.TaqI/AETA crystals.

There is always a slight lengthening of the lifetime values across the chosen wavelengths (which characterise a distribution of similarly quenching conditions) but for this decay set the lengthening is unusually large and significant. It is known from the results reported in Chapter 6 that 2AP-base heterodimers have longer lifetimes than their 2AP monomer counterparts. Evidently the same effect is found for the 2AP-tyrosine heterodimer. The 390- and 410-nm decays from the 2AP-T/M.TaqI/AETA crystals have become complicated by the presence of heterodimer fluorescence as well as 2AP monomer fluorescence thus caution must be exercised in the choice of decays used to calculate global decay parameters for 2AP monomer emission alone.

On addition of the different cofactors to the 2AP-T/M.TaqI binary complex discussed in Chapter 3, steady state measurements show that formation of the ternary complex incorporating SAM causes the intensity of 2AP emission to fall, whereas Φ_{rel} calculated from lifetime parameters suggests that the emission intensity should increase (Figure 3.16). Figure 7.2 presents emission profiles for the 2AP-T/M.TaqI binary and 2AP-T/M.TaqI/SAM ternary complex at three excitation wavelengths. The 2AP-heterodimer species has been selectively excited at 360 nm and its emission spectrum is shown alongside the corresponding 2AP monomer emission spectra for

these complexes. In solution λ_{max} of emission for the red-shifted band comes at 440 nm.

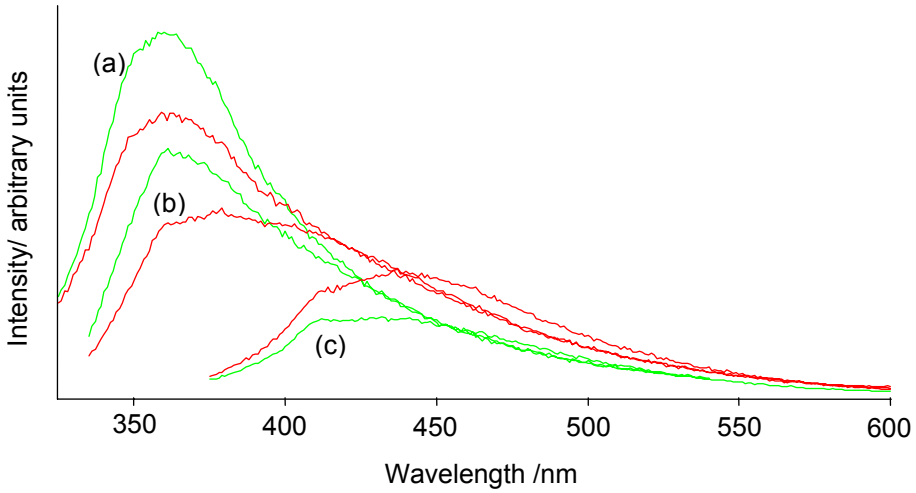


Figure 7.2 Emission spectra for binary 2AP-T/M.TaqI (green) and ternary 2AP-T/M.TaqI/SAM (red) in solution, excited at (a) 310 nm (b) 320 nm and (c) 360 nm wavelengths (note that there is always a Raman band present at the λ_{max} of 2AP in these enzyme complexes when 320-nm excitation is employed).

As can be seen, when the DNA-M.TaqI complex is converted from binary to ternary complex, 2AP-monomer emission intensity falls (Figure 7.2(a) and (b)) but 2AP-heterodimer intensity increases (Figure 7.2(c)). It is now evident that the Φ_{rel} for 2AP-T/M.TaqI/SAM is an overestimation since 2AP-heterodimer emission contributes to the fluorescence decay observed at 390 nm. Table 7.2 reproduces the lifetime parameters for this complex (previously presented in Table 3.6) and the 2AP-T/MTaqI binary complex (from Table 3.4).

Solution composition	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	Φ_r
2AP-T/M.TaqI	-	0.41	1.9	7.3	-	0.21	0.66	0.13	0.21
2AP-T/M.TaqI/SAM	0.12	0.94	3.3	9.9	0.32	0.32	0.20	0.16	0.24

Table 7.2 Lifetimes and A factors, reported for the decay observed at 390 nm, for the 2AP-T duplex in binary complex with M.TaqI (global parameters) and in ternary complex with M.TaqI and SAM (individual decay parameters). Also reported are the calculated quantum yields relative to free 2AP riboside $\Phi_{rel.}$.

Recall that for the ternary complexes with SAM some are in conformations where the target base is intrahelical whereas all the binary complexes are fully flipped. As can be seen, the lifetimes that characterise the flipped conformations for the ternary complex are all substantially longer than those for the binary complex. Given the 390-nm decays from the crystal complex, it is reasonable to assume that the lifetimes calculated for the 2AP monomer species for the ternary complex with SAM have also been lengthened as a result of additional dimer fluorescence in the decay at 390 nm. Presuming that enhancement of the red-shifted emission signifies an increase in population of 2AP-tyrosine heterodimers, it appears that the flipped 2AP becomes more constrained in the complex when M.TaqI binds SAM and that the 2AP and tyrosine adopt a geometry favourable for stable dimer formation.

Figure 7.3 shows the corresponding binary to ternary complex evolution of 2AP emission when AETA is employed as cofactor in solution. Of the set of different cofactors used (Chapter 3), AETA causes the smallest change in 2AP emission intensity (Figure 3.16).

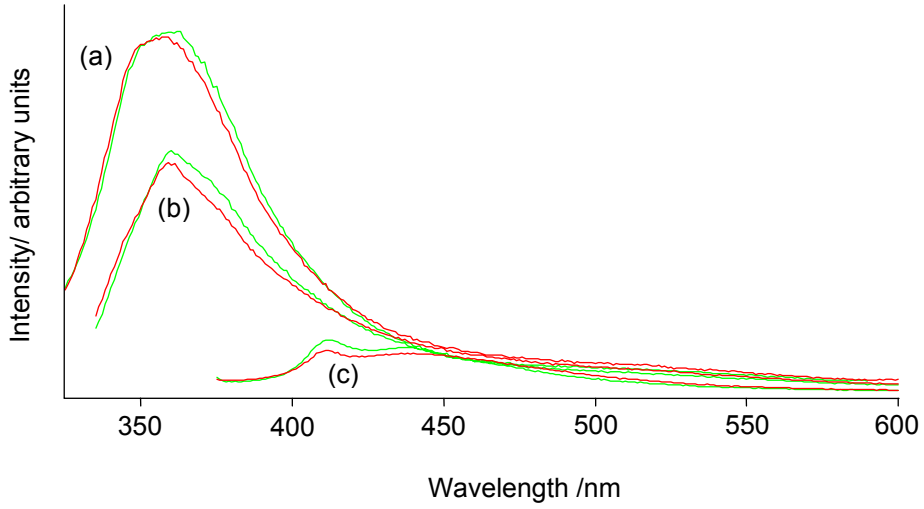


Figure 7.3 Emission spectra for binary 2APT/M.TaqI (green) and ternary 2APT/M.TaqI/AETA (red) in solution, excited at (a) 310 nm (b) 320 nm and (c) 360 nm wavelengths (note that there is always a Raman band present at the λ_{max} of 2AP in these enzyme complexes when 320-nm excitation is employed).

The long-wavelength emission band from 2AP-T/M.TaqI/AETA is much weaker (Figure 7.3 (c) red) than the short-wavelength emission band (Figure 7.3 (a) red). In the emission profile from 320-nm excitation (Figure 7.3 (b) red) dimer emission has caused little distortion to the principal 360-nm emission band, unlike for SAM (Figure 7.2 (b) red). So the calculated Φ_{rel} changes for the AETA complex formation agree with the 310 nm-excited steady state intensity measurements because the red-shifted emission does not contribute significantly to the 320 nm-excited fluorescence decays.

It is interesting that the DNA-M.TaqI ternary complex formed with AETA shows strong red-shifted emission in the crystal phase but not in solution. Based on 2AP monomer decay characteristics (Table 3.4) the flipped 2AP target in solution phase complex is also deemed to be in the catalytic NPPY motif inside M.TaqI but weak 415-nm emission reports that the population of 2AP-tyrosine dimers in complex is low. It is therefore thought that 2AP-tyrosine dimers do not form if the 2AP is too mobile inside the M.TaqI active site pocket. Thus lasting 2AP-tyrosine contacts cannot form in solution but when crystallised the 2AP becomes more constrained which then locks down dimer-favouring geometries. For the ternary SAM complex, where red-shifted emission was high, the large 2AP-tyrosine dimer population inferred supports the interpretation made from monomer emission, which is that 2AP resides in the NPPY active site in the 2AP-T/M.TaqI/SAM even though the solution phase lifetimes from this complex (Tables 7.2 and 3.6) differ from the crystal lifetimes (Table 3.4). The calculated fluorescence lifetimes for the 2AP-T/M.TaqI/SAM are merely longer as a consequence of included dimer fluorescence.

Chapter 3 showed that from unbound to bound DNA, and from DNA-M.TaqI binary to ternary complex formation, the emission intensity of 2AP at 360 (or 370 nm) changes. Table 7.3 lists the intensity of emission at 440 nm relative to the 360-nm emission intensity for each of these stages of complexation and for the ternary complex made with each cofactor. 2AP in the unbound 2AP-T duplex shows such strong 440-nm emission relative to other duplexes that the emission profile is broad, even when 300-nm excitation is employed (spectrum not shown). Perhaps this is not

surprising since the lifetime data report that this duplex adopts highly stacked conformations (87% of the 2AP in the sample have sub-nanosecond lifetimes, Table 3.4). Therefore the reason for the high long/short ratio for this duplex is that the short-wavelength emission of 2AP is very highly quenched in conjunction with strong long-wavelength emission.

In some proportion of the ternary complexes made with SAM, SAH and sinefungin 2AP is unflipped 2AP but, since intrahelical 2AP is highly quenched, it can be assumed that most of the emission at 360 or 370 nm in the enzyme complexes comes from extrahelical 2AP. Similarly, the absolute intensity of the 2AP-base heterodimer species in the free duplex is low compared with the strong long-wavelength emission from the 2AP-tyrosine heterodimer. Thus the ratios of Table 7.3 give an indication of dimer/monomer emission for the extrahelical 2AP population.

Cofactor	Intensity at 440 nm relative to 360 nm intensity
2AP-T	1.3
2AP-T/M.TaqI	0.24
AETA	0.19
SAM	0.69
SAH	0.40
sinefungin	0.43

Table 7.3 The emission intensity at 440 nm, excited at 360 nm, relative to that at 360 nm, excited at 320 nm, for the 2AP-T/M.TaqI/cofactor complexes. The value for 2AP-T/M.TaqI is an average.

Like the M.EcoKI system, strong long-wavelength emission occurs when the ternary complex forms[1]. The ratio of dimer/monomer emission is low for the AETA cofactor, high for SAM and intermediate for SAH and sinefungin. Figure 7.4 displays the normalised emission profiles that result from 320-nm excitation of these ternary complexes (using this excitation wavelength, emission profiles will include both monomer and 2AP-heterodimer fluorescence).

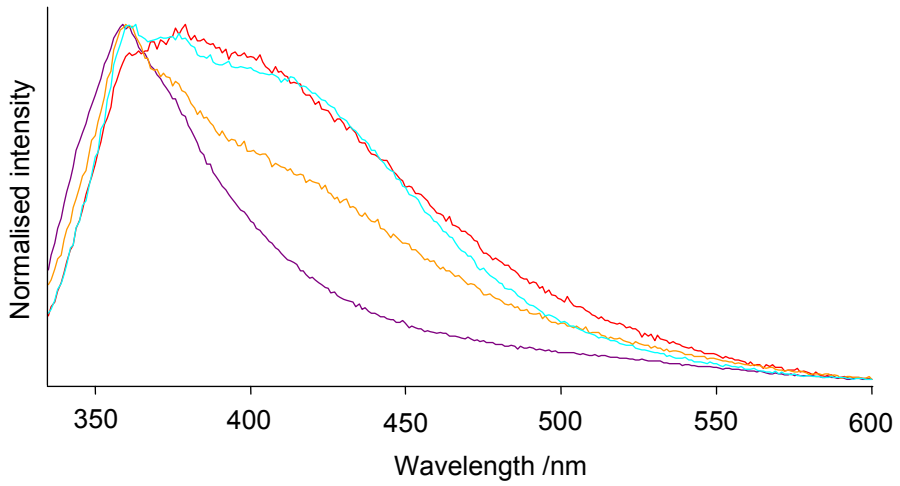


Figure 7.4 Normalised emission spectra, excited at 320 nm, for the 2AP-T/M.TaqI/cofactor complexes in solution with AETA (purple), SAM (red), SAH (orange) and sinefungin (turquoise). Note that there is always a Raman band present at the λ_{max} of 2AP in these enzyme complexes when 320-nm excitation is employed.

It can be seen from Figure 7.4 that the low 440/360 nm ratio of 0.2 in the AETA complex is not sufficient to bring about a noticeable change in the shape of the 360-nm emission profile whereas the other cofactors, with ratios above 0.4, cause an obvious change to spectral shape. The 2AP decay parameters for monomer emission support the notion that the flipped base does interact with the NPPY motif for all of these cofactors in complex. Some long-wavelength emission is evident in all cases however red-shifted emission enhancement is clearly strongest for the SAM cofactor.

For the 2AP-T duplex in complex with the M.TaqI mutants and SAM, lifetime data for 2AP monomer emission suggest that the flipped 2AP experiences little or no stabilisation in the M.TaqI catalytic motif when tyrosine 108 is replaced with an alanine or phenylalanine (section 3.3.5). This was inferred since the calculated 2AP lifetimes in the mutant complexes are larger than the sub- or near-nanosecond lifetimes known to describe a catalytic NPPY-stabilised 2AP. It is not possible to form a 2AP-alanine π-stacked heterodimer in Y108A but there is the potential for a 2AP-phenylalanine heterodimer in both mutants through phenylalanine 196 and additionally with phenylalanine 108 in the Y108F mutant. When the DNA-mutant-SAM ternary complexes are directly excited with 360-nm excitation to probe for

2AP-heterodimer emission, the subsequent 440-nm emission band is extremely weak. 440-nm emission intensity relative to the monomer 360-nm emission level is 6.0 x 10^{-3} for the Y108A complex and 0.03 for the Y108F complex (spectra not shown). These low relative populations of the 2AP-heterodimer species, compared with that of the complex made with wild type M.TaqI and SAM, further support that 2AP is not stabilised in NPPF in an equivalent way to NPPY and also that there is no π-stacking dimerisation with the phenylalanine196 in either of the mutants. Evidently, the lifetime data for the mutant complexes (Tables 3.9 and 3.10) are completely characteristic of monomer emission.

7.3.2. Long-wavelength emission from DNA-M.EcoRI complexes

Figure 7.5 presents emission spectra from the DNA/M.EcoRI/sinefungin complex where 2AP is in the target position for base flipping and has been excited with 320 and 360 nm wavelengths. As can be seen there is some long-wavelength emission from 2AP in this complex but the emission is weak.

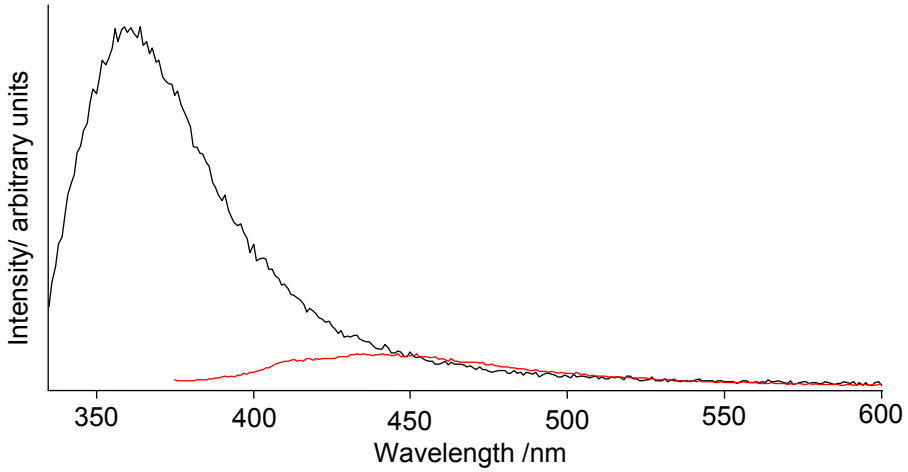


Figure 7.5 Emission spectra from the CBM/M.EcoRI/sinefungin complex in solution excited at 320 nm (black) and 360 nm (red) wavelengths.

The absolute intensity of the long-wavelength emission changes very little between bound and unbound DNA; long-wavelength emission of the enzyme-bound DNA is half as intense as that of the unbound DNA (data not shown). Given that the lifetime

data presented in Chapter 4 report that 25% of these DNA/M.EcoRI/sinefungin complexes remain unflipped, the red-shifted emission in Figure 7.5 is proposed to come mainly from intrahelical 2AP-base heterodimers.

Figure 7.6 presents the corresponding emission spectra from a DNA/M.EcoRI/SAM complex and compared to Figure 7.5, the emission profiles are quite different.

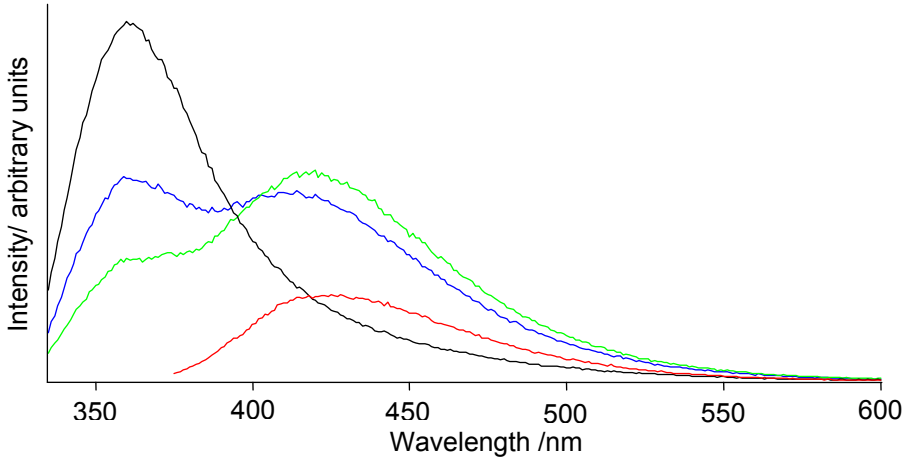


Figure 7.6 Emission spectra for the CBM/M.EcoRI/SAM complex in solution excited at 320-nm (black, blue and green) and 360-nm (red) wavelengths. Of the 320-nm excited spectra, the black spectrum was measured first, the blue spectrum occurred after 10 minutes and the same profile persisted for at least 20 more minutes. The green spectrum was measured 45 minutes after the black spectra.

There is a visibly larger red-shifted emission band from 360-nm excitation, with λ_{max} at 420 nm. Like in M.EcoKI[1], red-shifted emission occurs only with the natural cofactor SAM and, similarly, the emission profile that ensues from 320-nm excitation changes over the course of successive scans. With M.EcoRI, long-wavelength fluorescence from 2AP excited at 320 nm becomes distinctly visible and eventually dominates the emission profile (Figure 7.6 blue and green). As the 420-nm emission in Figure 7.6 is so strong it is reasonable to assume that a 2AP-M.EcoRI interaction is directly responsible and that a 2AP-tyrosine heterodimer has formed from stabilisation in the DPPY catalytic motif of M.EcoRI. If this is the case, then the absence of significant red-shifted emission from the ternary complex with sinefungin suggests that the flipped 2AP is not stabilised in a catalytic

conformation when sinefungin is employed as cofactor. The same conclusion is made from the lifetime data of the 2AP monomer species in Chapter 4.

7.3.3. Long-wavelength emission from DNA-M.EcoRV complexes

Chapter 5 discussed the use of 2AP to probe the two adenine positions inside the recognition sequence of M.EcoRV. Lifetime data support that the target base is flipped and held in the DPPY catalytic motif whereas the non-target adenine is merely unstacked as a consequence of duplex bending. The emission profiles from these DNA/M.EcoRV/SAH complexes directly excited with 360 nm are shown in Figures 7.7 and 7.8 (target and non-target positions respectively). Though both show large changes to the monomer emission on enzyme binding, little change occurs to the heterodimer emission levels. Figure 7.9 magnifies the long-wavelength emission region.

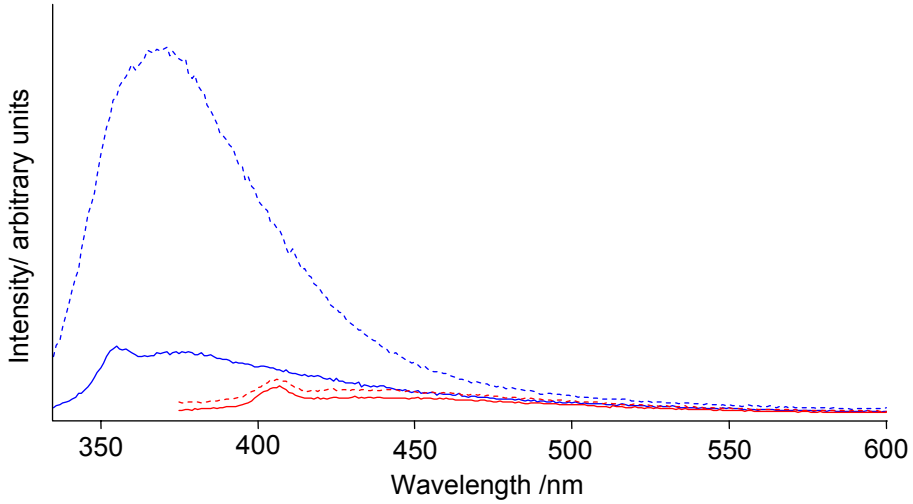


Figure 7.7 Emission spectra for the unbound AP1 (2AP at target site) duplex (solid line) and AP1/M.EcoRV/SAH complex (broken line) in solution, excited at 320-nm (blue) and 360-nm wavelengths (red).

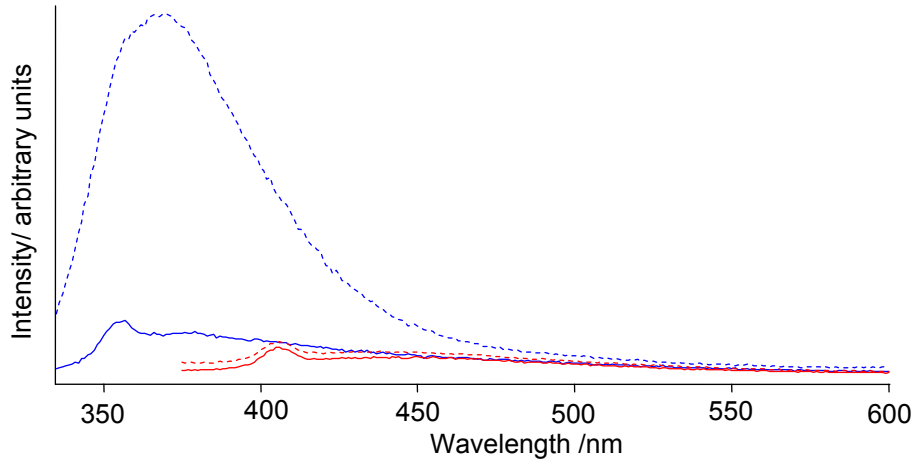


Figure 7.8 Emission spectra for the unbound AP2 (non-target 2AP) duplex (solid line) and AP2/M.EcoRV/SAH complex (broken line) in solution, excited at 320-nm (blue) and 360-nm wavelengths (red).

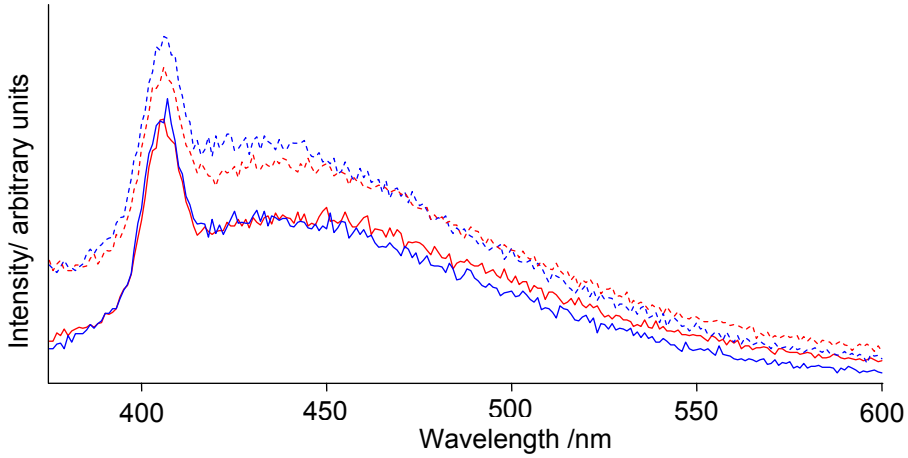


Figure 7.9 Emission spectra, from 360-nm excitation, for the AP1 (2AP at target site) duplex (blue) and the AP2 (non-target 2AP) duplex (red), unbound (solid line) and in ternary complex with M.EcoRV and SAH (broken line).

Additionally, the use of the natural cofactor SAM, in place of SAH, does little to change the shape and intensity of the red-shifted emission band (spectrum not shown) so this lack of effect in M.EcoRV is not cofactor dependent.

Weak red-shifted emission in the M.EcoRV system suggests a poor association between the flipped base and the aromatic tyrosine in the active site of this enzyme. This is a reasonable interpretation since M.EcoRV is generally regarded as a poor

methyltransferase. Compared with the other methyltransferases that have been investigated in this thesis, for example, the rate of M.EcoRV methylation (k_{cat}) is 3 orders of magnitude lower; 16.7 $x10^{-5}$ s^{-1} compared with 0.73 s^{-1} for M.TaqI and 0.105 s^{-1} for M.EcoRI (on substrates with a near-equivalent number of base pairs)[2]. Coupled with DNA binding affinity (K_m), the specificity of methylation of the M.EcoRV target sequence (k_{cat}/K_m) is 0.184 x 10^3 $M^{-1}s^{-1}$ compared with 1.21 x 10^6 $M^{-1}s^{-1}$ for M.TaqI and 2.45 $x10^6$ $M^{-1}s^{-1}$ for M.EcoRI[2].

Since there is no change in 2AP long-wavelength emission for the M.EcoRV substrate duplexes, red-shifted fluorescence from 2AP is not instructive in the current DNA-M.EcoRV investigation.

7.4. Conclusion

2AP-labelled DNA substituted at the target for base-flipping and bound to the M.TaqI and M.EcoRI methyltransferases give out strong 440-nm emission. The long-wavelength emission band is believed to come from a 2AP-tyrosine heterodimer species formed when the 2AP is flipped and stabilised inside the catalytic motif of these enzymes. As was the case for the adenine methyltransferase M.EcoKI, red-shifted emission from the M.TaqI and M.EcoRI complexes only arises when the natural cofactor SAM is bound. The long-wavelength emission band can dominate the emission excited at 320 nm and, for M.EcoRI, continually changes the shape of the emission profile. This study has revealed another way in which 2AP might be employed to probe DNA-methyltransferase interactions.

7.5. References

(1) Su, T. J.; Connolly, B. A.; Darlington, C.; Mallin, R.; Dryden, D. T. F. *Nucleic Acids Research* **2004**, *32*, 2223-2230.

(2) Bheemanaik, S.; Reddy, Y. V. R.; Rao, D. N. *Biochemical Journal* **2006**, *399*, 177-190.

Chapter 8

Conclusion

The aim of this thesis was to investigate base flipping in the adenine methyltransferases using the time-resolved fluorescence of 2-aminopurine (2AP) and to build on a previously devised strategy that was applied to a cytosine methyltransferase with much success.

2AP is an excellent substitute for the natural adenine target of the adenine methyltransferases for, in M.TaqI, the 2AP is delivered into the same position inside the enzyme as adenine and with the same orientation that prepares the adenine for enzyme catalysis. 2AP lifetimes have the sensitivity to detect local conformational changes to the DNA duplex at a specific sequence position. Yet again, the parameters of fluorescence decay respond uniquely to base flipping, without the ambiguity that has plagued earlier efforts to identify the occurrence of this physical mechanism. Furthermore the time-resolved fluorescence method lends itself to a completely solution-based investigative approach.

Essentially, it is the fluorescence of 2AP in DNA that is being monitored therefore it is important to understand how this fluorophore behaves in the context of the DNA duplex. 2AP gives out two types of emission when in DNA. The first is the well-known 370-nm emission, which emanates from 2AP as a monomer. The second is 450-nm emission and comes from a 2AP-base heterodimer.

This heterodimer species is a 2AP that is π-stacked with one of the neighbouring natural DNA bases, thus 2AP-heterodimers are present in single strand oligonucleotides as well as in duplex DNA. 2AP in a heterodimer is maximally excited with a 360-nm wavelength but its intensity is low compared with the 2AP monomer emission. Nevertheless, the fluorescence decay of the heterodimer is 4-exponential and so also reflects the conformational heterogeneity of DNA structure.

One of the ways in which the adenine methyltransferases stabilise a flipped adenine is by face-to-face π-stacking with the aromatic amino acid from the DPPY catalytic motif. This π-stacking interaction creates similar 2AP-heterodimers in DNA-methyltransferase complexes when 2AP is the base that is flipped. In this instance, and in the presence of cofactor, a 2AP-tyrosine or 2AP-phenylalanine heterodimer forms and also gives rise to emission at 450 nm. In DNA-enzyme complexes, strong 450-nm emission from 2AP-amino acid heterodimers confirms that the flipped 2AP is located within the catalytic motif of the methyltransferase.

Principally it is the fluorescence from 2AP monomers that has been monitored for lifetime analysis. If the emission from 2AP-amino acid heterodimers is significantly intense, then the decays measured from some DNA-enzyme complexes include both monomer and dimer fluorescence decay. In this case, the 2AP monomer lifetimes that are calculated are longer than expected. This effect is easily noticeable when individually fitting decays that have been measured at different emission wavelengths.

The methyltransferases in this thesis, M.TaqI, M.EcoRI and M.EcoRV, have been investigated using a combination of time-resolved fluorescence with the 2AP monomer species and steady state fluorescence with the 2AP-heterodimer species. Combined crystal- and solution-phase studies on the M.TaqI methyltransferase have allowed a greater understanding of 2AP lifetime behaviour relating to the adenine methyltransferases. Subsequently, an increased confidence has been gained in the interpretation of results from solution-phase studies alone.

For M.TaqI, fluorescence lifetimes show that all of the DNA-M.TaqI complexes are flipped and highlight the efficiency of this enzyme at extruding its target base. When 2AP is flipped into the active site of M.TaqI its 370-nm fluorescence is enhanced compared with the unbound DNA, however, there is still significant quenching of the 2AP via the stabilising enzyme interactions that hold the base ready for catalysis. Specifically, it is the face-to-face π-stacking between 2AP and the tyrosine from the NPPY catalytic motif that has the greatest influence on quenching. This interaction

creates a 2AP-tyrosine dimer and so the DNA-M.TaqI complex also gives out 450-nm emission.

It is evident that there is a degree of heterogeneity in the conformation adopted by the DNA-M.TaqI complex, in both crystal- and solution-phase. The decay parameters of the DNA-M.TaqI complex show that not all of the flipped bases are held inside NPPY (i.e. in a position that would favour catalysis). In a minority of complexes 2AP fluorescence is quenched very little and the base has minimal direct contact with the enzyme when buried in the enzyme interior.

The quenching conditions for 2AP inside M.TaqI differ from those of the cytosine methyltransferase M.HhaI, where there are no π-stacking interactions with the 2AP. Thus, fluorescence lifetimes reflect the nature of the environment that the enzyme provides for the flipped base in addition to revealing that base flipping has occurred.

In the sequence of bases recognised by M.TaqI, the thymine opposite the target adenine is not essential to the machinery of base flipping. It does, however, play some role delivering the target base to the active site motif. With this partner thymine removed, a third of the DNA-M.TaqI complexes hold the flipped target 2AP in a position of minimal enzyme interaction, in a geometry that is assumed not to allow catalysis. Finally, with this particular DNA substrate, that lacks the partner thymine, the DNA-M.TaqI conformations that crystallise are not wholly representative of the conformations that occur in solution.

The time-resolved fluorescence method has allowed a study on the effect of cofactor molecule on base-flipping behaviour. With the natural cofactor S-adenosyl-L-methionine (SAM) and with the cofactor product S-adenosyl homocysteine (SAH), a third of the DNA-M.TaqI-cofactor ternary complexes remain unflipped. In a biological context, these are the most relevant complexes to consider. When the synthetic cofactor 5'-[2(amino)ethylthio]-5'-deoxyadenosine (AETA) is employed the all the M.TaqI ternary complexes are flipped, and with sinefungin, a well-known methyltransferase inhibitor, only 50% of the complexes are flipped. Using AETA or

sinefungin as cofactor in the M.TaqI complex causes the enzyme to behave unnaturally. Steady state fluorescence measurements show that the flipped 2AP in the DNA-M.TaqI-SAM complex is held the most tightly within the catalytic NPPY motif.

M.TaqI is sensitive to the engineering of its NPPY catalytic motif. When the tyrosine is replaced with phenylalanine (F), another aromatic amino acid, the NPPF motif does not engage with the flipped base as NPPY does. The 2AP does not π-stack in a face-to-face manner with the NPPF phenylalanine. With an NPPF motif or an NPPA motif, where tyrosine is replaced with the non-aromatic alanine (A), the M.TaqI mutant is more reliant on the other amino acids interactions that stabilise the flipped base in the adenine binding pocket. These are edge-to-face π-stacking with phenylalanine 196 and available hydrogen-bonding. Such interactions are not as effective as the original tyrosine at quenching the 2AP fluorescence when inside M.TaqI.

The M.EcoRI methyltransferase also flips its target base. In the DNA-enzyme complex with sinefungin, the majority of the flipped 2APs are unquenched and thus not held in the NPPF catalytic motif. Hence most of complexes that form under these conditions hold conformations that do not feature on the catalysed reaction coordinate. This is purely a consequence of cofactor choice, however, since when the natural SAM cofactor is used, the flipped 2AP strongly associates with the NPPF catalytic motif of M.EcoRI.

The H235N mutant of M.EcoRI carries a single mutation in the conserved motif VIII that is part of the adenine binding pocket. When using sinefungin as cofactor, approximately half of the DNA-H235N complexes hold a conformation in which the flipped 2AP is not arranged within NPPF. In this regard, the mutant and wild-type enzyme behave similarly. There is some evidence that the mutant acts differently to the wild-type M.EcoRI. With the current resolution, however, it is unclear as to whether the mutant provides a different environment for the flipped base or whether less of the DNA-mutant complexes are flipped.

When M.EcoRI binds to a DNA base sequence that resembles its recognition sequence in all but one base, a non-cognate sequence, the enzyme perturbs a 2AP in the equivalent-to-target position, but only if that 2AP is not tightly buried inside the DNA duplex. The 2AP-enzyme interactions for the near-similar target base are not the same as those for the actual target base. Thus, M.EcoRI makes a distinction between a cognate and a non-cognate base sequence through the terminal cytosine of its GAATTC recognition sequence.

When the H235N mutant encounters the same non-cognate sequence the mutant does not disturb any of the 2APs, whether the base is highly stacked or not. Evidently the asparagine 235 (N) of the mutant specifically uses the terminal cytosine in substrate recognition.

The M.EcoRV methyltransferase is not as efficient at flipping its target base compared with M.TaqI or M.EcoRI. At equilibrium, at least 45% of the DNA-M.EcoRV complexes are unflipped. Three conformations are resolved for the flipped complexes; the 2AP in most of these complexes is significantly quenched. In contrast, when 2AP is placed at a non-target position in the M.EcoRV recognition sequence, the 2AP in a third of the DNA-enzyme complexes are unquenched. The base stacking of the 2AP at this position is alleviated by enzyme-binding. This shows how a larger emission signal arises when 2AP is placed at the non-target site compared with the target site. When M.EcoRV binds to its GATATC recognition sequence, the base-enzyme interactions between each adenine position are clearly quite different.

This thesis has shown that a time-resolved fluorescence approach has uncovered a wealth of information about the particular adenine methyltransferases studied. It has demonstrated that many aspects of the methyltransferase interaction can be investigated with such simplicity, namely; how the enzyme interacts with its target base; how the enzyme interacts specifically with each of the other bases within its recognition sequence; how the enzyme behaves towards its target if another base in

the recognition sequence is changed or removed; how the enzyme behaves when particular cofactors or ligands are bound; how important particular amino acids are, to base flipping or in target base stabilisation, and what happens when certain amino acids are mutated. These strategies easily can be applied to other DNA-binding proteins.

Another emissive species of 2AP has emerged, the 2AP π-stacked heterodimer, which provides additional means of probing the DNA-methyltransferase interaction. This species can be used to enhance or complement investigations based on the conventional 370-nm emission from 2AP.

Appendix A

Fit Parameters for Fluorescence Decay Curves

Global lifetimes /ns

$\tau_2 = 0.55$ $\tau_3 = 1.88$ $\tau_4 = 7.28$

Emission λ /nm	A_2	A_3	A_4	Local χ^2
370	0.55	0.39	0.06	1.135
390	0.54	0.38	0.08	1.072
410	0.54	0.36	0.10	1.176

Global $\chi^2 = 1.127$

Table A1 Global decay parameters for the 2AP-T/M.TaqI/AETA ternary complex in solution.

Global lifetimes /ns

$\tau_2 = 0.41$ $\tau_3 = 1.89$ $\tau_4 = 7.33$

Emission λ /nm	A_2	A_3	A_4	Local χ^2
370	0.23	0.58	0.19	1.111
390	0.21	0.66	0.13	1.050

Global $\chi^2 = 1.086$

Table A2 Global decay parameters for the 2AP-T/M.TaqI binary complex in solution, reproduced with permission from R.K. Neely.

Global lifetimes /ns

$\tau_1 = 0.03$ $\tau_2 = 0.57$ $\tau_3 = 2.38$ $\tau_4 = 7.82$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
370	0.89	0.03	0.06	0.02	1.141
390	0.81	0.06	0.09	0.04	1.038
410	0.88	0.05	0.03	0.03	1.070

Global $\chi^2 = 1.082$

Table A3 Global decay parameter for the free 2AP-T duplex.

Global lifetimes /ns

$\tau_2 = 0.92$ $\tau_3 = 2.34$ $\tau_4 = 7.57$

Emission λ /nm	A_2	A_3	A_4	Local χ^2
360	0.47	0.48	0.05	1.104
370	0.46	0.48	0.06	1.097
390	0.44	0.48	0.08	1.053
410	0.44	0.46	0.10	1.104

Global $\chi^2 = 1.090$

Table A4 Global decay parameters for the 2AP-D/M.TaqI/AETA ternary complex in the crystal phase.

Global lifetimes /ns

$\tau_2 = 0.28$ $\tau_3 = 2.81$ $\tau_4 = 9.64$

Emission λ /nm	A_2	A_3	A_4	Local χ^2
370	0.27	0.28	0.45	1.048
390	0.33	0.30	0.37	1.070

Global $\chi^2 = 1.059$

Table A5 Global decay parameters for the 2AP-D/M.TaqI/AETA ternary complex in solution.

Global lifetimes /ns

$\tau_2 = 0.23$ $\tau_3 = 3.31$ $\tau_4 = 9.67$

Emission λ /nm	A_2	A_3	A_4	Local χ^2
370	0.18	0.22	0.60	1.035
390	0.26	0.18	0.56	1.045

Global $\chi^2 = 1.040$

Table A6 Global decay parameters for the 2AP-D/M.TaqI binary complex in solution.

Global lifetimes /ns

$\tau_1 = 0.04$ $\tau_2 = 0.26$ $\tau_3 = 2.48$ $\tau_4 = 8.29$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
370	0.75	0.08	0.07	0.10	1.092
390	0.62	0.18	0.09	0.11	1.048
410	0.63	0.22	0.08	0.07	1.078

Global $\chi^2 = 1.072$

Table A7 Global decay parameters for the free 2AP-D duplex.

Wavelength /nm	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	χ^2
370	0.11	0.91	3.02	9.41	0.34	0.30	0.21	0.15	1.041
390	0.12	0.94	3.26	9.87	0.32	0.32	0.20	0.16	1.062
410	0.13	0.90	3.29	10.20	0.32	0.32	0.20	0.16	1.033

Table A8 Individual decay parameters for the 2AP-T/M.TaqI/SAM ternary complex.

Wavelength /nm	τ_1/ ns	τ_2/ ns	τ_3/ ns	τ_4/ ns	A_1	A_2	A_3	A_4	χ^2
370	-	1.03	4.80	9.43	-	0.08	0.29	0.63	1.074
390	-	0.99	5.29	9.68	-	0.06	0.28	0.66	1.003
410	-	0.91	5.12	9.66	-	0.07	0.26	0.67	1.005

Table A9 Individual decay parameters for the 2AP-D/M.TaqI/SAM ternary complex.

Global lifetimes /ns

$\tau_1 = 0.14$ $\tau_2 = 0.54$ $\tau_3 = 2.55$ $\tau_4 = 8.73$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
370	0.51	0.35	0.09	0.05	1.056
390	0.50	0.35	0.09	0.06	1.036
410	0.46	0.36	0.11	0.07	0.994

Global $\chi^2 = 1.028$

Table A10 Global decay parameters for the free CBM duplex.

Global lifetimes /ns

$\tau_1 = 0.17$ $\tau_2 = 1.27$ $\tau_3 = 4.36$ $\tau_4 = 8.40$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
370	0.24	0.09	0.11	0.56	1.014
380	0.24	0.14	0.09	0.53	1.050

Global $\chi^2 = 1.032$

Table A11 Global decay parameters for the CBM/M.EcoRI/sinefungin complex.

Global lifetimes /ns

$\tau_1 = 0.16$ $\tau_2 = 0.88$ $\tau_3 = 3.87$ $\tau_4 = 8.04$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
370	0.24	0.11	0.08	0.57	1.001
380	0.25	0.13	0.08	0.54	1.030

Global $\chi^2 = 1.015$

Table A12 Global decay parameters for the CBM/H235N/sinefungin ternary complex, the M.EcoRI mutant.

Global lifetimes /ns

$\tau_1 = 0.15$ $\tau_2 = 0.46$ $\tau_3 = 2.54$ $\tau_4 = 9.91$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
370	0.52	0.36	0.04	0.08	1.029
390	0.51	0.36	0.04	0.09	1.027
410	0.50	0.35	0.05	0.10	1.024

Global $\chi^2 = 1.026$

Table A13 Global decay parameters for the free A4 duplex and for the APT(T) duplex excited with a 318-nm wavelength.

Global lifetimes /ns

$\tau_1 = 0.17$ $\tau_2 = 0.74$ $\tau_3 = 3.62$ $\tau_4 = 9.21$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
370	0.47	0.20	0.11	0.22	1.001
380	0.48	0.21	0.10	0.21	1.028
390	0.49	0.23	0.09	0.19	1.016

Global $\chi^2 = 1.015$

Table A14 Global decay parameters for the A4/M.EcoRI/sinefungin complex.

Global lifetimes /ns

$\tau_1 = 0.15$ $\tau_2 = 0.49$ $\tau_3 = 2.69$ $\tau_4 = 9.64$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
370	0.55	0.30	0.06	0.09	1.029
380	0.53	0.31	0.07	0.09	1.010
390	0.51	0.33	0.06	0.10	1.025

Global $\chi^2 = 1.022$

Table A15 Global decay parameters for the A4/H235N/sinefungin ternary complex, the M.EcoRI mutant.

Global lifetimes /ns

$\tau_1 = 0.02$ $\tau_2 = 0.25$ $\tau_3 = 1.64$ $\tau_4 = 5.88$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
370	0.96	0.02	0.01	0.01	1.009
390	0.94	0.03	0.01	0.02	1.074

Global $\chi^2 = 1.042$

Table A16 Global decay parameters for the free AP1 duplex.

Lifetimes /ns

$\tau_1 = 0.11$ $\tau_2 = 0.40$ $\tau_3 = 4.54$

Emission λ /nm	A_1	A_2	A_3	χ^2
390	0.46	0.53	0.01	1.108

Table A17 Individual decay parameters for the free AP2 duplex.

Global lifetimes /ns

$\tau_1 = 0.09$ $\tau_2 = 0.62$ $\tau_3 = 2.86$ $\tau_4 = 7.57$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
380	0.44	0.23	0.20	0.13	1.050
390	0.45	0.23	0.19	0.13	1.022

Global $\chi^2 = 1.036$

Table A18 Global decay parameters for the AP1/M.EcoRV/SAH ternary complex.

Global lifetimes /ns

$\tau_1 = 0.07$ $\tau_2 = 0.86$ $\tau_3 = 3.16$ $\tau_4 = 9.32$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
370	0. 34	0.15	0.26	0.25	1.026
380	0.22	0.18	0.30	0.30	1.021
390	0.20	0.18	0.30	0.32	1.041
				Global χ^2	= 1.029

Table A19 Global decay parameters for the AP2/M.EcoRV/SAH ternary complex.

Global lifetimes /ns

$\tau_1 = 0.15$ $\tau_2 = 1.03$ $\tau_3 = 3.55$ $\tau_4 = 11.5$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
430	0.55	0.25	0.15	0.05	1.058
450	0.49	0.29	0.17	0.05	1.036
470	0.44	0.32	0.19	0.05	1.058
				Global χ^2	=1.051

Table A20 Global decay parameters for the APT(T) duplex excited with a 360-nm wavelength.

Global lifetimes /ns

$\tau_1 = 0.12$ $\tau_2 = 0.37$ $\tau_3 = 2.73$ $\tau_4 = 10.78$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
370	0.88	0.08	0.02	0.02	1.111
390	0.83	0.12	0.03	0.02	1.099
410	0.75	0.16	0.06	0.03	1.131
				Global χ^2	=1.114

Table A21 Global decay parameters for the APC(T) duplex excited with a 318-nm wavelength, reproduced with permission from R.K. Neely.

Global lifetimes /ns

$\tau_1 = 0.17$ $\tau_2 = 0.74$ $\tau_3 = 3.60$ $\tau_4 = 12.57$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
430	0.47	0.36	0.13	0.04	1.055
450	0.44	0.38	0.13	0.05	1.032
470	0.44	0.37	0.14	0.05	1.030

Global χ^2 =1.039

Table A22 Global decay parameters for the APC(T) duplex excited with a 360-nm excitation wavelength.

Global lifetimes /ns

$\tau_1 = 0.04$ $\tau_2 = 0.44$ $\tau_3 = 2.73$ $\tau_4 = 10.68$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
370	0.69	0.11	0.09	0.11	1.055
390	0.60	0.17	0.11	0.12	1.099
410	0.60	0.20	0.12	0.08	1.043

Global χ^2 = 1.066

Table A23 Global decay parameters for the GPG(T) duplex excited with a 318-nm wavelength, reproduced with permission from R.K. Neely.

Global lifetimes /ns

$\tau_1 = 0.19$ $\tau_2 = 0.97$ $\tau_3 = 3.64$ $\tau_4 = 9.17$

Emission λ /nm	A_1	A_2	A_3	A_4	Local χ^2
430	0.50	0.31	0.15	0.04	1.055
450	0.53	0.26	0.16	0.05	1.059
470	0.57	0.22	0.16	0.05	1.058

Global χ^2 = 1.057

Table A24 Global decay parameters for the GPG(T) duplex excited with a 360-nm excitation wavelength.

Appendix B

Lectures and Conferences

Postgraduate Lectures and Courses

Biophysical and chemical biology section seminars

School of Chemistry Colloquia

Word Intermediate

Reference Manager

EaStCHEM Academic Paper Writing Workshop

EaStCHEM Thesis Writing Workshop

Career Strategies for PhD Students

Effective Writing

Conferences and Meetings

7th International Meeting on Recognition Studies in Nucleic Acids (NACON VII)
Tapton Hall, University of Sheffield, UK, April 2007

10th Conference on Methods and Applications of Fluorescence (MAF 10)
Salzburg, Austria, September 2007

School of Chemistry Physical Chemistry Section Meetings,
Firbush Point, April 2005, 2006, 2007

Appendix C

Reprints of Publications

2-Aminopurine Flipped into the Active Site of the Adenine-Specific DNA Methyltransferase M.TaqI: Crystal Structures and Time-Resolved Fluorescence
Thomas Lenz, Eleanor Y. M. Bonnist, Goran Pljevaljčić, Robert K. Neely, David T. F. Dryden, Axel J. Scheidig, Anita C. Jones, and Elmar Weinhold,
J.Am.Chem. Soc. **2007**, *129*, 6240-6248