# Characterizing The N Terminus of the *Escherichia coli* DNA Repair Protein Vsr Endonuclease

Stavroula Monastiriakos

A Thesis

in

The Department

of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University

October 2002

## Abstract

### Characterizing The N Terminus of the *Escherichia coli* DNA Repair Protein Vsr Endonuclease

Stavroula K. Monastiriakos, Ph.D.
Concordia University, 2002

Spontaneous deamination of 5-methylcytosine in *Escherichia coli* and in human cells causes T/G mismatches. In *E. coli*, these mismatches are corrected by very short patch (VSP) repair, which requires the product of the *vsr* gene. Vsr nicks 5' of the mismatched T in a sequence dependent manner (C (T/G) WGG where W = A or T). DNA polymerase I removes the T, along with several bases on its 3' side, and resynthesizes the DNA, restoring the correct C.G base pair. Overexpression of *vsr* is mutagenic, causing a genome-wide increase in frameshifts and base substitutions. This mutational spectrum is consistent with saturation of mismatch repair.

The published structure of Vsr bound to DNA shows that the N terminus penetrates the minor groove of the DNA, possibly acting as a clamp to aid binding, while three aromatic residues intercalate the DNA from the major groove side. We have analysed five Vsr mutants that alter these above-mentioned areas that bind to the DNA: two N terminus deletion mutants, Δ14 and Δ19, along with three alanine amino acid substitutions for Phe67, Trp68 and Trp86. The alanine amino acid substitutions for F67, W68 and W86 are not active either *in vivo* or *in vitro*. The loss of activity may be due to conformational changes that alter the metal binding capabilities required for DNA binding.

Both N terminus mutants show decreased activity *in vitro* and *in vivo*, and decreased DNA binding in bandshift and fluorescence assays. In contrast to the wildtype Vsr, MutL does not stimulate the endonuclease activity of these mutants, indicating that an N terminus function is necessary for endonucleolytic stimulation. While the Δ14 mutant is half as mutagenic as Vsr when overproduced in *vivo*, the Δ19 mutant is considerably less mutagenic than the wild-type protein. The two N terminus mutants are no longer growth phase dependent indicating a requirement for the N terminus for growth phase regulation. Furthermore, overexpression of MutL *in vivo* was found to stabilize Vsr levels in log phase cells.

**Acknowledgements**

I would like to thank my supervisor, Dr. Claire Cupples, for her support and encouragement over the past four years. You really helped me see the biologist's point of view and look past the test tube. You always made time for me, and were willing to listen – thank you.

A special thank you goes out to my committee members, Dr. Justin Powlowski and Dr. Joanne Turnbull. Dr. Powlowski, in my eyes you are a biochemistry textbook. You have been an invaluable source of information to me over the entire course of my graduate studies. Dr. Turnbull, you have always extended yourself to me, not only did you help me with analysing my data, but you were always there when I needed advice. Both of you went above and beyond your call of duty as committee members. I really appreciate it – thank you so very much.

I would also like to extend gratitude to my external examiners, Dr. Robert Boushel, and Dr. Iain Lambert. I realize how precious your time is, especially between grant proposals, and I appreciate that you are taking the time to be on my examining committee.

I have been very fortunate with my time spent at Concordia, and made many good friends; many of them have been co-workers. I especially want to thank Dr. Nathalie Agar. Nathalie, not only are you a good friend, and no matter how much you had on your plate, there was always room for me. I appreciate you showing me how to do fourth

derivatives and how to do data simulations. Dr. Elisabeth Cadieux, you were an excellent role model when I started my graduate studies, and were always generous with your time; Derek Bell, not only are you my computer expert, but you made my time in the lab fly – we had a lot of good laughs together.

Last but not least, I would like to thank my family; my parents for their love and support, my cousins, especially Niki Marinos, who has had to sit through many painstaking science conversations for my benefit. And the biggest thank you goes to my husband Giovanni, who for four years worked two full-time jobs, just to put me through school, while enduring my different temperaments. You are an angel.

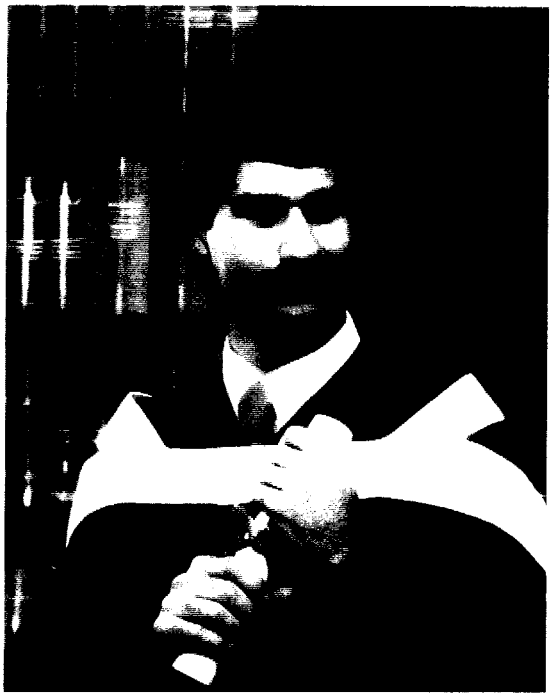
## Dedication

To my brother and best friend, Peter,

It has been a little over five years, and I see your smile and hear your laughter. Although our time together was cut short, I am blessed to have had you in my life; you made us all so proud. Until the day we meet again, in my heart is where I keep you friend; fond memories give me the strength I need to proceed. When this life is over, I know I will see your face.

Αδελφούλι μου γλυκέ, μακάρι να σε είχα εδώ.

Αιώνια σου η μνήμη.



## Table of Contents

## List of Figures

*List of Tables*

## Contribution of Authors

All the data presented in this thesis is the work of the author, except for the following: Gina Macintyre was the first one to show that MutL stimulates the endonuclease activity of Vsr. Marina Siponen constructed the F67A and W68A mutants. She analysed these mutants *in vivo* for VSP repair and mutagenesis, and *in vitro* for endonuclease activity. Stéphane Chappaz constructed the W86A mutant. Marina Siponen characterized this mutant *in vivo* for VSP repair and mutagenesis.

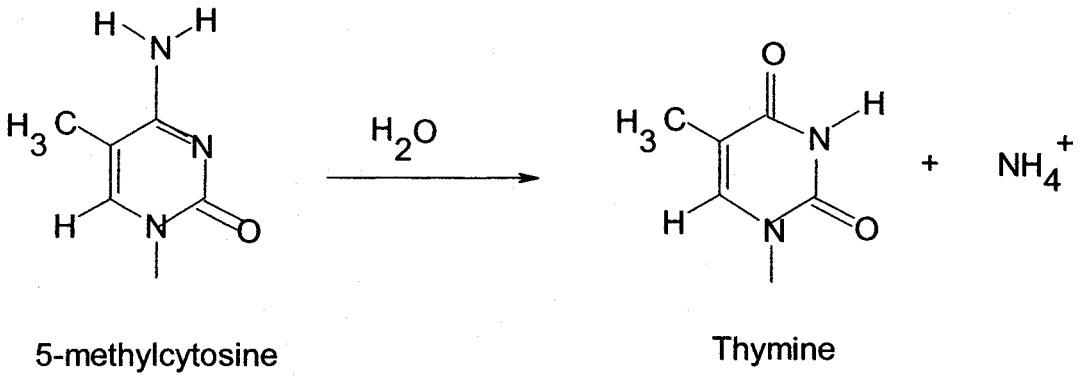
## Introduction

*Repair of DNA lesions*

The long-term survival of a species may be enhanced by genetic changes but the survival of the individual demands genetic stability (Alberts, 1994). Living cells require DNA integrity for the correct function of thousands of proteins, any of which may be damaged by a mutation at many different sites of the corresponding gene. When damage changes the coding properties of a gene, the organism may not survive without mechanisms to repair damaged DNA. Unrepaired DNA or inaccurate repair could lead to cell death, or to the generation of cells with tumorigenic potential. The evolutionary response to DNA damage has been the development of DNA repair systems, many of which are highly specific for the types of DNA lesions that they recognize and repair. In addition to mediating repair, these systems coordinate repair to other cellular processes such as transcription, DNA replication, and cell cycle progression.

Damage to DNA arises from a variety of environmental or intrinsic factors. Radiation, and mutagenic chemicals are environmental factors that result in altered bases, strand breaks, deletions, and insertions into the DNA sequence. Intrinsic factors arise from the chemical nature of the DNA itself or from enzymatic processes that occur in the cell. For example, during replication DNA polymerase may introduce mistakes into the DNA sequence. Also, the thermal decomposition of nucleosides can lead to the introduction of abasic DNA sites. Likewise, spontaneous deamination of bases can lead to base mismatches.

Many organisms, from enteric bacteria to multicellular eucaryotes, methylate selected cytosines in their genomes. Methylation has a variety of important biological functions; in prokaryotes it coordinates DNA replication, directs methyl-directed mismatch repair, and discriminates between self and non-self DNA (Reviewed in Neidhardt, 1993). In eucaryotes, it is responsible for normal development, gene repression, X chromosome inactivation, and replication timing (Reviewed in Jeltsch, 2002). However 5-methylcytosines may, under physiological conditions, spontaneously deaminate to thymine (Figure 1). In *Escherichia coli* and in human cells, this hydrolytic reaction leads to T/G mismatches. T/G mismatches are premutagenic lesions that consist of natural DNA constituents and are devoid of an obvious marker for the damaged DNA strand. When left unrepaired, T/G mismatches result in C.G to T.A transition mutations (Figure 2).




5-methylcytosine Thymine

**Figure 1: The Spontaneous Deamination of 5-Methylcytosine to Thymine**

In the presence of water, 5-methylcytosine may deaminate to form thymine, which would pair with a G on the opposite strand, forming a mismatched base pair.

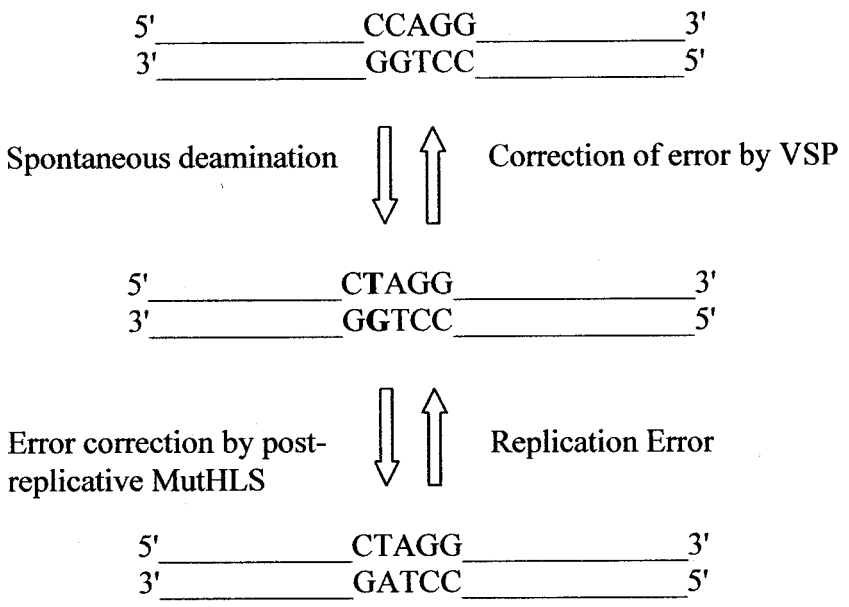



**Figure 2: Competition Between Methyl-Directed Mismatch Repair and VSP Repair**

If a methylated C residue spontaneously deaminates to form a thymine (top structure), the error is corrected by VSP repair. However, if a G residue is misincorporated opposite a T residue as indicated (middle structure), and escapes proofreading, the T/G mismatch is processed by either methyl-directed mismatch repair (MutHLS) or by the VSP repair pathway. Here, both pathways act in opposite directions. An increase in VSP repair caused by overexpression of Vsr or a decrease in mismatch repair should increase the yield of the top DNA structure. (Figure modified from Glasner *et al.,* 1995 and Doiron *et al.,* 1996).

*Very Short Patch Repair in Escherichia coli.*

In *Escherichia coli*, deamination events leading to T/G mismatches are corrected by very short patch (VSP) repair, which requires Vsr, the product of the *vsr* gene. The Vsr endonuclease, (Mr 18000 kDa) cleaves the phosphodiester bond in double stranded DNA 5' of the mismatched T, leaving 5' phosphate and 3' hydroxyl termini. The incision is strand-specific and mismatch dependent at C (T/G) WGG where W = A or T (Henneke *et al.,* 1991), a deaminated target site for Dcm, the sole DNA (cytosine-5) methyltransferase of *E. coli* K-12. It is not surprising that T/G mismatches arising within a Dcm recognition sequence are preferred substrates for Vsr endonuclease. DNA polymerase I, with its 5' to 3' exonuclease activity, removes the T along with several bases to its 3' side, and resynthesizes the DNA, restoring the correct C.G base pair. DNA ligase completes VSP repair by resealing the nick left by the endonuclease.

*VSP Repair and Methyl-Directed Mismatch Repair: Competing At CCAGG Sites*

T/G mismatches in the DNA sequence at Vsr recognition sites (C (T/G) WGG) are also potential substrates for repair by methyl-directed mismatch repair (Figure 2). If a T/G mismatch arises from misincorporation of a G residue opposite a T residue in the template DNA strand, correction by VSP repair would result in a T.A to C.G transition mutation. Published evidence suggests that VSP repair out-competes methyl-directed mismatch repair in correcting T/G lesions (Zell and Fritz, 1987, Doiron *et al.,* 1997). Dominance of VSP repair over methyl-directed mismatch repair might account for the over-representation of CCWGG sequences in *E. coli,* and the under-representation of CTWGG (W = A or T) (Bhagwat and McClelland, 1992, Merkl *et al.,* 1992).

Over-expression of *vsr* is highly mutagenic, causing a genome-wide increase in frameshifts and base substitutions (Doiron *et al.,* 1996). This mutational spectrum is consistent with saturation of mismatch repair, which corrects DNA errors that arise during chromosomal replication, and discourages recombination between substantially diverged DNA sequences. Initiation of mismatch repair involves the action of several proteins, MutL, MutS and MutH (discussed in more detail in a subsequent section). MutL has been shown to interact with Vsr, and to catalytically stimulate Vsr's ability to bind substrate DNA (Drotschmann *et al.,* 1998) and Vsr endonuclease activity (Macintyre, unpublished). The mechanism whereby MutL stimulates Vsr binding and endonuclease activity remains unclear. A physical interaction *in vivo* has been established between Vsr and MutL that inhibits MutL dimerization, and abolishes methyl-directed repair MutL interactions (Mansour *et al.,* 2001). It is possible that the interaction between Vsr and MutL is necessary for Vsr-stimulated mutagenesis.

*Growth Phase Regulation of Vsr*

The gene for the cytosine methyltransferase, *dcm*, and *vsr,* have an unusual chromosomal arrangement. Both genes lie at 43 minutes on the chromosomal map of *E. coli*, with the 5' end of *vsr* overlapping the 3' end of *dcm* in a plus one reading frame (Sohail *et al.,* 1990). Vsr and Dcm are thought to be cotranscribed from a promoter 5' of *dcm* as a single mRNA, with translation of the endonuclease being dependent on translation of the methyltransferase (Dar *et al.*, 1993). This coupled transcription mechanism would assure the presence of DNA repair specifically targeted at lesions caused by 5-methylcytosine

deamination. However despite VSP repair, 5-methylcytocines remain hot spots for mutation (Coulondre *et al.,* 1978).

Vsr endonuclease levels are growth phase dependent (Macintyre *et al.,* 1999). If stationary-phase dependence of Vsr were transcriptional, one would expect Dcm production to follow the same pattern as Vsr. This is not the case as cellular levels of Dcm are independent of growth phase, suggesting that regulation of Vsr is post-transcriptional (Macintyre *et al.,* 1999). Decreased production of Vsr during log phase could result in sub-optimal VSP repair and thus contribute to the mutability of 5-methylcytosines. However, keeping levels of Vsr low during periods of DNA replication would minimize mutagenesis resulting from the saturation of mismatch repair. In effect the cell may be sacrificing the efficiency of VSP repair to alleviate the stress that would be caused by saturation of mismatch repair (Macintyre *et al.,* 1999).

*Crystallization of Vsr Endonuclease – A Type II Restriction Endonuclease*

The work reported in this thesis carries on the biochemical characterization of Vsr endonuclease and several mutant Vsr proteins. Therefore it is appropriate to review the structure of Vsr, a member of the Type II restriction endonuclease family, and to summarize what is known about the biochemistry of Vsr endonuclease.

Initial crystallization of Vsr without DNA required the use of an N terminus truncation (Tsutakawa *et al.*, 1999a). Proteolysis with trypsin and chymotrypsin defined a stable core domain in the endonuclease, suggesting the N terminus contains an exposed turn

around residues fourteen and fifteen. Comparison of the one dimensional NMR spectra of the intact protein with a Δ20 fragment indicates that the main core completely retains its tertiary structure (Tsutakawa *et al.,* 1999a).

The crystal structure of the truncated protein was determined to 1.8 Å by the multiple isomorphous replacement method. Its overall topology resembles that of members of the Type II restriction endonuclease family. Although Vsr is monomeric, Type II restriction endonucleases are homodimeric enzymes that require magnesium to catalyse the hydrolysis of DNA. Type II restriction endonucleases generally recognize DNA sequences varying between four to eight base pairs, and introduce either staggered or blunt-ended double stranded breaks into DNA. The canonical restriction endonuclease fold is characterized by a five-stranded β-sheet and two α-helices (Aggrawal, 1995). Thousands of Type II restriction endonucleases are known, and all have very different sequences; yet all of the structures determined to date reveal a common fold in their catalytic domains. The common fold suggests divergent evolution from a common ancestor to acquire distinct biological functions (Burgess-Hickman *et al.,* 2000).

Type II restriction endonucleases typically have rigid active sites that are deeply located. The catalytic motif generally involves several acidic and one basic residue (Aggrawal, 1995). The most conserved residue is an aspartate always located on the second strand of the five-stranded β-sheet. The aspartate is usually involved in coordinating two metal ions at the active site. The other essential catalytic residues - usually one acidic and one

basic - show some variation, and the structural elements from which these residues arise are not always the same.

In most hydrolytic nucleases, active site residues are required either to interact directly with the water molecule that attacks the scissile phosphate of the DNA backbone, or to precisely orient two metal ions with respect to the phosphodiester backbone and the attacking water molecule(s) (Aggrawal, 1995). The requirement to bind metals or to abstract protons from a water molecule to generate a nucleophile to attack the P-O3' bond, limits the choice of amino acid residues that can function at the active site; histidine and cysteine residues have been implicated in both metal binding and in the activation of water molecules. Residues with carboxylic acid side-chains such as aspartate and glutamate are commonly associated with divalent cation binding. Acidic residues may also act to stabilize the transition state by accepting a proton from a transiently positively charged histidine residue (Saenger, 1991).

Vsr endonuclease is composed of a single domain β-sheet flanked by three α-helices (Figure 3). A deep cleft on one side of the protein contains a structural zinc-binding module; Cys66, His71, Cys73 and Cys117, are responsible for zinc co-ordination. Structural elements in the catalytic site, such as several acidic residues (Asp51, Glu25, Asp97), are conserved with the Type II restriction endonuclease family; however the identification of a critical histidine (His69) and evidence for an active site metal-binding coordination not found in restriction endonucleases, indicates a distinct catalytic mechanism for Vsr endonuclease.

*Crystallization of Vsr Bound to DNA Substrate*

Despite exhaustive screening, full-length Vsr endonuclease could not be crystallised in the absence of its DNA substrate (Morikawa and Shirakawa, 2000). Co-crystallization was achieved in the presence of magnesium along with a twelve base pair DNA duplex containing a T/G mismatch, within a Dcm recognition sequence. The crystal structure was resolved to 2.3 Å by isomorphous molecular replacement using the truncated Vsr structure as a model (Figure 4). The co-crystal is essentially identical to that of the previously solved structure with small shifts to optimise DNA contacts. The N terminus of the endonuclease forms a short alpha helix (Lys7-Arg15), followed by a loop that ends at Ile17. There is a $110^0$ rotation of the Trp68 indole ring out of the plane of the main protein surface, which is not present in the initial truncated Vsr crystal structure. With the presence of a cleaved DNA intermediate in the active center, the structure of the co-crystal allows some insight into the *novel* catalytic and recognition mechanisms for endonuclease activity.

Vsr contains a highly basic interface that contacts with the DNA duplex. Within this interface lies an acidic patch that corresponds to the catalytic centre of Vsr, and is necessary for divalent metal cation binding. Metal ion co-ordination in this site alleviates some of the repulsive forces formed between the negatively-charged active centre and negatively charged phosphate backbone (Morikawa and Sharikawa, 2000). A histidine (H69) and two water clusters catalyse phosphodiester cleavage and Vsr recognises the T/G mispair as a wobble base pair (Tsutakawa *et al.*, 1998b).

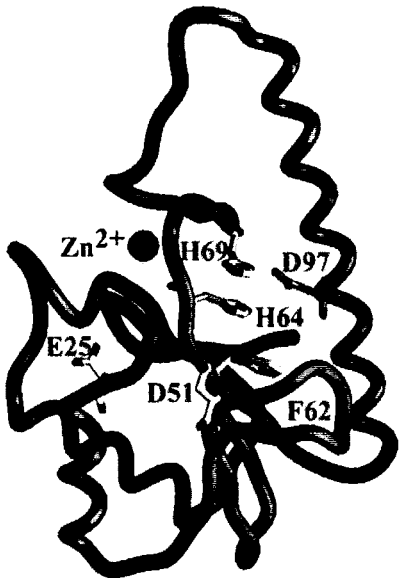



**Figure 3: The Structure of Truncated Vsr Endonuclease**

Crystal studies of an N terminus proteolysed (Δ20) Vsr endonuclease (1.8 Å), reveal a central β-sheet with α-helices on either side. The main chain of Vsr is shown in worm representation. Scanning alanine mutagenesis of conserved acidic residues identified Asp51, Glu25 and Asp97 as important for catalysis. His69 is absolutely required for activity, while His64 is important for activity. Colouring of white, blue and red stand for carbon, nitrogen and oxygen atoms. The zinc atom, shown as a purple sphere, plays a structural role. (This figure was taken from Tsutakawa and Morikawa, 2001).


**Figure 4: Co-Crystal Structure of Vsr Endonuclease in The Presence of DNA**

Crystal structure of the full-length Vsr (2.3 Å), bound to a twelve base pair DNA duplex. The main chain shown as a worm model, clamps the DNA with an N-terminus helical arm. Three aromatic residues, Phe67, Trp68 and Trp86 (green) intercalate into the DNA duplex from the major groove side of the TG mismatch. Catalytically essential residues, Asp51 and His69, (magenta) are proximal to the cleaved DNA termini. The major groove is flattened out, while the minor groove is expanded accommodating N-terminus binding within the minor groove side of the DNA. Blue and green spheres show the location of the zinc and magnesium atoms, respectively. DNA atoms are coloured as in Figure 3 with the addition of phosphate as yellow. (This figure was taken from Tsutakawa and Morikawa, 2001).

*Mismatch Recognition Normally Requires Base-Flipping*

The mode of mismatch recognition for Vsr is novel because it does not require base-flipping but rather base-pair disruption between the mispaired T/G and an adjacent AT base pair. All other DNA repair enzymes crystallised in the Type II restriction endonuclease family, recognize mismatches through base flipping. For example, T4 endonuclease V (T4 endo V) scans nontarget sequences to search for lesions in DNA duplexes, catalysing the excision of pyrimidine dimers produced within the DNA by UV irradiation (Dodson and Lloyd, 1989). T4 endo V carries two distinct catalytic activities - cleavage of the glycosyl bond at the 5' side of the pyrimidine dimer and subsequent scission of the phosphodiester bond at the 3' position of an abasic site.

T4 endo V has been crystallized, and despite the two catalytic activities, it possesses a single compact domain, consisting of three α-helices and connecting loops (Morikawa *et al.,* 1989). The electrostatic surface of the protein displays a 50 by 40 $Å^2$ region, which exceeds the diameter of the DNA duplex and appears to be responsible for scanning nontarget DNA sequences (Dodson and Lloyd, 1989). The DNA duplex bound to the concave basic surface of T4 endo V is sharply kinked ($60^0$), and in the crystal structure, the adenine base complementary to the 5' side of the photo-dimer is completely flipped out of the DNA duplex interior and accommodated into a cavity on T4 endo V (Morikawa *et al.,* 1989). While scanning DNA duplexes through the positively-charged concave surface, T4 endo V may pause at pyrimidine dimer sites, possibly because of an electrostatic disturbance surrounding them (Morikawa and Shirikawa, 2000). The proposed base-flipping model for T4 endo V DNA scanning assumes a direct coupling

between the kinked DNA duplex and the flipped-out base. One of the functional roles for base-flipping would be to alleviate the local tension generated by the kink induced by T4 endo V binding to the DNA (Morikawa and Shirikawa, 2000). However it is thought that base-flipping is required for several functional reasons – mismatch recognition, and generating a separate pocket for catalytic space (Morikawa and Shirkawa, 2000).

There are structural elements contributing to T/G mismatch recognition in Vsr endonuclease that preclude base flipping. Both the T and the G in the mismatch form direct hydrogen bonds with Asn93 and Met14, along with water-mediated hydrogen bond residues formed along the N terminus alpha helix. Steric forces also play a crucial role in recognition; namely, a slightly deformed DNA backbone with respect to a normal CG base pair, along with extensive protein DNA contacts from both the major and minor grooves of the DNA. All five nucleotides in the recognition sequence form direct and indirect polar interactions with protein atoms. Taken together, Vsr does not need to flip out bases to effect repair, as recognition of T/G mismatches relies on wobble base pairing and there is no need for generating catalytic space.

*T/G Mismatch Recognition by Vsr Endonuclease*

A distinguishing feature of the co-crystal Vsr-DNA complex is the intercalation of three aromatic residues, Phe67, Trp68, and Trp86, from the major groove side of the DNA, next to the T/G mismatch, on the side opposite to the cleavage site. These residues may play a major role in the binding and recognition of the mismatch (Figure 5). Phe67 and Trp68 stack with the AT base pair and with the mispaired thymidine, while Trp86 stacks

with two sequential sugar rings on the strand opposite the cleaved strand. Therefore, the cleaved T remains on one side of the intercalating residues and is separated from its formerly covalently linked strand. An amino terminal α-helix is accommodated into the expanded minor groove of the DNA. As if pinching the DNA, Met14 and Ile17 enter from the minor groove side and are responsible for mediating contact between the TG mismatched base and a neighbouring guanine through peptide atoms, leaving only enough space for the phosphate backbone to pass on either side. The extensive protein/DNA contacts cause dramatic changes in the DNA geometry; specifically an overall bend of 44° is produced in the DNA. Despite this conformational strain, all the above-mentioned base pairs remain intact within the DNA duplex (Tsutakawa *et al.*, 1999b).

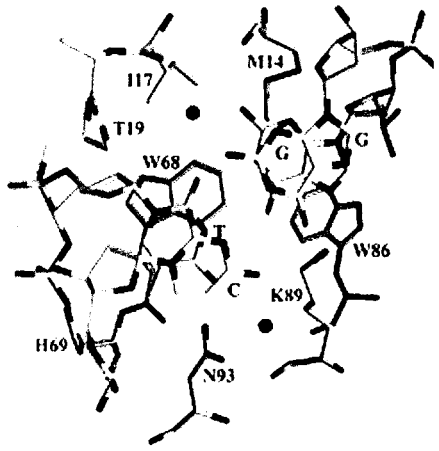



**Figure 5: Recognition of The T/G Mismatch**

Recognition occurs through H-bond interactions, surface complementarities with the shifted wobble base conformation, and intercalation with aromatic residues. Lys89 and the carbonyl of Met14 recognize the guanine base from either side, while Asn93 recognizes the lowered thymine base. The main chain of the N-terminus residues Ile17 to Thr19 push down onto the mismatched thymidine, leaving an interface in which only the TG wobble base pair can be accommodated. Trp68 stacks directly onto the mismatched thymine base, while Trp86 wedges between the sugars on the opposite side of the DNA duplex. The side chains of Ile17 and Met14 meet those of Trp68 and Trp86 to completely prevent base pair stacking from one side of the TG mismatch. Atoms are coloured as previously indicated in Figures 3 and 4, with sulphur coloured in green. (This figure was taken from Tsutakawa and Morikawa, 2001).

*Biochemical Characterization of Vsr Endonuclease*

The crystallographers limited their biochemical analysis of Vsr to identifying critical metal-binding and catalytic residues. Through the use of an endonuclease assay they show that intact Vsr, and proteolysed Δ14, Δ15 and Δ20 Vsr cleave T/G containing DNA (Tsutakawa *et al.,* 1999a). The relative activities of the proteolysed Vsr proteins are not specified. Vsr endonuclease activity is dependent on $Mg^{+2}$ or $Mn^{+2}$, with activity being quenched upon the addition of EDTA to endonuclease assays (Tsutakawa *et al.,* 1999a).

Fox *et al.,* 2000, found it necessary to include BSA in single turnover endonuclease assays to ensure more than 20% cleavage. No difference occurs upon the addition of a reducing agent such as DTT. The first-order rate constant for Vsr was determined to be 0.23 +/- 0.04 $min^{-1}$. An attempt was made to estimate a dissociation constant for Vsr through gel shift analysis, however enzyme concentrations in excess of 1 μM were required to produce a retarded DNA species, consistent with weak DNA binding, as also independently determined by another group (Drotschmann *et al.,* 1998). Removal of the 5-methyl group from the thymine and the 2 amino group from the guanine causes a 60% reduction in activity, indicating that both these groups are recognized by the endonuclease, and that Vsr has a preference for a hemi-methylated DNA complex (Fox *et al.,* 2000, Turner and Connolly, 2000).

Turner and Connolly (2000) also did single turnover kinetics for Vsr and determined a $k_{cat}$ of 1.4 +/- 0.1 $min^{-1}$ for a hemi-methylated substrate, and 0.39 +/- 0.02 $min^{-1}$ for a non-methylated substrate. By doing gel retardation assays in the presence of $Ca^{+2}$, which

does not support endonuclease activity, they determined dissociation constants for fully active Vsr to a hemi-methylated and non-methylated substrate to be 82 +/- 12 nM and 255 +/- 33 nM, respectively. No bandshift was observed for an oligonucleotide that lacked a T/G mismatch or did not possess a Vsr recognition sequence.

Surprisingly, an active site titration of Vsr indicated that less than 1% of the endonuclease was able to bind to the substrate. This result was confirmed with kinetic burst analysis that indicated an initial burst, followed by a slower steady state turnover. Since MutL had been shown to stimulate Vsr binding to DNA (Drotschmann *et al.*, 1998), it was hypothesised that MutL may function to convert Vsr from an inactive to an active form (Turner and Connolly, 2000).

Bacterial Vsr contains a highly conserved methionine 14 amino acid residues from the amino terminus methionine of the full-length protein (Choi and Leach, 1994, Lieb and Bhagwat, 1996). Putting a stop codon in *vsr* between these two methionines abolishes VSP activity, making it seem that translation initiation from the second start site results in a non-functional protein (Dar and Bhagwat, 1993). Translation from this methionine would result in a truncated Vsr. Therefore, Turner and Connolly, (2000), made a Δ14 mutant to show that the 1% activity in their Vsr preparations was not due to contamination of an inactive full length protein with an active Vsr truncation. The truncated mutant was unable to shift DNA in bandshift assays, but was able to hydrolyse a hemi-methylated substrate with a $k_{cat}$ of 0.62 $min^{-1}$.

Vsr can also bind and hydrolyse sequences that differ from *dcm* target sites by one base pair in any of the four bases surrounding the mismatched T ("star substrates"), with reduced binding and slower turnover kinetics (specificity constants 0.01- 4 % of wildtype sequence) (Gonzalez-Nicieza *et al.,* 2001). Vsr has a lower selectivity for its target site than other Type II restriction endonucleases (specificity constants $10^{-5} - 10^{-8}$ % of those observed with wildtype sequences). Lowered specificity may arise from intercalation of the three aromatic amino acids into the DNA helix; the contacts between protein and bases flanking the mismatch provide a relatively limited degree of selectivity for the target site. Dcm may also methylate non-canonical substrates, so the physiological significance of Vsr's relaxed specificity can be to repair deamination events occuring at *dcm* star sequences (Gonzalez – Nicieza *et al.,* 2001).

*DNA Glycosylases*

Vsr homologs are limited to several bacterial species (Lieb and Bhagwat, 1996). Other organisms utilize a sequence-dependent glycosylase to initiate repair at T/G mismatches. When the G is in the CpG sequence, thymine DNA glycosylase (TDG) is required for repair (Wiebauer and Jiricny, 1990). The rate of action of TGD is strongly affected by the base pair 5' to the mismatched guanine (Neddermann and Jiricny, 1994).

In vertebrate genomes, CpG islands are associated with regulatory regions of genes, and 5-methylcytosine mediates gene silencing (reviewed in Ng and Bird, 1999). CpG islands are susceptible to mutation, contributing to almost 30% of all germline mutations (Cooper and Youssoufian, 1988), and 23% of all human hereditary diseases have a G to

A transition mutation at CpG sites (Krawczak *et al.*, 1998). Uncorrected G/T and G/U mispairs from deamination of 5-methylcytosine and cytosine may alter both coding and regulatory sequences – repair being crucial for maintaining stable patterns of gene expression. TDG recognises and excises the mispaired thymine and uracil moieties from the deoxyribose or ribose ring, generating an abasic site (Reviewed by Schärer and Jiricny, 2001). The creation of a single nucleotide gap is accomplished by apurinic/apyrimidinic endonuclease I (APE) that cleaves the phosphodiester bond 5' of the missing base. DNA polymerase β replaces the missing nucleotide and completion of the repair process occurs when DNA ligase reseals the DNA.

There is currently no crystal structure of TDG, however there is a structure of uracil DNA glycosylase (UDG), the bacterial homologue of thymine DNA glycosylase (Barrett *et al.*, 1998). The structure was solved with the glycosylase bound to DNA containing an abasic site opposite a guanine. The protein inserts a "wedge" of amino acids (GLSR), into the DNA to push the uracil out of the helix into a separate binding pocket where excision occurs (Barrett *et al.*, 1998). Although the uracil is excised in the crystal structure, an abasic sugar remains flipped out of the helix while the wedge occupies the gap left in the DNA by the excised uracil. Although the wedge sequence in TDG, (SSAR), is not the same as in UDG, TGD is expected to insert a "wedge" into the DNA, flip-out the thymine and make hydrogen bonds with the mismatched guanine. The binding pocket for the mismatched thymine is proposed to contain a tyrosine residue that can stack with the thymine.

The rate of action of TDG is limited by its extremely low $k_{cat}$ of 0.026 $min^{-1}$, along with the fact that it binds to its product with a half-life of dissociation of several hours (Waters *et al.,* 1999). Therefore unless the glycosylase is displaced, each molecule of TGD can remove only one thymine. It has been shown that APE, which catalyses the second step in the repair pathway, displaces TDG from the DNA lesion (Waters, *et al.,* 1999), thereby co-ordinating repair and ensuring that the nick is not re-sealed by DNA ligase.

*Methyl-Directed Mismatch Repair*

A portion of this thesis deals with the interactions between MutL and Vsr; thus it is appropriate to review methyl-directed mismatch repair.

DNA replication is a highly accurate process resulting in less than one error per $10^{10}$ nucleotides synthesized (Kunkel, 1992). However, the fidelity of base insertion by replicative polymerases is lower. *In vitro* measurements have shown that DNA polymerases insert one incorrect nucleotide for every $10^4$ to $10^5$ correct one (Alberts, 1994). The fidelity of DNA replication is therefore attributed to DNA polymerase proofreading activity and post-replicative mismatch repair (MMR) correction. After polymerase editing carried out by the 3' to 5' exonuclease activity, MMR corrects any remaining errors. In order to do so, MMR must be able to discriminate between the parental strand and the daughter strand that contains the incorrect nucleotide. Repair then proceeds by the degradation of a long stretch of DNA, followed by DNA resynthesis, and finally DNA ligation.

The best-characterized MMR pathway is the methyl directed mismatch repair system of *Escherichia coli* (Figure 6). MMR initiation requires three proteins, MutH, MutL and MutS; MutS for mismatch recognition, MutH for introduction of a nick in the newly synthesized DNA strand, and MutL for mediating interactions between MutH and MutS. MMR recognises and repairs small insertion/deletion loops up to four nucleotides in length and all base-base mismatches except C/C. Strand discrimination is achieved through recognition of the DNA strand that is undermethylated in the GATC sequence by MutH.

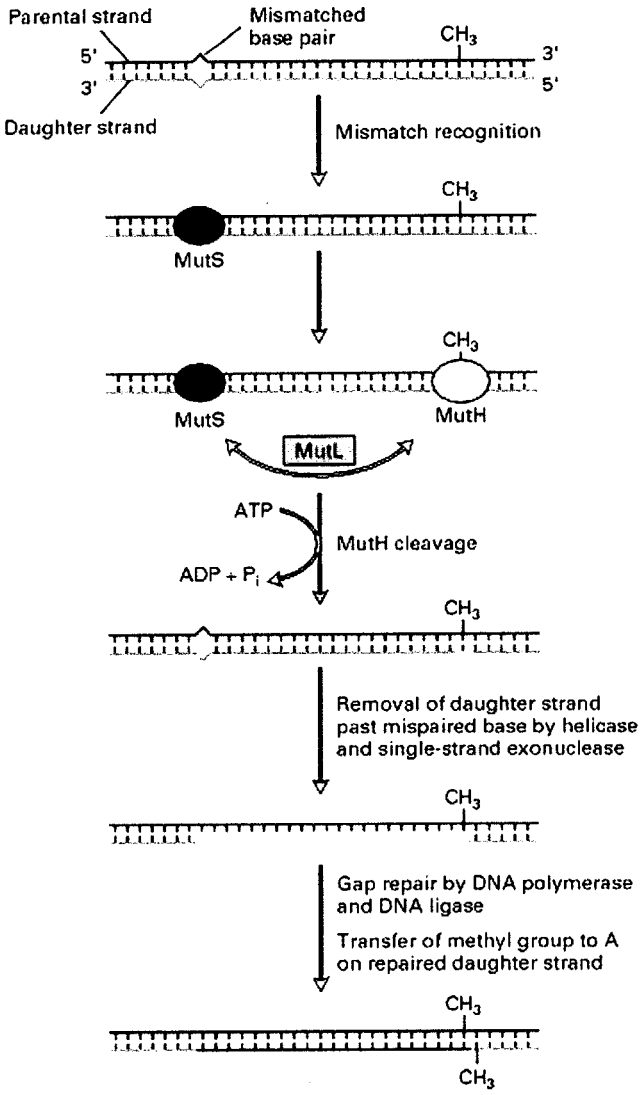



**Figure 6: Methyl -Directed Mismatch Repair**

MutH binds specifically to a hemimethylated GATC sequence, and MutS binds to the site of a mismatch. Binding of MutL protein simultaneously to MutS and to a nearby MutH activates the endonuclease activity of MutH, which then cuts the unmethylated (daughter) strand in the GATC sequence. A stretch of the daughter strand containing the mispaired base is excised followed by gap repair, ligation and methylation of the daughter strand. (Adapted from Kolodner, 1996)

*Recognition of a Mismatch - MutS*

MMR is initiated by binding of a MutS homodimer to a mismatch. It is poorly understood how MutS recognises many types of structurally different alterations in the DNA. Insights have become possible due to MutS having been co-crystallized in the presence of a G/T mismatch (Lamers *et al.*, 2000), and with heteroduplex DNA containing one unpaired thymidine (Obmolova *et al.*, 2000). Although two different mismatches were used, the resolved crystal structures are similar. In both crystal structures, binding of a MutS homodimer to the mismatch induces a 60° kink in the DNA, which leads to a widening of the minor groove, while narrowing the major groove on the opposite side.

In the absence of DNA, MutS might adopt an open "U-shaped" conformation, since loading onto the DNA requires an open protein structure. In the presence of DNA, two MutS subunits come together on either side of the DNA, forming an oval disk that contains two large holes in the centre. The DNA passes only through the upper channel of the disk; the asymmetry is due to interaction with the mismatch. Interaction with the mismatch is mediated by a highly conserved N terminus motif GXFY(E) in MutS. The phenylalanine, in the middle of the motif, is inserted into the DNA double helix and stacks with the mispaired or unpaired thymine present in both DNA substrates. MutS senses mismatches by inserting the phenylalanine residue into the double helix, trying to induce a kink at the site of a mismatch. Only at sites where the DNA can be bent to a certain extent will the MutS dimer attain an optimal conformation, suggesting that recognition relies on structural changes in the DNA and protein dimer.

*Cleavage at a Hemimethylated GATC Site - MutH*

After recognition of a mismatch by MutS, downstream repair events are initiated. The strand discrimination signal is MutH, bound to a hemimethylated (GATC) site (Modrich and Fang, 1991). The MutS dimer has two ATPase sites at the bottom of the disk. When ADP is bound, the protein dimer forms a stable association with the mismatch (Lamers *et al.,* 2000). Exchanging ADP for ATP leads to MutS being released from the mismatch in search of enzymes required for the incision step (Lamers *et al.,* 2000). MutS translocation, along with MutL, also promotes DNA loop formation providing a simple mechanism for interaction of the mismatch and GATC sites (Allen *et al.,* 1997).

MutH is a latent sequence - and methylation-specific endonuclease. *Escherichia coli* MutH is similar to Vsr in that it is a magnesium-dependent, monomeric endonuclease (Welsh *et al.,* 1987). The MutH structure closely resembles that of Type II restriction endonucleases (Ban and Yang, 1998). MutH resembles a clamp with two arms separated by a large cleft. All structures except three loops surrounding the cleft are well-ordered in MutH. The two arms share a hydrophobic interface, and are connected by three polypeptide linkers. Five residues distributed between the two arms are thought to make the active site, Glu56, Asp70, Glu77, Lys79 and Lys116. The interface between the arms is flexible; therefore the active site is also flexible, capable of adopting different conformations.

When activated by MutS and MutL, MutH cleaves 5' of an unmethylated (GATC) sequence in a hemimethylated duplex, thereby targeting the daughter strand (Welsh *et al.,*

1987). Even in the presence of a mismatch and ATP, MutS alone cannot activate MutH. MutL is required to mediate the interaction between MutS and MutH (Lu *et al.,* 1984) stimulating its endonuclease activity (Hall and Matson, 1999). It was suggested that the fifty-fold activation in endonuclease activity occurs via interaction with a C terminal alpha helix molecular lever in MutH that directs opening and closing of its two subdomains (Ban and Yang, 1998).

*Molecular Switch For Mismatch Repair -MutL*

MutL and its homologues form a large protein family, which share extensive sequence homology in the N terminus 300-400 residues, while the C terminal halves of their sequences are very diverse. All MutL homologues form homo or heterodimers via the C terminal region (Ban and Yang, 1998b). It was initially thought that MutL did not possess any enzymatic activity (Modrich and Lahue, 1996). However, when the crystal structure of the N terminus portion of MutL was solved (LN40), it became apparent that it was similar to the ATPase fragment of DNA gyrase (Wigley *et al.,* 1991). MutL is part of the GHKL ATPase superfamily, characterised by a Bergerat fold in the ATP binding pockets instead of the more common Walker motifs. However, GHKL ATPases are characterised by four sequence motifs, and all the residues required for ATP binding and hydrolysis are conserved in the MutL family (Ban and Yang, 1998a).

LN40 (crystal structure of a 40 kDa N terminus fragment of MutL) consists of two domains, residues 20-200, and 207-331, which are of mixed $\alpha/\beta$ folds (Ban and Yang, 1998b). In the absence of nucleotide substrate, LN40 is monomeric and partially

unstructured, with 20% of its residues disordered in the crystal structure. The ATP binding pocket is occupied by protein side chains that must move in order to accommodate ATP (Ban and Yang, 1998b). A non–hydrolysable ATP analog (ADPnP) transforms LN40 from a monomeric to a dimeric form. Five disordered loops (70 residues) become well-ordered in the presence of ADPnP. Almost one third of these residues are directly involved in nucleotide binding, with the rest being essential for dimer formation. Upon ATP hydrolysis, the LN40-ADP complex quickly releases the nucleotide, and LN40 becomes monomeric and partially unstructured (Ban and Yang, 1998b).

Proteins that can hydrolyse nucleoside triphosphate are generally characterized into two classes – motor proteins, such as myosin and actin, or signalling proteins, such as G-proteins. MutL does not seem to be a motor protein, as ATP hydrolysis is 100-fold slower than any known motor protein (Ban and Yang, 1998b). The fact that it mediates interactions between MutS and MutH suggests that it is acting as an ATP operated signalling molecule. Therefore it is thought that MutL serves as a molecular switch that recruits different proteins at various steps in the pathway to coordinate MMR. As a molecular switch, the ATP/ADP exchange activities of MutL are crucial. MutL mutants defective in ATP hydrolysis but not in nucleotide binding can activate MutH better than wildtype MutL, causing MutH to cleave DNA independent of a mismatch, thereby interfering with normal DNA replication. Also MutL mutants lacking ATPase activity fail to mediate interactions between MutS and MutH and completely block MMR (Ban and Yang, 1998b).

MutL also interacts with helicase II (UvrD) (Hall *et al.,* 1998). MutL may enhance loading of helicase II onto the DNA at the site of the nick, which then begins to unwind the DNA. A single helicase II protein translocates approximately 50 base pairs before it dissociates, but MMR tracts can be as long as 1 kb (Modrich, 1991). Therefore, without additional molecules of helicase II being loaded onto the DNA, the unwound DNA would re-anneal. Thus MutL may also be responsible for loading helicase II onto the appropriate DNA strand to ensure that the DNA is unwound in the proper direction (Mechanic *et al.,* 2000).

*Completion of Methyl-Directed Mismatch Repair*

The unwound nicked single stranded DNA is then degraded by one of four exonucleases (RecJ, ExoI, ExoVII, ExoX), depending on which side of the mismatch nicking occurred (Cooper *et al.,* 1993). Repair synthesis of the resulting gap is carried out by DNA polymerase III holoenzyme, which also requires single strand binding protein (Modrich and Lahue, 1996). The specific requirement for the replication polymerase provides a link between MMR and components of the replication machinery. The remaining nick is resealed by DNA ligase, completing MMR.

## Purpose of This Study

Vsr contains both a *novel* recognition mechanism and a *novel* catalytic mechanism. Not only is Vsr able to carry out efficient DNA repair, but when over-expressed it is responsible for mutagenesis. Therefore the main objective of this study is to characterize areas of the protein that interact with the DNA, and MutL, to help elucidate which regions of Vsr are involved in repair and/or mutagenesis. To this end, two N terminus deletion mutants were constructed, one with 14 and the other with 19 amino acids missing to help clarify the role of the N terminus "clamp" in Vsr endonuclease, along with three amino acid substitutions, Phe67Ala, Trp68Ala and Trp86Ala, for the three intercalating residues that are involved in mismatch recognition.

The biochemical characterization of Vsr has been limited because researchers used a mixture of full-length Vsr and a form that has lost 12-14 amino acids from its N terminus (Glasner *et al.,* 1995); others used protein that could only cleave 20% of substrate DNA (Fox *et al.,* 2000); and still others, protein that is apparently only 1% active (Turner and Connolly, 2000). Much of the earlier *in vivo* data was also limited, as VSP repair was monitored by transfecting lambda heteroduplexes prepared *in vitro* containing T/G mismatches into *E. coli* K12 bacteria, scoring repair via genetic analysis of progeny phages from such heteroduplexes (Lieb, 1987, Jones *et al.,* 1987, Meselsen, 1988).

We have used a combination of *in vivo* and *in vitro* techniques to monitor VSP repair, mutagenicity, DNA binding, and endonuclease activity coupled with *in vivo* and *in vitro*

interactions with MutL. Highly active purified preparations of Vsr endonuclease were obtained and used for our studies.

Progress has been made in determining the role of the N terminus, which appears to be required for activation by MutL. MutL may aid in forming an N terminus alpha helical clamp. Ordering of the N terminus appears to occur after DNA cleavage. The N terminus is also required for log phase regulation of Vsr. Furthermore, overexpression of MutL stabilizes Vsr levels in log phase cells causing a loss of growth phase regulation.

## Materials and Methods

*Materials and Chemicals*

All chemicals were reagent grade or higher.

*Preparation of Vsr Oligonucleotide Substrates*

Oligonucleotides were quantified from the absorbance at 260 nm. For fluorescence assays, enough of each oligonucleotide was used to make 100 μL of a 100 μM double-stranded substrate. For radioactive assays, the oligonucleotides were prepared as previously described (Cupples and Macintyre, 2000). The oligonucleotides were annealed in 1XSSC by placing them in a water bath at 85 $^{0}$C for five minutes, followed by cooling slowly to room temperature. Annealed oligonucleotides were desalted using Nick Spin Columns, and stored at -20 $^{0}$C until further use.

*Plasmid Constructs*

DNA manipulations were performed according to standard techniques (Sambrook *et al.*, 1989). Plasmids and strains used in this study are listed in Table 1. In order to construct plasmids for the overexpression of Vsr mutant proteins, *Δ14vsr* and *Δ19vsr* N terminus truncations, were first amplified using the polymerase chain reaction designed to incorporate *BamHI* and *NdeI* restriction enzyme sites. The amplified product was cloned into pCRII (Invitrogen). Δ14Vsr and Δ19Vsr coding regions were transferred on *NdeI-BamHI* fragments from pCRII constructs into the T7 promoter based pET15b. *Δ14vsr* and *Δ19vsr* N terminus truncations, were also amplified using the polymerase chain reaction designed to incorporate *NcoI* and *HindIII* restriction enzyme sites to make

plasmids for mutagenesis studies and Western blot analysis. The PCR product was cloned into pCRII and transferred on *NcoI-HindIII* fragments into pKK233-2 (Clontech) under control of the constitutive *trc* promoter. To create plasmids for *in vivo* repair, Δ14Vsr and Δ19Vsr were subcloned from pKK233-2 into pACYC184. The variants F67A, W68A and W86A were made by site directed mutagenesis (Marina Siponen, Biology 490 Thesis). Briefly, the oligonucleotides for F67A and W68A removed the internal *NcoI* site in *vsr*. These mutants were transferred on *NdeI-BamHI* fragments into pET15b (Novagen) and *NcoI-HindIII* fragments into pKK233-2 (Clontech). For the *in vivo* repair assay, a fragment of pKK233-2 clones containing *dcm, vsr* and the shared promoter 5' of *dcm* was subcloned into the *SphI* and *BamHI* sites of pACYC184 (Fermentas).

*Culture Growth and Induction of Vsr and Mutant Proteins*

All plasmids derived from pET15b (Novagen), which contained a coding sequence to add an N-terminus hexahistidine tag, were transformed into *E. coli Bl21(DE3)* for expression. One colony from a fresh transformation was used to inoculate LB containing ampicillin (100 μg/mL). Cultures were then grown overnight at 37 $^0$C with aeration. The next morning cells were pelleted at 7520 X g and diluted 1/100 in fresh LB supplemented with ampicillin. Cultures were allowed to grow to $OD_{650}$ 0.6-0.8 at which point, IPTG (0.5 mM) was added to induce the expression of the host-encoded T7 polymerase that is under control of the $P_{tac}$ promoter. After an additional three hours of growth, cells were harvested at 15300 X g, washed with 20 mM Tris-HCl pH 7.5, and stored as paste at -86 $^0$C until further use.

*Protein Purification*

Cells containing Vsr and mutant proteins were disrupted using B-Per (Pierce), following the directions supplied by the manufacturer. The purification of Vsr and mutant Vsr proteins utilized a cobalt affinity resin (Clontech) directed towards the histidine tag. After washing with 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10% glycerol, and either 25 mM or 100 mM imidazole, to remove unbound proteins, the proteins were eluted in the same buffer, containing 300 mM imidazole. Typically, from 100 mL of induced culture, 2 mg of Vsr, 2 mg of Δ14Vsr, 5 mg of Δ19Vsr, 1mg F67A, 1 mg W68A, and 5 mg of W86A were recovered.

*Dialysis*

Dialysis was performed with Slide-A-Lyzers (Pierce) in order to remove the imidazole. The buffer chosen for dialysis was 50 mM Tris-HCl pH 7.5 with 100mM NaCl and 10% glycerol, with or without 10 mM $MgCl_2$ or $CaCl_2$, depending on the application. Dialysis was performed in the cold over a minimum of 16 hours, and involved at least four buffer changes. Samples were removed from the cassettes with a syringe, and centrifuged at 4 $^0$C at 7520 X g to remove any precipitate before further use.

*Protein Estimation*

An extinction coefficient of 31000 $M^{-1}cm^{-1}$, at 280 nm, was determined for Vsr and N terminus truncations, 30800 $M^{-1}cm^{-1}$ for F67A, and 27000 $M^{-1}cm^{-1}$ for the W86A and W86A mutant protein as previously described by Gill and von Hippel, 1989. To ensure

accurate extinction coefficients, equivalent amounts of Vsr and mutant proteins were electrophoresed by 16% SDS PAGE and quantitated using the LI-COR infrared scanner.

*Removal of Histidine Tag*

Removal of the hexa-histidine tag was accomplished using a thrombin cleavage capture kit (Novagen). Reactions were carried out as indicated in the protocol supplied by the manufacturer. The extent of cleavage was monitored by electrophoresing an aliquot of the reaction mixture by 16% SDS-PAGE and staining with Coomassie blue.

*In Vitro Endonuclease Assay*

Prior to structural studies all proteins were tested for activity utilizing an *in vitro* endonuclease assay (Figure 7A). The assay consisted of two annealed oligonucleotides that form a 31 base pair fragment containing an asymmetrical C(T/G)AGG site. The T containing strand was 5' end-labelled with 32PγATP or LI-COR. Oligonucleotides were labelled as previously described (Cupples and Macintyre, 2000). Several substrates have been synthesized, two heteroduplexes, one with a T/G mismatch and the other with a C/A mismatch, and two homoduplexes, one with a TA base pair and the other with a CG base pair. Each reaction consisted of the following - 100 fmol of DNA substrate in reaction buffer (20 mM Tris pH 7.5, 10 mM $MgCl_2$, 100 mM NaCl, 100μg/mL acetylated bovine serum albumin, and 0.5 mM DTT), which was incubated with Vsr (or mutant protein) for 30 minutes at 30 $^0$C. The endonuclease reactions were stopped upon the addition of 2X loading dye (80% formamide and 0.1% w/v bromophenol blue), boiled, and loaded onto

a pre-run 15% denaturing polyacrylamide gel containing 6 N urea. The gels were covered in plastic wrap, and visualized by autoradiography or LI-COR infrared scanner.

*In Vitro MutL Stimulation Assay*

His-tagged MutL was expressed, purified and stored at –86 $^{0}$C as previously described (Feng and Winkler, 1995). Typically a 50 μM concentration of purified MutL was thawed on ice, diluted 1/10 in 1X Vsr reaction buffer. This 5 μM MutL was then either diluted again, or added directly to a Vsr endonuclease assay as outlined above.

*In Vitro Single Turnover Kinetics*

Vsr endonuclease or mutant proteins were pre-incubated with the reaction buffer for twenty minutes on ice. Cleavage was initiated by adding 1 μM protein to 1 nM radioloabeled DNA. Samples (10 μL) were removed from the reaction at various times and added to an equivalent volume of 2X loading dye. The gels were covered in plastic wrap and visualized by autoradiography. The gels were then exposed to phosphorimaging and the intensity of the bands was quantitated using Molecular Analyst software. The amount of digested material at each time point was expressed as a fraction of the total radioactivity (cut and uncut species) in each sample. The data was fitted to a first order rate equation using GraFit (Erithacus software) version 3.09.

*In Vitro DNA Bandshift Assay*

The bandshift (Figure 7B) reaction conditions were the same as the endonuclease reaction conditions, except the reactions were incubated on ice and 200 fmol of substrate was used

per reaction mixture. Glycerol was added to 5 % final concentration in each reaction, and the samples were loaded onto a pre-cooled 6% polyacrylamide/TB (89 mM Tris-HCl, 89 mM boric acid, pH 8.0) gel. 0.5 X TB was used as the running buffer, electrophoresis was stopped when the bromophenol blue dye has migrated 5-6 cm. Bands were visualized by autoradiography or by the LI-COR infrared scanner.

*Electrophoresis of Protein Samples*

Protein samples were subjected to 16 % SDS-PAGE. Gels were electrophoresed at 100 V as the samples migrated through the stacking gel and then the voltage was increased to 200 V as the samples entered the resolving portion of the gel. Once the bromophenol blue dye had migrated off the resolving portion of the gel, samples were electrophoresed for an additional fifteen minutes. Bands were visualized by staining with Coomassie Brilliant Blue.

*Western Blot Analysis Using Chemiluminescence and LI-COR*

Proteins were electrophoresed on SDS PAGE, and transferred to PVDF membranes for Western analysis (Sambrook *et al.,* 1989). Antigen antibody complexes were either detected with a chemiluminescent substrate, (NEN Life Science Products), or by the use of an infrared scanner using a secondary antibody coupled to a 700 nm fluorescing IR dye (LI-COR), following the procedure outlined by the manufacturer.

*Circular Dichroism Spectroscopy*

CD spectra were recorded at room temperature on a Jasco spectopolarimiter interfaced with an IBM computer. Far-UV CD (190-240 nm) spectra of 8 μM Vsr and mutant proteins in 50 mM Tris-HCl buffer + 10 mM $MgCl_2$, pH 7.5 were taken using a circular quartz cuvette with a 0.1 cm pathlength cell. All measurements were recorded at a scan rate of 50 nm/minute and averaged over ten accumulations. Any contributions to the signal from the buffer were subtracted and the data were smoothed using the software supplied by Jasco.

*Fluorescence Spectroscopy*

1.5 μM Vsr and mutant proteins were pre-equilibrated with metal ions on ice for 30 minutes in Vsr reaction buffer, and then pre-equilibrated for 15 minutes at 25 $^0$C in a Shimadzu or Aminco spectrofluorophotometer. To eliminate any contribution arising from the inner filter effect, the absorbance of these proteins at 280 nm never exceeded 0.1 for any of the fluorescence assays carried out in this thesis, and for titration experiments, the absorbance of the DNA was kept below 0.01 at 280 nm. Excitation was set at 280 nm and the bandwidth for both excitation and emission at 2 nm. Fluorescence emission spectra were recorded from 300 to 400 nm to determine $\lambda_{max}$ emission.

Samples containing 0.5 or 1 μM enzyme were pre-equilibrated with calcium ions at the desired temperature, as indicated above. Titration with 50 μM or 100 μM stock DNA substrate was accomplished by adding substrate to a final concentration range of 0.1 μM to 2 μM and monitoring the fluorescence change at 350 nm for ten seconds. Baselines

were subtracted and the average relative fluorescence intensity was recorded. To ensure equilibrium measurements, samples were re-scanned at half hour intervals for up to two hours, to ensure a stable reading. The raw data were then plotted and fitted to nonlinear regression according to the following equation (Clarke, 1996):

$$F=((([L]_T + [E]_T + K_d) - (([L]_T + [E]_T + K_d)^2 - 4.[L]_T . [E]_T)^{0.5}) / (2. [E]_T)). (F_E - F_S) + F_s$$

where F is the measured fluorescence, $[L]_T$ the total ligand concentration, $[E]_T$ the total binding sites in the sample, $F_E$ the fluorescence endpoint and $F_S$ the fluorescence starting point.

*In Vivo Repair Assay*

The *in vivo* repair assay (Figure 7C) for Vsr activity measured the ability of the enzyme to prevent C to T mutations caused by deamination of 5-methylcytosine to thymine (Ruiz *et al.,* 1993, Petropulos *et al.,* 1994, Wyszynski *et al.,* 1994). The strain CC503 had a kanamycin resistance cassette in *vsr*, which eliminated Vsr activity (Macintyre *et al.,* 2001). A CCAGG to CTAGG mutation in *lacZ* converted a glutamine codon (CAG) at position 461 to a nonsense codon (TAG). In the presence of an amber suppressor gene carried on a multicopy plasmid, this codon was translated as glutamic acid, the only amino acid at that site which would give a functional β-galactosidase (Cupples and Miller, 1988). Therefore if a deamination event occurred, the result was a change in the phenotype of the cell from $Lac^-$ (prior to deamination, C.G) to $Lac^+$ (deamination, T/A). If Vsr repaired the lesion, the phenotype of the cell remained $Lac^-$ (C.G) (Cupples and Miller, 1989).

*In Vivo Mutagenesis Assay*

CC108 transformants were plated on minimal lactose plates to score increased or decreased Lac reversion (Figure 7D). When a frameshift [$(G)_6$ to $(G)_5$ ] occurred in CC108 that restored the correct amino acid sequence in *lacZ,* the phenotype of the cell changed from $Lac^-$ to $Lac^+$ (Cupples *et al.,* 1990). An increase in $Lac^+$ meant an increase in mutation frequency.

| Strain | Properties/Genes | Reference |
|---|---|---|
| BL21(DE3) | F⁻*ompT [lon] hsdS*$_B$ ($r_B^- m_B^-$; an *E. coli* B strain) with DE3, a λ prophage carrying the T7 RNA polymerase gene | Studier and Moffatt, 1986 |
| CC108 | CHS142, *ara Δ9(gpt-lac)*5 *thi* F' *lacZ*⁻YA, *pro*AB | Cupples *et al.*, 1989 |
| CC503 | CHS142, *ara Δ9(gpt-lac)*5 *thi* F' *lacZ*⁻YA, *pro*AB, *vsr::kan* | Macintyre *et al.*, 2001 |
| CAG12033 | F-, LAM-, *potI*3058*::Tn10, rph*-1 | Singer *et al.*, 1989 |
| C600 | F-, *thi*-1, *thr*-1, *leuB*6, *lacY*1, *tonA*21, *supE*44, λ⁻ | Appleyard, 1954 |
| SG12060 | Derived from C600 with *clpP::kanΔlon510* | Gottesman, 1996 |
| Plasmid | | |
| pCRII | $Amp^R$ $Kan^R$ cloning and sequencing vector. | Invitrogen |
| pET15b | $Amp^R$ T7 promoter expression vector with N terminus hexahistidine tag. | Novagen |
| pKK233-2 | $Amp^R$ cloning and expression vector. | Clontech |
| pACYC184 | $Tet^R$ $Cat^R$ cloning vector. | Fermentas |

**Table 1: Strains and Plasmids Used in This Study**

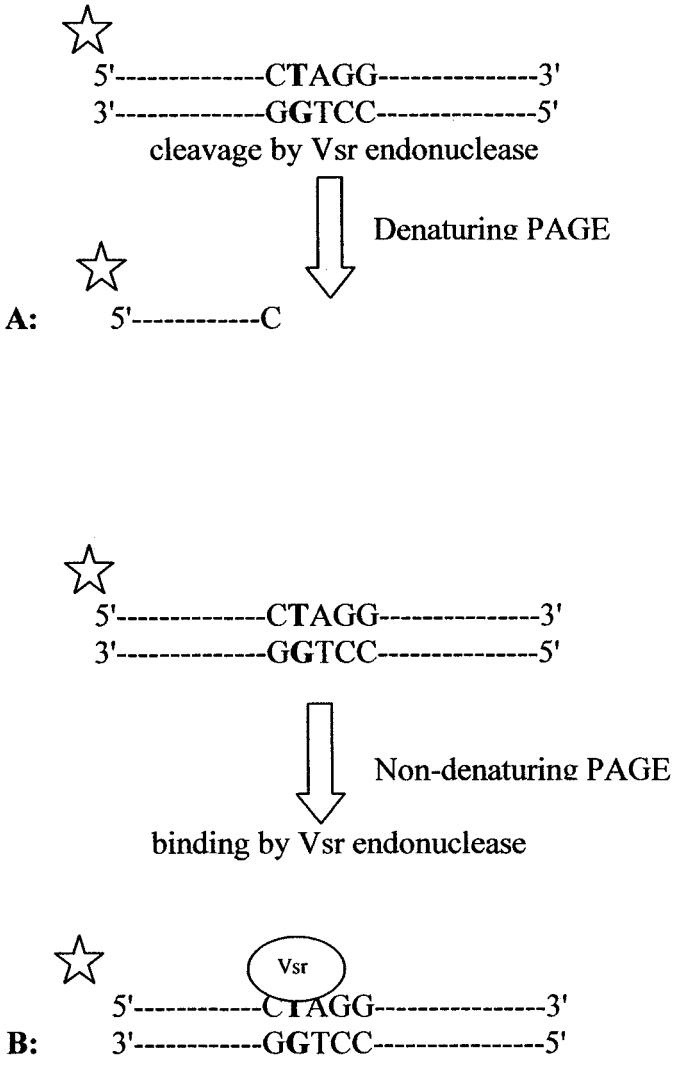



**Figure 7: Schematic Representation of Assays Used in This Study**

Assays were carried out as described in *Materials and Methods*. Endonuclease (**A**) and bandshift (**B**) assays have a star on the 5' end of the DNA, to represent the radioloabeled end of the oligonucleotide. The circle sitting atop the T/G mismatch represents Vsr bound to the DNA in a bandshift assay.

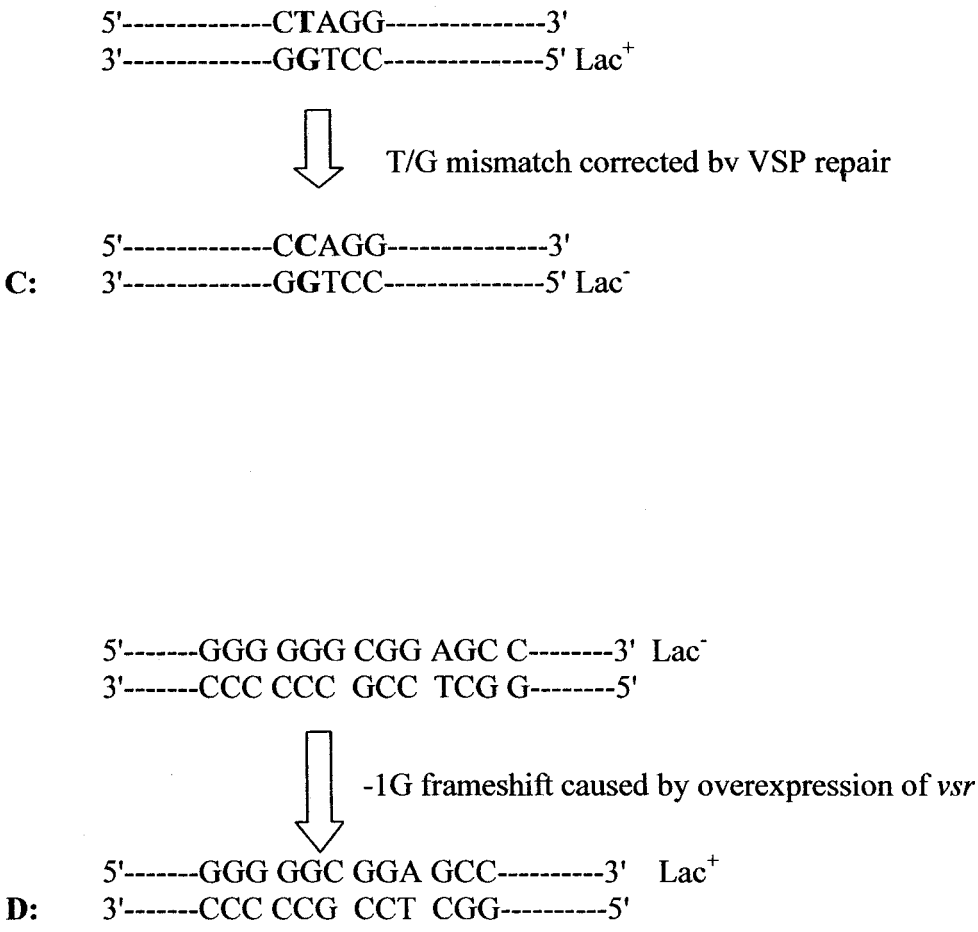



**Figure 7: Schematic Representation of Assays Used in This Study**

Assays were carried out as described in *Materials and Methods*. $Lac^+$ and $Lac^-$ in the *in vivo* repair assay (**C**) describe the phenotype of the cell, which changes with VSP repair. $Lac^-$ and $Lac^+$ in the *in vivo* mutagenesis assay (**D**) describe the phenotype of the cell, which changes as a result of Vsr-stimulated mutagenesis.

## Results

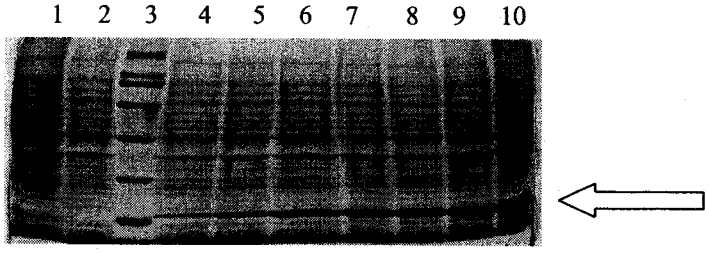
As mentioned in the *Introduction*, Vsr was crystallized in the presence (Tsutakawa *et al.*, 1999b) and absence of its DNA substrate (Tsutakawa *et al.*, 1999b). Several key features were shown in each structure. Crystallization of Vsr required the use of an N terminus truncation removing 20 amino acids, suggesting this region is unstructured and flexible in solution (Tsutakawa *et al.*, 1999a). In the DNA-bound form, the N terminus region of Vsr, specifically residues 7-15, formed an alpha helix spanning the minor groove of the DNA. From the major groove side of the DNA, three aromatic residues, Phe67, Trp68 and Trp86, intercalated into the DNA and distorted it. Met14 and Ile17 contacted the aromatics, pinching the DNA, leaving only enough space for the backbone to pass from either side (Tsutakawa *et al.*, 1999b).

On the basis of structural information, it was hypothesized (Tsutakawa *et al.*, 1999b) that the three aromatic residues were primarily involved in mismatch recognition; they stacked with the mismatch, and residues in the recognition sequence. The N terminus was thought to act as a clamp aiding DNA binding. One limitation of the crystal structure was that the DNA was cleaved, making it impossible to tell the exact mechanism and events prior to cleavage. Also, there was no information about Vsr interactions with other proteins, such as MutL, which may be important in VSP repair. We therefore constructed five mutant proteins, two N terminus deletions Δ14 and Δ19, along with three point mutations F67A, W68A, W86A, and examined their overall solution conformation, their proficiency *in vivo* for VSP repair, *in vitro* DNA binding, DNA cutting, and their ability to be stimulated by MutL.

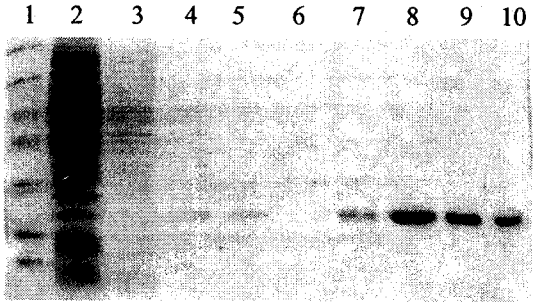
*Expression and Purification of Vsr and Mutant Proteins*

To aid purification of the wild type and mutant Vsr proteins, as outlined in *Materials and Methods* the genes were first amplified and cloned into pET15b, a vector that contains a coding sequence for an N terminus hexa-histidine tag, and a T7 polymerase-dependent promoter for protein overexpression. Expression in this system (Figure 8A) was used to obtain preparations of >95% purity as judged by SDS-PAGE (Figure 8B). The purity of all five proteins was similar, but with different yields of protein obtained in the soluble lysed cell fraction (Figure 8C). The Δ14 mutant had the same yield in this expression system as the wildtype. F67A and W68A (not shown) mutants had the lowest yields, and the Δ19 and W86A (not shown) mutants had the highest yields (see *Materials and Methods).*

A:




B:




C:

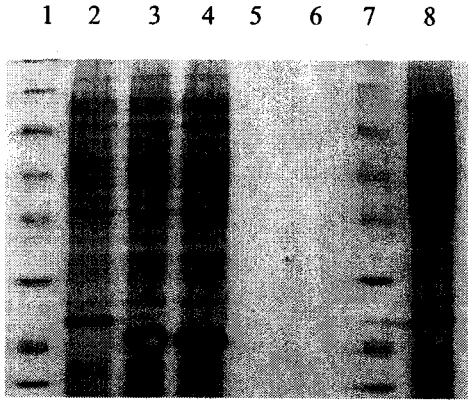


**Figure 8: Expression and Purification of Vsr in *BL21(DE3)***

**(A)** 12.5% SDS PAGE gel of soluble proteins from *E. coli BL21*(DE3) harbouring pETV. Lane 1: strain background; Lane 2: cells prior to induction; Lane 3: molecular weight markers, (200, 116, 77, 66, 45, 31, 21.5, and 14.4 kDa). Vsr protein is His-tagged, therefore it migrates slightly above the 21.5 kDa marker (arrow). Lanes 4-10 are different time points post IPTG induction: 15, 30, 60, 120, 180, 140 min, and a 16 hour sample. **(B)** 16% SDS PAGE gel of Vsr purification. Lane 1: molecular weight markers, (116, 77, 66, 45, 31, 21.5, and 14.4 kDa); Lane 2: a fraction of whole cells containing Vsr, prior to cell lysis; Lane 3: after cell lysis, proteins that did not bind to the resin; Lanes 4-6: washes (25X times resin volume) with 25 mM imidazole; Lane 7: wash with 100 mM imidazole; Lanes 8-10, elutions with 300 mM imidazole. **(C)** 12.5% SDS PAGE of soluble proteins from *E. coli BL21*(DE3) harbouring, Lane 2: pETV; Lane 3: pETΔ14; Lane 4: pETΔ19; Lane 8: pET F67A. Lanes 1 and 7 contain molecular weight markers (116, 77, 66, 45, 31, 21.5, and 14.4 kDa).

*Conformational Analysis of Vsr and Mutant Proteins*

When a mutant protein is made, it must be ascertained that any change in activity is not attributed to gross differences in protein structure. Therefore, wildtype Vsr and mutant proteins were subjected to secondary and overall tertiary structural analysis using circular dichroism and fluorescence spectroscopies. These two spectroscopic methods were chosen on the basis of the crystal structure (Tsutakawa *et al.,* 1999a and 1999b). Vsr is composed of three alpha helices, and Vsr contains chromophores within key structural regions that contact the DNA. A brief overview of the two spectroscopic methods follows.

*Circular Dichroism Spectroscopy of Vsr and Vsr Mutant Proteins*

Proteins absorb light because of peptide groups, aromatic amino acids and disulfide bonds and emit radiation in the UV region of the light spectrum (Schmid, 1998). When a chromophore is present in an asymmetric unit, or when it is immobilized in an asymmetric environment, left-handed and right-handed circularly polarized light is absorbed to different extents. This phenomenon is called circular dichroism (CD) (Schmid, 1998). The CD in the amide region (170-250 nm) is used to estimate the content of secondary structure; the α-helix displays a strong and characteristic CD absorption in the far-UV region (at 222 nm and 208 nm), while the contribution of other elements of secondary structure in this region are less well defined (Schmid, 1998).

In the absence of DNA, Vsr is composed of a central β-sheet buttressed by three α-helices (Tsutakawa *et al.,* 1999a). The CD signal from the three alpha helices was used

in this study to compare wildtype Vsr and mutant proteins. There were no detectable differences found in the secondary structure spectra of Vsr and the mutant proteins (Figure 9). The α-minimum inflection observed at 222 nm and 208 nm, indicative of alpha helicity, was superimposable for all five proteins. In the co-crystal structure of Vsr bound to DNA, the N terminus formed an alpha helix between residues 7-15 (Tsutakawa *et al.,* 1999b). The fact that Vsr itself did not give an increased alpha helical CD signal when compared to the N terminus deletion mutants indicated that in solution, the N terminus of Vsr was un-ordered and in the absence of DNA substrate, it did not form an alpha helix. This result confirmed those obtained by X ray crystallography and NMR spectroscopy (Tsutakawa *et al.,* 1999a).

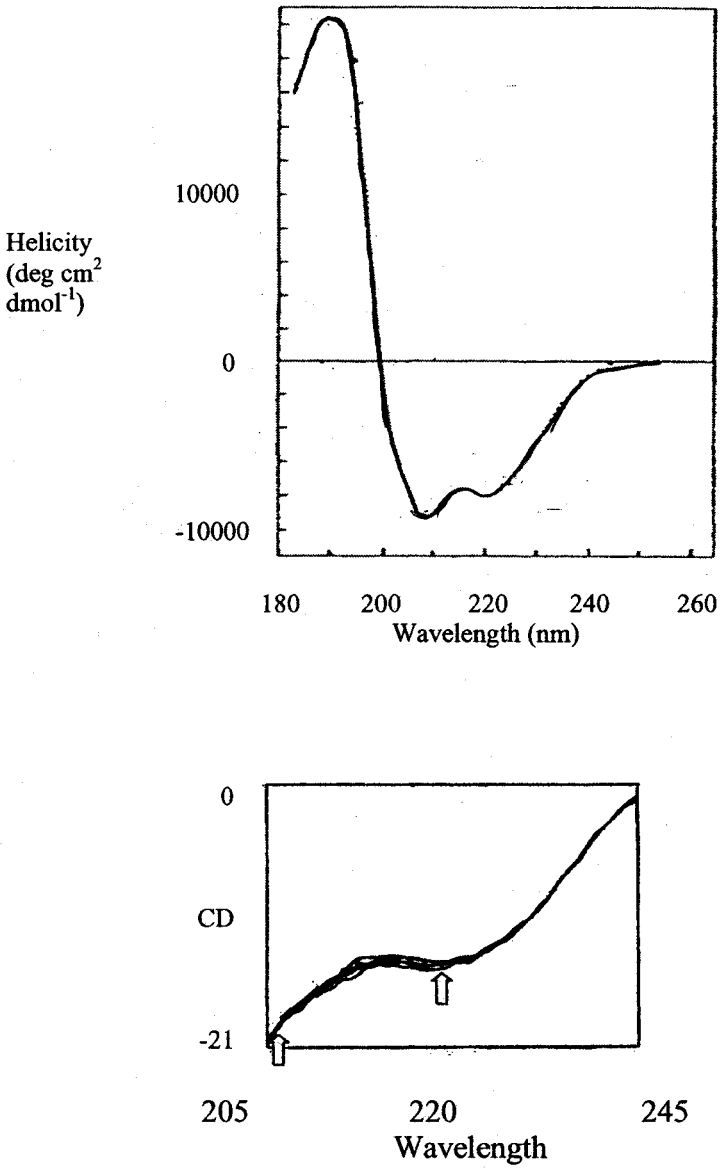



**Figure 9: CD Spectra of Vsr and Mutant Proteins**

The CD spectra were obtained and analysed as outlined in *Materials and Methods.* The top spectrum represents 10 smoothed accumulations of 8 μM Vsr. The bottom overlaid raw spectra contain 10 μM concentrations of Vsr, F67A, W86A and Δ19 N terminus truncation. The positions of the arrows indicate spectral overlaps at 208 and 222 nm.

*Fluorescence Spectroscopy of Vsr and Vsr Mutant Proteins*

During light absorption, light energy promotes electrons from the ground state to an excited state. Fluorescence emission is observed when the excited electrons revert back to ground state (Schmid, 1998). In the excited state, some energy is lost by non-radiative processes, such as vibrational transitions; therefore, the light emitted is always less than the absorbed light and occurs at higher wavelengths (Schmid, 1998).

The fluorescence spectra of proteins originate from Phe, Tyr and Trp residues. However the fluorescence spectrum is usually dominated by the contribution of Trp residues; their quantum yields of emission are considerably greater than the respective values for Tyr and Phe (Schmid, 1998). In a hydrophobic environment, such as in the interior of a protein, a completely buried Trp can emit at 320 nm (indole shows an emission maximum of 320 nm in hexane), while in an aqueous environment, a completely solvent exposed Trp emits at 360 nm (Schmid, 1998). Therefore, fluorescence emission maxima are often used to assess overall conformation in wildtype and mutant proteins.

Vsr contains five tryptophans, two tyrosines and two phenylalanines, with none of these chromophores present in the N terminus. Therefore fluorescence spectroscopy was used to compare the overall conformations of Vsr and mutant Vsr proteins. Phe fluorescence is disregarded and is not observed in native proteins because its sensitivity is very low and energy transfer to the other two aromatic amino acids quenches its emission; Tyr and Trp residues absorb strongly at 280 nm where Phe emits (Schmid, 1998).

Excitation was set at 281 nm, which excited both Tyr and Trp residues. There was a small shoulder in the fluorescence spectrum of Vsr and N terminus truncations at 310 nm (Figure 10A) representing the contribution of the two tyrosine residues in the observed Vsr fluorescence spectrum. Tyr was barely detectable in the overall fluorescence spectrum because Trp emission was much stronger. Also, in folded proteins Trp emission is shifted to smaller wavelengths towards that of Tyr, due to non-radiative energy transfer from Tyr to Trp residues (Schmid, 1998). Selectively exciting Trp residues at 295 eliminated the small shoulder at 310 nm that was observed when excitation was set at 281 nm, while the remainder of the fluorescence spectrum was identical (Figure 10B); therefore as expected, the measured fluorescence emission spectrum originated almost exclusively from Trp residues in Vsr.

There were several differences in overall conformation among the different proteins as measured by fluorescence spectroscopy in the presence and absence of magnesium chloride (Table 2). The $\lambda_{max}$ excitation was 281 nm for all the proteins. However when Vsr was in solution in the absence of added metal ions, the $\lambda_{max}$ emission was 342 nm. Upon addition of either calcium chloride or magnesium chloride, the $\lambda_{max}$ emission shifted to 348 nm. The N terminus truncations also had the same spectra as Vsr in the absence of metal ions, however N terminus truncations consistently underwent a shift upon metal ion addition to 345 nm. Studies have shown that a completely buried Trp residue emits at 320 nm, while a completely solvent-exposed Trp residue emits at 360 nm (Schmid, 1998); therefore the shifts in wavelength accompanied by metal binding indicated an exposure of tryptophan residues in Vsr and N terminus truncation mutants to

solvent. The 3 nm shift in wavelength was present in all the enzyme preparations comparing Vsr and N terminus truncation mutants. The 3 nm difference in maximum emission wavelength between Vsr and N terminus truncations meant that the positions of tryptophan residues in the N terminus mutants were slightly less solvent-exposed than the wildtype protein (Figure 10A). Although solvent exposure was increased in the presence of metal ion, the chromophores in both Vsr and N terminus truncations were not completely solvent exposed. Denatured Vsr and N terminus truncations in 6M guanidine hydrochloride had a $\lambda_{max}$ emission at 355 nm (Figure 10B). As expected, under denaturing conditions, the shoulder for tyrosine fluorescence at 310 nm became more pronounced as the tryptophan fluorescence shifted in wavelength.

The fluorescence spectra of F67A, W68A and W86A mutants were similar to Vsr in the absence of added metal ions since they all had a $\lambda_{max}$ emission at 342 nm (Table 2). However, one limitation to fluorescence spectroscopy is that the spectrum generated for each protein represents the average overall contribution of all the tryptophan residues present in a protein. Individual tryptophan residues may be in environments with different solvent polarities, with respect to wildtype protein, but these changes may be masked by polarity changes occurring at another tryptophan residue, cancelling out the changes caused by the first.

Evidence from several different protein preparations for F67A, W68A and W86A mutants suggested that these proteins had altered conformations when compared to the wildtype "active" Vsr protein in the presence of magnesium ions. Vsr is inactive in the

absence of magnesium or manganese ions (Tsutakawa *et al.,* 1999a). The tryptophan residues in the F67A, W68A and W86A mutants were less solvent-exposed than metal bound wildtype Vsr and N terminus truncations as indicated by the lack of red shift in the presence of magnesium chloride (Table 2).

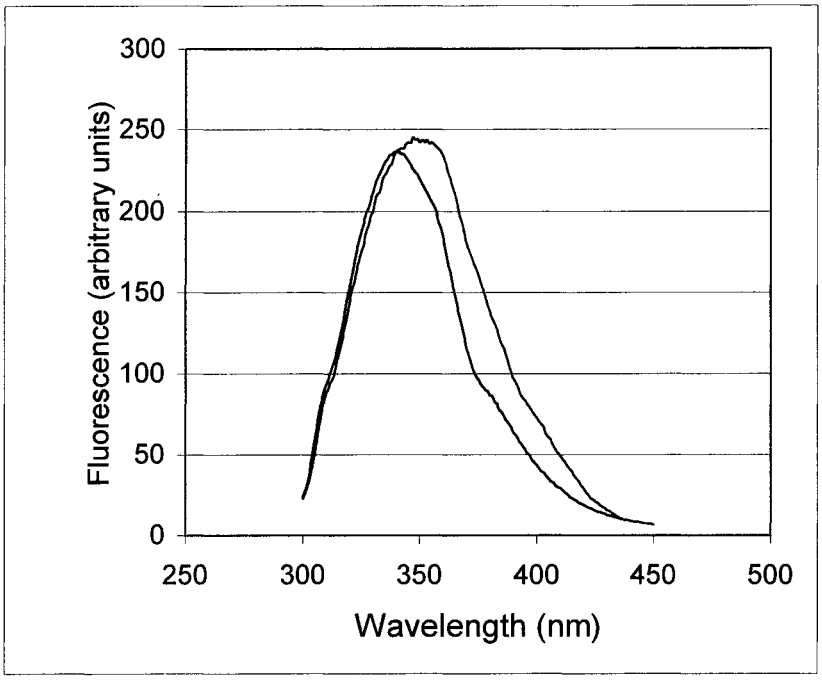
Fluorescence spectra of F67A, W68A and W86A mutants did not undergo any changes (shifts in wavelength or quenching) upon the addition of divalent metal ions, indirectly suggesting that they could no longer coordinate metal ions. None of these amino acids was directly involved in metal coordination; according to the crystal structure, His69, His64 and the backbone atoms of Thr63, Gly65 and Glu136 re-arrange to co-ordinate the metal ion (Tsutakawa *et al.,* 1999a). With an inability to bind metal ions and adopt proper "active" Vsr confomation it seems quite reasonable that these proteins were inactive *in vitro* and *in vivo* (see later section).

*Metal Dependence of Vsr*

Having determined the structural importance of metal binding for Vsr using fluorescence spectroscopy (Table 2), metal ion requirements were also examined *in vitro* with an endonuclease assay to monitor Vsr cleavage of a T/G mismatch in a *dcm* recognition sequence. The T-containing strand was 5' end-labelled with 32PγATP. Electrophoresis of endonuclease assays on a denaturing gel allowed resolution between uncut substrate (upper band) and nicked DNA (lower band) (Figure 11). 10 mM magnesium chloride in 10X Vsr reaction buffer (1 mM final metal ion concentration) was sufficient to carry out

optimal cleavage under our assay conditions, with some cleavage occurring with 1 mM magnesium chloride. Calcium and zinc were not used as catalytic metal ions (Figure 11).

A:


300
250
200
150
100
50
0
Fluorescence (arbitrary units)
250
300
350
400
450
500
Wavelength (nm)

B:

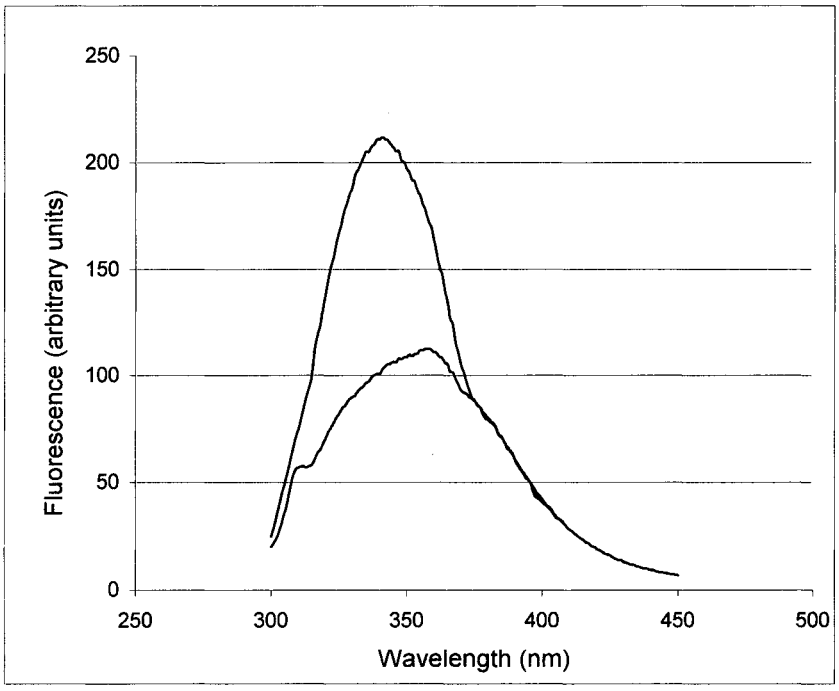

250
200
150
100
50
0
Fluorescence (arbitrary units)
250
300
350
400
450
500
Wavelength (nm)

**Figure 10: Fluorescence Emission Spectra for Vsr and Mutant Proteins**

**(A)** Spectra were collected as indicated in *Materials and Methods* and averaged for triplicate samples of 1.5 μM Δ19 (blue) and 1.5 μM Vsr (pink). **(B)** Spectra were collected as indicated in *Materials and Methods* and averaged for triplicate samples of 1.5 μM Vsr. The blue spectrum shows the selective excitation of Trp residues with an excitation wavelength set at 295 nm. The black spectrum is an average of triplicate samples of 2 μM denatured Vsr in 6 M guanidine hydrochloride. Excitation was set at 281 to show both tyrosine and tryptophan residues.

| Protein | λmax Excitation (nm) | λmax Emission (nm) – in absence of metal ion | λmax Emission (nm) – in presence of metal ion |
|---|---|---|---|
| VSR | 281 | 342 | 348 |
| Δ14 Vsr mutant | 281 | 342 | 345 |
| Δ19 Vsr mutant | 281 | 342 | 345 |
| F67A mutant | 281 | 342 | 342 |
| W68A mutant | 281 | 342 | 342 |
| W86A mutant | 281 | 342 | 342 |

**Table 2: Fluorescence Emission Maxima in the Absence and Presence of 10 mM Magnesium Chloride**

Emission maxima were determined by finding the wavelength with the highest relative fluorescence intensity. In instances where there was a broad range, the middle part of the peak was chosen as the emission maxima.

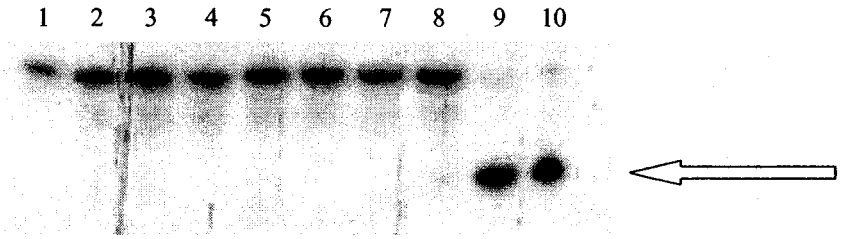



**Figure 11: Metal Ion Requirements for Endonuclease Activity of Vsr**

Endonuclease assays were carried out as described in *Materials and Methods*, however the 10X reaction buffer differed in metal and metal ion concentration. Lane 1: DNA substrate; Lane 2: 100 mM calcium chloride in Tris-HCl buffer; Lane 3: 100 mM zinc chloride in HEPES buffer; Lane 4: 100 mM calcium chloride in Tris-HCl; Lane 5: 0.1 μM magnesium chloride in Tris-HCl buffer; Lane 6: 10 μM magnesium chloride in Tris-HCl; Lane 7: 0.1 mM magnesium chloride in Tris-HCl; Lane 8: 1 mM magnesium chloride; in Tris-HCl buffer; Lane 9: 10 mM magnesium chloride in Tris-HCl buffer; Lane 10: 100 mM magnesium chloride in Tris-HCl buffer. The arrow indicates the size of the 15 base product.

*Kinetics for Vsr and Mutant Proteins*

Through the use of the *in vitro* endonuclease assay we found that the N terminus was not required for Vsr mismatch specificity. A C/A mismatch and homoduplex DNA were used as possible substrates for Vsr and N terminus truncations and only the T/G mismatch gave rise to cleaved product (Figure 12).

Single turnover kinetic experiments were used to determine the $k_{cat}$ of Vsr and N terminus truncation mutant proteins, with a variety of protein preparations, in triplicate, with the data plotted and fitted (Figures 13-15). The enzyme concentration (1 μM) was always kept in excess of substrate (1 nM) so as to achieve single turnover kinetics; it was important to utilize single turnover conditions to assay Vsr, as all the published kinetic data for Vsr was determined under comparable conditions (Fox *et al.,* 2000 and Turner and Connolly, 2000). Turner and Connolly also used single turnover kinetic assays to show that their Vsr preparation was only 1% active. Therefore it was imperative to examine our Vsr preparations for comparison with published activities.

Under single turnover conditions the cleavage rate should be independent of enzyme concentration. However, when a protein/DNA substrate interaction is weak, multiple turnovers are measured instead of single turnover kinetics. Therefore, the rate of the reaction was also examined as a function of enzyme concentration, from 100 nM to 30 μM protein; the rate changed by less than fifteen percent (within the standard deviation of the assay) with a three-hundred-fold difference in protein concentration. Since the experiments described here were done in the μM range for Vsr and mutant proteins it

seemed sufficient to ensure single turnover kinetics with the DNA substrate. By fitting the data to a first order rate equation the $k_{cat}$ were determined as follows: 2.9 +/- 0.43 $min^{-1}$ for Vsr (Figure 13), 0.089 +/- 0.0044 $min^{-1}$ for the Δ14 N terminus truncation (Figure 14), and 0.069 +/- 0.0041 $min^{-1}$ for the Δ19 N terminus truncation (Figure 15). The F67A, W68A and W86A mutants had no measurable enzymatic activity (Figure 21D).

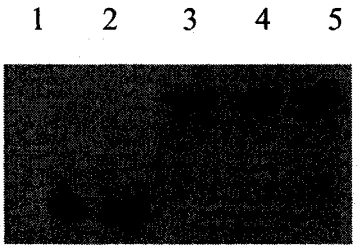
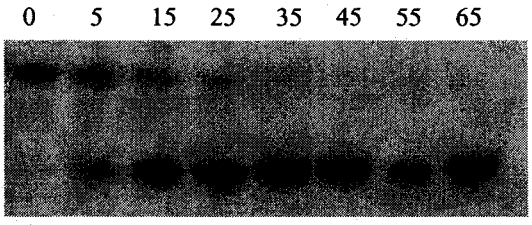



**Figure 12: Specificity of Endonuclease Assay**

Endonuclease assays were carried out as described in *Materials and Methods*. Lane 1: 1 μM Δ19 with a T/G mismatch; lane 2: 1 μM Δ14 with a T/G mismatch; lane 3: 1 μM Δ19 with a C/A mismatch; lane 4: 1 μM Δ14 with a C/A mismatch; lane 5: 1 μM Δ19 with a homoduplex substrate.

A:




B:

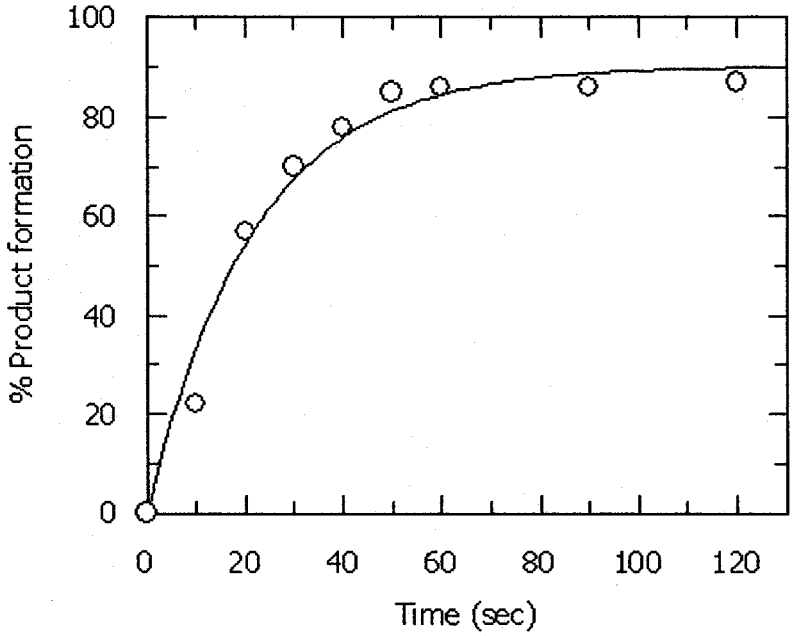
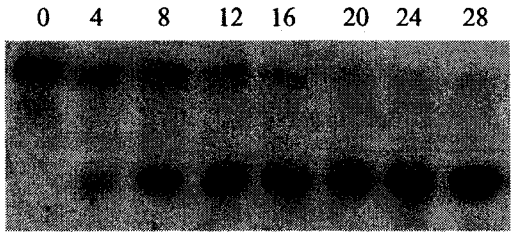
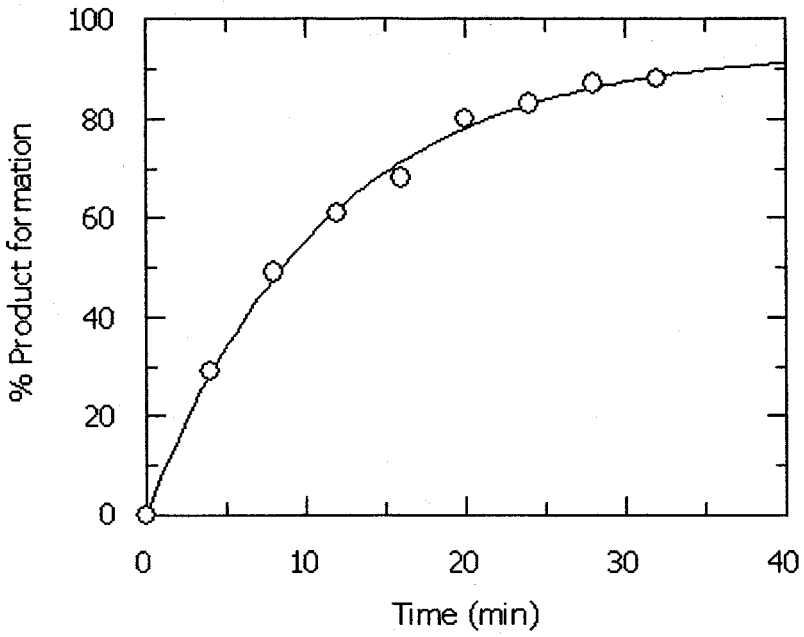



**Figure 13: Single Turnover Kinetics of Vsr Endonuclease**

**(A)** The top panel shows an autoradiogram of the gel produced when the substrate DNA was incubated with Vsr for the times given in seconds. **(B)** The graph shows an analysis of the data fitted from the average of three independent assays to a first order rate equation. The solid line represents the data fitted to a first order rate equation using GraFit (Erithacus Software) with $k_{obs}$ of 2.9 +/- 0.43 $min^{-1}$.

A:



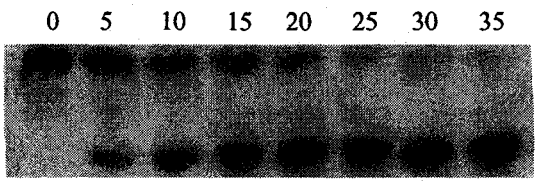




B:

**Figure 14: Single Turnover Kinetics for Δ14 N terminus Truncation**

**(A)** The top panel shows an autoradiogram of the gel produced when the substrate DNA was incubated with Vsr for the times given in seconds. **(B)** The graph shows an analysis of the data fitted from the average of three independent assays to a first order rate equation. The solid line represents the data fitted to a first order rate equation using GraFit (Erithacus Software) with $k_{obs}$ of 0.089 +/- 0.0044 $min^{-1}$.

A:




B:

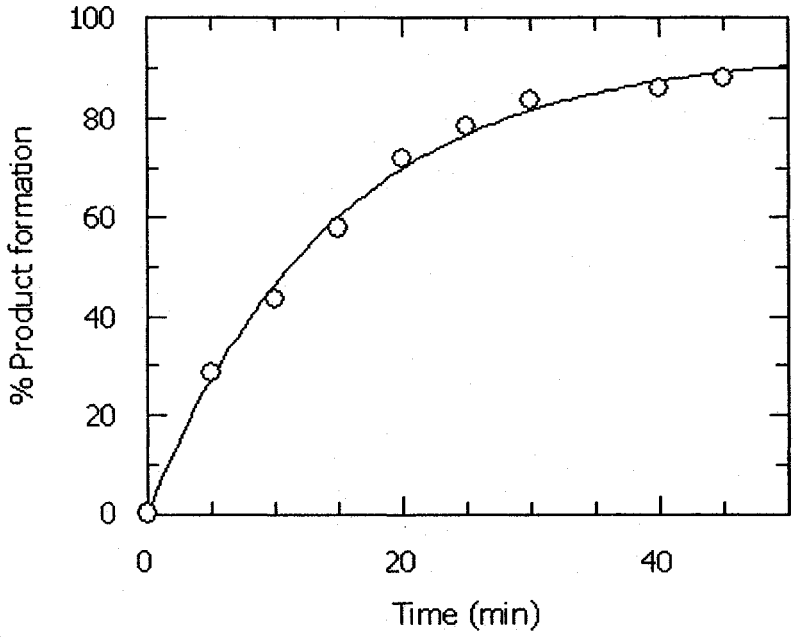



**Figure 15: Single Turnover Kinetics for Δ19 N terminus Truncation**

**(A)** The top panel shows an autoradiogram of the gel produced when the substrate DNA was incubated with Vsr for the times given in seconds. **(B)** The graph shows an analysis of the data fitted from the average of three independent assays to a first order rate equation. The solid line represents the data fitted to a first order rate equation using GraFit (Erithacus Software) with $k_{obs}$ of 0.069 +/- 0.0041$min^{-1}$.

*DNA Binding Assays for Vsr and Mutant Proteins*

An informative method for comparing the catalytic efficiency of enzymes is to compare the specificity constant $k_{cat}/K_d$ (Turner and Connolly, 2000, Gonzalez-Nicieza *et al.,* 2001, Fox *et al.,* 2000). Bandshift analysis is often used as a measure of DNA/protein binding (Clarke, 1996). In this case, Vsr or mutant protein is incubated with the same substrate as in the endonuclease assay, and then run on a gel. The reactions are electrophoresed on a non-denaturing gel so that DNA bound to protein migrates more slowly than that of free DNA alone.

Experiments to monitor protein binding using bandshift analysis with a non-methylated T/G substrate were initially modelled after the Drotschmann protocol, which resulted in smeary Vsr shifts, similar to those already published by that group (Drotschmann *et al.,* 1998). Turner and Connolly's bandshift protocol was similar to that of Drotshmann, but Turner and Connolly had clean, distinct, Vsr bandshifts in the presence of a calcium-containing buffer (Turner and Connolly, 2000). It is generally accepted that sub-nanomolar dissociation constants result in "clean" gel shifts (Carey, 1991). Turner and Connolly (2000), obtained a dissociation constant of 0.74 +/- 0.1 nM for a hemi-methylated T/G substrate (clean shifts), and 2.3 nM +/- 0.3 for a non-methylated substrate (Turner and Connolly, 2000).

Although the conditions to estimate $K_d$ by bandshift analysis were identical to those presented by Turner and Connolly, bandshift analysis proved futile, as enzyme concentrations in excess of 1 μM were required to shift all the DNA in the presence of

magnesium ions (Figure 16A) and greater than 8 μM protein was required to shift the DNA in calcium containing buffers (Figure 16A). In addition, there were no discrete bands that could be assigned to a protein/DNA complex. The free DNA band smeared out as the concentration of protein increased (Figure 16A and 16C); high protein concentrations and smeary bands suggested "weak" binding.

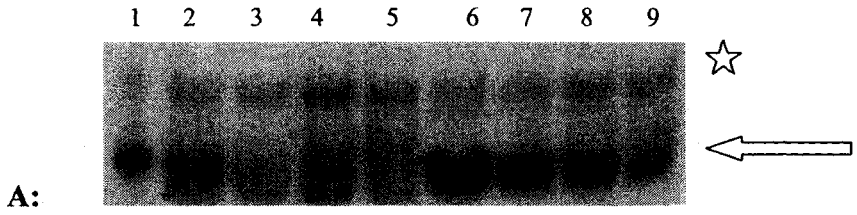
An additional complication to the bandshift assay arose in assays that contained magnesium ions, which supported endonuclease activity. Due to the micromolar concentrations of Vsr required to visualize shifted smears in bandshift assays, electrophoresing a portion of a bandshift reaction containing magnesium ions on a denaturing gel showed that the DNA was entirely cleaved (Figure 16B); therefore a calculated dissociation constant for these reactions would be for the cleaved complex. Unpublished gel shifts by Marina Siponen and Gina Macintyre showed that Vsr could also bind to a pre-nicked substrate, consistent with the result observed in Figure 16A and 16C. Preliminary fluorescence evidence also suggested that Vsr could bind a pre-nicked substrate. However, due to time constraints, and a non-funtional fluorimeter, a full analysis was not possible.

Calcium is a non-reactive equivalent to magnesium (Tsutakawa *et al.,* 1999a, Turner and Connolly, 2000). Using fluorescence spectroscopy, the overall tertiary structure of Vsr was identical in either magnesium or calcium containing buffers ($\lambda_{max}$ 348 nm). However, note the difference in gel shifts between the reactions taking place in magnesium-containing buffer, which supported endonuclease activity, and in calcium-

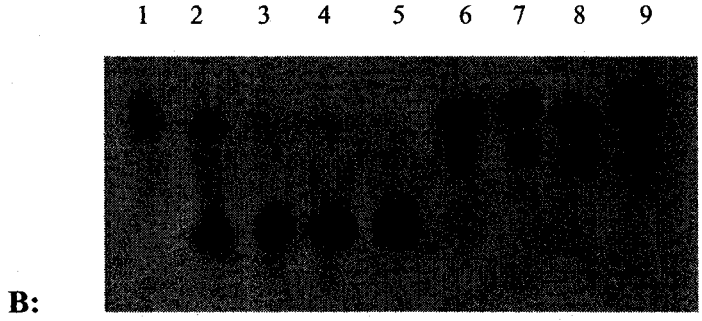
containing buffer that did not (Figure 16A). Higher protein concentrations were required in calcium-containing buffer to form an equivalent degree of shift in comparison to those in magnesium containing buffer - this also indicated that bandshifts visualized at lower concentration in magnesium buffer were due to protein-product complexes, suggesting that Vsr dissociated slowly from its product. This data suggested that substrate/Vsr dissociation proceeded more quickly than product/Vsr dissociation.

Dissociation of DNA/protein complexes during electrophoresis, which led to smeary bands, was not in and of itself a problem, because the disappearance of free DNA could have been quantitated instead of the appearances of complexes (Carey, 1991). However, the shifted products (Figure 16A and 16C) were not well resolved as they were running right above the free DNA band. Another complication arising from this assay was that a non-specific DNA band was present in the negative control sample (Figure 16A, lane 1) making it impossible to judge where the actual shifts were taking place. Bandshifts in the presence gel purified oligonucleotides lacked this non-specific band, but were comparably smeary (Figure 16C). Changing the acrylamide concentration, adding metals to the running buffer, running the gels in the cold, and loading gels at a higher voltage (Carey, 1991), did not help minimize Vsr/DNA dissociation. The N terminus truncation mutants cut the DNA and obviously bound to the DNA in order to do so, however they did not shift the DNA at all in a bandshift assay (Figure 16C), suggesting weak or transient binding. The smears do not correspond to nicked DNA, as a gel shift reaction with either Δ14 or Δ19 N terminus truncation, lacks smearing, even though, these

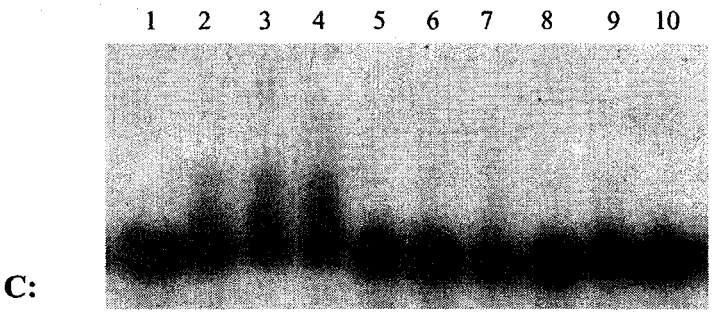
proteins cut the DNA (Figure 16C). These limitations prompted us to look for an alternative way to determine substrate binding.




A:




B:




C:

**Figure 16: Bandshift Analysis of Vsr Endonuclease**

**(A)** Lane 1: DNA substrate only; lanes 2-5: DNA substrate plus increasing concentrations of Vsr (0.25, 1, 2.5 and 8 μM) in magnesium chloride containing reaction buffer; lanes 6-9: increasing concentrations of Vsr (0.25, 1, 2.5 and 8 μM) in calcium chloride containing reaction buffer. The arrow denotes where the shifts are taking place. The star denotes the non-specific band present in every sample. In lanes 1, 6 and 7 the non-specific DNA band corresponds to 5 % of the total DNA sample. For lanes 2-5, 8-9, the amount of radioactivity in the non-specific band corresponds to 8 % of the total DNA in each fraction. The remainder of the DNA in each lane is either at the DNA only position (lane 2, lane 6, lane 7 and partially lane 8) or as a smear (lane 3, lane 4, lane 5, lane 9 and partially lane 8). **(B)** Denaturing gel containing an aliquot of each bandshift reaction. The protein concentrations from panel A are maintained in this figure. This gel is not as clean as other endonuclease assays presented in this thesis, because of the presence of glycerol in the bandshift reactions (Carey, 1991). **(C)** Bandshifts with gel purified DNA. Lane 1: DNA substrate only; lanes 2-4: DNA substrate plus increasing concentrations of Vsr (0.25, 0.5 and 0.75 μM); lanes 5-7: increasing concentrations of Δ14 N terminus truncations (1, 4 and 8 μM); lanes 8-10, increasing concentrations of Δ19 N terminus truncation (1, 4 and 8 μM).

The intrinsic fluorescence of Vsr was exploited to develop a substrate-binding assay. Vsr contains two tyrosine and five tryptophan residues, which are in, or proximal to, the DNA binding area of the protein. Upon the addition of DNA substrate to the wildtype protein and N terminus mutants, a quenching in the fluorescence spectrum was observed (Figure 17). Quenching was not caused by the inner filter effect, as the DNA did not absorb at either 280 nm (abs < 0.01) where the protein was being excited, or at 350 nm where the emission was recorded (abs < 0.01). Binding to the T/G containing substrate was accomplished in calcium-containing reaction buffer, to avoid complications arising from endonuclease activity. In this assay, a fixed protein concentration was used to titrate in substrate, making the substrate limiting for most of the data points. An estimate of binding sites for the DNA (Clarke, 1996) was also made, upon fitting with the equation for binding as indicated in *Materials and Methods*; one binding site per protein molecule was observed for both Vsr and N terminus truncations.

Generally, when a spectroscopic change originates from a macromolecule (protein) and not a ligand (DNA), titration curves are more precise (Lohman and Bujalowski, 1991). In order to use a spectroscopic change to calculate binding data, it is commonly assumed that a linear relationship exists between the signal change and the fractional saturation of protein. However if a linear relationship does not hold, then the calculated dissociation constant will be incorrect (Lohman and Bujalowski, 1991). This becomes more of a concern when binding can exist in several modes, each possessing a different spectroscopic signal change. This seems unlikely with Vsr since only one binding site

for DNA was found in the co-crystal structure (Tsutakawa *et al.,* 1999b), and the fluorescence changes themselves indicate a 1:1 binding ratio.

The stability of an enzyme-ligand complex is an intrinsic property of the atomic contacts made between partners and has a constant value irrespective of their respective concentrations (Clarke, 1996). Spectroscopic changes, such as fluorescence intensity changes, measure the average degree of saturation of the macromolecule with ligand (Lohman and Bujalowski, 1991). In order to calculate accurate dissociation constants, the transition must occur smoothly; in general, a smooth hyperbolic curve gives a good estimate of $K_d$, while an angular response, with a distinct corner gives a good estimate of binding sites (Clarke, 1996). The titration curves with Vsr and N terminus truncations are good candidates for determining dissociation constants, as they are smooth hyperbolic curves (Figure 17).

Ideally the untransformed data should be fitted by non-linear regression, as it is statistically more accurate than transforming the data (Clarke, 1996). With the equation used to fit the data (see *Materials and Methods*), the untransformed data was used with only the starting fluorescence being substituted into the equation. The end point (fluorescence intensity), the $K_d$, and the concentration of binding sites were then simultaneously fitted by non-linear regression, with $R^2$ values between 0.998 to 0.999. One binding site for the DNA was obtained from the fits, and the fluorescence endpoints were reasonable, within one or two fluorescence units of the endpoints observed at a 2:1 DNA to protein ratio. Data simulations using tenfold differences in dissociation

constants, with identical starting/ending points as those originally fitted by non-linear regression also indicated that the fluorescence assay was sensitive enough to calculate dissociation constants (Figure 18). The dissociation constants were determined to be 25 +/- 8 nM for Vsr, and 52 +/- 10 nM for Δ14 and for Δ19 N terminus truncations. The catalytic efficiencies of these proteins ($k_{cat}/K_d$ ($min^{-1}/\mu M$)) were thus calculated to be 53 for Vsr, 1.7 for Δ14 and 1.3 for the Δ19 mutant (Table 4).

There appeared to be a discrepancy between bandshift analysis and DNA binding monitored by fluorescence; the N terminus mutants were unable to shift the DNA in a bandshift assay, and in the fluorescence assays these mutants appeared to bind almost as effectively as the wildtype Vsr. However, examination of the gel shifts in Figure 16 helped explain these differences. All the fluorescence assays were done under non-cutting conditions, in the presence of calcium-containing reaction buffer. The calcium-containing reactions appeared to have a different shifting pattern than the magnesium-containing reactions. Even with 8 μM concentrations of Vsr in lane 9, the DNA appeared not to have completely shifted in the calcium-containing buffer. If the N terminus truncations had a slightly weaker binding relative to Vsr as indicated by the fluorescence assay, then it seemed reasonable that they may not shift in a bandshift assay.

The mutants F67A, W68A and W86A were unable to bind to the DNA – quenching with T/G substrate was not observed (Figure 17). The addition of homoduplex (C.G) DNA did not lead to quenching either (Figure 17), suggesting that Vsr does not scan the DNA to locate a mismatch. This was not surprising as DNA scanning proteins such as T4 endo

V have large basic faces, which Vsr lacks, allowing DNA scanning proteins to translocate along the DNA (Morikawa and Shirakawa, 2000).

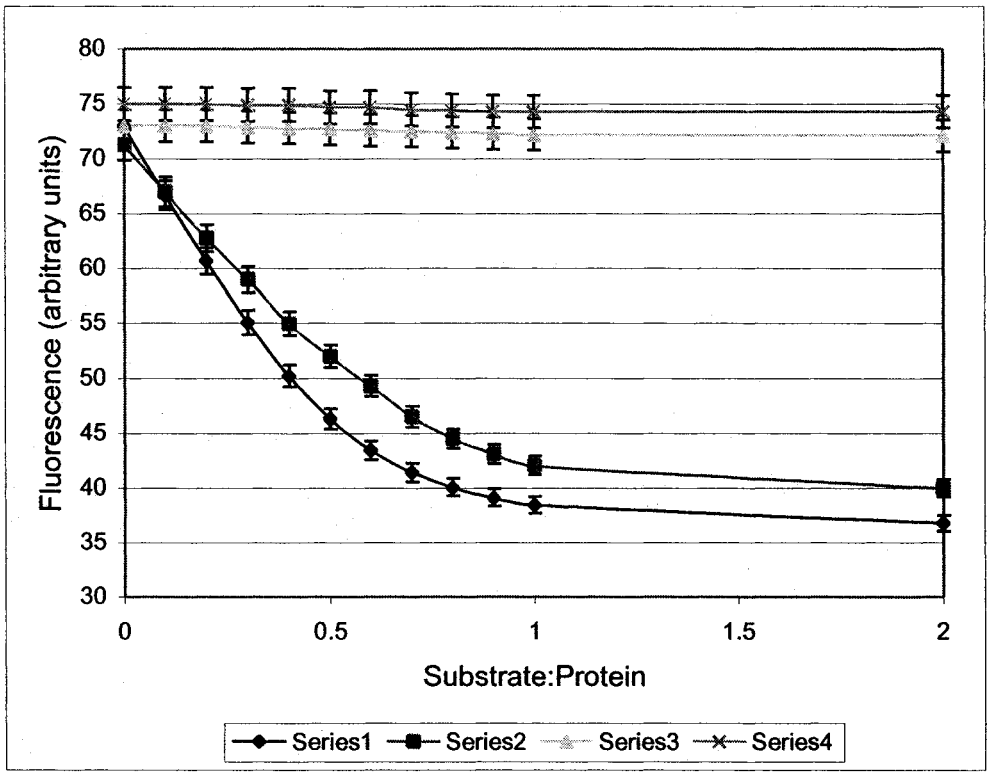



**Figure 17: Fluorescence Quenching as a Measure of Substrate Binding**

The average fluorescence measurements from triplicate samples of 0.5 μM protein were carried out and analysed as described in *Materials and Methods.* Series 1: Vsr titrated with T/G substrate; Series 2: Δ14 N terminus truncation titrated with T/G substrate; Series 3: Vsr titrated with homoduplex DNA; Series 4, 0.75 μM W68A mutant protein titrated with a T/G mismatch. Parameters used to fit the dissociation constant were 1 binding site, an initial fluorescence of 72 (arbitrary units) and an end point of 39 (arbitrary units) for Vsr and 42 (arbitrary units) for N terminus truncations.

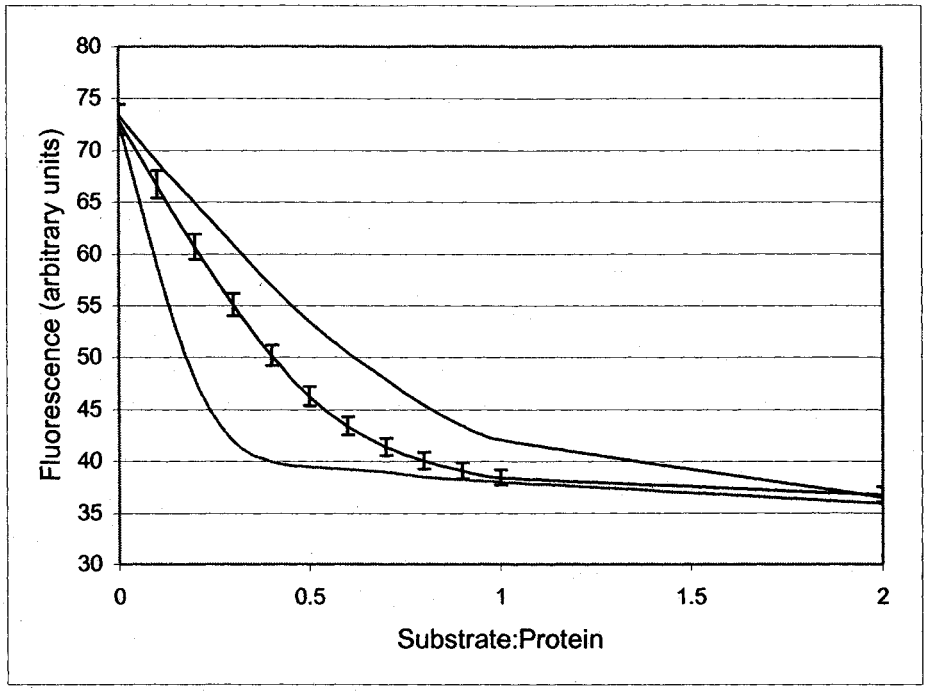



**Figure 18: Data Simulation With Vsr as Model**

The titration of Vsr with a T/G mismatch (middle curve) has a calculated dissociation constant of 25 nM. The parameters used for the simulations were 1 binding site, an initial fluorescence of 72 (arbitrary units) and an end point of 39 (arbitrary units). In the top curve, a dissociation constant of 255 nM was substituted into the fluorescence equation to generate a curve for 10 fold weaker binding. In the bottom curve, a dissociation constant of 2.5 nM was substituted into the fluorescence equation to generate a curve for 10 fold stronger binding.

| Protein | $k_{cat}$ ($min^{-1}$) | $K_D$ (nM) | $k_{cat}/K_D$ ($min^{-1}/\mu M$) |
|---|---|---|---|
| Vsr | 2.9 +/- 0.43 | 25 +/- 8 nM | 53 |
| Δ14 | 0.089 +/- 0.0044 | 52 +/- 10 nM | 1.7 |
| Δ19 | 0.069 +/- 0.0041 | 52 +/- 10 nM | 1.3 |

**Table 3: Summary of Catalytic Constants**

This table presents a summary of the catalytic constants obtained for Vsr and N terminus truncations. Catalytic constants were determined as outlined in *Materials and Methods.*

*Conditions Affecting DNA Binding and Endonuclease Activities of Vsr Endonuclease*

It was recently suggested that Vsr was only 1% active (Turner and Connolly, 2000). This was determined by kinetic burst experiments in which substrate (50 nM) was turned over in the presence of limiting protein. Similar experiments (Figure 19) with our preparations yielded different results. The single turnovers experiments were tried with 50 nM, 100 nM, 150 nM and 300 nM substrate (300 nM is shown in Figure 19), without indication of a burst. The fitted data consistently resulted in smooth hyperbolic curves.

There are several factors that could be contributing to the 1% active enzyme found by the Connolly group:

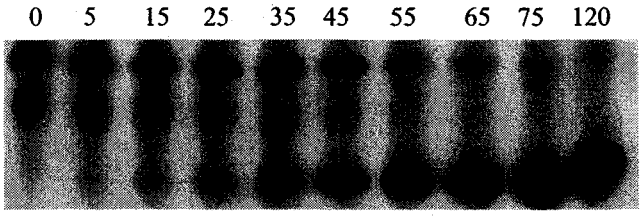
1. Connolly's group purified Vsr in the presence of 1 mM EDTA, which could have resulted in a loss of the structural zinc metal ion. Atomic absorption spectroscopy was not used to confirm the presence of zinc in their Vsr preparations. Although we did not confirm the presence of zinc, adding back additional zinc had no effect on Vsr activity.
2. Second, their Vsr was stored for two weeks in the cold and used for the entire period in endonuclease and bandshift analysis. By monitoring Vsr stored under identical conditions we found a decrease in endonuclease activity within this two-week time frame (Table 4). Furthermore we also found that Vsr aggregated with time as indicated by the bandshift experiments shown in Figure 20.
3. Third, in Turner and Connolly's experiments, Vsr was not left to pre-equilibrate with magnesium ion prior to an endonuclease assay. By examining the fluorescence spectrum of Vsr, with and without added metal, we found that the

metal ion caused structural changes that took several minutes to reach equilibrium. Non-equilibration of protein with metal prior to endonuclease assays can lead to curves that appear to have a burst, most probably due to a small fraction of Vsr already in the proper metal bound conformation while the rest of the molecules are undergoing a slower structural change.

4. Finally, the DNA substrate used by Turner and Connolly was not desalted, yet the reactions containing fifty times more substrate were kept in the same volume, thereby also increasing the SSC concentration by a factor of fifty. When we repeat this experiment with increased salt concentrations, the curves appear to have an initial burst, and then slower steady state kinetics (Figure 19B).

Combined with our fluorescence data that indicated a 1:1 binding ratio with the DNA substrate, the results of these experiments confirm that Vsr obtained and assayed under our conditions was not 1% active, but fully active.

A:




B:

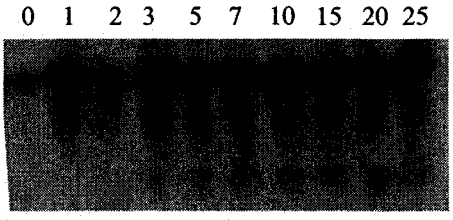



**Figure 19: Burst Experiments for Vsr Endonuclease**

**(A)** Assays were carried out as described in *Materials and Methods* with 1 μM Vsr and the exception that 300 nM substrate was used instead of 1 nM in an effort to view a burst. The top panel shows an autoradiogram of the gel produced when the substrate DNA desalted in nick spin columns is incubated with Vsr for times indicated (in seconds). **(B)** The bottom panel shows an autoradiogram of the gel produced when the substrate is not desalted. Times are given in minutes.

| Duration of Vsr storage at 4 °C | $k_{cat}$ ($min^{-1}$) |
|---|---|
| Fresh | 2.8 |
| Three days | 1.2 |
| One week | 0.58 |
| Two weeks | 0.13 |

**Table 4: Loss of Vsr Activity Upon Storage in the Fridge**

Single turnover reactions were carried out with 1 nM substrate and 1 μM protein, and analysed as described in *Materials and Methods.*

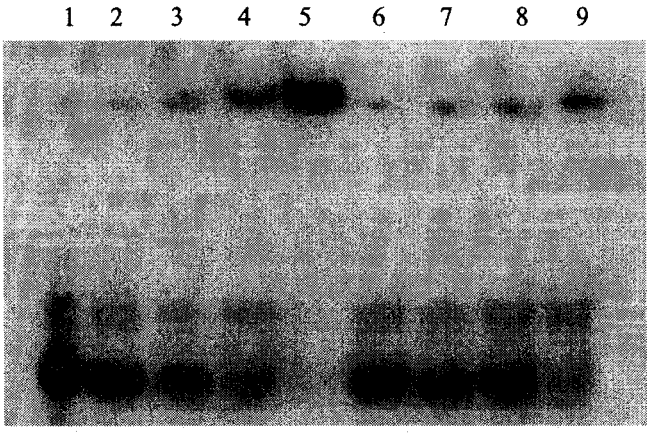



**Figure 20: Stored Vsr Aggregates With Time**

Vsr bandshifts were carried out as indicated in *Materials and Methods*. Lane 1: DNA only: lanes 2-5: increasing concentrations of Vsr (0.25, 1, 2.5 and 8 μM) that was stored for two weeks in the fridge. Note the protein/DNA complexes stuck in the wells of lanes 4 and 5. Lanes 6-9: increasing concentrations of Vsr (0.25, 1, 2.5 and 8 μM) that was thawed out and used fresh for this experiment.

*MutL Stimulation of Vsr Endonuclease Activity*

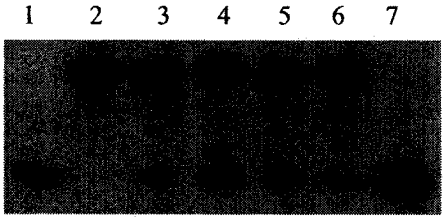
MutL was shown to stimulate the binding of Vsr to substrate DNA containing a T/G mismatch (Drotschmann *et al.,* 1998). Macintyre (unpublished) demonstrated that MutL stimulated the endonuclease activity of Vsr. MutL incubated with DNA in the absence of Vsr had no effect (Macintyre, unpublished). Figure 21A shows that MutL stimulated the activity of wildtype Vsr when the DNA substrate was in excess relative to Vsr, and Vsr would undergo multiple turnovers to cleave all the substrate. However MutL had no effect under single turnover conditions (Figure 21B and 21C). Although MutL can still associate with the N terminus truncations in a two hybrid assay (Mansour *et al.,* 2000), it could no longer stimulate the endonuclease activity of the two deletion mutants (Figure 21A), nor did it restore any activity to F67A, W68A and W86A under the same conditions (Figure 21D).

*Mutagenic Spectrum of Vsr and Mutant Proteins*

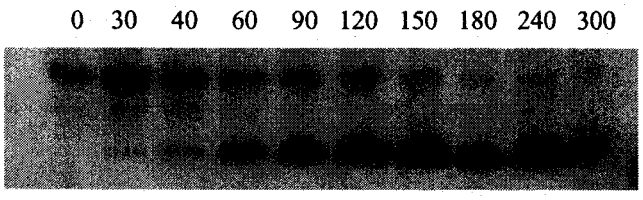
Mutagenesis of Vsr was monitored in CC108, a strain that monitors frameshift mutations (a high incidence of Lac reversion is indicative of a higher mutation rate). High transition mutation rates are diagnostic of defective methyl-directed mismatch repair. All mutant proteins were compared to Vsr. F67A, W68A and W86A mutants were as mutagenic as Vsr (Marina Siponen, Biol 490 Thesis), the Δ14 mutant half as mutagenic, and the Δ19 mutant fifteen fold less mutagenic (Figure 22). The remaining CC strains (101-107, 109-111) which monitored either base substitutions or frameshift mutations were co-transformed with either Vsr, Δ14, or Δ19 and the spectrum of mutation was identical for all three proteins (Table 5). However the absolute numbers of mutations

were decreased for the N terminus truncations. The loss in mutagenicity does not reflect a loss in the interaction between MutL and N terminus truncations (Mansour *et al.*, 2001).

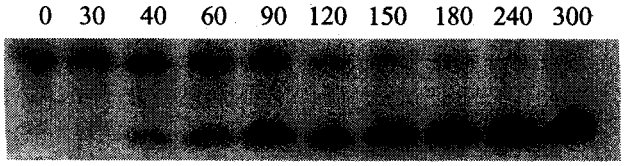
A:


1 2 3 4 5 6 7

B:


0 30 40 60 90 120 150 180 240 300

C:


0 30 40 60 90 120 150 180 240 300

D:

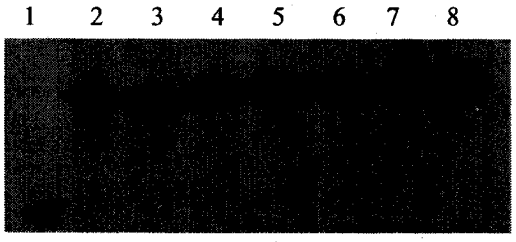

1 2 3 4 5 6 7 8

**Figure 21: MutL Stimulation of Vsr Endonuclease Activity**

**(A)** Autoradiogram of endonuclease assay. Lane 1: is a base size marker for the cleaved fragment; Lane 2: 0.3 nM Vsr; Lane 3: 0.3 nM Vsr with 0.05 nM MutL; Lane 4, 0.3 nM Vsr with 0.5 μM mutL; Lane 5: 1 nM Δ14; Lane 6: 1 nM Δ14 with 0.5 μM mutL; Lane 7: 2 μM Δ14. **(B)** Autoradiogram of single turnover experiment of 1 μM Vsr at room temperature, reaction time is given in seconds. **(C)** Autoradiogram of single turnover experiment of 1 μM Vsr in the presence of 5 μM MutL, also at room temperature, reaction times are given in seconds. **(D)** Endonuclease assay with F67A, W68A and W86A mutants, in the absence (lanes 2, 4, 6) or in the presence (lanes 3, 5, 7) of 5 μM MutL. Lane 1 is a size marker for the cut DNA, and lane 8 is substrate DNA. This endonuclease assay was carried out at the same time as the one in panel A.

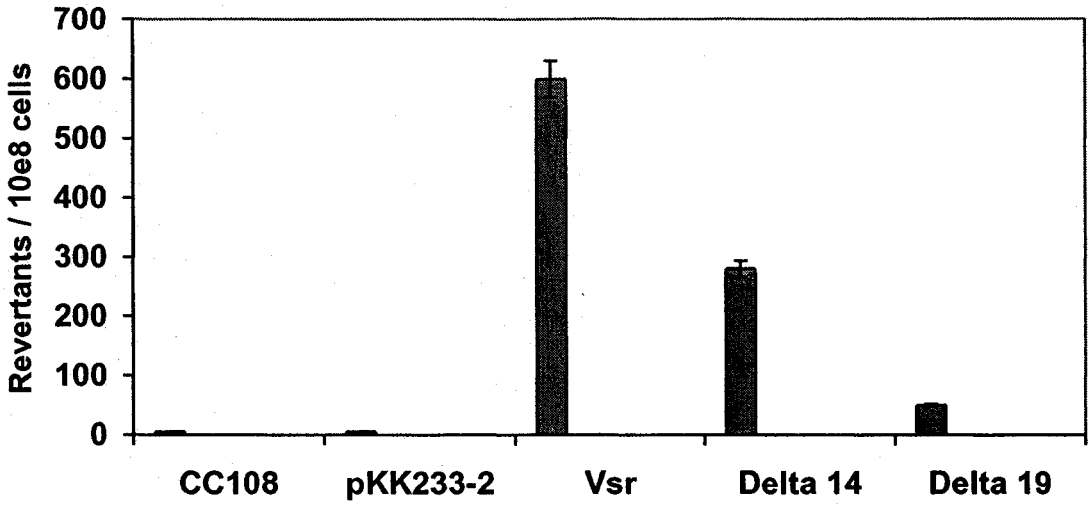



**Figure 22: *In Vivo* Mutagenesis by Vsr and N terminus Truncations**

Plating CC108 transformants on minimal lactose plates allows for monitoring increased or decreased Lac reversion. In CC108 a frameshift must occur to restore the correct amino acid sequence in *lacZ,* and the phenotype of the cell changes from $Lac^-$ to $Lac^+$ (Cupples *et al.,* 1990). For this particular assay, triplicate samples from nine different transformants were used to score Lac reversion. From left to right, CC108 strain background, pKK233-2 plasmid background, cells transformed with plasmids encoding Vsr, Δ14 and Δ19.

| Strain | Mutation | Number of Lac Revertants / $10^8$ Viable Cells | | | |
|---|---|---|---|---|---|
| | | pKK233-2 | pKKV | Published Vsr Data | pKKΔ14 |
| CC101 | AT to CG | 0.5 | 0.5 | 0.1 | 0 |
| CC102 | GC to AT | 0.5 | 24 | 27 | 15 |
| CC103 | GC to CG | 1.9 | 0 | 0 | 0 |
| CC104 | GC to TA | 1.7 | 1 | 2 | 0.85 |
| CC105 | AT to TA | 0.2 | 0.9 | 0.3 | 0.3 |
| CC106 | AT to GC | 0 | 0.3 | 12 | 0 |
| CC107 | +G | 30 | 9752 | 10942 | 4057 |
| CC108 | -G | 10 | 1955 | 1124 | 1284 |
| CC109 | -(GC) | 85 | 1643 | 1556 | 847 |
| CC110 | +A | 1.8 | 52 | 68 | 18 |
| CC111 | -A | 1.8 | 989 | 907 | 565 |
| CC113 | CTAGG to CCAGG | 4 | 270 | 223 | 104 |

**Table 5: Mutagenic Spectrum of Vsr and Δ14 N Terminus Truncation**

Mutagenesis was scored in the CC strains, which monitor specific base substitutions and frameshift mutations. For this particular assay, triplicate samples from three different transformants of each strain were used to score Lac reversion. The published results for pKKV are taken from Doiron *et al.*, 1996.

*In Vivo Assays for VSP Repair*

The mutant proteins were also assayed for VSP repair *in vivo*. As shown in Figure 23 the Δ14 mutant had a twofold difference relative to Vsr in VSP repair, while the Δ19 had a seven fold difference in activity. As *in vitro*, F67A, W68A and W86A were no longer able to carry out VSP repair (Marina Siponen, Biol 490 Thesis).

To determine whether the decrease in activity of these proteins relative to wildtype was due to reduced production *in vivo*, proteins from stationary phase cells were subjected to Western analysis and probed with anti-Vsr antibody (Figure 24). The levels of the Δ14 mutant were reduced in comparison to wildtype protein and F67A and W86A were also reduced (Figure 24A). F67A, W68A and the Δ19 mutant were present at chromosomal Vsr levels, which was apparent on the blot, as CC108 contains chromosomal *vsr*. The decrease in protein levels was unexpected; all these proteins were produced from a plasmid under control of the *trc* promoter. Purified protein probed with Vsr antibody showed that the mutant proteins react with the antibody almost as well as the wildtype protein (Figure 24B). The difference was probably not large enough to account for the decreased protein levels observed in stationary phase cells.

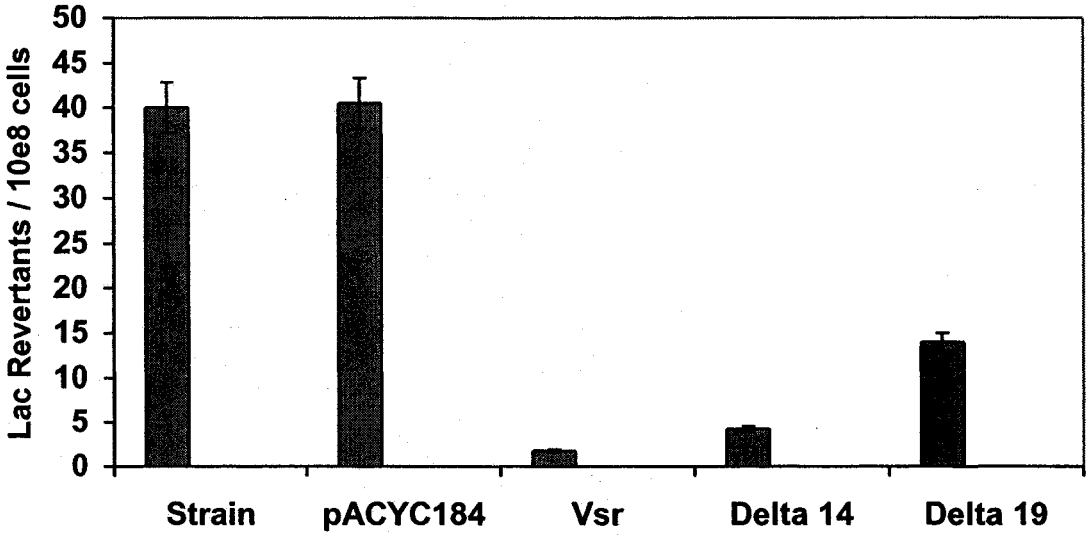



**Figure 23: VSP Repair Activity *In Vivo***

To assay Vsr mutants *in vivo* for VSP repair, mutant DNA fragments are cloned into pACYC184 with the *trc* promoter, and expressed in CC503, that contains a kanamycin cassette in *vsr*. This strain contains *dcm*, because Dcm is required to produce the T/G lesion for Vsr to repair. A CCAGG to CTAGG mutation in *lacZ* converts a glutamine codon (CAG) at position 461 to a nonsense codon (TAG). In the presence of an amber suppressor gene carried on a multicopy plasmid, this codon is translated as glutamic acid, the only amino acid at that site which would give a functional β-galactosidase (Cupples and Miller, 1988). A high level of mutation is indicative of non-functional VSP repair. By quantitatively measuring Lac reversion it is possible to assess the repair capabilities of mutant proteins compared to wild-type Vsr endonuclease. For this particular assay, triplicates from nine different transformants were used to score Lac reversion. From left to right, strain background, empty plasmid background, cells that contain Vsr, cells that contain Δ14 and cells that contain Δ19.

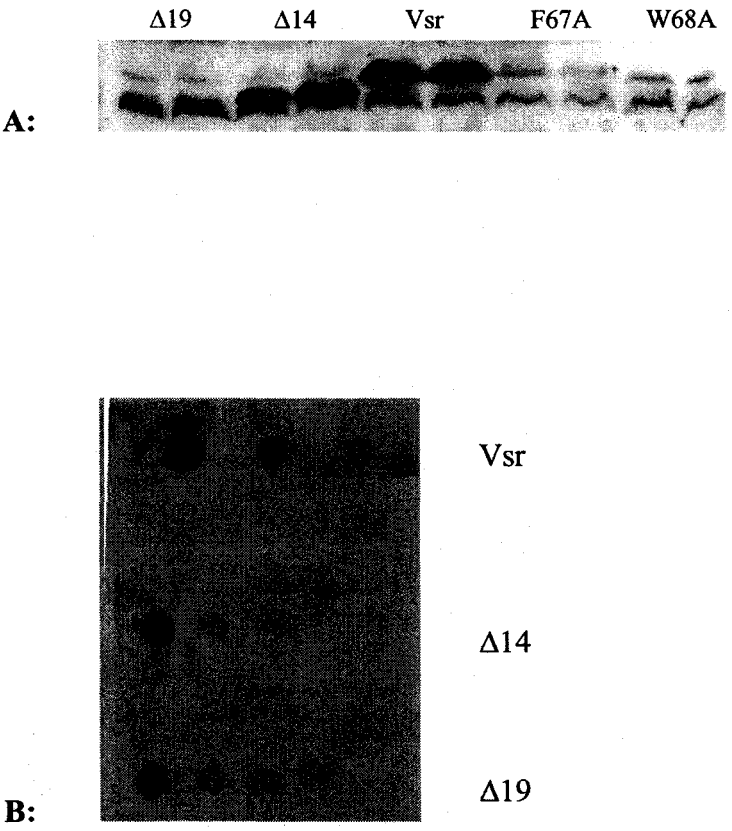



**Figure 24: Stationary Phase levels of Vsr and Mutant Proteins**

**(A)** CC108 transformants of Δ19, Δ14 and Vsr, F67A and W68A were grown to stationary phase $O.D_{600}$=1.7, cells were harvested, resuspended in 20 mM Tris-HCl pH 7.5 and electrophoresed on a 16% SDS PAGE, proteins from whole cells were transferred to a membrane and subjected to Western blot analysis. **(B)** Purified Vsr (top panel), Δ14 N terminus truncation (middle panel), and Δ19 N terminus truncation (bottom panel), spotted onto a PVDF membrane at successive 10 fold dilutions (0.5 μg, 50 ng and 5 ng), and probed with anti-Vsr. These spots were quantitated with the LI-COR infrared scanner, and have 15 % variability between equimolar spots of Vsr and N terminus truncations.

*Excess Bands On Western Blots*

Almost every Western blot presented in this thesis (Figures 24A, Figure 30, Figure 31, Figure 32, Figure 33) contains bands that are smaller in molecular weight than Vsr. We are unsure of what these smaller species are, but there are several possibilities:

1. These bands may have been Vsr-specific, such as Vsr protein being translated from the methionine at position 14. This seems unlikely as a stop codon placed between the first and second start site of Vsr abolished VSP repair (Dar and Bhagwat, 1993); however that same study used lambda phage crosses to score VSP repair and is not as sensitive as the one we use to monitor VSP repair.
2. These lower bands may have been Vsr-specific degradation products. This possibility seems more likely as we have seen shorter Vsr forms *in vivo* and *in vitro* (discussed in more detail in a subsequent section).
3. Alternatively, these bands may be non-specific. The rabbit used to make the polyclonal Vsr antibody almost certainly had pre-existing antibodies to *E coli.* proteins. The Vsr antibody was made with His-tagged Vsr purified with a cobalt resin. The final preparation may have been impure so that antibodies could have been generated to recognize other proteins. This also seemed less likely as bands recognized are consistently shorter than Vsr and sometimes the intensity rivals that of full length Vsr. Preliminary evidence in our laboratory with a *vsr* deficient strain indicate that these bands are Vsr-specific.

Shorter forms of Vsr had been observed *in vivo*, upon Western blot analysis, under conditions where Vsr was over-expressed (Pitsikas and Cupples, unpublished). One published report claimed that Vsr was unstable *in vitro,* with a purified preparation of Vsr losing 12-14 amino acids from its N terminus (Glasner *et al.,* 1995). Due to the flexibility of the N terminus, X-ray crystallography of Vsr in the absence of DNA required the use of a trypsinized, Δ20 proteolysed Vsr fragment. The proteolysed Δ20 Vsr was still active *in vitro* (Tsutakawa, 1999a). Why is it that under certain conditions shorter forms of the endonuclease are present, and is there any biological significance? The remaining results section deals with experiments aimed at answering these questions.

*Vsr Truncation In Vitro*

As mentioned earlier, while doing control bandshift experiments to assay Vsr, it was noted that upon storage at –20 °C, or at 4 °C, Vsr aggregated with time (Figure 20). However electrophoresing Vsr stored in the fridge for several days on SDS PAGE indicated that shorter forms were also becoming prevalent in the preparation (Figure 25A). Short form Vsr could have proceeded by either proteolysis or autocatalysis. A brief description of the two mechanisms along with experiments to delineate between the two follows.

At least forty enzymes capable of hydrolysing peptide bonds have been discovered in *E. coli* (Neidhardt, 1996). The rules that govern susceptibility to proteolysis remain unclear. Bacteria appear to have no general mechanism for marking proteins for degradation and there is no evidence to suggest that substrates for proteolysis are sequestered in any way.

A common property shared by all of these proteases is the requirement of energy for *in vivo* protein degradation. However ATP hydrolysis is not always required; ATP may function as an allosteric effector as in the case with Clp protease, or by keeping a protease in an open conformation, as is the case of Lon protease (Neidhardt, 1996). The majority of these enzymes are classified as serine proteases, however many of them are atypical and are placed in distinct classes (Neidhardt, 1996).

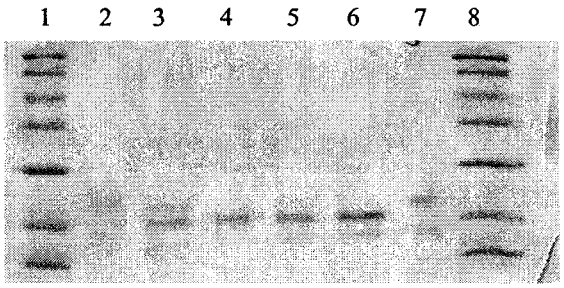
If Vsr truncation proceeds by proteolysis it has to be a site-specific protease, unlike Clp and Lon that are highly processive, and would result in degradation. The easiest experiment to prove or disprove proteolysis is to add protease inhibitors to an enzyme preparation and confirm whether or not truncation is taking place.

If an enzyme is co-purifying with a protease, then dilution of the enzyme should result in dilution of the protease, so that the proportion of short form with respect to long form should decrease at lower protein concentrations. Reactions that are independent of protein concentration are suggestive of autocatalysis. Establishing that a reaction is autocatalytic is difficult (reviewed in Little, 1993). Any self-processing reaction involves an active site and a substrate site in the same molecule, with the rate being limited by the chemistry of the reaction as the local concentration of active site to substrate can be high (Fersht, 1985). Sometimes self-processing reactions are observed only under special conditions i.e. high pH, high salt, high metal concentrations, or reaction rates may vary in the presence of effectors (Little, 1993). The discovery of autocatalysis in many cases has been a matter of luck in trying the right reaction conditions (Little, 1993). Intermolecular

reactions can often be confused with proteolysis; however, the hallmark of an intramolecular reaction is that its rate constant should be insensitive to dilution (Little, 1993). In any case, proteolysis must be rigidly excluded as an appropriate mechanism of truncated protein formation, before autocatalysis can be declared.

Many different conditions were investigated for Vsr truncation. Prolonged storage of thrombin cleaved Vsr led to 100% Vsr truncation. This preparation remained stable and further incubation for up to several months showed no other products being formed from the shorter form, nor a loss of protein (Figure 28 lane 9). N terminus truncations stored under identical conditions remained stable (Figure 25B lanes 6 and 7). Incubation of F67A (Figure 25B lane 4) and W68A (not shown), were inconclusive because of pH differences (pH 2 versus pH 7).

A:




B:

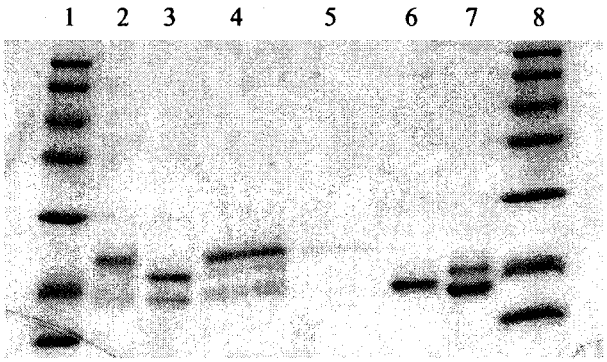



**Figure 25: Thrombin Cleavage and Observance of Vsr Truncation**

**(A)** 16% SDS PAGE depicting aliquots of protein removed at different time points after cleavage initiation by thrombin for lanes 2-4. These proteins were then stored in the fridge for several days, lanes 5-6; lanes 1 and 8: molecular weight markers (116, 77, 66, 45, 31, 21.5, and 14.4 kDa), lane 2: Vsr prior to thrombin cleavage; lane 3 and 4: 2 and 6 hours after addition of thrombin; lane 5 and 6: thrombin cleaved Vsr incubated in fridge for lane 36 and 50 hours; lane 7: Vsr non-thrombin cleaved incubated in fridge for a week. **(B)** Lanes 1 and 8: molecular weight markers (as above); lane 2: non-thrombin cleaved Vsr incubated in fridge for one week; lane 3: thrombin cleaved Vsr one week in fridge; lane 4: F67A two hours post thrombin cleavage; lane 5; F67A thrombin cleaved, stored in fridge for two days; lane 6: thrombin cleaved Δ14; lane 7: thrombin cleaved Δ19 (incomplete cleavage).

To determine the identity of the N terminus of the Vsr truncation, purified thrombin cleaved protein was incubated for several days in the fridge, electrophoresed by 16% SDS PAGE, transferred to a nylon membrane and Coomassie stained. The short band was excised, and N terminally sequenced by Sheldon Biotechnology Centre (Montreal, Quebec), revealing a Δ15 Vsr protein with an N terminus arginine residue. Truncated Vsr *in vivo* would be equivalent to a Δ14 N terminus truncation; it has been shown that like Vsr (Tsutakawa *et al.,* 1999a), a Δ14 *in vivo* also loses its N terminus methionine, thus becoming a Δ15 Vsr (Turner and Connolly, 2000).

*Investigation of Mechanism Leading to Short Form Vsr*

Although no other bands were apparent on the gel with Coomassie staining other than the two forms of Vsr protein, proteolysis and autocatalysis were investigated as mechanisms for Vsr truncation. In an effort to determine whether a protease was co-purifying with Vsr, protease inhibitors were added to Vsr immediately after thrombin cleavage of the Histidine tag; an almost equivalent amount of short form relative to full length was present in the treated versus non-treated sample (Figure 26). This did not entirely preclude proteolysis, since there is no protease inhibitor that can eliminate the activity of every possible protease in the cell. In another experiment, thrombin-cleaved Vsr was diluted, incubated overnight at room temperature, and subjected to Western analysis to quantitate the protein forms present in each dilution. Almost no difference was observed when the enzyme was diluted tenfold (Figure 27). Further dilution was impossible since the amount of protein would have been below the level of detection of the Vsr antibody.

The experiment in Figure 27 was carried out several times, and was inconclusive. However it is not necessarily inconsistent with intramolecular autocatalysis.

The pH dependence for the conversion to short form (pH 4.5 to 9.5) was also monitored with no difference observed amongst the different pHs (Figure 28). Many other *in vitro* combinations of Vsr protein incubated with homoduplex DNA, T/G mismatch substrate, high metal concentrations and with MutL, showed no significant difference between the two Vsr forms under the different reaction conditions (Figure 29) – making it impossible to determine the precise mechanism leading to Vsr truncation *in vivo*.

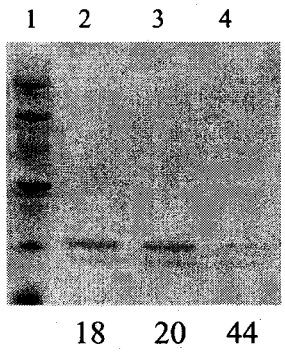



**Figure 26: Vsr With and Without Protease Inhibitors**

16 % SDS PAGE, stained with Coomassie Brilliant Blue. This experiment represents one of several independent determinations. Freshly thrombin cleaved Vsr incubated overnight at room temperature in either the absence (lane 2) or presence (lane 3) of an EDTA free protease inhibitor, (Complete tablet, Boehringer Mannheim). The starting protein did not contain any short form and was equivalent to the protein in lane 4, in Figure 22A. Lane 1 and 4 are included as size markers. Lane 1: molecular weight markers (77, 66, 45, 31, 21.5, and 14.4 kDa); Lane 4: Vsr that has been stored in the fridge for two weeks, allowed to form both species. The amount of short form quantitated with the LI-COR scanner as a percentage of protein in each lane is given under the gel.

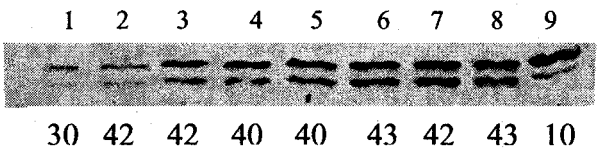



**Figure 27: Western Blot Analysis of Vsr Dilutions**

This experiment represents one of three independent determinations. A thrombin cleavage reaction was set up for 8 hours on a Vsr preparation. Different dilutions of the protein were made up to 20 μL in 1X Vsr reaction buffer. From left to right, 25, 75, 120, 150, 175, 200 and 220 ng of Vsr, incubated for three days at room temperature. The last lane is an aliquot of Vsr, following thrombin cleavage but prior to dilution, which was frozen away in 1XSDS buffer to use as a marker. These proteins were electrophoresed on a 16 % SDS PAGE, and transferred to a PVDF membrane for Western analysis. The percentage of Vsr present as a truncation was determined with the LI-COR infrared scanner the values are shown under the membrane.

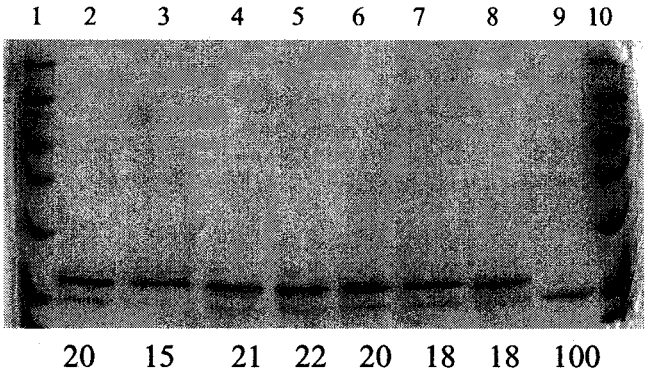



**Figure 28: Cleavage to Short Form is Not pH Dependant**

16 % SDS PAGE, stained with Coomassie Brilliant Blue. This experiment represents one of several independent determinations. Vsr was subjected to thrombin cleavage. Aliquots of Vsr were incubated overnight in different pH reaction buffers. The starting protein did not contain any short form and was equivalent to the protein in lane 4, in Figure 22A. Lanes 1 and 10: 116, 77, 66, 45, 31, 21.5, and 14.4 kDa molecular weight markers; lane 2: size marker for Vsr (mixture of two forms); lane 9: three month stored, thrombin cleaved Vsr, size marker for truncation; lanes 3-8: pH 4.5 Sodium Acetate buffer, pH 5.5 MES-NaOH buffer and pH 6.5, 7.5, 8.5, and 9.5 Tris-HCl buffers. The percentage of Vsr present as a truncation was determined with the LI-COR infrared scanner and the values are shown under the gel.

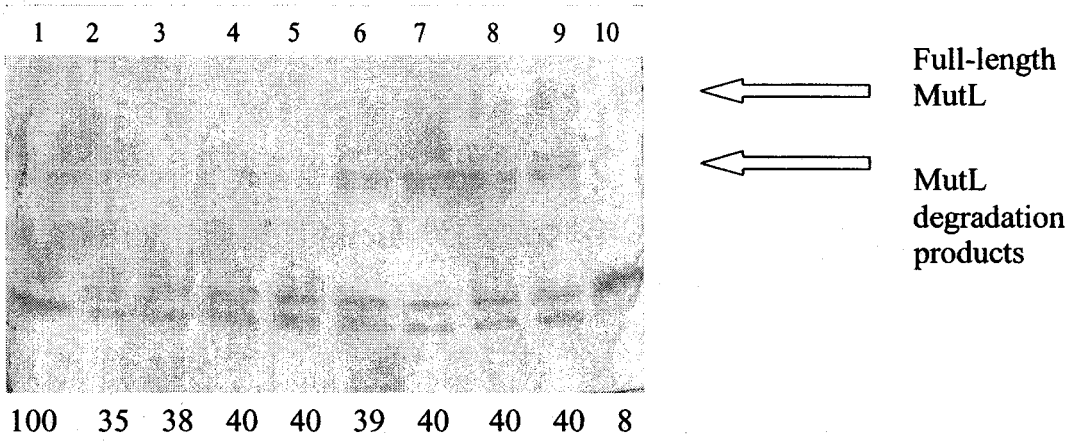



**Figure 29: Vsr Truncations Under Varied Conditions**

Vsr was thrombin cleaved, incubated at 37°C over a three-day period under different reaction conditions. Lane 1: three month stored, thrombin-cleaved Vsr; lane 2: Vsr in reaction buffer; lane 3: Vsr in the presence of homoduplex DNA and reaction buffer; lane 4: Vsr in the presence of a T/G mismatch and reaction buffer; lane 5: Vsr in the presence of 100 mM magnesium chloride; lane 6: Vsr in the presence of 1 μM MutL; lane 7; Vsr and MutL in the presence of a homoduplex; lane 8: Vsr and MutL in the presence of a T/G substrate; lane 9: Vsr and MutL in the presence of 100 mM magnesium chloride, Lane 10: starting Vsr material. The percentage of Vsr present as a truncation was determined with the LI-COR infrared scanner and the values are shown under the gel.

*Growth Phase Regulation of Vsr and Mutant Proteins*

It has been previously shown that levels of Vsr are low in growing cells, and then increase in stationary phase cells (Macintyre *et al.,* 1999). In fact, the sensitivity of the LI-COR infrared scanner allows us to see that the amount of Vsr increases throughout log phase and does not just increase as cells enter stationary phase as suggested by Macintyre (Figure 33A).

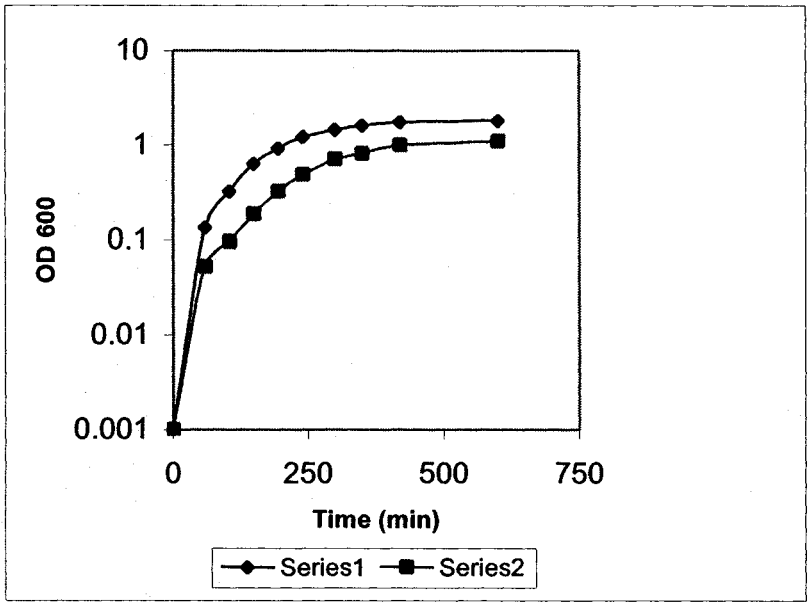
In Figure 30, samples were removed at different time points, and protein levels were monitored through Western blot analysis. pKKV was transformed into a wildtype strain, C600, as well as a C600 derivative, SG12060 that was *clp* and *lon* deficient, to see if growth phase regulation would be abolished by lack of either of these proteases. The data suggest that Clp and Lon do not play a role in Vsr growth phase regulation. Surprisingly, removal of the N terminus abolished Vsr growth phase regulation in a wildtype CAG strain (Figures 31-32); the cells made stationary phase levels of Vsr in log phase. Analysis of F67A and W68A in the same wildtype strain demonstrated that they retained their growth phase dependence (Figure 35A).

*Overexpression of MutL and its Effect on Vsr and Mutant Proteins*

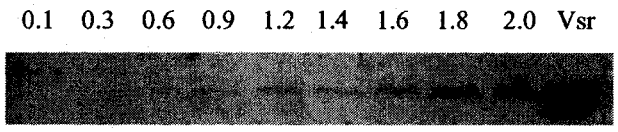
Chromosomal Vsr levels were monitored during culture growth by Western blot analysis of whole cells extracts transformed with a plasmid containing MutL. Chromosomal evidence indicated loss of growth phase regulation of Vsr (data not shown). Experiments were also carried out using Vsr expressed from a plasmid to amplify the signal. A stabilization of Vsr was observed (Figure 33B). This was a MutL specific response as

Vsr levels were unchanged when the overexpressed protein was MutS (Figure 34). F67A and W68A were also stabilized by MutL overexpression (Figure 35). As indicated above, the N terminus truncations were stabilized in log phase in the absence of MutL, and no change in the pattern was observed in the presence of MutL.

A:




B:




C:

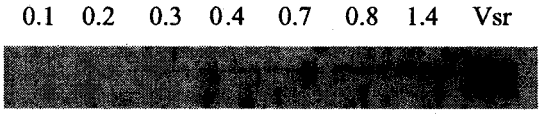
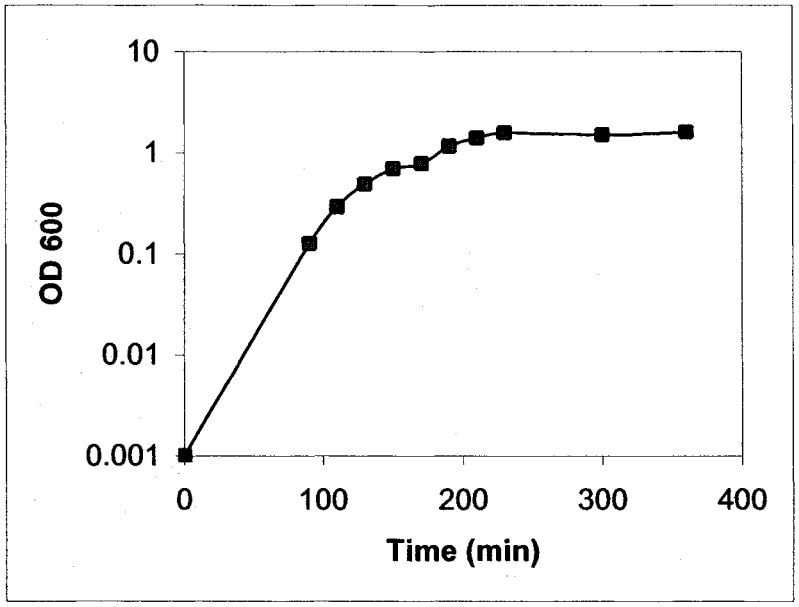



**Figure 30: Production of Vsr in Growing and Stationary Phase *E. coli* Cells**

**(A)** Cultures were initiated from a 1:100 dilution of an overnight-saturated culture. Culture growth was measured by optical density at 600 nm ($OD_{600}$) Samples were removed from C600 (series 1) **(B)** and SG12060 (series 2) **(C)** at $OD_{600}$ shown and subjected to Western analysis. Each Vsr Western blot also had a positive control that contains both Vsr forms. These samples could not be quantitated, however they represent two of eight Western blots that depict identical results.

A:




B:

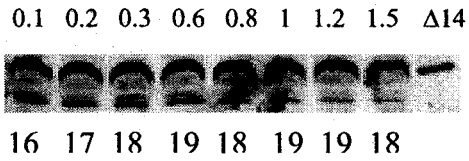
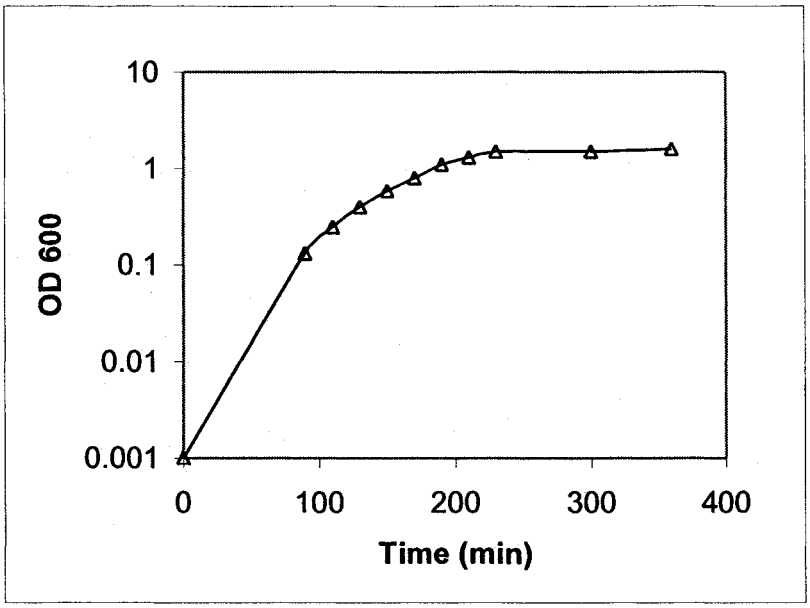



**Figure 31: Production of Δ14 in Growing and Stationary Phase *E. coli* Cells**

Cultures of CAG12033 were initiated from a 1:100 dilution of an overnight-saturated culture. Culture growth was measured by optical density at 600 nm ($OD_{600}$)**(A)**. Samples were removed from cells transformed with pKKΔ14 at the $OD_{600}$ shown, and subjected to Western analysis **(B)**. The positive control on this Western blot was purified Δ14 protein that was thrombin cleaved and stored in the fridge. The amount of antibody-bound complexes in each lane (top band) was quantitated with the LI-COR infrared scanner, and is shown under the blot.

A:




B:

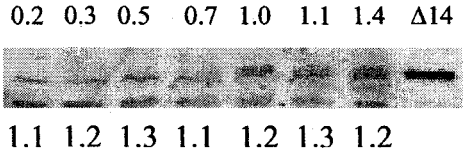
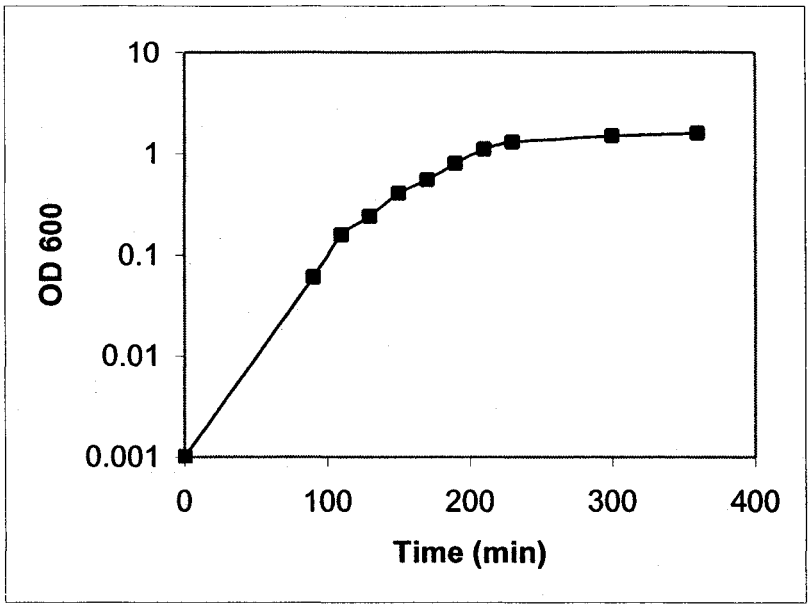


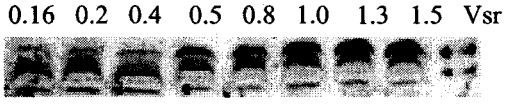

**Figure 32: Production of Δ19 in Growing and Stationary Phase *E. coli* Cells**

Cultures of CAG12033 were initiated from a 1:100 dilution of an overnight-saturated culture. Culture growth was measured by optical density at 600 nm ($OD_{600}$)**(A)**. Samples were removed from cells transformed with pKKΔ19 at $OD_{600}$ shown, and subjected to Western analysis **(B)**. The positive control on this Western blot was purified Δ14 protein that was thrombin cleaved, and stored in the fridge. The amount of antibody-bound complexes in each lane was quantitated with the LI-COR infrared scanner, and is shown under the blot.

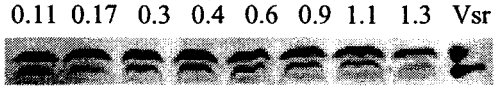
A:


10
1
0.1
0.01
0.001
OD 600
0
100
200
300
400
Time (min)

B:


0.16 0.2 0.4 0.5 0.8 1.0 1.3 1.5 Vsr
1.6 1.7 1.5 10 11 13 19 22

C:


0.11 0.17 0.3 0.4 0.6 0.9 1.1 1.3 Vsr
11 11 10 11 11 12 13 11

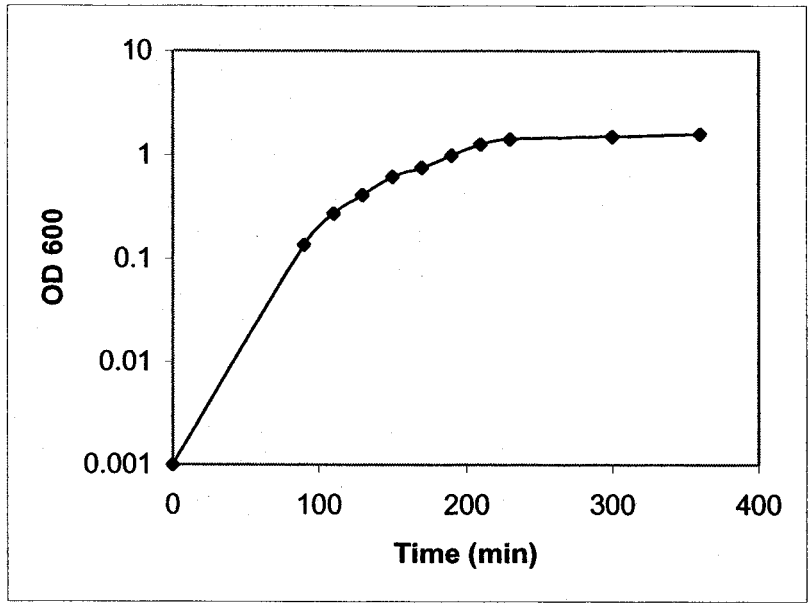
D:


4.6 4.7 5.1 5.3 5.0 4.9 4.7 4.5

**Figure 33: Production of Vsr in Growing and Stationary Phase *E. coli* Cells With Over-Expressed MutL**

Cultures of CAG12033 were initiated from a 1:100 dilution of an overnight-saturated culture. **(A)** Culture growth was measured by optical density at 600 nm ($OD_{600}$). Samples were taken from cells transformed with pKKDV at $OD_{600}$ shown, and subjected to Western analysis. **(B)** The membrane was probed with Vsr antibody. **(C)** When co-transformed with pMutL, the Vsr was no longer growth phase dependent. **(D)** MutL Western, membrane in (C), stripped with 0.2M HCl, over a two-hour period, re-blocked and re-probed with anti-MutL. The amount of antibody-bound complexes in each lane was quantitated with the LI-COR infrared scanner, and is shown under the blot. Each Vsr Western blot had a positive control of purified protein that contained both Vsr forms.

A:




B:

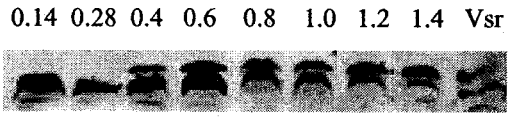
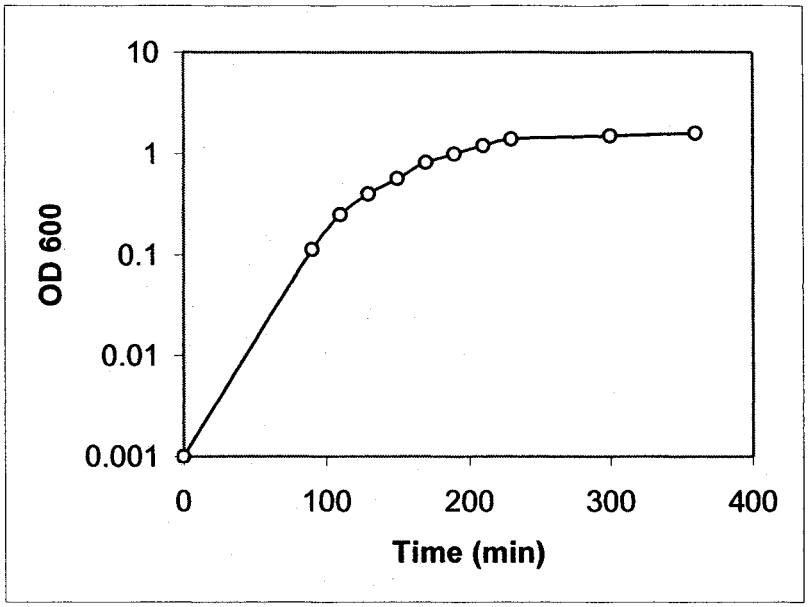


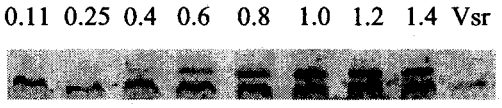

**Figure 34: Production of Vsr in Growing and Stationary Phase *E. coli* Cells With Over-Expressed MutS**

Cultures of CAG12033 were initiated from a 1:100 dilution of an overnight-saturated culture. Culture growth was measured by optical density at 600 nm ($OD_{600}$)**(A)**. Samples were removed from cultures co transformed with pMutS at $OD_{600}$ shown and subjected to Western analysis **(B)**. The Western blot had a positive control of purified protein that contained both Vsr forms. The amount of antibody-bound complexes in each lane was quantitated with the LI-COR infrared scanner, and is shown under the blot.

A:




B:




C:

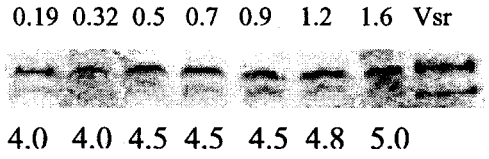



**Figure 35: Production of F67A in Growing and Stationary Phase *E. coli* Cells**

Cultures of CAG12033 were initiated from a 1:100 dilution of an overnight-saturated culture. Culture growth was measured by optical density at 600 nm ($OD_{600}$)**(A)**. Samples were taken from cells transformed with pKKF67A at $OD_{600}$ shown, and subjected to Western analysis. **(B)**. When co-transformed with pMutL, F67A was no longer growth phase dependent **(C)**. Each Vsr Western blot had a positive control of purified protein that contained both Vsr forms. The amount of antibody-bound complexes in each lane was quantitated with the LI-COR infrared scanner, and is shown under the blot.

## Discussion

*Fluorescence Spectroscopy of Vsr and Vsr Mutant Proteins*

The fluorescence spectrum of a protein is very sensitive to changes in its physical and chemical environment (Clarke, 1996). Therefore fluorescence emission of tryptophan and tyrosine residues was used to monitor overall conformation of Vsr and mutant Vsr proteins. It was apparent from the fluorescence data that the Vsr truncations and native Vsr had identical spectra in the absence of metal ions ($\lambda_{max}$ emission 342 nm). The crystal structure showed that Vsr contained two $Mg^{+2}$ in the proposed active site, however it was unclear whether or not these metals were just cofactors that were required for enzymatic activity or whether they were structural ions as well (Tsutakawa *et al.,* 1999a).

From the fluorescence data it appeared that metal ion coordination changed the structure of the native protein, causing a red shift in wavelength, $\lambda_{max}$ emission 348 nm, which signified increased solvent-exposure of the tryptophan residues that gave rise to the fluorescence spectrum (Table 2). An identical shift occurred upon $Ca^{+2}$ binding, which indicated that calcium ions may be used as non-reactive magnesium ion equivalents (not shown). This was consistent with data that showed that $Ca^{+2}$ stimulated DNA binding but not endonuclease activity (Turner and Connolly, 2000, Tsutakawa, unpublished). It has also been demonstrated that chelation of the metal ion by EDTA abolished DNA binding (Tsutakawa *et al.,* 1999a). These results taken altogether suggest that the magnesium ions are not only catalytic, as they are involved in the mechanism for hydrolysis, but are

structurally important, with proper conformation an absolute requirement for substrate binding.

When the fluorescence spectra of Vsr and N terminus truncated mutants in the presence of metal ion were compared to spectra of native Vsr (Figure 10A), a difference in the two spectra was observed; the overall tertiary structure of the N terminus truncation mutants was slightly different relative to native Vsr. Fluorescence emission maxima of the N terminus mutants only shifted to 345 nm rather than 348 nm (Table 2). Tryptophan residues emit in this region of the fluorescence spectrum and a 3 nm difference between Vsr and N terminus truncations indicate that the N terminus truncations have less solvent-exposed tryptophans relative to Vsr. Although the average solvent exposure of all the tryptophan residues in Vsr determines its fluorescence spectrum, the 3 nm shift in wavelength indicates very small differences in overall conformation.

On the basis of the crystal structure, Trp68 becomes a likely candidate for the 3 nm shift seen with Vsr and N terminus truncations. When the crystal structure of truncated Vsr ($\Delta$20), was compared to the co-crystal (full length) Vsr structure, a possible structural basis for the fluorescence changes was suggested. In the co-crystal, Trp68 is turned out of the plane of the main protein surface, whereas in the structure of Vsr alone, it is not. From the crystallization studies alone it seemed that the presence of bound DNA was causing Trp68 to turn outward. However the fluorescence data suggested that this shift could be attributed to the presence of the N terminus. The N terminus is flexible and may allow the tryptophan residues to adopt a slightly different, more solvent exposed

conformation. Furthermore, Trp68 stacks with the mispaired T in the DNA recognition sequence; this may also explain why the N terminus truncations were not as active as Vsr *in vitro* and *in vivo*.

The crystal structure of Vsr endonuclease in the presence of a DNA substrate showed that three aromatic residues Phe67, W68 and W86 intercalated into the major grooves of a cleaved DNA intermediate, distorting and widening both the major and minor groove of the DNA. The literature indicated that steric recognition was important in T/G mismatch recognition by Vsr as most contacts were directed towards the backbone and not the mismatch itself (Morikawa and Shirakawa, 2000). This was a novel form of recognition that allowed the N terminus alpha helix to bind, which normally would have been a steric impossibility (Morikawa and Shirakawa, 2000). F67 stacked with the A or T base in the recognition sequence (CT/GWGG where W =A or T), W68 stacked with the mispaired T, and W86 stacked with the sugar opposite the T. Since the DNA in the crystal structure was cleaved, it is not clear whether the placement of the aromatic residues precedes or follows the enzymatic reaction.

The alanine amino acid substitutions for F67, W68, and W86 caused conformation changes. In the presence of metal ions, tryptophan residues in these mutants were less solvent exposed than Vsr endonuclease and N terminus truncations (Table 2). F67, W68, and W86 were shown to stack with the recognition sequence and mismatch; if these residues are less solvent exposed, access to DNA binding may be limited or denied. A definitive conclusion cannot be made. A most curious attribute of Vsr variants with these

aromatic amino acid substitutions was that they did not undergo a structural change upon the addition of metal ion, suggesting a lack of metal binding (Table 2). In the crystal structure, His69, His64 and the backbone atoms of Thr63, Gly65 and Glu136 may slightly re-arrange to co-ordinate with the metal ion (Tsutakawa *et al.,* 1999a). Although none of these aromatic residues (F67, W68, W86) is directly involved in metal binding, F67 and W68 and W86 are proximal to His69, in the three-dimensional crystal structure. His69 is directly involved in metal co-ordination - a conformational shift in these mutants may have caused His69, His64, Thr63, Gly65 and Glu136, that coordinate metal ion to become improperly spaced for metal ion co-ordination. Changing these amino acids to more conservative residues, i.e., phenylalanine to tyrosine or tryptophan (both are larger and tyrosine would add a hydroxyl), and tryptophan to either phenylalanine or tyrosine (both are smaller, and tyrosine would add a hydroxyl), in the future may help reconstitute proper metal bound structure and DNA recognition, by allowing the protein to restack, with its steric partner. This would also give a better measure of the individual amino acid residue's function.

*Kinetics of Vsr and Mutant Proteins*

Catalytic and dissociation constants have been previously cited for Vsr endonuclease (see *Introduction*). However, for a variety of technical reasons, the proteins used were a mixture of N terminus truncations and wildtype protein (Glasner *et al.,* 1993), protein that cleaved only 20% of the DNA in the absence of BSA (Fox *et al.,* 2000) or protein that was found to be only 1% active (Turner and Connolly, 2000). Thus it was important to investigate the kinetics for our enzyme preparations. It was found (Figure 13, Table 3)

that our preparations had a higher catalytic turnover 2.9 +/- 0.43 $min^{-1}$ for Vsr in comparison to 0.39 $min^{-1}$ +/- 0.02 $min^{-1}$ (Turner and Connolly, 2000) for a non-methylated substrate. Our dissociation constant for fully active Vsr was determined to be 25 +/- 8 nM (Figure 17, Table 3), while their dissociation constant was 255 +/- 33 mM for fully active Vsr, or 2.3 +/- 0.3 nM for 1% active Vsr, with a non-methylated substrate. Turnover was found to decrease dramatically under storage and assay conditions used by other workers (Figure 19). This may have been due to protein aggregation, and Vsr conversion to short form, which we observed under these conditions using our preparations (Figure 20, Figure 25).

Using the intrinsic fluorescence of Vsr as a probe for DNA binding (Figure 17) had several advantages. The fluorescence spectrum of an aromatic group is extremely sensitive to its environment (Clarke, 1996), making the fluorescence assay more sensitive than that of gel shifts. Increased sensitivity allows for analysis without foreknowledge of the signal change at saturation (Lohman and Bujalowski, 1991). Bandshifts were smeary and required higher concentrations of protein (Figure 16). Quantitative data was obtained for the N terminus truncations that did not shift the DNA in a bandshift assay (Figure 16). Furthermore, non-canonical substrates (substrates that differ in the recognition sequence by one base pair) can also be used for quantitative analysis. Fluorescence occurred directly in solution, with no complication introduced by the method itself. This assay was quick to perform, occurred in real-time and was non-destructive.

A solution assay can also provide some information about the catalytic mechanism, which cannot be inferred from a bandshift. For instance the fact that only a quenching in the tryptophan fluorescence signal was observed upon DNA binding and not a shift in emission wavelength, indicated that water was included in the DNA-protein interaction (Lakowicz, 1983). This corroborates the proposed catalytic mechanism, which involves magnesium water clusters in the active site activated by His69 for nucleophillic attack (Tsutakawa *et al.*, 1999b).

Although not apparent from gel shift analysis, N terminus truncations bound to the DNA (Figure 16). This was reasonable, as 20% of Vsr was estimated to be buried in the DNA via the aromatic residues in the major groove (Tsutakawa *et al.*, 1999b) and that the interactions were stabilized by extensive steric protein/DNA interactions. However given the decreased catalytic turnover and slightly higher dissociation constant of the truncated Vsr's, these interactions were clearly sub optimal, likely due to the altered conformation of these mutant proteins in the presence of magnesium ion. The decreased catalytic turnover was possibly due to the sensitivity of catalysis to the exact positioning of His69 relative to the scissile phosphate, as His69 is thought to activate the water required for nucleophillic attack (Tsutakawa and Morikawa, 2001). Since the fluorescence spectra suggested that both N terminus truncations had a less solvent exposed tryptophan, perhaps Trp68, relative to wildtype, this may have caused positioning of His69 to also be slightly less accessible to solvent. Both N terminus truncations exhibited similar kinetic constants *in vitro*. The Δ19 N terminus truncation was present at approximately chromosomal Vsr levels in stationary phase cells, one would assume that its VSP repair

activity *in vivo* should have been comparable to the Δ14 mutant. Only one chromosomal copy of *vsr* is required for efficient VSP repair *in vivo* (Macintyre *et al.*, 1999). However, given the decreased kinetic parameters of the N terminal deletion mutants (Table 3), the differences in VSP repair activity could be due to the fact that in log phase cells, the Δ14 is present at a higher concentration than the Δ19 mutant.

Based on our results and indirect data presented by other groups, a function of the N terminus *in vivo* could be to help anchor the protein to the DNA once cleavage has occurred. The biological significance would be to protect the nicked DNA from non-specific repair enzymes, such as DNA ligase, until the arrival of the next protein in the specific repair pathway. Vsr has a low turnover rate in comparison to other endonucleases in its class, which suggests it remains bound to product once it has cleaved (Tsutakawa and Morikawa, 2001). There are several lines of experimental evidence suggesting that this is so. The DNA isolated from bandshift assays in magnesium-containing reaction buffer resulted in fully cleaved DNA, indicating that the shifts were entirely due to protein-product complexes (this study and Connolly and Turner, 2000). A smaller proportion of Vsr may form a pre-ordered DNA bound N terminus prior to cleavage, which was suggested by the fact that the gel shifts in buffer containing calcium ions required far more protein than those obtained using magnesium ions. Therefore, it appears that formation of the N terminus clamp is concomitant with cleavage; in support of this theory, despite exhaustive screening formation of a co-crystal with an ordered DNA bound N terminus was *not detected* without the presence of a *cleaved* DNA intermediate (Tsutakawa *et al.*, 1999b).

Other examples of DNA repair enzymes, such as MutY glycosylase (Zharkov and Grollman, 1998) and AP endonuclease (Waters *et al.,* 1999) remain associated to the DNA for substrate protection. Therefore the differences observed in VSP repair activity *in vivo* for the two N terminus truncations could be attributed to the fact that the Δ14 mutant retained a partial N terminus; for instance, in the crystal structure, amino acids Ile17 and Thr19 contacted the cleaved DNA intermediate (Tsutakawa *et al.,* 1999b). Therefore, the Δ14 mutant may have remained partly associated with the DNA after cleavage, thus protecting the nick against re-sealing by DNA ligase, while the Δ19 mutant could not remain associated with the DNA.

*Stability of Stationary Phase Levels of Vsr and Mutant Proteins*

Most of the Vsr mutants were less stable than wildtype *in vivo,* which was apparent from Western blot analysis (Figure 24). These proteins should not have been made at different levels, as they were under the control of the *trc* promoter. The altered conformation of the F67A, W68A and W86A mutants in the presence of magnesium ion (Table 2) may have left them susceptible to "housekeeping proteases", which degrade abnormally folded proteins (Neidhardt, 1996); Misfolded proteins in the cell may refold, aggregate, or be degraded (Neidhardt, 1996).

In terms of the truncations, the identity of the N terminus amino acid *in vivo* can affect the degradation rate of some proteins. All N terminus amino acids are stabilizing except for Phe, Leu, Trp, Arg, and Lys (Neidhardt, 1996). It has been shown that all N terminus

amino acids are stable in mutants lacking *clpAP* (Tobias *et al.*, 1991). Similar stability experiments for Vsr and Δ14 executed in *clp* and *lon* protease deficient strains yielded the same result as represented in Figure 23, in which stationary phase levels of these proteins were still different relative to one another – mutants were present at lower concentrations in respect to Vsr. The Δ19 mutant had an N terminus glycine residue, its half-life should have been even longer than wildtype Vsr. Taken altogether, it appears that these mutants are not being degraded because of the identity of their N terminus amino acid.

*Vsr Truncation In Vitro*

Unfortunately, the mechanism of Vsr truncation *in vitro* could not be determined. However, the data suggests that it is not mediated by a second protein (Figure 27), had there been no limitation due to the antibody, we may have been able to discriminate - looking closely at lane 1, it appeared that there was slightly less short form Vsr relative to long form, 30 %, while the rest of the samples are approximately 40 %. One more dilution may have completely indicated a loss of truncation, thereby confirming this hypothesis.

Proteolysis of the *E. coli* sigma factor, $\sigma^s$, during log phase and its stabilization during stationary phase (Zhou and Gottesman, 1998) provided an attractive model for Vsr regulation. Proteolysis of Vsr in log phase cells would have ensured the levels of Vsr were kept lower in growing cells, where MMR was functioning to repair errors from DNA replication, thus minimizing the competition between MMR and VSP repair at CTAGG sites, thereby lowering T.A to C.G mutations (Zell and Fritz, 1987, Doiron *et*

*al.*1997). The protease responsible for growth phase regulation of the sigma factor is ClpXP; since Vsr was not stabilized in a *clp* nor a *lon* deficient strain, a different protease would have to be associated with Vsr regulation.

The data presented here suggests that Vsr truncation proceeds by a mechanism that is independent of pH changes, making it impossible to tell which amino acid residues are involved. The N terminus His tag also imparts some stability to the full-length form (Figure 25B, lanes 2 and 3) – perhaps the His tag keeps the flexible N terminus from adopting certain conformations that leads to exposure of a proteolytic site. Whatever the mechanism that leads to Vsr truncation, it appears that Vsr in the cell is either present as full-length form, which is growth phase regulated or as an N terminus truncation, which is no longer growth phase regulated. Although it is unclear whether low molecular weight Vsr bands come from full-length Vsr or from translation initiation from an alternate start site, preliminary evidence in our laboratory indicates that these bands are Vsr specific; these bands are no longer present on a Western blot, in a strain lacking *vsr*.

*The Role of MutL in VSP Repair*

The role of MutL in MMR is much better understood than the function of MutL in VSP repair. In MMR MutL mediates interactions between MutS and MutH in an ATP dependent manner. It is thought that MutL serves as a molecular switch that recruits different proteins at various steps in the pathway to coordinate MMR.

The function of MutL in VSP repair is not as well defined. There are several proposed models, an indirect and a direct role for MutL in VSP repair. The indirect models (Lieb and Bhagwat, 1987, Tsutakawa *et al.,* 1999b) are based on the fact that MutS-MutL complexes form loops in the DNA via an ATP dependent translocation away from the mismatch (Allen *et al.,* 1997). Loop formation may facilitate Vsr binding to a mismatch (Tsutakawa *et al.,* 1999a). This model is inconsistent with the fact that MutL catalytically stimulates binding of Vsr to DNA in a bandshift assay (Drotschmann *et al.,* 1998) and endonuclease activity of Vsr *in vitro* (Macintyre and Cupples, unpublished and this thesis, Figure 21), independent of MutS. It has also been suggested that MutL converts an inactive Vsr protein into an active protein (Turner and Connolly, 2000). Based on our wildtype enzyme being fully active (Figure 17), and the fact that VSP repair is partially functional in a MutL deficient strain (Zell and Fritz, 1987, Jones *et al.,* 1987, Lieb, 1987, Bell and Cupples, 2001) it appears unlikely that the role of MutL is to convert inactive protein to active protein as suggested by Turner and Connolly, 2000. An alternative model is that MutL acts directly on Vsr through protein-protein interactions (Mansour *et al.,* 2001), as it acts on MutH (Ban and Yang 1998) and UvrD (Hall *et al.,* 1998). This model is supported by existing data that shows MutL stimulates Vsr in a MutS independent manner, and data presented in this thesis (Figure 21).

The data in Figure 21 shows that MutL enhances the activity of Vsr under low Vsr concentrations, when the DNA is in excess, but does not stimulate the N terminus truncations. This stimulation of wildtype protein is no longer apparent under single turnover conditions. This suggests that MutL enhances Vsr endonuclease activity by

helping it associate with the DNA through a process mediated by the N terminus. The N terminus of Vsr is associated with a cleaved DNA intermediate in the co-crystal (Tsutakawa *et al.,* 1999b). MutL may facilitate N terminus ordering of Vsr. The activation of MutH by MutL provides an analogous model. Binding of MutL to MutH through a C terminal lever is thought to cause a conformational shift in MutH leading to a fifty-fold activation of endonuclease activity (Ban and Yang, 1998). Likewise MutL interacting with the C terminal of Vsr (Mansour *et al.,* 2001) may cause an ordering of the N terminus.

It is thought that mutagenesis of Vsr comes from saturation of mismatch repair, in particular, by taking away MutL from interacting in MMR. Is mutagenesis due to competition between MutH and Vsr for limiting amounts of MutL? Several pieces of data argue against this. In the bacterial two-hybrid system, MutL interacts with both Vsr N terminus truncations (Mansour *et al.,* 2001). Δ14 and Δ19 N terminus Vsr proteins are less mutagenic; the simple interaction of Vsr with MutL therefore, is not responsible for mutagenesis. Furthermore, Vsr, and MutH interact with MutL from opposite sides of the molecule (N versus C terminal) acting as non-competitive inhibitors for MutL (Mansour *et al.,* 2001). Also, the interaction between Vsr and MutL inhibits the ability of MutL to dimerize, and to interact with MutH and MutS (Mansour *et al.,* 2001), which completely blocks MMR. As a molecular switch for mismatch repair, the ATPase activity of MutL is crucial. ATP binding and MutL dimerization is thought to be the rate limiting step in MMR (Ban and Yang 1999b). If the functional interaction between Vsr and MutL requires ATP hydrolysis, for example, by ordering the N terminus of Vsr, it would keep

MutL from participating in MMR. However the fact that we do not need to add exogenous ATP to Vsr stimulation experiments argues against this possibility.

The fact that the N terminus truncated Vsr proteins are not as mutagenic as Vsr (Figure 22), suggests a direct role for the N terminus of Vsr in mutagenesis. The N terminus binds to a cleaved DNA intermediate (Tsutakawa *et al.,* 1999b). It has been shown that Vsr binds to non-canonical substrates and thus had a lower selectivity for the mismatch in comparison to other endonucleases in its class (Gonzalez-Niciera *et al.,* 2001). If Vsr associates with lesions that should be corrected by MMR, the N terminus truncations would be less mutagenic, because they are not as efficient endonucleases as Vsr; they have no N terminus to remain bound to nicked product and their catalytic turnover rates are much lower. It is thought that VSP repair out-competes MMR at CTAGG sites by a factor of nine (Zell and Fritz, 1987). The competition may be eliminated in the presence of a truncated enzyme whose catalytic turnover is forty times slower and binding is two times weaker.

The theory of Vsr associating with lesions that should be corrected by MMR is consistent with the fact that MutL is not essential for VSP repair (Zell and Fritz, 1987, Jones *et al.,* 1987, Lieb, 1987, Bell and Cupples, 2001), and also explains why adding more MutS when Vsr levels are over-expressed increases mutagenesis (Macintyre *et al.,* 1997). MutL interacts with MutS after mismatch binding and is thought to coordinate MutS with MutH. Cells with over-expressed Vsr may contain MutL-MutS complexes that interact with Vsr, instead of MutH. Therefore increased levels of MutS cause an accumulation of

MutL-Vsr complexes at mismatches that should be corrected by MMR, increasing the rate of mutagenesis. Also, providing additional MutH to cells overexpressing *vsr* causes a decrease in the rate of mutation (Macintyre *et al.,* 1997).

There can be several ways that overexpressing MutL *in vivo* leads to a stabilization of Vsr in log phase cells. Preliminary evidence in our laboratory indicates that in a MutL deficient strain, Vsr appears to be growth phase regulated, and present as a full-length protein. MutL may stabilize Vsr by mimicking DNA binding, causing burial of a proteolytic site. Since MutL binds at the C terminal of Vsr, and growth phase regulation occurs at the N terminus, this indicates a conformational change brought about by MutL binding to the N terminus of Vsr, which is also required for stimulation of endonuclease activity *in vitro*. Stabilization of F67A and W68A mutants proves that the MutL interaction with Vsr does not require DNA and that the physical interaction with MutL alone is sufficient to cause stabilization. Alternatively, MutL could physically protect Vsr against proteolysis.

Although these experiments have provided several alternatives for the role of MutL in VSP repair, and are in support of a direct MutL effect, future experiments should be aimed at clarifying the role of MutL in VSP repair.

## Significance of Experiments

The Vsr endonuclease is a unique DNA repair enzyme, as it utilizes a *novel* recognition and catalytic mechanism. Although the N terminus is not involved in mismatch specificity, it is required for growth phase regulation of Vsr, and for functional interaction with MutL, thereby directly involved in mutagenesis. The gel shifts also suggest that the N terminus is required for the endonuclease to remain bound to its cleaved DNA substrate - physiologically this plays an important role, ensuring substrate protection from DNA ligase. Overexpression of MutL leads to a stabilization of Vsr log phase levels, also mediated through an N terminus interaction, possibly causing burial of a proteolytic site. Therefore, a crucial structure-function relationship between the N terminus of the Vsr endonuclease is emerging.

Information on structure-function provides powerful insight into other proteins i.e DNA binding nucleases, DNA restriction enzymes, DNA repair proteins, and even proteins involved in recombination to help formulate relevant models for activity and regulation. Structural analysis may help in the design of more useful enzymes by protein engineering techniques. Structure-function relationships also play crucial roles in commercial processes such as drug design, etc.

It is a well-documented fact that DNA lesions could lead to cancer, and it would be of great benefit to learn how the cell potentially regulates and coordinates DNA repair enzyme turnover, while keeping a delicate balance between potentially harmful DNA damaging mutagenesis and the beneficial DNA repair process.

## References


Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J.D. (1994) *Molecular Biology of the Cell,* pp 242-262, Garland Publishing, New York

Allen, D.J., Mahkov, A., Grilley, M., Taylor, J., Thresher, R., Modrich, P. and Griffith, J.D. (1997) *EMBO J.* **13**, 4467-4476

Aggrawal, A.K. (1995) *Curr. Opin. Struct. Biol.* **5**, 11-19

Appleyard, R.K. (1954) *Genetics*, **39**, 440-452

Ban, C. and Yang, W. (1998) *EMBO J.* **17**, 1526-1534

Ban, C. and Yang, W (1998b) *Cell,* **95**, 541-552

Barrett, T.E., Savva, R., Panayotou, G., Barlow, T., Brown, T., Jiricny, J. and Pearl, L.H. (1998) *Cell,* **92**, 117-129

Bell, D.C. and Cupples, C.G. (2001) *J. Bacteriol.* **183**, 3631-3635

Bhagwat, A.S. and McClelland, M. (1992) *Nucleic Acids Res.* **20**, 1663-1668

Brégeon, D., Matic, I., Radman, M. and Taddei, F. (1998) *J. Genet.* **78**, 21-28

Burgess-Hickman, A., Li, Y., Varghese-Mathew, S., May, E.W., Craig, N.L., and Dyda, F. (2000) *Mol. Cell* **5**, 1025-1034

Carey, J. (1991) *Methods in Enzymology, Protein DNA Interactions,* pp 103-109, Academic Press, San Diego

Choi, S. H. and Leach, J.E. (1996) *Mol. Gen. Genet.* **244**, 383-390

Clarke, A.R. (1996) *Enzymology Lab Fax*, pp 199-221, Academic Press, San Diego

Cupples, C.G. and Miller, J.H (1988) *Genetics*, **12**, 637-44

Cupples, C.G. and Miller, J.H. (1989) *Proc. Natl. Acad. Sci.* **86**, 5345-5349

Cupples, C.G., Cabrera M., Cruz, C. and Miller, J.H. (1990) *Genetics* **125**, 275-280

Cooper, D.N. and Youssoufian, H. (1988) *Hum. Genet.* **78**, 151-155

Cooper, D.L., Lahue, R.S. and Modrich, P. (1993) *J. Biol. Chem.* **268**, 11823-11829

Coulondre, C., Miller, J.H., Farabaugh, P.J. and Gilbert, W. (1978) *Nature* **274**, 775-780

Creighton, T.E. (1993) *Proteins Structures and Molecular Properties,* pp 463-473, W.H. Freeman and Company, United States of America

Dar, M.E. and Bhagwat, A.S. (1993) *Mol. Micro.* **9**, 823-833

Dodson, M.L. and Lloyd, R.S. (1989) *Mutat. Res.* **218**, 49-65

Doiron, K.M.J, Viau, S., Koutroumanis, M. and Cupples, C.G. (1996) *J. Bacteriol.* **178**, 4294-4296

Drotschmann, K., Aronshtam, A., Fritz, H.J., and Marinus M.G. (1998) *Nucleic Acids Res.* **26**, 948-953

Dzidic, S., and Radman, M. (1989) *Mol. Gen. Genet.* **217**, 254-256

Feng, G., and Winkler, M.E. (1995) *Biotechniques.* **19** (6), 956-965

Fersht, T. (1985) *Enzyme Structure and mechanism 2nd edition*, W.H. Freeman and Company, New York

Fox, K.R., Allinson, S.L., Sahagun-Krause, H., and Brown, T. (2000) *Nucleic Acids Res.* **28**, 2535-2540

Harris, R.S., Feng, G., Ross, K.J., Sidhu, R., Thulin, C., Longerich, S., Szigety, S.K., Winkler, M.E. and Rosenberg, S.M. (1997) *Genes Dev.* **11**, 2426-2437

Gabbara, S., Wyszynski, M. and Bhagwat, A.S. (1994) *Mol. Gen. Genet.* **243**, 244-248

Gill, S.C. and von Hippel, P.H. (1989) *Anal. Biochem.* **182**, 319-326

Glasner, W., Merkl, R., Schellenberger, V. and Fritz, H.J. (1995) *J. Mol. Biol.* **245**, 1-7

GraFit, (1992) (version 3.09a), Erithacus Software, UK, Staines

Grilley, M., Welsh, K.M., Su, S. and Modrich, P. (1989) *J. Biol. Chem.* **264**, 1000-1004

Gonzales, M., Frank, E.G., Levine, A.S., and Woodgate, R. (1998) *Gen. Devel.* **12**, 3889-3899

Gonzalez-Nicieza, R., Turner, D.P., and Connolly, B.A. (2001) *J. Mol. Biol.* **3**, 501-508

Gottesman S. (1996) *Annu. Rev. Genet.* **30**, 465-506

Hall, M.C., Jordan, J.R., and Matson, S.W. (1998) *EMBO J.* **5**, 1535-1541

Hall, M.C. and Matson, S.W. (1999) *J. Biol. Chem.* **274**, 1306-1312

Harlow, E. and Lane, D. (1999) *Using Antibodies, a laboratory manual,* pp 221-256, Cold Spring Harbor Laboratory Press, New York

Hennecke, F., Kolmar, H., Brundl, K. and Fritz, H.J. (1991) *Nature* **353**, 776-778

Hwang, B.J., Woo, K.M., Goldberg, A.L. and Chung, C.H. (1988) *J. Biol. Chem.* **263**, 8727-8734

Jeltsch, A. (2002) *Chembiochem.* **4**, 274-293

Jones, M., Wagner, R. and Radman, M. (1987) *J. Mol. Biol.* **193**, 155-159

Kolodner, R. (1996) *Genes and Develop.* **10**:1433-1442

Krawczak, E., Ball, V. and Cooper, D.N. (1998) *Am. J. Hum. Genet.* **63**, 474-488

Kunkel, T.A. (1992) *J. Biol. Chem.* **267**,18251-18254

Lakowicz, J.R. (1983) *Principles of Fluorescence,* Plenum Publishing Corporation, New York

Lamers M.H., Perrakis A., Enzlin J.H., Winterwerp H.H., de Wind N., and Sixma T.K. (2000) *Nature.* **407,** 711-717

Lieb, M. (1987) *J. Bacteriol.* **169**, 5241-5246

Lieb, M. (1991) *Genetics.* **128**, 23-27

Lieb, M. and Bhagwat, A.S. (1996) *Mol. Mircobiol.* **20**, 467-473

Lieb, M., and Rehmat, S. (1995) *J. Bacteriol.* **177**, 660-666

Lindhal, T. and Nyberg, B. (1974) *Biochemistry.* **16**, 3405-3410

Lohman, M. and Bujalowski, W. (1991) *Methods in Enzymology, Protein DNA interactions,* pp 258-290, Academic Press, San Diego

Lu, A.L., Clark, S. and Modrich, P. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4639-4643

Lu, A.L., Welsh, K., Clark, S., Su, S. and Modrich P. (1984) *Cold Spring Harbor Symp, Quant. Biol.* **49**, 589-596

Macintyre, G., Doiron, K.M.J. and Cupples, C.G. (1997) *J. Bacteriol.* **179**, 6048-6052

Macintyre, G., Pitsikas, P. and Cupples, C.G. (1999) *J. Bacteriol.* **181**, 4435-4436

Macintyre, G., Atwood, C.V. and Cupples, C.G. (2001) *J. Bacteriol.* **183**, 921-927

Mansour, C.A., Doiron, K.M.J. and Cupples, C.G. (2001) *Mut. Res.* **485**, 331-338

Merkl R, Kroger M, Rice P, Fritz H.J. (1992) *Nucleic Acids Res.* **20**, 1657-1662

Martin, S.R. (1996) *Proteins Lab Fax,* pp 195-204, Academic Press, San Diego

Mechanic L.E., Frankel, B.A. and Matson, S.W. (2000) *J. Biol. Chem.* **275**, 38337-38346

Michaels, M.L. and Miller, J.H. (1992) *J. Bacteriol.* **174**, 6321-6325

Modrich, P. and Fang, W.H. (1991) *J. Biol. Chem.* **16**, 11838-11844

Modrich, P. and Lahue, R. (1996) *Annu. Rev. Biochem.* **65**, 101-133

Morkiwaw, K., Tsujimoto, M., Ikehara, M., Inaoka, E., and Ohtsuka, E. (1988) *J. Mol. Biol.* **202**, 683-684

Morikawa, K. and Shirakawa, M. (2000) *Mut. Res.* **460**, 257-275

Neddermann, P. and Jiricny, J. (1994) *Proc. Natl. Acad. Sci.* **91**, 1642-1646

Neidhardt, (1996) *Escherichia coli and Salmonella, Cellular and Molecular Biology,* Second Edition, pp 938-950, ASM press, Washington, DC

Ng, H.H. and Bird, A. (1999) *Curr. Op. Genet. Dev.* **9**, 158-163

Obmolova G., Ban C., Hsieh P., and Yang W. (2000) *Nature.* **407**, 703-710

Petropoulos, L., Vidmar, J.J., Pasii, E. and Cupples, C.G. (1994) *Mutat. Res.* **304**, 181-185

Ruiz, S. M., Letourneau, S. and Cupples, C.G. (1993) *J. Bacteriol.* **175**, 4985-4989

Saenger, W. (1991) *Curr. Opin. Struct. Biol.* **1**, 130-138

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning a Laboratory Manual,* Cold Spring Harbor Laboratory Press, New York

Schärer, O.D. and Jiricny, J. (2001) *Bioessays* **23**, 270-281

Schmid, F. (1998) *Protein Structure A Practical Approach,* Creighton, New York

Sohail, A., Lieb, M., Dar, M., and Bhagwat, A.S. (1990) *J. Bacteriol.* **172**, 4214-4221

Singer, M., Baker, T.A., Schnitzler, G., Deischel, S.M., Goel, M., Dove, W.F., Jaacks, K.J., Grossman, A.D., Erickson, J. and Gross C.A. (1989) *Microbiol. Rev.* **53**, 1-24

Stanford, N.P., Halford, S.E., and Baldwin, G.S. (1999) *J. Mol. Biol.* **288**, 105-116

Studier, F.W. and Moffatt, B.A. (1986*) J Mol Biol.* **189**, 113-130

Teo, I., Sedgwick, B., Demple, B., Li, B. and Lindahl, T. (1984) **3**, 2151-2157

Tsutakawa, S.E., Muto, T., Kawate, T., Jingami, H., Kunishima, N., Ariyoshi, M., Kohda, D., Nakagawa, M., and Morikawa, K. (1999a) *Mol. Cell* **3**, 621-628

Tsutakawa, S.E., Jingami, H., and Morikawa, K. (1999b) *Cell* **99**, 615-623

Tsutakawa, S.E. and Morikawa, K. (2001) *Nucl. Acids. Res.* **29**, 3775-3783

Turner, D.P. and Connolly, B.A. (2000) *J. Mol. Biol.* **00**, 1-14

Voellmy, R.W and Goldber, A.C. (1981) *Nature* **290**, 419-421

Wagner, R., and Meselson, M. (1976) *Proc. Natl. Acad. Sci.* **73**, 4135-4139

Waters, T.R., Gallinari, P., Jiricny, J., and Swann, P.F. (1999) *J. Biol. Chem.* **274**, 67-74

Wang, L., Wilson, S., and Elliott, T. (1999) *J. Bacteriol.* **181**, 6033-6041

Welsh, K.M., Lu, A.L., Clark, P. and Modrich, P. (1987) *J. Biol. Chem.* **262**, 15624-15629

Wiebauer, K. and Jiricny, J. (1990) *Proc. Natl. Acad. Sci.* **87**, 5842-5845

Wigley, D.B., Davies, G.J., Dodson, E.J., Maxwell, A., and Dodson, G. (1991) *Nature.* **351**, 624-629

Woo, K.M., Chung, W.J., Ha, D.B., Goldberg, A.C. and Chung, C.H. (1989) *J. Biol. Chem.* **264**, 2088-91

Wyszynski, S., Gabbara, S. and Bhagwat, A.S. (1994) *Proc. Natl. Acad. Sci.* **91**, 1574-1578

Zell, R., and Fritz, H.J (1987) *EMBO J.* **6**, 1809-1815

Zhou, Y. and Gottesman, S. (1998) *J. Bacteriol.* **180**, 1154-1158

Zharkov, D.O. and Grollman, A.P. (1998) *Biochemsitry.* **36**, 12384-12389

## Appendix 1 - Technical Note

*Novel Application for Infrared Fluorescence Imaging*

Fluorescent technology allows for sensitive detection of proteins. LI-COR has developed an IR laser imaging system designed to image membranes and arrays for protein and nucleic acid applications, which employs longer wavelength near infrared fluorophores for protein detection. Their system for nucleic acid labelling and detection uses infrared (IR) fluorescent dyes with emissions at approximately 700 and 800 nm that are coupled to a carbodiimide reactive group, which is attached directly to the DNA through a thymine or guanine group. These dyes can bind either single or double stranded nucleic acids. These DNA labelled oligonucleotides were demonstrated for Southerns, dot blots, northern membranes arrays and gels. Detection involves an IR laser imaging system designed to simultaneously detect two distinct wavelengths. Two laser diodes generate excitation light at 680 and 780 nm, with image resolution ranging from 25-200 μm. This high sensitivity is due to the lower background present in the infrared region as compared to the lower wavelength dyes used in other fluorescent detection systems.

Wanting to eliminate the use for radioisotopes in the laboratory for DNA gel shift and endonuclease analysis, a 50mer was ordered which contained the recognition sequence for Vsr along with a T/G mismatch at position 20. The T-containing strand was 5' end was labelled with a 700 nm fluorescing dye. Endonucleolytic cleavage resulted in two fragments, a 30mer and a 20mer. To test the sensitivity of the instrument, we ran different concentrations of oligonucleotides to test our detection limits. Figure 36A showed that even with a tenth of the oligonucleotide used in the radioactive assay,

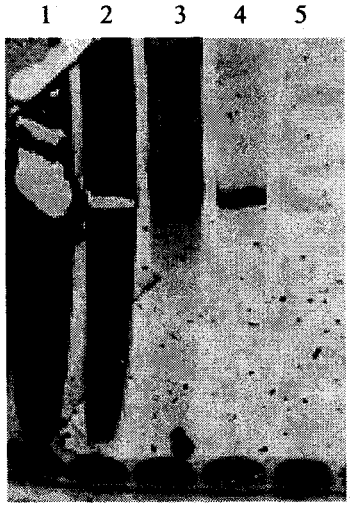
allowed for substrate detection. Therefore the oligonucleotides were annealed and desalted to test in an endonuclease assay. Figure 36B shows a single turnover experiment with Vsr in the presence of this substrate. Catalytic turnover was comparable to that with the radioactive 30mer substrate. Unfortunately the addition of this larger substrate did not stabilize Vsr binding to DNA, gel mobility shifts were comparably smeary with those observed with the 30mer under radioactive conditions (data not shown).

With radioactive assays, the oligonucleotide DNA is labelled and purified. A portion is electrophoresed to monitor the amount of annealing, and to quantify how "hot" the strand is. The day after, endonuclease reactions are carried out behind a protective screen, incubated and electrophoresed at a space in the lab designated for radioactive work; the radioactive gels are dried under vacuum and exposed to film overnight, which may result in over or under exposure depending on how "hot" the isotope is. Swipe checks of the radioactive area must be performed to check for contamination. The following day, the films are developed and original gels are exposed to a $^{32}P$ screen. Signals are quantified after several days; this often leads to band smearing along with a decreased signal. After a week of using the initial radiolabeled substrate, the oligonucleotides must be re-labelled and re-annealed due to breakdown products present in the radioactive substrate.

With the LI-COR system, the oligonucleotide is pre-labelled, so the oligonucleotides are annealed, a gel is run, without drying, scanned for signal. Assays need not be conducted behind a protective screen, or in a designated part of the laboratory. Within minutes of an assay, bands are visualized and quantified. Therefore, after having obtained quantitative

results from an initial experiment, multiple reactions may be run in the course of a day. Note the tightness of the bands in the LI-COR scan, which can not be rivalled with radioactive substrates. The signal is never over-exposed, and re-scanning using different resolutions is also possible. This method is easy, quick and extremely convenient. Cost is much lower than repeatedly ordering $^{32}P$, labelling enzymes, and columns to remove unincorporated radioisotope. One tube of pre-labelled oligonucleotide costs $200 US and is enough for hundreds of thousands of reactions, thereby theoretically also improving quality control. A reaction using the annealed substrate may be comparable to a reaction six months later using the same substrate (data not shown). LI-COR dyes are extremely stable and stored as aliquots in the fridge.

A:




B:

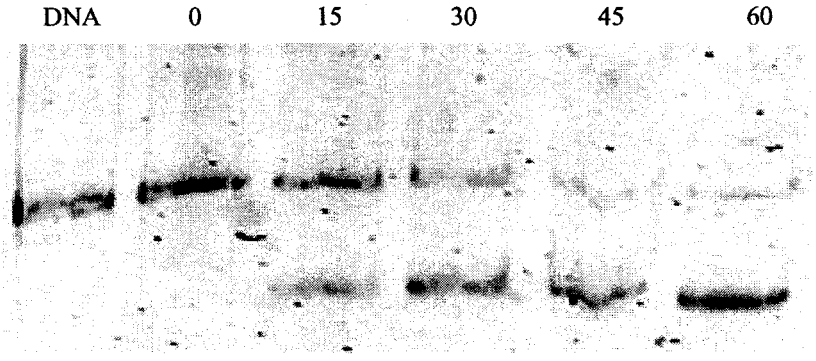



**Figure 36: Elimination of Radioactivity**

16% PAGE depicting differing amounts of oligonucleotides. Lane 1, 10 pmol, lane 2, 1 pmol, lane 3, 0.1 pmol, lane 4, 0.01 pmol, (normally used in radioactive assay) and lane 5, 0.001 pmol **(A)**. Single turnover assay of Vsr incubated with IR labelled oligonucleotide for the times given in seconds **(B)**.