# Studying Chemical Modification and Limited Proteolysis of *E. coli* Citrate Synthase by Mass Spectrometry

by

Gillian Sadler

A Thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree of

Master of Science

Department of Chemistry
University of Manitoba
Winnipeg, Manitoba

**THE UNIVERSITY OF MANITOBA**

**FACULTY OF GRADUATE STUDIES**

*****

**COPYRIGHT PERMISSION PAGE**

**Studying Chemical Modification and Limited Proteolysis of E. coli Citrate Synthase by Mass Spectrometry**

**BY**

**Gillian Sadler**

**A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirements of the degree of**

**Master of Science**

# *Acknowledgements*

I would like to thank Dr. Duckworth for his patient support and guidance throughout my stay as a student in his lab, both as an undergraduate and graduate student. His patience and encouragement during the writing of this thesis were particularly appreciated.

Thanks to the members of my committee, Dr. Court and Dr. Perreault for reading and providing a critical assessment of this thesis.

Dr. Lynda Donald provided much advice and encouragement, and provided the ideas for the online limited proteolysis and allosteric activation experiments. Many thanks for running and assisting in the analysis of the ESI TOF MS samples for this thesis (and for teaching me how to make flaming snowballs, orange juice clocks, and elephant's toothpaste).

Thanks to past and present students of the Duckworth and Jamieson labs for their friendship and assistance in the lab.

I am grateful to the members of the Time-of-Flight lab in the Department of Physics for access to their mass spectrometry facilities and expertise, and for running samples.

Thanks to Mark McComb for mass spectrometry advice, and for running and analyzing CE/MS samples. Thanks also to Colin Lee for running MALDI MS samples.

Thanks to Dr. Philip Hultin, Manjula Sudarsham, and Marion Earle for providing equipment and a great deal of advice during my attempts to synthesize photoaffinity labels. Thanks also to Dr. Jim Charlton for helpful conversations on light sources for photoaffinity labeling.

Finally, I thank my family and friends for their constant love, support, and encouragement.

# *Table of Contents*

# *List of Figures*

# *List of Tables*

# *Abbreviations*

AMP, ADP, ATP: Adenosine mono, di, or triphosphate

LB: Luria – Burrous bacterial culture media

CE/MS: Capillary electrophoresis/mass spectrometry

CoA: Coenzyme A

CS: Citrate synthase

CZE: Capillary zone electrophoresis

Da: Daltons

DPDS: 4, 4'-dithiodipyridine

DTNB: 5, 5'-dithiobis-2-nitrobenzoic acid

EDTA: Ethylenediaminetetraacetic acid

ESI MS: Electrospray ionization mass spectrometry

HPLC: High performance liquid chromatography

$K_M$: Michaelis constant (enzyme-substrate affinity)

$K_D$: Dissociation constant

MALDI MS: Matrix assisted laser desorption/ionization

MS/MS: Tandem mass spectrometry

m/z: mass to charge ratio

NADH: Nicotinamide adenine dinucleotide

NMR: Nuclear magnetic resonance

OAA: oxaloacetate

P4T: Pyridine-4-thione

SDS PAGE: Sodium dodecyl sulfate polyacrylamide gel electrophoresis

TCA cycle: Tricarboxylic acid cycle

TFBA: 1, 1, 1-trifluorobromoacetone

TNB: Nitrophenolate

TOF: Time-of-flight

Tris: Tris (hydroxymethyl) amino methane

UV: ultraviolet

WTCS: Wild type citrate synthase

*Note:*

*This thesis uses the one-letter code for abbreviating amino acids.

*The naming convention used to describe mutant CS makes use of the one-letter code, where the first letter of the name is the residue being mutated, the following number is the sequence position corresponding to that particular amino acid, and the final letter is the amino acid which now exists in this position (eg: P205A: proline, residue #205 mutated to alanine).

# *Abstract*

Citrate synthase (CS) of *E. coli* is strongly inhibited by the binding of NADH to an allosteric site distinct from the enzyme's active site. Previous chemical modification and mutagenesis studies have implicated a reactive cysteine, C206 as being at or near the allosteric site. This thesis details an attempt to clarify the location of the allosteric site using the photoaffinity label 2-azidoATP. Electrospray ionization time-of-flight mass spectrometry (ESI TOFMS), and capillary electrophoresis interfaced with mass spectrometry were used to map the locations of photolabel incorporation in unfractionated tryptic digests of CS. Although several putatively labeled peptides were noted, the best candidate was a peptide of mass 1666.8 Da, which could arise from one or both of tryptic fragments T17 (1146.48 Da, corresponding to residues 168-177) and T15-16 (1147.68 Da, corresponding to residues 158-167). Both of these peptides are structurally close to C206.

Mass spectrometry was also applied to the study of several other characteristics of *E. coli* CS previously examined by more conventional means. Chemical modifications targeted at the two reactive sulfhydryl groups of CS were mapped in unfractionated digests by matrix assisted laser desorption ionization (MALDI) TOFMS. Limited proteolysis of CS by trypsin and subtilisin was monitored by ESI TOFMS. An unexpected result of these experiments was the "all or nothing" breakdown of CS, with the production of mostly small peptides and very few large intermediate proteolytic fragments. The behaviour of the subunit equilibrium of CS in the presence of an

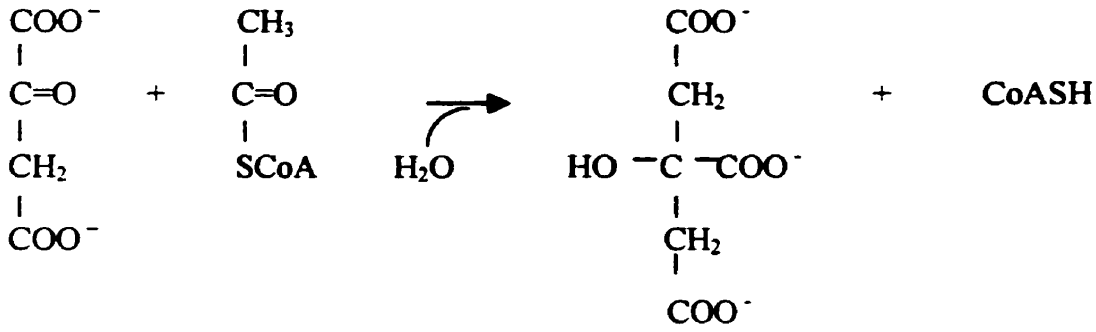
allosteric activator was also studied using ESI TOFMS.

Finally, the results of undergraduate project course work carried out on two allosteric site mutants and the manner in which these results relate to those previously obtained for other CS mutants, are also discussed.

# *Introduction*

# *An Introduction to Citrate Synthase*

Citrate synthase (EC 4.1.3.7) catalyzes the condensation of oxaloacetate (OAA) and acetyl coenzyme A (acetyl CoA) as shown below:




**OAA** **acetyl CoA** **Citrate**

The reaction is believed to take place in the manner of an ordered bisubstrate reaction, such that oxaloacetate binds first to CS, followed by acetyl CoA. The reaction produces citryl CoA as an intermediate, which is then hydrolyzed to give citrate and CoA.

This is the first reaction of the tricarboxylic acid (TCA) cycle, which as an amphibolic pathway produces both energy (ATP via NADH and $FADH_2$) and biosynthetic intermediates (as shown in Figure 1). When facultative anaerobes such as *E. coli* derive their energy via fermentation instead of oxidation, a non-cyclic variation of the TCA cycle provides the biosynthetic intermediates α-ketoglutarate and succinyl-CoA (Figure 2). In addition, it is also the first reaction of the glyoxylate shunt, a pathway found in bacteria, yeasts, plants, and certain invertebrates which branches from the TCA cycle and allows them to use acetate as a carbon source (Figure 3). CS is in a position to control the entry of carbon into these three pathways.

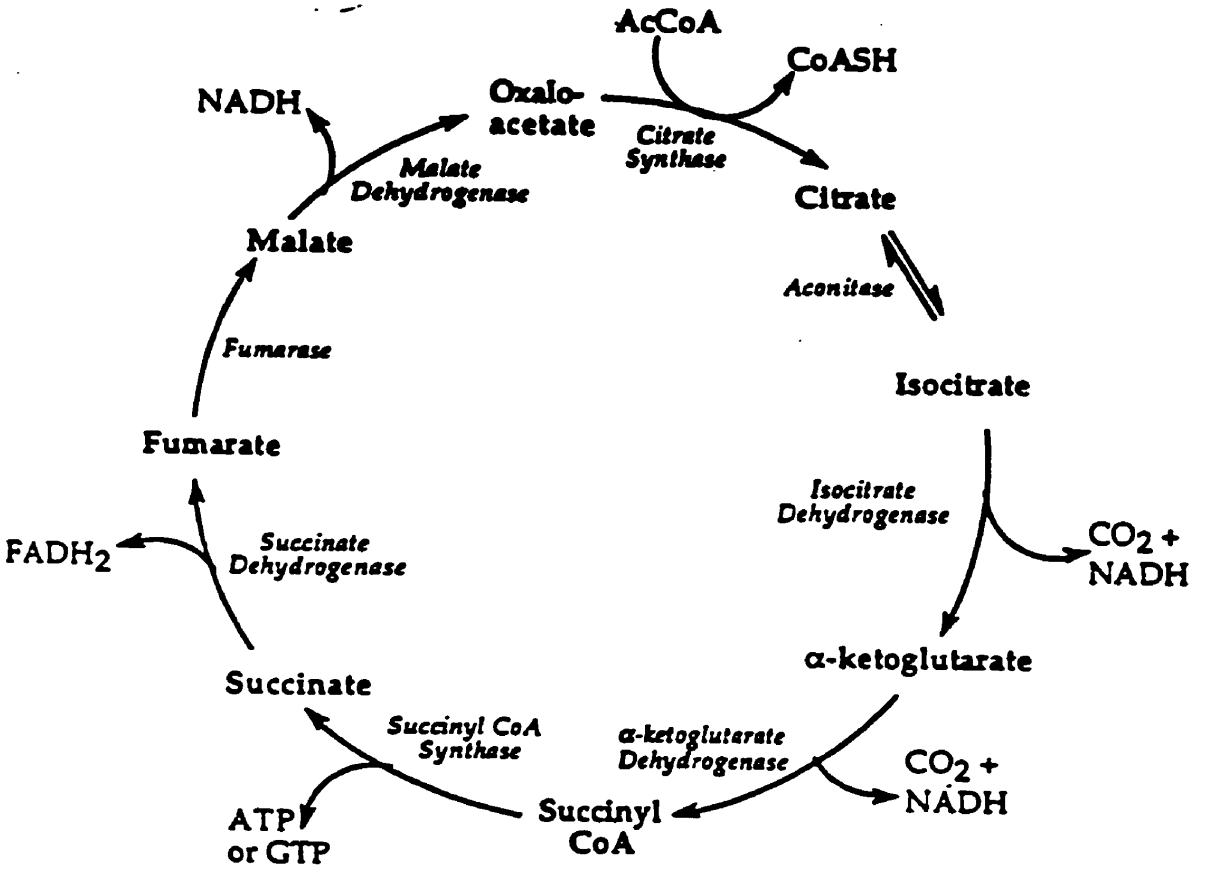



Figure 1: The tricarboxylic acid (TCA) cycle, with enzyme names shown in italics.

*Adapted from Ayed, 1998 with permission*

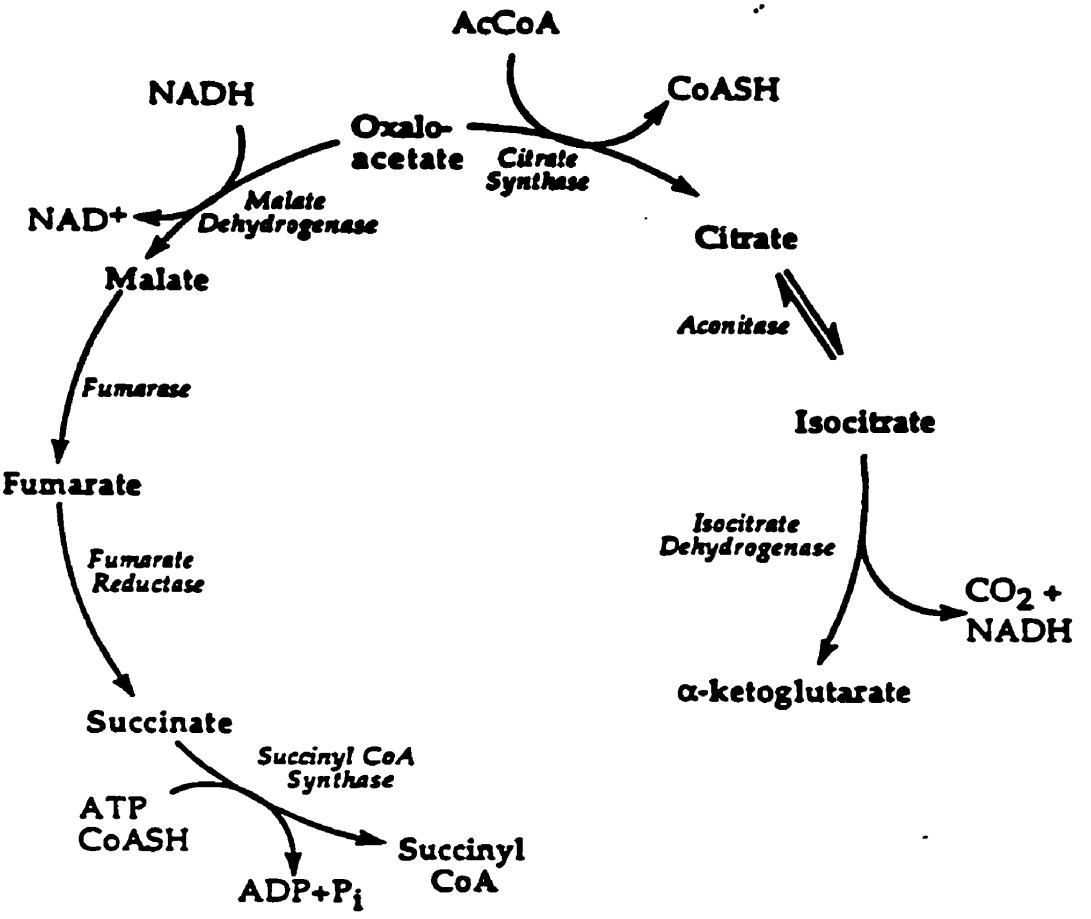



Figure 2: The non-cyclic TCA cycle.

*Adapted from Ayed, 1998 with permission*

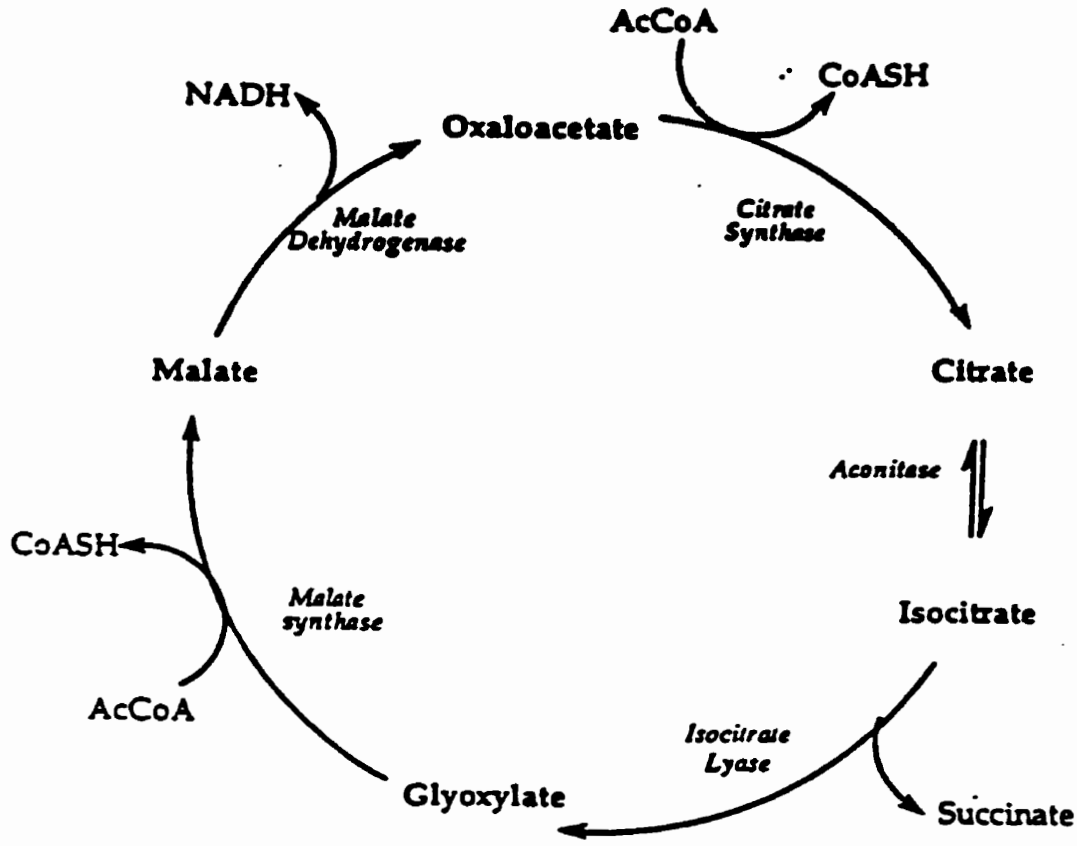



Figure 3: The glyoxylate cycle.

*Adapted from Ayed, 1998 with permission*

As a ubiquitous enzyme from a key metabolic cycle, CS from a wide variety of both higher and lower organisms have been isolated and studied. Thus far, all have been found to consist of identical subunits of approximately 48 000 Da. The 12 amino acids identified by crystal structure studies and site-directed mutagenesis (e.g. Remington et al., 1982; Zhi et al., 1991 and references therein) as playing a role in substrate recognition and catalysis are reasonably well conserved across all known CS (Henneke et al., 1989). Based on their physical properties and the manner in which their activity is regulated, CS have been classified into three groups (Weitzman & Jones, 1968). Citrate synthases from eukaryotes, Gram-positive eubacteria, and the archaea are dimeric, having molecular weights of around 100 000 Da. These "small" CS are isosterically inhibited by ATP. By contrast, the CS of Gram-negative eubacteria form hexamers of around 250-280 000 Da, and are allosterically inhibited by NADH. This class of "large" CS is further subdivided along the lines of bacterial physiology. In strict aerobes, such as species of *Pseudomonas*, NADH inhibition of CS activity is relieved in the presence of low concentrations of AMP, whereas no such relief is observed in facultative anaerobes like *Escherichia coli*. This difference is rationalized by the dependence of strict aerobes on the TCA cycle, making it a logical location for a response to lowered energy levels (signalled by AMP) to be initiated (Weitzman & Jones, 1968). Facultative anaerobes possess alternate methods of glucose catabolism (eg: glycolysis), and so increasing AMP levels may signal more efficiently to another pathway.

The crystal structures of dimeric CS from pig heart (Remington et al., 1982; Wiegand et al., 1984; Karpusas et al., 1991), chicken heart (Liao et al., 1991), *Pyrococcus furiosus* (Russell et al., 1997), *Thermoplasma acidophilum* (Russell et al., 1994), and the

psychrophilic Antarctic bacterial strain DS2-3R (Russell et al., 1998) have been determined. In general, CS subunits are largely α-helical, and organized into a large and a small domain. The active site lies in the cleft between the two domains, and requires amino acids from both subunits of a dimer to be complete, such that when subunits are paired in an antiparallel fashion there are two active sites per dimer. The pig and chicken CS structures are available in both two different conformations of the dimer shown in Figure 4. The "open" conformation corresponds to unliganded CS, whereas the "closed" form is found in the presence of CoA, and substrate. Analysis of these two structures found that the difference between them can be accounted for by an 18.5° rotation of the small domain relative to the large domain (Remington et al., 1982). Thus it appears that a large conformational change is induced upon binding of substrates to CS, and is required for catalysis.

***Isoforms of CS***

Several organisms have more than one CS gene, raising questions as to the structural differences between these CS enzymes with regards to their regulation and their roles in metabolism. *Saccharomyces cerevisiae* expresses both a mitochondrial and a cytoplasmic CS, which are believed to participate in energy production and glutamate synthesis, respectively (Rosenkrantz et al., 1986). Two CS genes, *citA* and *citZ*, have been identified in *Bacillus subtilis*, where *citZ* encodes the major dimeric CS, and *citA* expresses lower levels of a less active CS (Jin & Sonenshein, 1996). The role of these two CS is thought to be tied to the growth phases of *B. subtilis*, but this is still under investigation, as *citA* and *citZ* demonstrate overlapping patterns of expression.

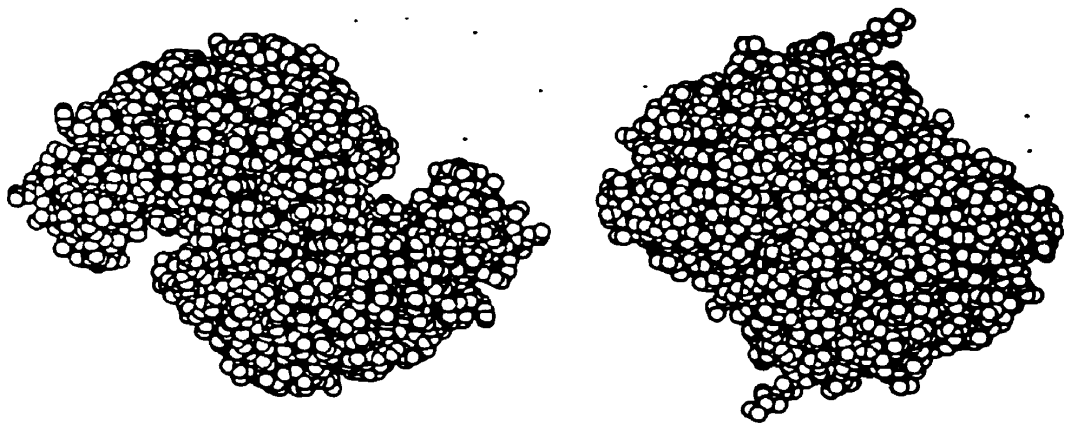



Figure 4: Space-filling model of the eukaryotic (pig and chicken heart) CS dimer as determined by X-ray crystallography.

a) Open (unliganded) conformation
b) Closed (substrate bound) conformation

*Adapted from Remington et al. 1982*

Multiple forms of CS have also been found among Gram-negative eubacteria, providing exceptions to the "large" and "small" CS classifications. Species of *Pseudomonas* can contain both large NADH-sensitive CS I and small NADH-insensitive CS II, but in some species, only CS I or CS II alone is expressed (Mitchell & Weitzman, 1983). It is believed that CS I and CS II arise from different genes, as they have different amino acid compositions, kinetics, immunological properties, and N-terminal sequences (Mitchell et al., 1995; Mitchell & Anderson, 1996). The N-terminal sequence of CS II shows greater similarity to that of a small dimeric CS of a *Bacillus* species (27%) than to CS I of *P. aeruginosa* (<5%). Where both CS I and CS II are present, the pattern of expression has been found to relate to the growth phase of the cells (CS I mainly in exponential phase, CS II in stationary phase) and the carbon source available (CS I from growth on glucose, CS II from acetate-grown cells) (Mitchell et al., 1995). Mitchell has raised the possibility that these two forms of CS play both catalytic and structural roles in a multienzyme complex of TCA cycle enzymes in *P. aeruginosa* (Mitchell, 1996).

Several mutant *E. coli* CS obtained from phenotypic revertants of a CS mutant were found to display the dimeric structure, NADH insensitivity, and ATP inhibition that are characteristic of the small enzyme (Harford & Weitzman, 1978; Danson et al., 1979). Similar results have been obtained in reversion studies of CS from *P. aeruginosa* (Mitchell et al., 1995) and *Acinetobacter lwoffi* (Weitzman et al., 1978). Speculation about a second, NADH-insensitive CS of *E. coli* encoded by a separate gene arose when it was discovered that one of the small mutant CS from a reversion study was in fact immunologically different from the wild type enzyme, and had a different N-terminal sequence (Patton et al., 1993). With the release of the complete *E. coli* genome in 1997

(Kroeger & Wahl, 1997), the presence of a gltA-like gene was demonstrated (A. Ayed, personal communication; Gerike et al. 1998). This second CS showed approximately 30% homology with the sequence of the primary, NADH-sensitive CS of *E. coli*, with conservation of key active site residues, and also possessed very weak CS activity. However, the second CS was induced in the presence of propionate instead of being constitutively expressed (Textor et al., 1997). Its primary function was as a methylcitrate synthase participating in the degredation of propionate to succinate and pyruvate via 2-methylcitrate. Nonetheless, the level of sequence and key amino acid conservation between the two enzymes suggests that they might have a common evolutionary origin. Analogues of other TCA cycle enzymes such as aconitase and fumarase have been found to catalyse the same or similar reactions as their parents under different metabolic conditions (Patton et al., 1993).

***CS of Archaea***

One rapidly expanding area of CS research is the isolation and characterization of archaeal CS, all of which have so far been found to fall into the "small" category. As with other studies comparing enzymes from mesophiles and the archaea with specially adapted phenotypes, these results could shed light on the functional significance of certain CS structural elements. Information regarding archaeal CS might also assist with constructing a phylogenetic tree of CS evolution. For instance, an interesting relationship was demonstrated by Gerike et al., who discovered that two thermophilic archaeal CS showed up to 40% sequence homology with the second, NADH-insensitive CS of *E. coli*, and that both exhibited low levels of methylcitrate synthase activity (Gerike et al., 1998).

## *Escherichia coli* CS

As previously mentioned, the CS of *E. coli* is a hexameric enzyme regulated in an allosteric fashion by NADH. The *gltA* gene, which encodes the protein, was sequenced in 1983 (Ner et al., 1983), and the majority of the amino acid sequence was reported a year later (Bhayana & Duckworth, 1984). The complete sequence of *E. coli* CS is found in Appendix 1, as part of a sequence alignment with CS sequences from other representative organisms. The insertion of the *gltA* gene into a multicopy, drug-resistant plasmid (pHSgltA, and later pESgltA) allows for the production of relatively large amounts of pure CS (Duckworth & Bell, 1982).

The *E. coli* CS subunit has a mass of 47 887 Da and consists of 426 amino acids (after removal of the N-terminal methionine residue during post-translational processing). Until recently, no crystal structure was available for any of the "large" hexameric CS, and the only available model for *E. coli* CS was based on the dimeric pig CS subunit (Duckworth et al., 1987). As shown in Figure 5, the putative *E. coli* structure consists of 20 α-helices and a short section of anti-parallel β-sheet organized into a small domain (helices N-R) and a large domain (helices A-M, S-T and β-sheet).

Very recently, preliminary crystal structure data has been obtained for the *E. coli* CS hexamer (H.W. Duckworth, personal communication). The structure, a trimer of dimers, is shown in Figure 6a.

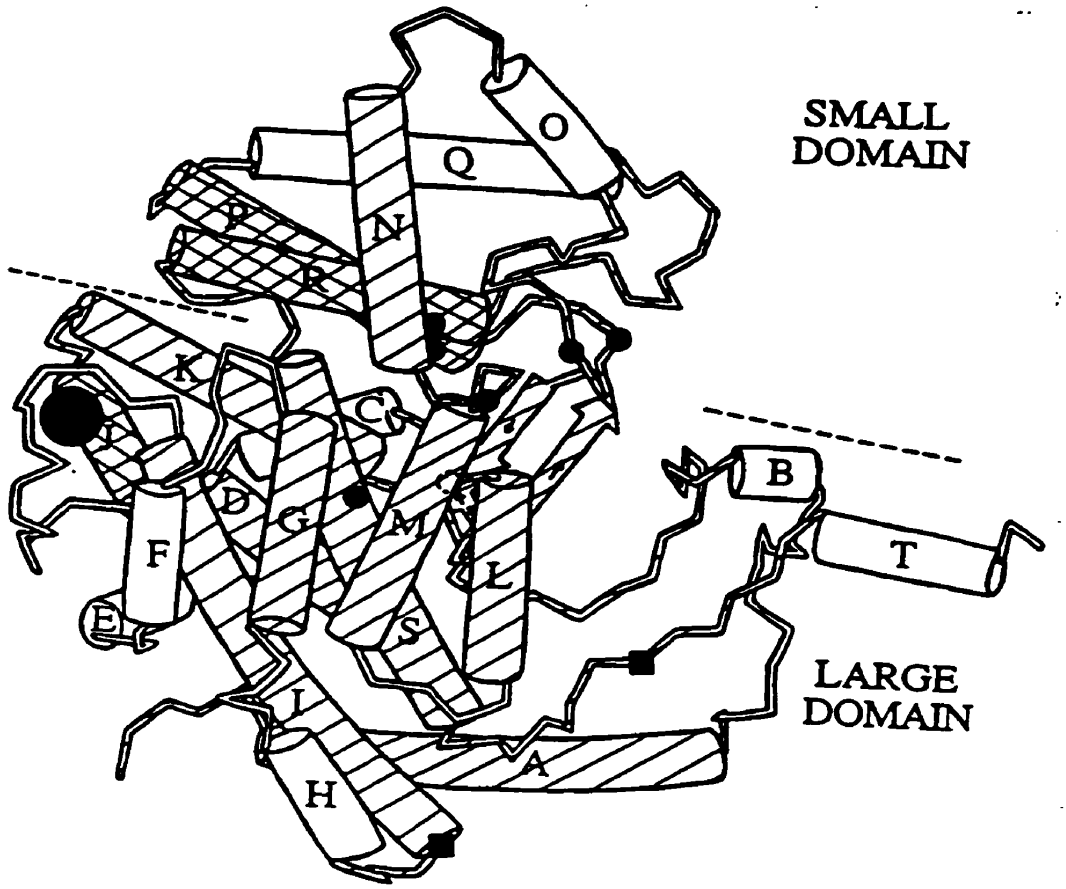



Figure 5: Model of an *E. coli* CS subunit, based on that of pig heart CS. Twenty alpha-helices and one section of anti-parallel beta-sheet are organized into two domains, large and small. Active site residues are shown as filled in circles. Filled in squares denote residues contributing to the active site of the other subunit of the CS dimer. A red dot represents C206, a marker for the allosteric site. Doubly shaded regions have greater than 50% homology with pig CS, singly shaded regions have between 25 and 50% homology, and unshaded regions have less than 25% homology.

*Adapted from Anderson, 1988*

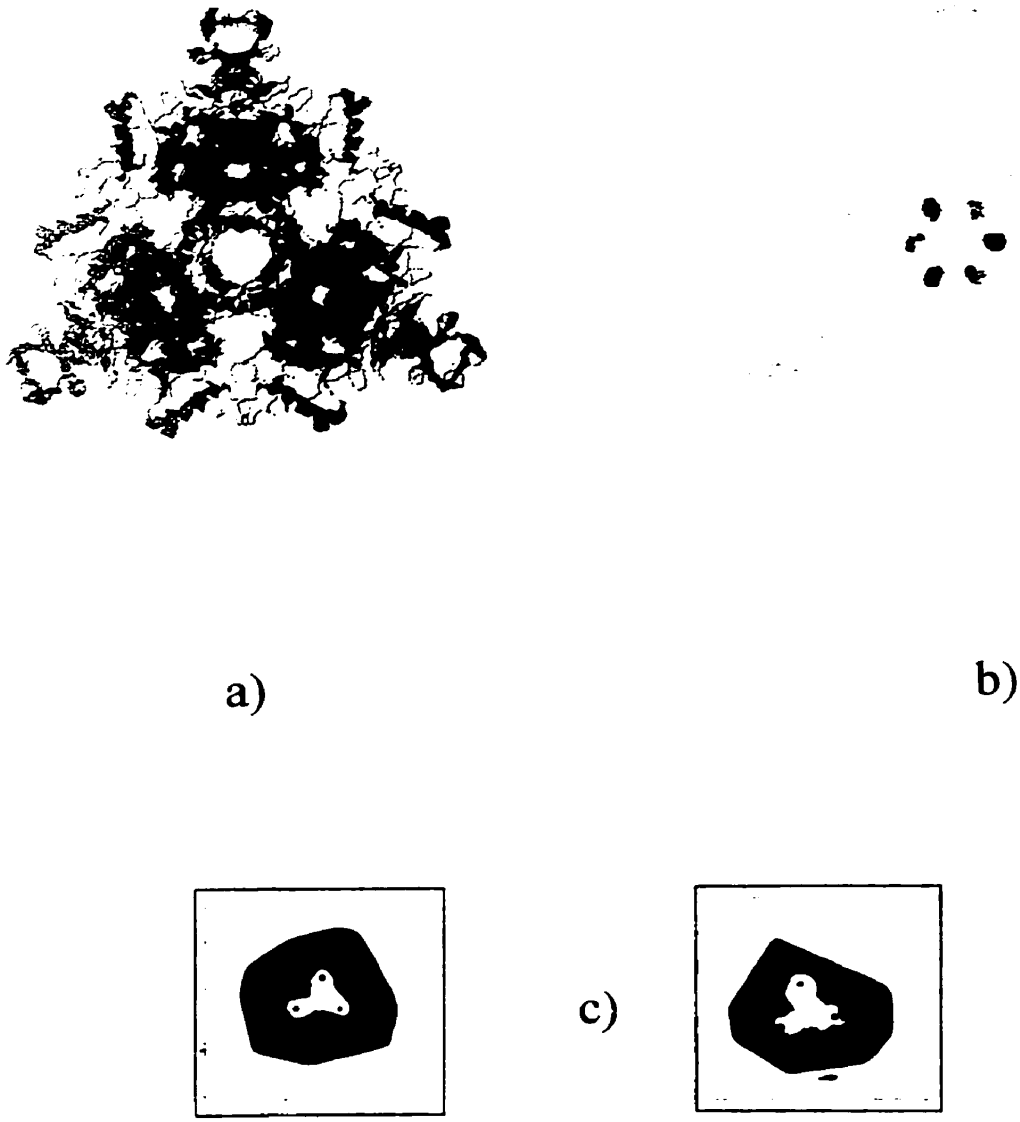



Figure 6: Structure of the *E. coli* CS hexamer

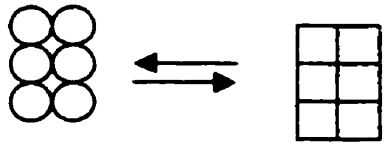
a) Preliminary X-ray crystal structure, consisting of a trimer of dimers (red, blue, and green) surrounding a central void.
b) X-ray crystal structure showing all six C206 residues enlarged in blue.
c) Electron micrographs of *Acinetobacter lwoffi* hexameric CS in the absence (left) and presence (right) of NADH (Rowe & Weitzman, 1969). These show essentially the same structural arrangement as in 6a).

***Regulation of E. coli CS***

The first CS to be isolated and characterized was from pig heart. This enzyme (and other dimeric CS) showed strong inhibition by ATP in a manner competitive with acetyl CoA (eg: Jangaard et al., 1968 and references therein). Weitzman discovered that *E. coli* CS was inhibited by NADH instead of ATP (Weitzman, 1966a). The effect was strong, such that 100 μM NADH was sufficient to eliminate virtually all CS activity, whereas 20 mM ATP was required to reduce enzyme activity to 25%. The inhibition was also very specific for NADH, as no inhibition was observed for $NAD^+$, NADPH, and $NADP^+$. The presence of acetyl CoA was observed to decrease the level of NADH inhibition. Weitzman thus proposed that NADH, as a product of the TCA cycle and high energy level indicator, was the feedback inhibitor of *E. coli* CS.

***Allostery and the Regulation of E. coli CS***

Allosteric enzymes are those whose activity is regulated through the reversible, non-covalent binding of certain ligands at sites distinct from the active site. Such proteins are generally multimeric, and, according to the Monod-Wyman-Changeux theory of allostery (Monod et al. 1965), can exist as two possible states in equilibrium as shown below:



**Relaxed** **Taut**

The T, or Taut state is the inactive form of the protein, while the R or Relaxed state represents the active form, to which substrates, for example, bind with greater affinity.

The identical subunits interact with one another in a cooperative fashion, such that when subunit conformation is changed to the R state by the binding of an activating molecule, the shift from T to R state occurs for all subunits in a concerted fashion.

Early studies of the NADH inhibition of *E. coli* CS indicated that this was likely an allosteric process (Weitzman, 1966b; Weitzman, 1967). The inhibition was found to be pH dependent, as the degree of NADH inhibition observed rapidly decreased when the pH was raised above 7.5. This was a different pH dependence than the bell-shaped curve of the effect of pH on CS activity, which showed that the optimal pH for CS was 8. The addition of KCl not only eliminated NADH inhibition, but also stabilized the enzyme against urea and heat denaturation at 55°C, based on measurements of activity. Faloona and Srere (1969) later determined that this effect was not limited to $K^+$, but that other monovalent cations could also increase CS activity to a lesser degree ($NH_4^+ > Na^+ > Li^+ = (CH_3)_4N^+$). Using the model of allostery proposed above, as a negative allosteric effector, NADH would bind to the T state, while KCl would stabilize the R, or active form of CS. The fact that at low or nonexistent levels of KCl, enzyme activity is low, acetyl CoA binding curves are sigmoid, and inhibition by NADH is favored suggest that in its purified form, *E. coli* CS is in T state (Duckworth & Tong, 1976). The allosteric nature of the NADH inhibition of *E. coli* and *P. aeruginosa* CS was confirmed by Harford and Weitzman (1975) using mutiple inhibition studies. The same paper also concluded that the inhibition of pig and *Bacillus megatarium* CS by ATP and bromoacetyl CoA was isosteric (competitive with substrate).

It was also observed that α-ketoglutarate, another TCA cycle product and a homologue of oxaloacetate, could inhibit *E. coli* CS, with a $K_i$ of approximately 60 μM

(Wright et al., 1967). Other ketoacids and TCA cycle-associated biosynthetic intermediates (eg: pyruvate, phosphoenolpyruvate, glutamate, aspartate, isocitrate, and α-ketovalerate) did not inhibit the enzyme even at a concentration of 5 mM. As inhibition by α-ketoglutarate was reversed by addition of KCl, and by raising the pH, it was initially proposed to be an allosteric inhibitor too (Wright et al., 1967). However, this theory was later disproved. Talgoy and Duckworth (1979) suggested that because α-ketoglutarate failed to strengthen NADH binding to the enzyme in the presence of either KCl or acetyl CoA, it could likely not act to stabilize the T-state as NADH did. Most convincingly, Anderson and Duckworth (1988) found that in two active site mutants of CS, the lowering of oxaloacetate binding affinity was mirrored by a loss in α-ketoglutarate binding affinity. This suggested that inhibition by α-ketoglutarate was in fact the result of competition for the oxaloacetate binding site. It was proposed by both of the above studies that the observed loss of this inhibition in the presence of KCl could be due to a KCl-initiated conformational change in the enzyme which would increase the specificity of the oxaloacetate binding site for the substrate alone.

***Conformational Changes and the Regulation of E. coli CS***

As the majority of CS are found in dimeric form, and residues from both subunits of the dimer are required to form the active site, it could be postulated that the basic active unit of CS is the dimer. In fact statistical analysis of limited proteolysis of the dimeric and hexameric enzymes has demonstrated that the hexamer does indeed behave as a trimer of dimers (Else et al., 1988). It is reasonable to assume that the subunit contacts

between these dimers, a feature absent in the small NADH-insensitive CS could play a role in the allosteric nature of large CS.

By using equilibrium centrifugation, and dimethylsuberimidate cross-linking of CS subunits, Tong and Duckworth (1975) demonstrated that over a wide range of low to high pH values, *E. coli* CS formed mixtures of monomers and aggregates, ranging up to decamers, except at pH 9, when only dimers were observed. There did not appear to be any preference or trend towards one particular form. However, upon addition of 50 mM KCl at pH 7.8, this random mixture was replaced by a dimer – hexamer equilibrium, which was then converted to hexamer alone in the presence of 100 mM KCl. A similar effect was observed using mass spectrometry, which showed increasing amounts of hexamer formed when CS was titrated with increasing amounts of NADH (Ayed et al., 1998). Thus it would appear that the NADH-bound T state and KCl-activated R state of *E. coli* CS are both hexameric. Electron microscopy of hexameric CS from *Acinetobacter lwoffi* (see Figure 6c) also showed that the average observed diameter of the hexamer molecules decreased in the presence of KCl and increased in the presence of NADH (Rowe & Weitzman, 1969). It remains to be determined if there are in fact two conformationally different hexameric forms of *E. coli* CS.

***Acetyl CoA and the Regulation of E. coli CS***

The link between AcCoA and the allosteric properties of *E. coli* CS has been shown in several ways. Firstly, and most importantly, the presence of the allosteric activator KCl decreases the $K_M$ of this substrate, and converts its saturation curve from sigmoid to hyperbolic, whereas the saturation curve of oxaloacetate remains hyperbolic with and

without KCl, indicating that this substrate likely binds to both T and R states equally well (Weitzman, 1966a; Weitzman, 1966b; Faloona & Srere, 1969; Anderson et al., 1991). The potential synergy between KCl and AcCoA in stabilizing the R state was demonstrated by the observation that while levels of KCl capable of eliminating NADH inhibition (0.2 M) still allowed binding of the inhibitor to the enzyme with a $K_D$ of 11 μM, the combination of 0.1 M KCl and 2 mM AcCoA also abolished NADH binding (Duckworth & Tong, 1976).

Mutagenesis studies of CS helped to further define the relationship between allostery and AcCoA (Anderson & Duckworth, 1988). The deletion mutant Δ(264-287)CS, was missing residues 264 to 287, and as such a part of the AcCoA binding site. Acetyl CoA was still able to inhibit NADH binding to the mutant enzyme as well as to wild type CS, indicating that it was possible for this substrate to interact with both the active and allosteric sites of CS. Nonetheless, three mutants of the AcCoA binding portion of the active site also affected the allosteric equilibrium. F383A CS appeared to destabilize both T and R states to some extent, with the T state being favored. (Anderson et al., 1991). In contrast, H264A and D362A were found to favor the R state (Pereira et al., 1994). These results indicate that there is also a link between the changes at the active site and the observed properties of the allosteric site.

***Characterization of Adenylate Binding to the Allosteric Site***

Quantitation of the enhancement of NADH fluorescence upon binding to *E. coli* CS allowed Duckworth and Tong to calculate a $K_D$ of 1.6 ● 0.1 μM for the binding of NADH to CS at pH 7.8 (Duckworth & Tong, 1976). At this pH, the number of NADH binding

sites per subunit was found to be approximately 0.5, meaning that only one NADH per dimer was required to inhibit the enzyme. As expected both the affinity of CS for NADH and the number of available NADH binding sites decreased (from 0.65 to 0.25) as the pH was raised from 6 to 9. KCl and acetyl CoA also weakened NADH binding, but the presence of acetyl CoA did not reduce the number of available NADH binding sites.

Kinetic analysis of the mass spectra acquired by Ayed et al. (1998) showing the hexamerization of CS as titration with NADH progressed revealed three types of NADH binding – a tight specific binding of six NADH/hexamer ($K_D$ = 1.1 ± 0.2 μM), a weak, non-specific interaction of 12 NADH/hexamer ($K_D$ = 155 ± 19 μM), and a weak, non-specific interaction of 2 NADH/dimer ($K_D$ = 28.3 ± 3.4 μM). The dissociation constant for the tight, specific binding of NADH is similar to that obtained by the fluorescence method. Whether the weak binding of NADH is real or an artifact produced by the electrospray ionization process is unknown (Ayed et al., 1998).

Characterization of the binding of other adenylates to CS was also studied by fluorescence methods. The adenylates were placed in competition with either NADH, or the fluorescent label 8-anilino-1-naphthalenesulfonate (ANS) complexed to CS (Talgoy & Duckworth, 1979). The results from this study are presented in Table 1. Essentially it was found that while the selected adenylates could bind CS and block the binding of NADH, the majority acted as activators, and not inhibitors of enzyme activity. One exception to this trend was 3'-AMP, which could conceivably interact with the acetyl CoA binding site as an analogue of this substrate, and competitively inhibit CS activity. Such observations suggest that the features of NADH which define its abilities as an

Table 1: Characteristics of the binding of adenylates to *E. coli* citrate synthase.

| Compound | $K_D$ for binding to CS (μM) | Concentration of compound tested (mM) | Reaction velocity* relative to control, no KCl | Reaction velocity* relative to control, 0.1 M KCl |
|---|---|---|---|---|
| NADH | 1.6 ± 0.3 | 0.1 | 0.03 | 0.95 |
| NAD+ | 1100 ± 200 | 1.0 | 4.1 | ND |
| NADPH | 53 ± 5 | 1.0 | 2.4 | ND |
| NADP+ | 1900 ± 300 | 1.0 | 2.8 | ND |
| 3'-AMP | 65 ± 4 | 1.0 | 0.54 | 0.60 |
| 2'-AMP | 260 ± 30 | 1.0 | 1.5 | ND |
| 5'-AMP | 83 ± 5 | 1.0 | 3.8 | ND |
| ADP-ribose | 150 ± 10 | 1.0 | 14 | 1.28 |
| ATP | 46 ± 5 | 1.0 | 2.3 | ND |
| 2',5'-ADP | 160 ± 20 | 1.0 | 4.7 | 1.01 |
| 3',5'-ADP | 68 ± 7 | 1.0 | 4.8 | 1.05 |

Adapted from Talgoy & Duckworth, 1979.
ND = not determined
*Reaction velocities were determined using the standard CS assay (Srere et al., 1963) with the concentration of acetyl CoA reduced to 20 μM. NADH is most effective at inhibiting the enzyme at low levels of acetyl CoA, and it was hoped that this would better allow for the effects of the adenylates on CS to be determined.

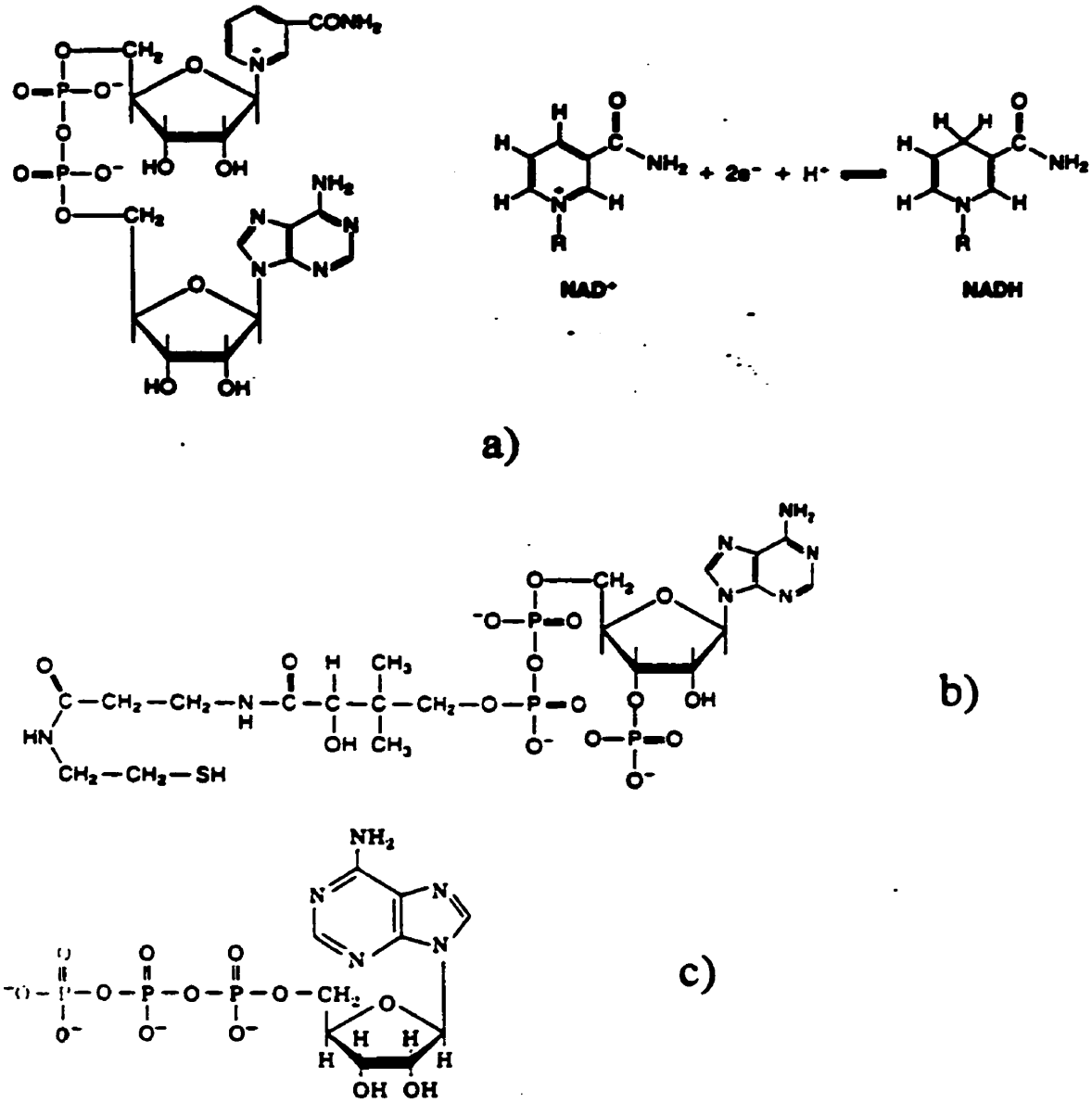



Figure 7: Structures of some adenylates capable of binding to *E. coli* CS.
a) NAD/NADH
b) Acetyl CoA
c) ATP

inhibitor reside in the nicotinamide portion of the molecule, as the adenine ring appears to be a feature generally recognized by the allosteric site (see Figure 7).

***Implication of a Cysteine Residue as Part of the Allosteric site***

Studies on the pH dependence of the photo-oxidation of *E. coli* CS in the presence of photo-sensitive dyes Methylene Blue (cationic) and Rose Bengal (anionic) suggested the involvement of a cysteine residue during NADH inhibition (Danson & Weitzman, 1973; Weitzman et al., 1974). Weitzman observed that the decrease in both CS activity and NADH inhibition above pH 8 corresponded to the pH where the sulfhydryl groups of cysteine would begin to dissociate (Weitzman, 1966b). Modification studies of the enzyme using thiol reagents provided additional evidence, as described below.

Treatment of *E. coli* CS with either N-ethyl-maleimide or $HgCl_2$ resulted in the loss of over 90% of enzyme activity, a condition which could be partially reversed by the addition of KCl (Weitzman, 1966b). No activity loss was reported when the enzyme was incubated with these reagents in the presence of KCl. The same author also reported inactivation of CS (and reactivation by KCl) when modified with 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB; also known as Ellman's reagent). This reagent targets the reactive thiol groups of cysteines as shown in Figure 8a. The nitrothiophenolate (TNP) anion released upon modification of the cysteine has an absorption maximum at 412 nm (extinction coefficient = 13 600 $M^{-1}$ $cm^{-1}$), and so can be quantitated by absorption spectroscopy.

Danson and Weitzman (1973) found that DTNB modified two cysteine residues per subunit in *E. coli* CS in a time-dependent manner, which resulted in a loss of both

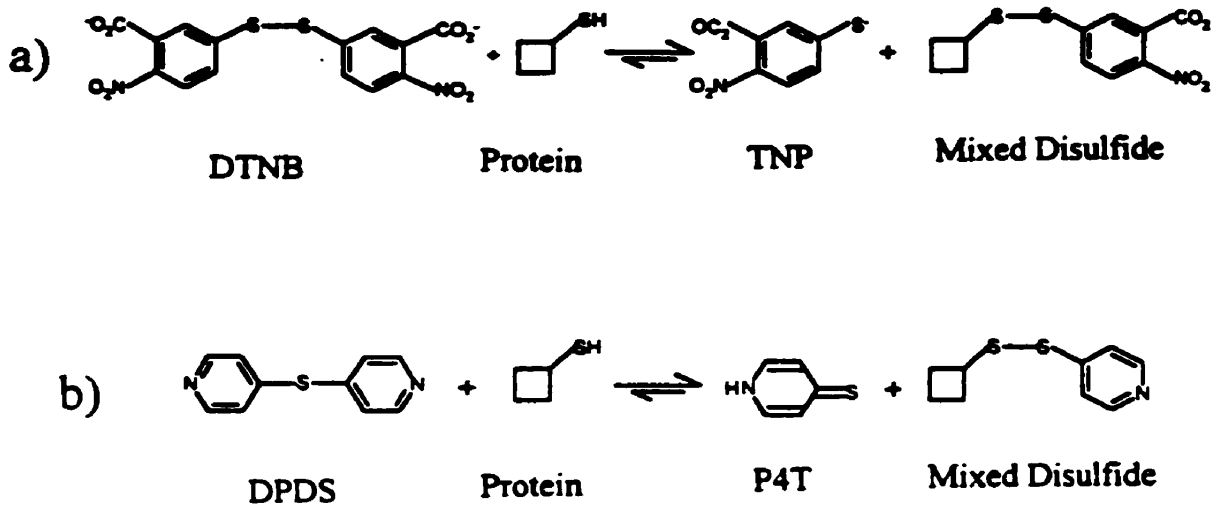



Figure 8: Reactions of thiol reagents with the sulfhydryl group of cysteine.
a) Reaction with DTNB
b) Reaction with DPDS

enzyme activity and NADH inhibition. One cysteine was observed to be more rapidly modified than the other, resulting in an initial loss of up to 50% of CS activity. Slower modification of the second cysteine completed the inactivation, and produced insensitivity to NADH inhibition.

However, a later study disputed these findings, determining instead that only one cysteine per subunit of *E. coli* CS was reacting with DTNB (Talgoy et al., 1979). Further, while this modification desensitized the enzyme to NADH inhibition, no loss of activity was observed. There was also no sign of the initial rapid phase of sulfhydryl modification reported by Danson and Weitzman. Talgoy et al. reported that reaction of CS with 4,4'-dithiodipyridine (DPDS), another thiol reagent similar to DTNB (see Figure 8b) did show modification of two cysteines, as opposed to one reported by Danson and Weitzman. More interestingly, loss of both enzyme activity and NADH was observed. Overall, these results suggested that there is a cysteine residue either at or near the NADH binding site, but there is another cysteine which was involved in an indirect manner in CS activity.

***Identification of C206 as a Marker of the Allosteric Site***

Modification of CS with photo-oxidative dyes, DTNB, and DPDS had implicated a single reactive cysteine as being at or near the binding site of NADH. An important breakthrough in the search for the allosteric site was the identification of this residue as cysteine 206 (C206), by reacting the sulfhydryl group with monobromobimane, and using this fluorescent tag to pick out the appropriate peptide from a cyanogen bromide digest (Donald et al., 1991). This result was confirmed by Edman degradation of the peptide.

At the time, the pig heart model of the CS subunit indicated that C206 was likely on or near the surface, between the J and K helices of the large domain and 25 to 30 Å from the active site (Figure 5). Later experiments with chimeric CS composed of interchanged large and small domains from the *E. coli* and *Acinetobacter anitratum* enzymes demonstrated that the entire NADH binding site was contained within the large domain of the CS subunit (Molgat et al., 1992). The current version of the hexameric *E. coli* CS crystal structure shows C206 to be located near the contact points between dimers (Figure 6b).

As with the other chemical modifications of C206, the reaction of 1,1,1-trifluorobromoacetone (TFBA) with CS, as shown in Figure 9, resulted in the loss of NADH binding without loss of CS activity (Donald et al., 1991). Interestingly, later mass spectrometry experiments also showed that TFBA-modified CS can only form dimers (Ayed, 1998; Ayed et al., 1998), perhaps, in light of the new crystal structure, due to physical hindrance of the interactions of the dimers necessary to assemble the hexamer. This fact was not known when $^{19}F$ NMR was used to monitor changes in the environment of TFBA-labelled C206 during titration with the allosteric activator KCl (Donald et al., 1991). These NMR spectra showed two distinct peaks, whose individual areas gradually changed as the titration progressed while maintaining the same total area. This change in the environment of C206 was interpreted as being the result of the structural changes undergone by CS when the allosteric equilibrium shifts from T to R state.

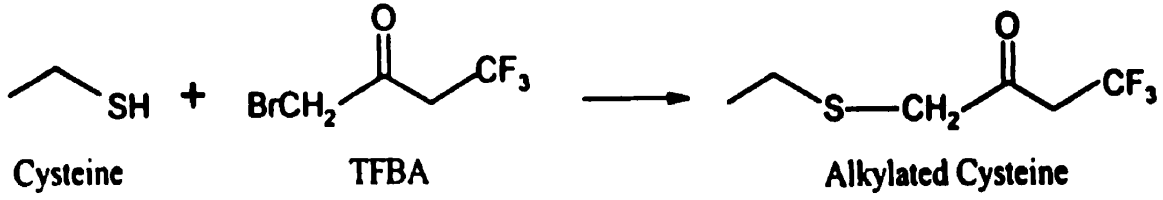



Figure 9: Reaction of TFBA with the sulfhydryl group of cysteine.

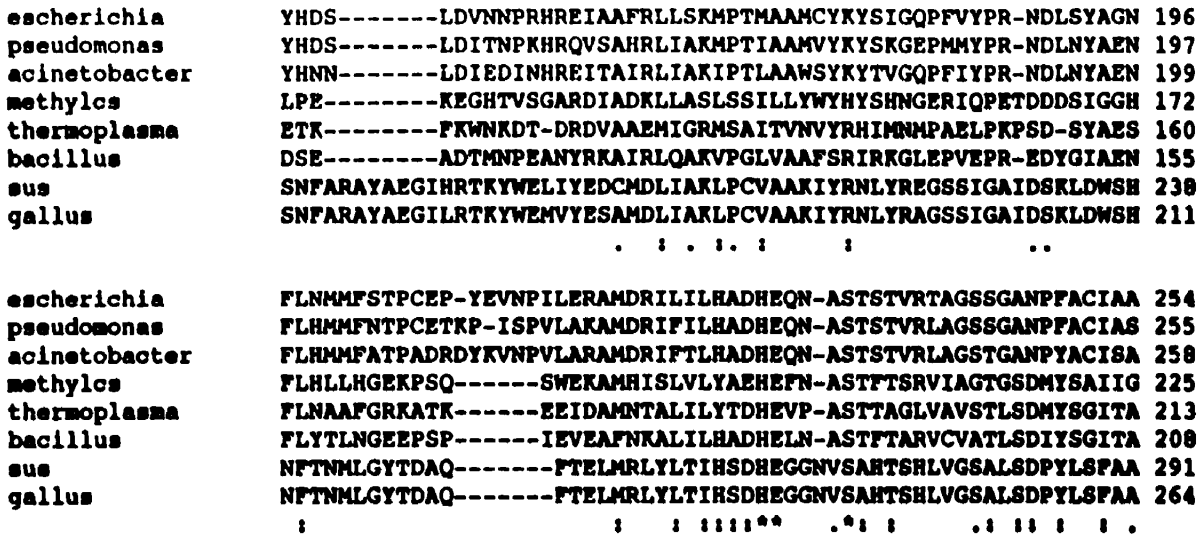



Figure 10: Sequence alignment of selected NADH sensitive and insensitive CS, showing the region surrounding C206, the marker for the allosteric site. See Appendix 1 for the complete alignment.

*Escherichia coli* (Gram negative bacteria, hexamer)
*Pseudomonas aeruginosa* (Gram negative bacteria, hexamer)
*Acinetobacter acidophilum* (Gram negative bacteria, hexamer)
The second (methyl) CS of *E. coli* (Gram negative bacteria, dimer)
*Thermoplasma acidophilum* (Archaea, dimer)
*Bacillus subtilis* (CS II) (Gram positive bacteria, dimer)
*Sus scrofa* (Pig, dimer)
*Gallus gallus* (Chicken, dimer)

*Alignment constructed using ClustalW (default settings) (Thompson et al., 1994)*

** indicates complete conservation of the residue at that position*
*: indicates good conservation of the residue*
*. indicates fair conservation of the residue*

***Site-Directed Mutagenesis of Putative Allosteric Site Residues***

Once C206 was established as a marker for the allosteric site, it was possible to examine the surrounding area in the enzyme by site-directed mutagenesis. An alignment of NADH sensitive and insensitive CS shown in Figure 10 demonstrates significant differences between the two main classes of enzyme in this particular region of sequence. It has been postulated that these differences might reflect structural features necessary for NADH inhibition of CS, assuming that such features would not be found where they were not in operation (ie: in non-allosteric CS) (Anderson, 1988). Thus, a co-requisite for the selection of amino acids for the following mutagenesis studies was that these residues should be conserved in allosteric CS only.

The first three allosteric site mutants produced were C206S, E207A, and a double mutant of both (Donald et al., 1991). As expected, all three mutant CS showed greatly weakened NADH binding, as well as altered substrate affinities and higher KCl activation ratios. The $^{19}F$ NMR experiment was performed on TFBA-modified forms of the mutants. The weak signal obtained for C206S was expected of an enzyme lacking the reactive sulfhydryl group which was typically modified by TFBA. The peak shift observed with wild type CS was not observed for the E207A, a result which seemed to indicate that this mutant was naturally found in the active R state. This theory was strengthened by the increased rate of reaction of this mutant with DTNB, something which usually occurs in wild type CS in the presence of KCl.

The next allosteric site mutants to be examined were selected on the basis of some common features of $NAD^+$ and FAD binding sites in other enzymes (Eklund et al, 1976; Biesecker & Wonacott, 1977; Birktoft et al., 1982; Anderson, 1988). Arginines were

previously found to form key ion pair interactions with the pyrophosphate portion of NAD in certain cases, so the conserved arginine residues predicted as being structurally closest to C206 were replaced with leucines. Unfortunately, mutants R188L, R217L, R221L, and R217/R221L (double mutant) showed no changes in NADH binding and inhibition, although they were found to have greater affinity for both substrates and the competitive inhibitor α-ketoglutarate (Anderson, 1988).

Sternberg and Taylor (1984) observed a G-X-G-X-X-G motif (X=any amino acid) in alignments of NAD and FAD binding sites from glyceraldehyde-3-phosphate dehydrogenase, lactate dehydrogenase, alcohol dehydrogenase, glutathione reductase, and p-hydroxybenzote hydroxylase. Anderson (1988) observed that while the allosteric CS do not contain this exact motif, a somewhat similar sequence of G-X-P-X-X-X-P (residues 181-187) is found in the same region as C206. In the original motif, the glycine residues serve primarily to hold the local conformation of the cofactor binding site in such a manner that specific hydrogen bonding and charge interactions with NAD or FAD are favorable (Sternberg & Taylor, 1984; Wierenga & Hol, 1983). Proline is a residue which, by virtue of its rigid imino ring side chain, is capable of inducing certain structural limitations (Williamson, 1994) which might also serve to define and maintain the allosteric site of CS. Prolines 205, 208, and 213 were thus suggested as candidates for mutagenesis, and mutants P205A and P208A were constructed, purified, and kinetically characterized (Ye, unpublished results).

A second look at the structural features surrounding C206 led to the selection of ten amino acids conserved among the allosteric CSs which at that time (based on the pig CS subunit model) were believed to be either at or near the enzyme surface, and capable of

forming hydrogen bonds with NADH (Hacking, unpublished results). These residues were Y82, Q99, H110, T111, K177, Y186, N189, D190, Y193, and N196. Of this list, both Y186A and D190A mutants were constructed and purified.

The mutants P205A, P208A and D190A demonstrated hyperbolic saturation curves for both substrate binding and KCl activation (Hacking, Ye, unpublished results). Steady state kinetics in the presence of 0.1 M KCl show that, as with C206S, they bind OAA considerably better than wild type enzyme, and all except P205A show enhanced AcCoA binding as well. It has been postulated by Ye (personal communication) that this could indicate greater surface exposure of the active site in these mutants. When considering NADH inhibition, the P205A mutant, like C206S, shows slightly lessened responsiveness to KCl activation, and virtually no NADH inhibition. P208A, by contrast, is more sensitive to both inhibition by NADH and activation by KCl, a strong indication that it may play a role in the allosteric site (Ye, unpublished results). A fluorescence study of NADH binding to D190A gave a $K_D$ for the enzyme-NADH complex of 3.3 μM, as compared to the wild type value at the same pH of 1.9 μM, demonstrating a weakened ability in the mutant to bind the allosteric inhibitor. Table 2 summarizes the characteristics of the best studied allosteric site mutants, and Figure 11 shows their locations in the new crystal structure of hexameric *E. coli* CS.

Table 2: Summary of the kinetic characteristics of some allosteric site mutants of CS

| | WTCS | C206S | E207A | C206S+ E207A | P205A | P208A | D190A | Y186A |
|---|---|---|---|---|---|---|---|---|
| $K_{OAA}$[a] | 26 ± 5 | 1 ± 5 | 160 ± 7 | 2 ± 3 | < 0.5 | < 1 | • | 28 ± 16 |
| $K_{AcCoA}$[a] | 120 ± 20 | 68 ± 9 | 170 ± 60 | 37 ± 9 | 250 ± 43 | 27 ± 2 | 51 ± 7 | 97 ± 31 |
| $K_{iOAA}$[a] | 33 ± 7 | 57 ± 8 | 140 ± 100 | 390 ± 90 | 10 ± 3 | 87 ± 23 | 68 ± 21 | 60 ± 22 |
| $K_m$ KCl | 28 ± 4 | 48 ± 6 | 20 ± 3 | 16 ± 3 | 57 ± 9 | 12 ± 2 | 33 ± 28 | • |
| KCl Act. Ratio | 39 ± 3 | 67 ± 6 | 280 ± 25 | 200 ± 10 | 65 ± 3 | 125 ± 5 | 330 ± 16 | • |
| $K_{iNADH}$ | 1.5 ± 0.4 | 14 ± 2 | 15 ± 4 | 380 ± 140 | 6 | 0.66 | 8.0 ± 0.6 | • |
| Max. % NADH Inhib. | 95 ± 3 | 92 ± 5 | 64 ± 6 | 82 ± 15 | 63 | 0.96 | 87 | • |

[a]Measured in the presence of 0.1 M KCl

WTCS, C206S, E207A, C206S+E207A data from Donald et al, 1991
P205A and P208A data from Ye, unpublished results
D190A and Y186A data from Hacking, unpublished results
• Data not available

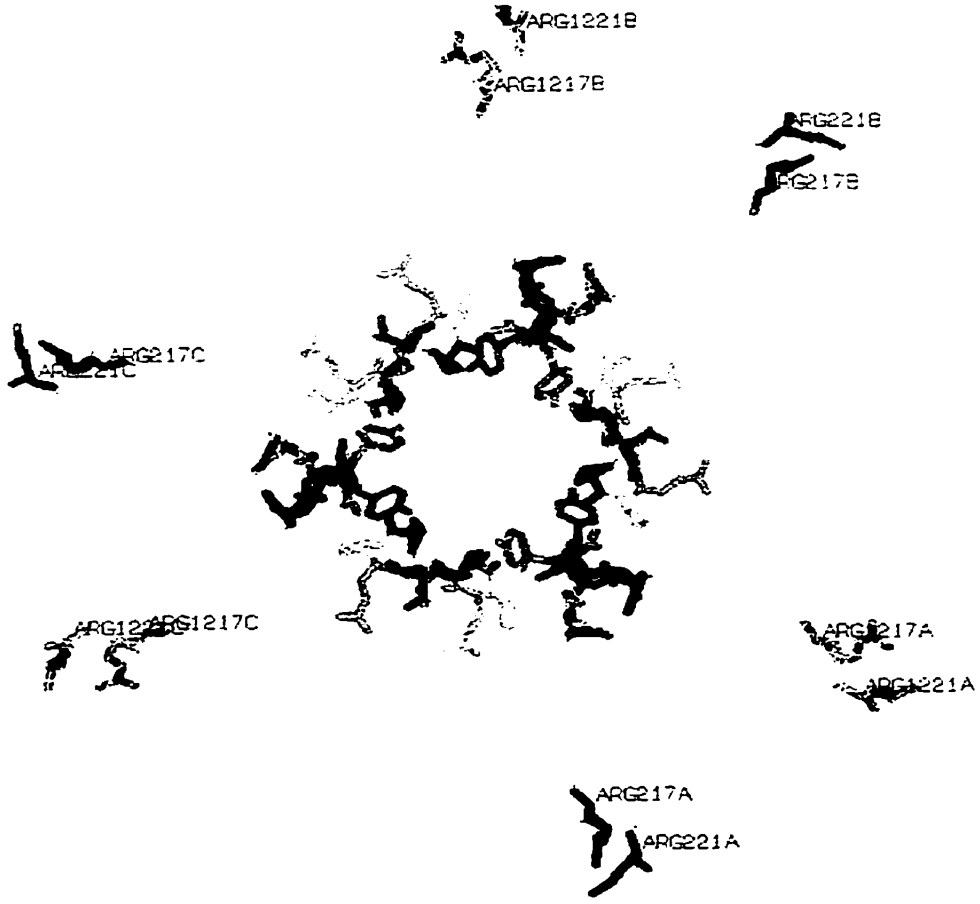



Figure 11: Enlargement of the central portion of the preliminary crystal structure of hexameric *E. coli* CS, showing the positions of residues involved in mutational studies of the allosteric site.
Yellow = C206
Orange = E207
Red = P205, P208
Green = Y186
Blue = D190
Purple = R188, R217, R221

## *Photoaffinity Labeling*

The basis of any type of affinity labeling is essentially that there is sufficient flexibility in the characteristics of most ligand binding sites on a biological molecule to accept the natural ligand derivatized with some specific chemically reactive group. These groups tend to be electrophilic in nature, and so limit the ability of the label to bind to targets which do not exhibit overtly nucleophilic character, such as some regulatory or allosteric sites. In addition, the reactive lifetimes of these labels tend to be relatively long, leading to nonspecific labeling. Organic solvents may be required to keep the chemically modified ligand in solution.

One alternative to classical affinity labeling is called photoaffinity labeling. Photoaffinity probes consist of the desired ligand or substrate to which a photoreactive group is attached. When exposed to a short burst of UV light (typically on the order of several minutes), these groups form highly reactive, short-lived intermediates which proceed to covalently link the ligand to the nearest available structure, theoretically regardless of its chemical nature (Figure 12). Thus the specificity of ligand-binding site interaction is combined with a method of nonspecific linkage.

In a standard photoaffinity labeling study, HPLC or SDS PAGE is used to separate peptides from a proteolytic digest of the modified protein. The peptide(s) bearing the photolabel are typically identified by a radioactive tag incorporated into the probe. Several different proteolytic enzymes producing series of overlapping peptides can be used to narrow down the region of the modification. Confirmation of identity of the labeled residue can then be obtained by amino acid sequencing.

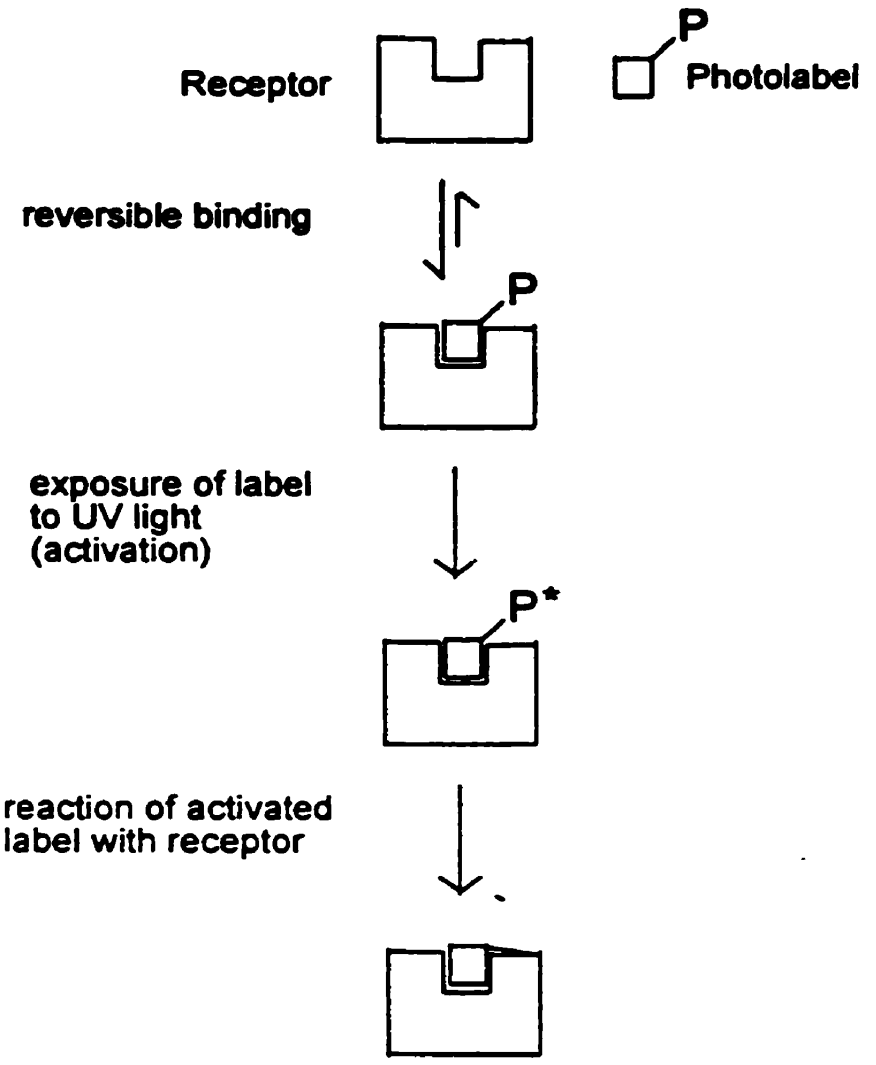



Figure 12: The basic photoaffinity labeling reaction. Initially inert, the probe is first allowed to bind to a specific site on the receptor molecule. The photoreactive group (**P**) is then cleaved by brief exposure to ultraviolet light, producing a highly reactive species (**P***) which indiscriminately forms a covalent linkage between the ligand and surrounding residues of the binding site.

*Adapted from Bayley, 1983*

Since the 1970's, photoaffinity labeling has been used to study a wide variety of biological molecules, including the active and regulatory sites of enzymes, receptors, transport proteins, protein-nucleic acid interactions, and antibodies (Chowdhry & Westheimer, 1979). The photoactive nature of the reactive group means that the binding characteristics of the affinity label can be established prior to covalent linkage. This ability to activate the reactive group of the affinity label at will also circumvents the problems previously encountered with efforts to identify receptors within molecular assemblies such as membranes.

Photoaffinity labeling was introduced as a technique by Westheimer and his co-workers in 1962 (Singh et al., 1962), who demonstrated the photolysis of diazoacetylchymotrypsin with UV light to produce a reactive carbene group. Unfortunately, carbene precursors often suffer from chemical instability and tended to undergo rearrangement reactions when photolysed (Guillory & Jeng, 1983). Aryl azido labels, introduced by Fleet et al. in 1969 give aryl nitrenes upon photolysis with UV light, which are less prone to rearrangements when reacting. In addition, this latter group of photoaffinity probes can be photolysed at wavelengths above 350 nm, which is outside the range typically damaging to most biomolecules. It should be noted that greater photoincorporation efficiencies for both carbene and nitrene type photolabels still occurs using activating light at 254 nm (Bayley & Knowles, 1977). As the photoaffinity labels used in this thesis work were of the aryl azide type, the focus of this section will be narrowed to concentrate on this class of nitrene precursor.

### *The Chemistry of Photoaffinity Labeling*

Nitrenes are very short-lived reactive groups (although slightly longer lived than carbenes), which allows them to covalently link to the target site at rates similar to that of the ligand interaction (Guillory & Jeng, 1983). However, while the basic concept of photoaffinity labeling with nitrenes is quite straightforward, the actual photochemical mechanism is complex and not well understood. Photoreaction products have been shown to depend not only on the substituent groups of the aryl azide, but also on the type of UV light source used (Guillory & Jeng, 1983; Bayley & Knowles, 1977). The singlet nitrenes typically produced by photolysis have an electrophilic character, and will preferentially insert into O-H or N-H bonds, while the less reactive triplet nitrenes formed by intersystem crossing from the singlet state tend towards insertion into C-H bonds, and hydrogen abstraction reactions forming free radicals (Bayley & Knowles, 1977). There is evidence to suggest that insertion of a singlet nitrene into the target molecule would be a faster process than conversion to the triplet state (Guillory & Jeng, 1983). Figure 13 shows some typical reactions of the aryl nitrenes.

One of the commonest types of arylazido photoaffinity probes used to study nucleotide binding sites are the 8-azidopurines. These probes are thought to be conformationally shifted about the N-glycosidic bond from anti to syn, which could reduce their affinity for the binding site architecture (Czarnecki, 1984; Kim & Haley, 1990). The 2-azidopurine analogues retain the anti conformation, but can undergo reversible tautomerization to two nonphotoreactive tetrazole isomers, as shown in Figure 14. The position of the azido-tetrazolo equilibrium in a labeling solution has been shown to be dependent on the pH, temperature, and polarity of the labeling reaction solvent (Czarnecki, 1984).

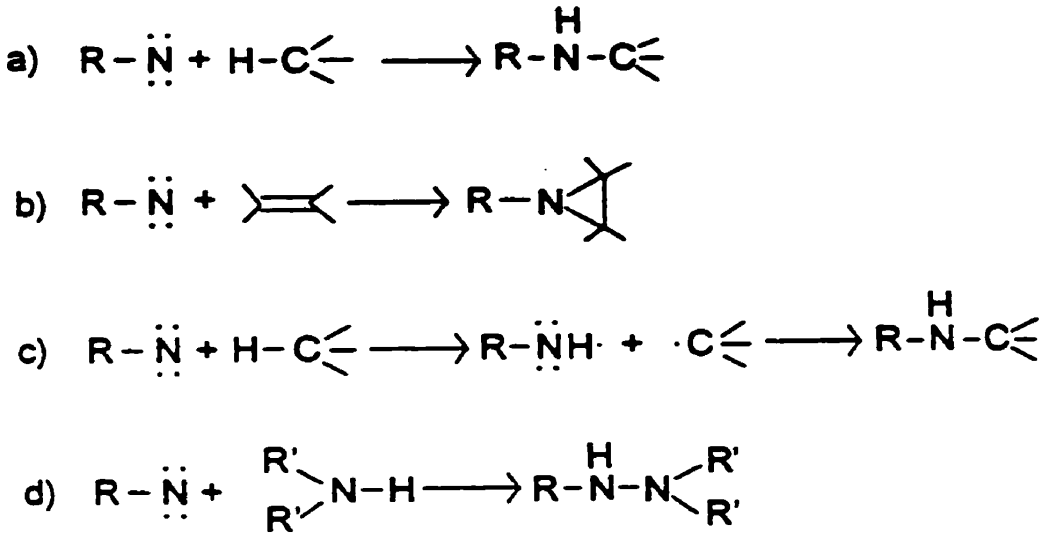


Figure 13: Some typical reactions of nitrenes.
a) Insertion into a single bond to give a secondary amine.
b) Cycloaddition to a multiple bond to give a three-membered imine ring.
c) Hydrogen abstraction to give two free radicals which then couple.
d) Addition to nucleophiles.

*Adapted from Gaillory & Jeng, 1983*

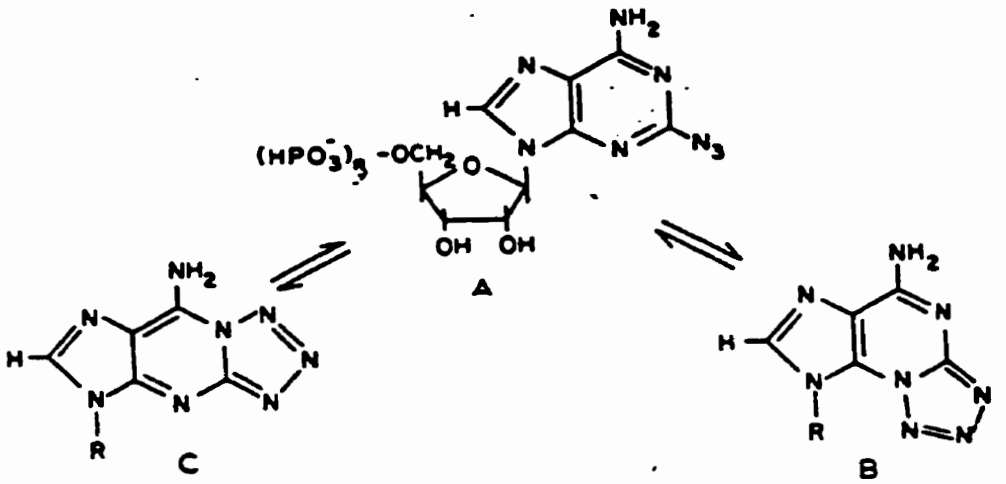



Figure 14: Tautomerization of a 2-azidopurine photoaffinity label.
A = 2-azidoadenosine – 5'triphosphate
B = tetrazolo [5, 1 – b] adenosine tautomer
C = tetrazolo [1, 5 – a] adenosine tautomer

*Adapted from Czarnecki, 1984*

***Experimental Design Considerations***

For all the advantages of photoaffinity labeling, there are a number of potential problems of which the experimenter must be aware. Key among these are low yields of photoincorporation, nonspecific and pseudophotoaffinity labeling, and uncertainty in identification of the covalent modification.

The most obvious cause of low rates of photoincorporation is the label itself, wherein the azido group might be expected to increase the binding constant for a particular receptor. Often, it is found that one type of photolabel is favored over another (e.g. 2- and 8-azidopurines), and this information may even serve to help define the binding site. Though the nitrene group is described as being indiscriminate in its reactivity, certain functional groups seem to react more efficiently, depending on the azido probe in use, although this claim is based purely on accumulated experimental observations (Bayley & Staros, 1984). Simply increasing the concentration of photolabel in the experiment usually increases the amount of nonspecific labeling (Bayley & Knowles, 1977; Czarnecki et al., 1979), and so for cases of low levels of photoincorporation, multiple rounds of labeling with saturating concentrations of probe are preferable.

Nonspecific photolabeling is more of a problem with weaker binding interactions (binding constants on the order of $10^{-5}$ to $10^{-6}$ M) and longer lived reactive groups which may diffuse in and out of the binding site several times before forming a covalent bond, a phenomenon called pseudoaffinity labeling (Chowdhury & Westheimer, 1979). There are several control experiments performed as part of any photolabeling study to determine the specificity of label incorporation. Labeling may be performed in the absence of activating light to demonstrate the stability of the probe, and increasing the

amount of probe should result in saturation of the targeted binding site. The level of nonspecific photoincorporation can be judged by comparing the results of photolysis in the presence and absence of the natural ligand, which should protect the binding site from labeling.

If nonspecific labeling is a problem, scavengers may be added to destroy unbound probe and prevent it from binding to other locations on the molecule (Figure 15). However, care must be taken that they do not unnecessarily lower specific photoincorporation further, and in cases where ligand affinity for the target site is already weak, they may be of little use (Bayley & Staros, 1984). Compounds that have been used as scavengers include β-mercaptoethanol, dithiothreitol, Tris, hydroquinone, p-aminobenzoic acid, and p-aminophenylalanine.

Another approach to eliminating nonspecific labeling is to protect the target receptor with the natural ligand, modify as many other sites on the molecule as possible using nonradioactive photolabel, remove the natural ligand, and then proceed with radioactively tagged photolabel (Guillory & Jeng, 1983).

### *Analysis of Labeled Proteins and Peptides*

Theoretically, locating the incorporated photolabel in a mixture of proteolytic peptides should be straightforward, even if the exact nature of the final photoproduct is not predictable. In addition to detection of radioactivity from the label, the electrophoretic and chromatographic mobilities of the peptides bearing the modification tend to differ from their unmodified counterparts (Bayley & Knowles, 1977). If there is a certain amount of flexibility or looseness in the interaction of probe and target site, or if a free

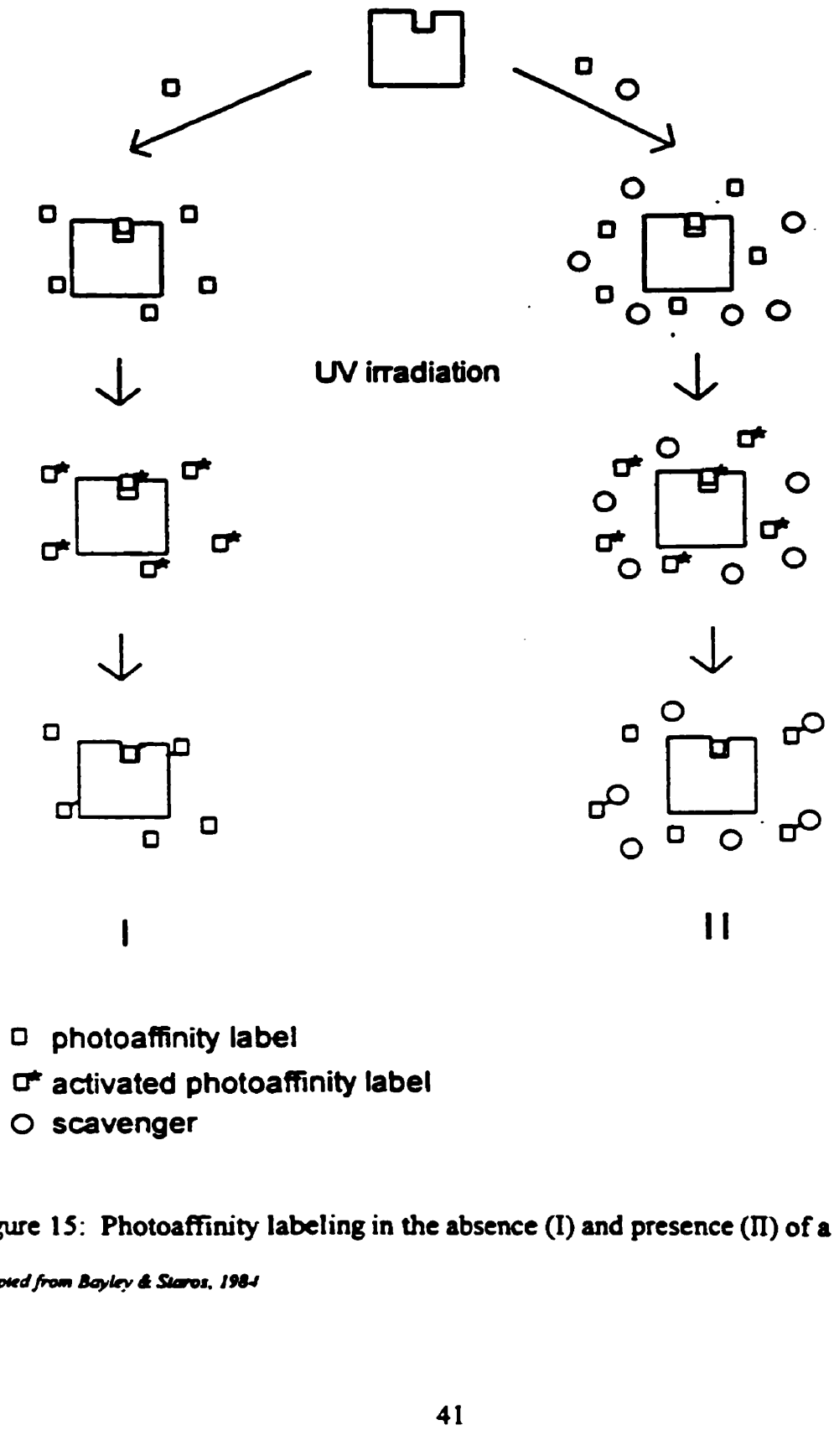



Figure 15: Photoaffinity labeling in the absence (I) and presence (II) of a scavenger.

*Adapted from Bayley & Staros, 1984*

radical species is produced by photolysis, there may not one, but several proteolytic peptides bearing the modification (King et al., 1991). It is then up to the investigator to reconstruct the conformation of the binding site based on this information.

The most serious difficulty is when no modification can be detected. An obvious explanation for this might well be that levels of photoincorporation are too low. However, it is also possible that labeling did occur, but was lost during analysis of the sample. There are conflicting reports as to the stability of the photoinserted bonds. For instance, it has been noted in the literature that some of these covalent bonds are stable to cyanogen bromide treatment, whereas others are cleaved (Bayley & Knowles, 1977; King et al., 1991). Haley found that the insertion of 8-azido cyclic AMP into membrane proteins was stable to both acid and boiling treatments, and states that "We have found nothing that will remove photoincorporated, covalently bound label without also destroying the primary protein chain." (Haley, 1977).

Another experimental finding is that typical reversed-phase HPLC conditions can result in loss of photolabel, such that many groups using photoaffinity labeling have switched to using immobilized metal affinity columns and lower flow rates for separation of proteolytic peptides (Chavan, 1996; King et al., 1991; Salvucci et al., 1994). The work described in this thesis attempts to circumvent this problem, and also avoid the larger amounts of protein and photoprobe required for HPLC separations by using mass spectrometry to examine unfractionated digests of photolabeled CS.

# *Mass Spectrometry and the Study of Proteins*

The development of MALDI (matrix assisted laser desorption ionization) and ESI (electrospray ionization), two "soft" methods of ionizing biological molecules first introduced in the late 1980's (Karas & Hillenkamp, 1988; Fenn et al., 1989), has led to a great increase in the usage of mass spectrometry (MS) in biological research, particularly in the study of proteins and peptides.

### *Sample Preparation and Ionization*

In MALDI, the protein or peptide solution to be analyzed is mixed with a matrix, which then crystallizes when layered and dried on a solid support (the target). The analyte is thought to be embedded in the surface of the crystals. When laser light (ultraviolet or infrared) is applied to the crystals, the result is a plume of ionized sample and matrix (Figure 16). While the exact mechanism by which ionization occurs is still debated, it is believed that the matrix assists in the process by keeping analyte molecules well separated, and by absorbing the laser light energy and transmitting it to the sample.

It may be necessary to test a number of matrices and matrix solvents in order to optimize the ionization of a sample. Typical matrices for ultraviolet lasers (eg: $N_2$ lasers emit at 337 nm) include sinapinic acid, α-cyano-4-hydroxy-*trans*-cinnamic acid, and 2,5-dihydroxybenzoic acid. Glycerol is a common matrix for infrared MALDI. Solvents in which the matrix is dissolved (usually at saturating levels), can include varying percentages of acetonitrile, methanol, acetone, and small amounts of acids such as

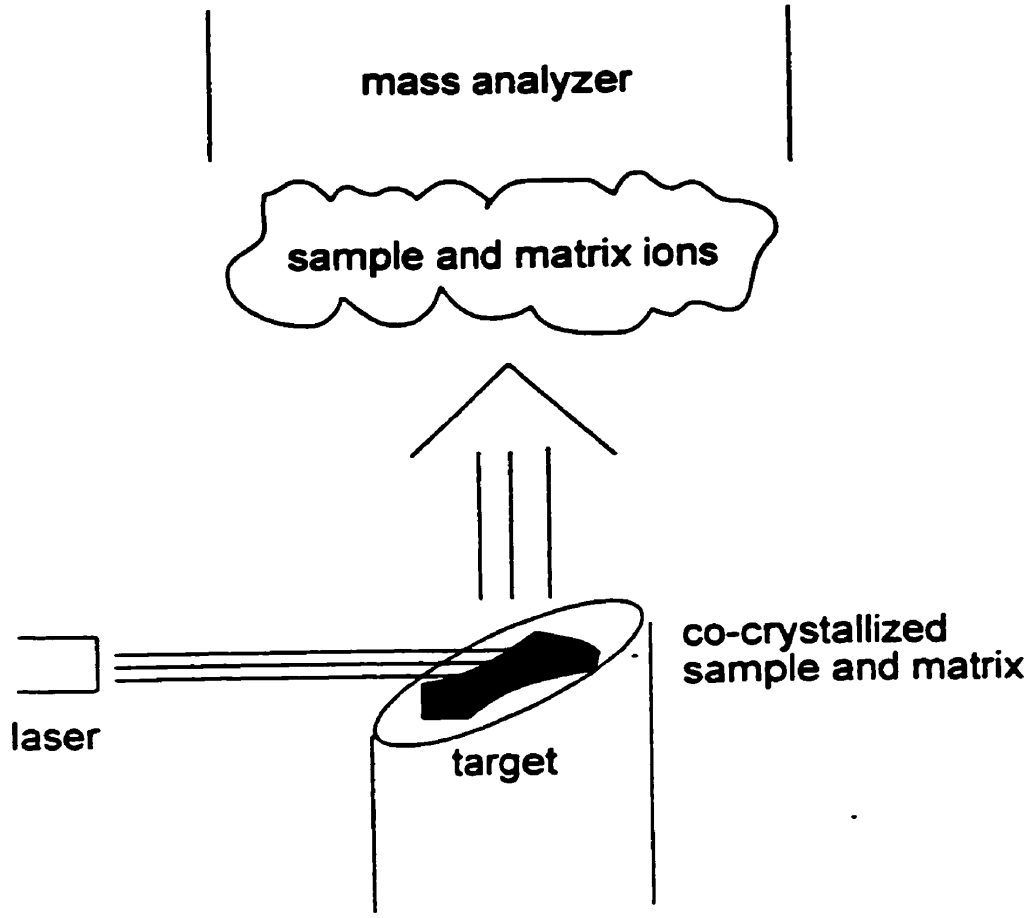



Figure 16: Diagram of a matrix-assisted laser desorption/ionization (MALDI) source. Laser light (infrared or UV) is focused on a metal target coated with a crystalline mixture of sample and matrix.

trifluoro- or trichloroacetic acid. Several methods also exist for the layering and drying the matrix/sample mixture onto the target.

Biological buffer components such as salts and detergents are usually present in far greater concentrations than the actual proteins or peptides to be analyzed, and so are largely removed prior to mass spectrometry lest their ions swamp the signal of the intended analyte. Such impurities can also form multiple adducts with the analyte itself, producing peaks in the spectrum which are broader and poorly resolved. Nonetheless, MALDI is more tolerant of buffer components when compared to ESI, which requires samples to be virtually free of salts and detergents, and in either an organic-based solvent or a low concentration of a volatile buffer such as ammonium bicarbonate or ammonium acetate.

An ESI source consists of a charged needle, through which a solution of proteins or peptides is infused. When this solution passes through the needle tip, a cloud of tiny, highly charged droplets is formed (Figure 17). The needle can produce either positive or negative charging, depending on the polarity of the applied electric field. Proteins are usually analyzed in positive mode. As the droplet cloud is drawn into the high vacuum internal region of the mass spectrometer, it passes through a flow of hot $N_2$ gas, which produces desolvated ions of the analyte. These ions can then be analyzed by their mass to charge ratio (m/z). This ratio tends to be lower for ions produced by ESI, which are generally more highly charged than those produced by MALDI. Instead of one or two peaks, then, an ESI spectrum of a protein ion will contain a charge "envelope", or distribution of multiply charged forms of the protein. A mathematical program is used to "deconvolute" the envelope into a single peak representative of the ion. Another notable

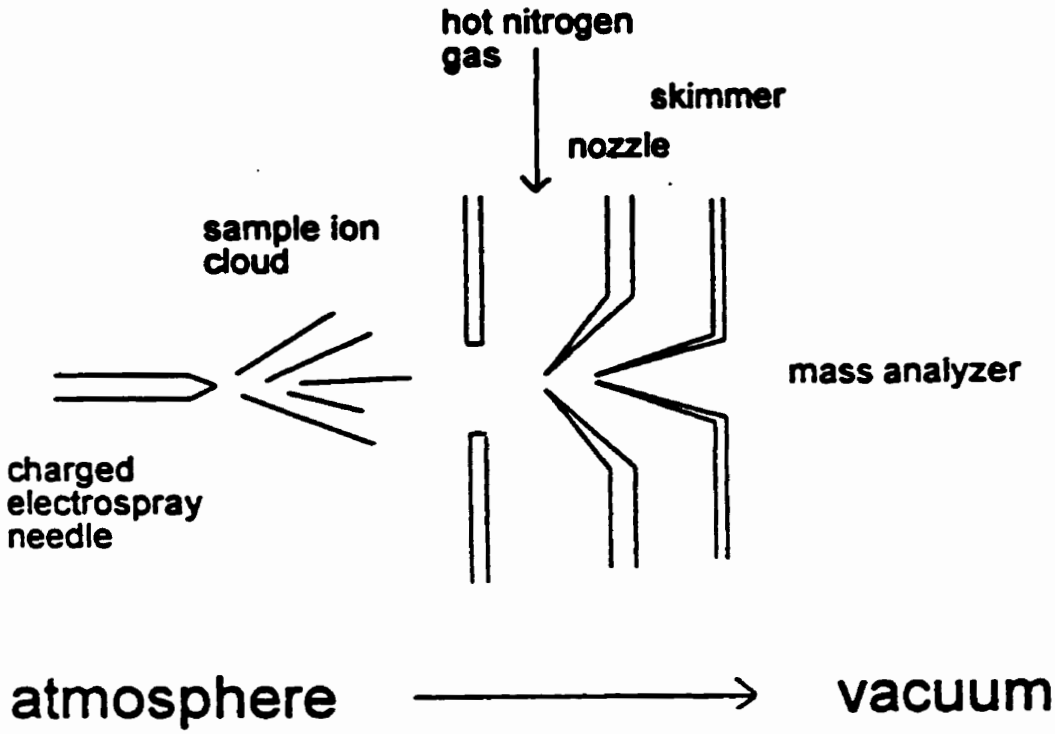



Figure 17: Diagram of an electrospray ionization (ESI) source. Liquid sample is applied through the charged needle to produce a cloud of charged droplets. As the droplets are drawn into the mass spectrometer, they are desolvated by passage through a region of heated nitrogen gas.

feature of ESI MS is that non-covalent interactions are conserved in the ionization process, allowing for the study of protein complexes. This is not the case with MALDI, although it has been recently reported that under certain carefully controlled circumstances, non-covalent complexes can be seen in MALDI spectra (e.g. Gruic-Solvuj et al., 1997, and references therein).

Finally, an alternate ion source for ESI-configured mass spectrometers mentioned in this thesis is the nanospray source (Wilm & Mann, 1994), which involves the use of an ultrafine capillary in place of the standard electrospray needle. The two key advantages of this source are the requirement for less sample than ESI, and the ability to use sample buffers of higher concentration. Unfortunately, the narrow capillary opening means that it clogs more easily.

### *Mass Analyzers*

Two of the most popular mass analyzers for ions produced by MALDI and ESI are time-of-flight (TOF) and quadrupole analyzers. A TOF analyzer separates ions as they pass down a high vacuum region of set length within the mass spectrometer and collide with a detector. Assuming that all ions are given the same initial kinetic energy when accelerated into the flight tube, separation occurs on the basis of the m/z ratio such that the larger the ratio, the longer the flight time. Theoretically then, TOF mass analyzers should have an unlimited mass range. Calibration using protein/peptide standards such as substance P, bovine insulin, myoglobin, and bovine serum albumin allows for translation of TOF data into m/z ratios. Two types of TOF analyzers exist: Linear, in which the flight path is straight and in one direction, and reflecting, where ions in flight are

reflected off an electrostatically charged surface ("mirror") and back towards a detector, increasing the length of the flight path and compensating for energy differences between ions.

Quadrupoles are bundles of four rods, electrically insulated from each other. Each diagonal pair carries either DC (direct current) or RF (radio frequency) potential. As ions pass down the space at the center of the quadrupole, the potentials in all four rods are increased uniformly until a given time, and then reduced to zero again. This process, or mass scan, is repeated, to give oscillating hyperbolic potentials on the inside surfaces of the quadrupole. The quadrupole thus acts as a mass filter, where the balance between RF and DC potentials in the quadrupole determines which ions may pass through to the detector, and which collide with the rods and are neutralized. This balance can be adjusted to select for a wide or extremely narrow range of m/z. However, most commercially available quadrupoles have an upper m/z limit in the region of 4000, which means that they cannot handle some larger biomolecular complexes.

The resolution obtained from a mass spectrometer is strongly influenced by the method used to introduce the ion clouds produced by MALDI or ESI into the high vacuum inner regions of a TOF mass analyzer. A TOF analyzer is a pulsed system, injecting packets of ions into the flight tube by the application of a set accelerating voltage at regular intervals. There are many energetic and kinetic factors to be considered in this process, and what follows here is an extreme simplification of the primary problem encountered and some solutions.

The ion clouds from MALDI and ESI tend to be diffuse, containing ions with a wide distribution of velocities and positions within the cloud. When this is injected directly

into a TOF mass analyzer, the ions are then starting their journey down the flight tube from slightly different positions, velocities and energies, although the accelerating voltage applied is the same for all. Collisions which occur as the ions are being accelerated further widen the velocity spread of the ion cloud. The result is variations in the actual flight times within an ion packet, and so reduced resolution in TOF mass spectrometers where such direct, or "axial" ion injection is used.

In MALDI mass spectrometers, a technique called delayed extraction or time-lag focusing is used to circumvent this problem. The ion cloud is allowed to drift in a field free region prior to the application of the accelerating pulse, and the resulting expansion of the plume along the axis of the applied potential means that ions receive this voltage in proportion to their position along the plume. This adjustment thus compensates somewhat for the spread of ions within the injected cloud, and increases resolution.

Another method of increasing spectral resolution is orthogonal injection, where the ion cloud is injected into the high vacuum region of the instrument along a plane perpendicular to the flight path. The ions arrive in a field free ion storage region defined by two parallel plates producing a narrowed ion packet. The application of the accelerating voltage from the bottom plate means that the ions in this packet are then extracted into the TOF tube all relatively at the same time (for full review see Ens & Standing, 1999).

Further improvements to resolution may be obtained by focusing and slowing down the ions before the injection process. This can be achieved by passing the ion cloud as it exits the ion source through a quadrupole called a collisional cooling ion guide. As previously described in this section, the balance of RF and DC fields in the quadrupole

serve to filter and focus the ion beam, while collisions between ions and collision gas molecules within the quadrupole slow the ions down ("collisional cooling"). The end result of these two processes is a narrower ion beam which is then injected into the mass spectrometer.

### *Application of Mass Spectrometry to Protein Biochemistry*

High accuracy mass determination by mass spectrometry has found many applications in the analysis of proteins and peptides. The presence of certain post-translational modifications such as phosphorylations and acetylations, alterations to the primary structure of a protein (e.g. amino acid deletions/additions/substitutions by site-directed mutagenesis), and the identity of a protein can be confirmed by comparing the theoretical and observed masses. Modifications can then be located using peptide mapping. This involves enzymatically or chemically digesting the protein into specific peptides, and looking for variations in these peptide masses from the expected masses calculated from the unmodified sequence. Once a potentially modified peptide is found, it can be sequenced either by conventional means (generally peptide isolation by liquid chromatography followed by amino acid sequencing), or directly by a mass spectrometer able to carry out MS/MS. In the latter case, the digest is first injected into the mass spectrometer, and peptides separated by m/z by the mass analyzer (typically a quadrupole). This is the first MS stage. The modified peptide sought is selected from the digest mixture by careful adjustment of the DC and RF voltages of a second quadrupole in the injection path. The peptide is then passed into another quadrupole where it is fragmented by collisions with gas molecules. Finally, the resulting fragments are

separated by m/z in another mass analyzer (quadrupole or TOF). This is the second MS stage. Knowing that peptides will fragment stage-by-stage in a certain manner, the sequence can then be reconstructed from this final mass spectrum. Any deviation in the pattern of expected peptide fragments thus represents the location of a modification and confirms the type of modification. MS/MS also has the advantage of being able to sequence past N-terminally blocked peptides, unlike conventional amino acid analyzers.

As mentioned briefly above, ESI (and possibly MALDI) MS can maintain non-covalent interactions such as protein-protein, protein-ligand, and protein-DNA interactions. Certain aspects of protein structure and function can thus be analyzed by MS. One example of such a study already mentioned in the introduction to Citrate Synthase is the characterization of the CS dimer-hexamer ratios during titration with the allosteric inhibitor NADH (Ayed et al., 1998). By careful calibration of the detector's differing sensitivity to dimer and hexamer, the dimer and hexamer peak areas in the presence and absence of NADH were used to calculate binding constants for this ligand.

Mass spectrometry is also being used to demonstrate changes in protein conformational states. One can easily see the difference between native (folded) and denatured protein by the change in their ESI charge envelopes. Native protein will produce a narrower distribution of charge states than denatured protein, as fewer amino acids with the potential for charging are directly accessible on its surface. The conformational changes of native proteins have been monitored in a deuterated solvent by observing levels of hydrogen-deuterium exchange (Robinson, 1996).

A current area of great interest is the development of mass spectrometry systems capable of dealing with the large numbers of samples generated by proteomic studies,

where not one, but the entire set of proteins expressed in a given tissue or organism is examined (eg: Jungblut & Thiede, 1997 and references therein). Characterization by mass of the contents of chromatography fractions or spots from two-dimensional gels allows for faster optimization of initial protein separation and purification. Once a purified protein sample is available, MS is used to determine the masses of peptides obtained under specific digest conditions (e.g. cyanogen bromide, trypsin, etc.). This information, and any peptide sequence data (perhaps obtained through MS/MS) is then fed into databases of enzymatically-derived peptide masses and sequences from known proteins, in order to determine if the protein has been previously isolated, or has any close sequence neighbors. Rapid, accurate, and efficient determination of protein and peptide masses by mass spectrometry in proteomic studies is achieved by new methods of low-volume high sample throughput (e.g. large MALDI targets containing many samples on one surface), increasing automation of sampling and data analysis, and the coupling of the mass spectrometer to other separation methods such as liquid chromatography and capillary electrophoresis to provide an extra "dimension" of sample separation.

## *Thesis Objectives*

The original intent of this thesis was to clarify the location of the allosteric site of *E. coli* CS by photoaffinity labeling, specifically by using azidoadenosine-type labels. Mass spectrometry would then be performed on unfractionated proteolytic digests of labeled and unlabeled CS and the results compared in order to identify the photolabeled peptides. It was hoped that the elimination of separation steps prior to MS of peptides (e.g. HPLC) would not only simplify the process of mapping photoaffinity labeling of enzyme binding sites, but also limit the loss of photolabel during analysis.

This thesis describes the attempts to map photoaffinity labeling of CS peptides by MALDI TOF and ESI TOFMS, using both unfractionated tryptic digests of CS and tryptic digests partially separated by capillary electrophoresis. It also describes the mapping of certain chemical modifications targeted at the two reactive sulfhydryl groups of CS.

The usefulness of mass spectrometry in the analysis of *E. coli* CS as a functioning enzyme has previously been demonstrated (Ayed et al., 1998). Information presented in this thesis continues to show the utility of ESI TOFMS in the analysis of certain characteristics of *E. coli* CS, including its behaviour in the presence of allosteric activators, and under limited proteolysis conditions. Mass spectrometry results are then compared to those previously obtained by more conventional means.

Finally, the results of project course work carried out on allosteric site mutants of CS, and the manner in which these results relate to those previously obtained for other CS mutants, are also discussed.

# *Materials And Methods*

## *Purification of CS*

The expression plasmid pESgltA, bearing the *E. coli* citrate synthase gene, was transformed into *E. coli* strain MOB 154, which contains a stable mutation in the gltA gene (Wood et al., 1987). The method of expression and purification of citrate synthase was adapted from that of Duckworth & Bell (1982). Briefly, cells were grown up in 2-3 L of LB broth with 25 μg/ml ampicillin (LBA), harvested by centrifugation, resuspended in 3 volumes of standard citrate synthase buffer (20 mM Tris-Cl, 1 mM EDTA, 50 mM KCl pH 7.8), and broken using an Aminco French Press. The lysate was then centrifuged to remove cellular debris, and the supernatant was loaded onto a 10 x 5 cm diethyl aminoethyl-cellulose (DEAE) anion exchange column equilibrated with standard CS buffer. Any unbound protein was removed by washing the column overnight with 2 L of buffer. An elution gradient of 50 to 300 mM KCl was then applied, and fractions of approximately 12 mL were collected. The presence of CS in these fractions was established by assaying first for protein using absorbance at 278 nm, and then for enzyme activity using the standard CS assay, which is described elsewhere in this chapter. Wild-type CS tended to elute approximately half way through the linear gradient, and constituted the major protein peak of the elution profile.

Fractions with the highest specific activities were pooled and concentrated to approximately 10 mL by either vacuum dialysis using 12 000-14 000 molecular weight cut-off Spectra/Por dialysis tubing or ultrafiltration through an Amicon cell with a YM 100 membrane. The concentrate was loaded onto a 110 x 4 cm Sephadex G-200 size exclusion column (separation range $5 \times 10^3 - 6 \times 10^5$), and eluted with CS buffer into 12

mL fractions. Once again, protein absorbance at 278 nm and CS activity were used to establish the elution profile, and fractions of high specific activity from the one major peak observed were pooled and concentrated as before. The final yield of CS was usually in the range of 50 to 100 mg. The enzyme was concentrated to at least 40 mg/ml, as CS remains stable for longer periods of time at higher concentrations, and stored at 4°C. Concentrations were determined using a molar extinction coefficient of 47 699 $M^{-1}$ $cm^{-1}$ (Ayed, 1998).

Purification of the CS mutants P205A and Y186A used in allosteric site studies was as described above, using the appropriate expression plasmids. Preparation of the deletion mutant Δ(264-287)CS and its variations from the wild type CS protocol has been mentioned in Anderson & Duckworth (1988). The method is essentially the same as for wild type, except that the enzyme assay cannot detect the mutant, as the loss of the key active site residue H264 within the 24 amino acid deletion destroys the enzyme's activity. The presence of Δ(264-287)CS was instead established by SDS PAGE and Ouchterlony double-diffusion plates (Ouchterlony, 1953) using rabbit antiserum raised against the wild type enzyme.

***The CS Assay***

All assays for CS activity were based on the method of Srere et al. (1963). The reaction of OAA and AcCoA catalyzed by CS forms both citrate, and CoASH, whose free sulfhydryl group can in turn react with DTNB (in the same manner as described for the modification of protein sulfhydryl groups in Figure 8a) to give the yellow nitrophenolate (TNP) ion which absorbs at 412 nm. It should be noted that as there is

generally so little CS present in the assay mixture, the contribution to the nitrophenolate concentration from protein sulfhydryl modification is negligible.

CS activity was recorded as U/ml or per mg protein, where 1 U is the amount of CS required to produce 1 µmol of TNP (assumed equivalent to CoASH) in 1 minute at room temperature. This value was determined by adjusting the rate of the assay reaction ($\Delta A_{412}$/minute) for dilution and unit conversion factors, and dividing by 13 600 $M^{-1}cm^{-1}$, the molar extinction coefficient of TNP.

## *Chemical Modification of CS*

### *Photoaffinity Labeling*

Photolabeling methodology was primarily derived from a technical bulletin published by the now-defunct Research Products International (RPI) Corp. (Chavan, 1996). Unless otherwise indicated, all photolabeling of CS was done in 20 mM ammonium bicarbonate buffer lacking both EDTA and KCl, mostly for reasons of compatibility with mass spectrometry, but also because in certain cases EDTA and KCl have been found to interfere with photolabeling (Chavan, 1996). As an activator of CS, KCl also relieves NADH inhibition of the enzyme, a situation undesirable in a study of the NADH binding site.

While the azido photoprobes are said to be stable in normal room lighting (Chavan, 1996), precautions such as the transport of samples in light-proof boxes, and turning off as many overhead lights as possible, were taken in order to shield the probes from exposure to light prior to photoactivation with the ultraviolet lamp.

The labels themselves, both radioactive and non-radioactive, were supplied as triethylammonium salts in absolute methanol and obtained from three different distributors (RPI Corp., Andotek Inc., and most recently ICN Pharmaceuticals). Attempts to synthesize 8-azidoATP in the lab using published methods (Haley, 1977; Czarnecki *et al.*, 1979) were time-consuming and gave unsatisfactory yields. Diagrams of 8- and 2-azidoATP, the two photolabels used in this thesis, are shown in Figure 18.

Preliminary experiments to characterize azidoATP binding to CS required the use of $\gamma$ $^{32}P$ radioactive photolabels. The labels were aliquotted into 1.5 mL microfuge tubes, dried down in a Savant Speedvac, and redissolved in a volume of CS solution in 20 mM ammonium bicarbonate (generally 0.4 mg/mL) to give the correct final concentration of azidoATP. To conserve radioactive photolabel, some experiments used a mixture of radioactive and non-radioactive label at a known molar ratio. Final volumes of this labeling solution ranged from 2.5 to 5 μL, containing 1 to 2 μg of CS. This mixture was incubated on ice for 2 minutes, and then exposed to UV light from a Mineralight R-52G lamp with the filter removed, at a distance of approximately 5 cm, for the desired length of time. The samples were then immediately mixed with SDS PAGE sample loading buffer (0.3 M Tris-Cl, 3.6 M 2-mercaptoethanol, 10 % SDS, 50 % glycerol, with bromophenol blue), vortexed, and loaded onto 8% SDS PAGE gels. As boiling samples can cause loss of radiolabel (Chavan, 1996), this step in the standard SDS PAGE sample preparation method was omitted. Electrophoresis was performed as per the standard SDS PAGE procedure.

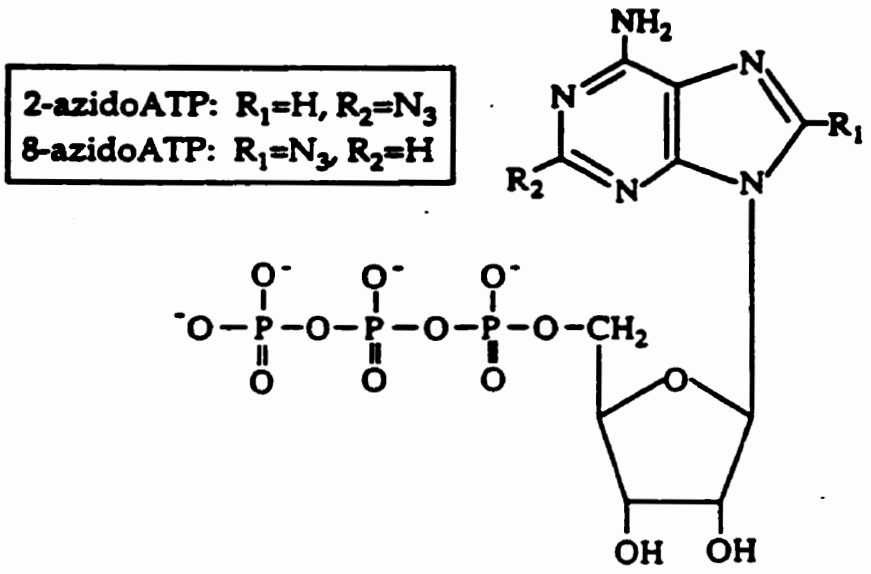



Figure 18: Structure of the azidoATP photolabels used in this thesis for photoaffinity labeling.

Autoradiography was used as a preliminary, qualitative method of judging the extent of photolabel incorporation into CS. The finished gels were wrapped in plastic wrap and exposed to X-ray film overnight.

Quantitation of photolabel incorporation was done by liquid scintillation counting of CS gel bands. Gels were fixed in 25% isopropanol/10% acetic acid for 20 minutes, stained for up to three hours with Coomassie brilliant blue in isopropanol/acetic acid, and rinsed briefly with fix solution. Two standards containing 75 to 100 μg of citrate synthase were run on the outside lanes of each gel, and showed prominently once stained, without the need for destaining. Using these visible standards as a guideline, sample bands were excised as gel slices approximately 0.5 x 2 cm in size. As controls, one standard band, and a comparable piece from an unused lane were also collected. Gel slices were immersed in 5 mL of Amersham Aqueous Counting Scintillant, and counted immediately in an LKB-Wallac 1215 Rackbeta II Liquid Scintillation Counter. Readings in counts per minute (CPM) were converted to percent photoincorporation by the following calculations (Chavan, 1996):

$$\mu Ci = CPM/2.2 \times 10^{-6}\ CPM \cdot \mu Ci^{-1}$$

$$\text{nmole photolabelled protein} = \mu Ci/\text{specific activity of probe on day of experiment}$$

$$\%\ \text{photoincorporation} = (\text{nmoles photolabelled CS monomer/nmoles CS monomer in sample})100$$

Once the ideal conditions for photoaffinity labeling of CS were established by preliminary experiments, modification of the enzyme with non-radioactive probe for tryptic mapping by mass spectrometry followed virtually the same procedure, with the elimination of the SDS PAGE step. Once it was determined that UV irradiation of CS

diminished its activity (see Results), both labeled and unlabeled CS samples used in the tryptic mapping experiments were irradiated alongside each other.

Photoaffinity labeling of creatine phosphokinase used a scaled down version of the method of Olcott et al (1994). The labeling mixture contained 34 μg of rabbit brain creatine kinase (Sigma) in 10 mM $NaH_2PO_4$, 1 mM $MgCl_2$, pH 7.0 with 5 mM creatine. The protein was subjected to two-30 second rounds of photolabeling, with either 50 μM 8-azidoATP or 10 μM 2-azidoATP, with the lamp at a distance of 4 cm from the sample.

***DTNB and DPDS Modification***

A reaction mixture of 700 μL 20 mM Tris-Cl/1 mM EDTA pH 7.8 buffer, 100 μL 1 M KCl and 100 μL of either 1.0 mM DTNB or 0.15 mM DPDS stock in the Tris/EDTA buffer was prepared in a 1 mL spectrophotometer cuvette. The modification reaction was started by the addition of 100 μL of 0.12 mM CS stock, and the absorbance at 412 nm (DTNB) or 324 nm (DPDS) monitored at two minute intervals for an hour. Careful note was made of the time lapse between addition of CS to the reaction mixture and the start of data acquisition, and the absorbance readings were extrapolated back to time=0. By determining the total change in absorbance since time=0, the molar concentration of nitrophenolate (TNP) or pyridine-4-thione (P4T) accumulated at each time point was determined as the following:

$[TNP]_{t=tn} = \Delta A_{412}/13600$ where 13600 $M^{-1}cm^{-1}$ is the molar extinction coefficient of TNP

$[P4T]_{t=tn} = \Delta A_{324}/19800$ where 19800 $M^{-1}cm^{-1}$ is the molar extinction coefficient of P4T

Dividing this concentration by the final subunit concentration of CS in the reaction mixture gives the number of moles of NTP or P4T released per mole of enzyme subunit, and so the number of modified sulfhydryl groups per CS monomer as a function of time.

### *β-Mercaptoethanol Modification*

Wild type CS at 2 mg/mL was stored overnight at 4°C in 20 mM ammonium bicarbonate containing 20 mM β-mercaptoethanol, conditions which mimic the original serendipitous experiment in this lab (Donald et al., 1998).

### *TFBA modification*

Wild type CS was dialysed overnight against 4 L of phosphate buffer (0.02 M $NaH_2PO_4$, 0.1 M KCl, 1 mM EDTA, pH 7.8) in SpectraPor 12 000-14 000 MWCO dialysis tubing. The dialysate was made up to 10 mL with phosphate buffer and heated at 37 °C for 5 minutes. 100 µL of 1 M TFBA in acetonitrile was then added, and the enzyme incubated at 37°C for 1 hour. The reaction was quenched with a few crystals of dithiothreitol (DTT), and the solution dialysed overnight against 4 L of citrate synthase buffer, or washed directly into the appropriate ammonium bicarbonate buffer for mass spectrometry using a Centricon unit (Amicon) with a 30 000 molecular weight cutoff (MWCO).

### *Dimethylsuberimidate Crosslinking of CS*

In a procedure adapted from Davies & Stark (1970), a stock of 12 mg/mL dimethylsuberimidate (DMS) in 0.4 M triethanolamine was prepared, and diluted such

that when combined with an equal volume of 3 mg/mL CS in standard buffer, it gave the desired ratio (by weight) of protein to DMS. Samples were prepared with CS/DMS ratios of 1:4, 1:3, 1:1, and 3:1, as well as a control containing only CS in an equal volume of 0.4 M triethanolamine. Reactions were then incubated at room temperature overnight, and run on SDS PAGE gels the following day.

## *Fluorescence Studies of NADH Binding*

NADH fluorescence binding studies were carried out using a method adapted from Duckworth and Tong (1976). The experiments were performed on a Gilford Fluoro IV spectrofluorimeter, using an excitation wavelength of 340 nm and an emission wavelength of 430 nm. The NADH stock used was made fresh the day of the experiment, and its concentration determined spectrophotometrically at 340 nm using an extinction coefficient of 6220 $M^{-1}cm^{-1}$.

A typical experiment was in three parts. The first determined E, the enhancement factor for bound NADH, and involved measuring the fluorescence of a small known quantity of NADH, and then titrating this with just enough CS to bind most of the NADH. The second required value was $F_f$ , the specific fluorescence of 1 μM free NADH. The values of E and $F_f$ were then used to calculate $F_b$, the specific fluorescence of bound NADH. Finally, the fluorescence of a small, known amount of CS titrated with increasing amounts of NADH was determined, and all values and concentrations corrected for volume changes during the titration. These measured fluorescence values (F) would arise from the fluorescence of both bound and free NADH such that:

$$F = [NADH]_{free} * F_f + [NADH]_{bound} * F_b$$

Given that: $$[NADH]_{bound} = [NADH]_{total} - [NADH]_{free}$$

These two equations can then be combined and rearranged to solve for $[NADH]_{free}$:

$$[NADH]_{free} = ([NADH]_{total} * F_b - F) / (F_b - F_f)$$

Thus $[NADH]_{free}$ was calculated in this way, and $[NADH]_{bound}$ was determined for every addition of NADH. The ratio of $[NADH]_{bound}$ / [CS] gave Y, the number of NADH bound per CS molecule. The Scatchard plot of Y/ $[NADH]_{free}$ vs. Y gave a reasonably straight line whose slope is $-1/K_D$ for the CS-NADH complex, and whose x-intercept gave the number of NADH binding sites per CS subunit.

## *Proteolytic Digests of CS*

Complete tryptic digests of CS (and other proteins) were generated by combining the protein at a ratio of 100:1 with TPCK trypsin (Sigma) in 20 mM ammonium bicarbonate. The digest was then incubated at 37°C for at least one hour (although most samples were left overnight). Digests were terminated either by addition of glacial acetic acid to a final concentration of 1 to 5%, or simply by freezing. The resulting mixture of peptides was then analyzed by either MALDI or ESI TOFMS as described in the following section.

Mass spectrometry was also used to obtain timecourses of the progress of limited proteolysis of CS by trypsin and subtilisin BPN (Sigma). Initially, sets of timed digest samples were generated in the traditional manner, by removing aliquots from the digest mixture and stopping them with 1-5% acetic acid/50% methanol at selected intervals. However, a method, referred to here as "online" limited proteolysis, was developed in collaboration with Lynda Donald, and was found to greatly simplify the timecourse experiment. A small (<20 μL) CS sample was mixed with the desired amount of proteolytic enzyme and immediately loaded into the injection syringe of the ESI mass spectrometer. Thus, proteolysis occurred as the sample was continuously sprayed into the machine. An addition to the in house Tofma data acquisition software called "Store and Restart" allowed for spectra to be recorded and automatically saved at set time intervals. Thus, the progress of proteolysis could be followed minute by minute if needed.

## *Mass Spectrometry*

As explained in the Introduction, protein samples for mass spectrometry must be carefully prepared to eliminate traces of non-volatile buffer components, detergents, and salts which may form adducts with the analyte and degrade resolution, or mask the analyte signal altogether. To this end, protein samples destined for mass spectrometry, including CS, creatine kinase, and both trypsin and subtilisin used in certain limited proteolysis experiments, were washed into freshly made ammonium bicarbonate buffer (usually 20 mM) using either Centricon units with the appropriate MWCO, or repeated

rounds of waterbug dialysis (Orr et al, 1995) through Spectra/Por 12 000-14 000 MWCO dialysis tubing. MALDI MS is somewhat more tolerant of impurities than ESI MS, and various on-target methods exist for further "washing" of a sample (see below). More stringent procedures were followed for ESI MS samples, including storage of all stock buffers in plastic rather than glass bottles, to reduce sodium contamination. NADH used in limited proteolysis experiments was provided as a disodium salt, and so was dialysed three times against 1 L changes of 20 mM ammonium bicarbonate using a waterbug with 500 MWCO Spectra/Por cellulose ester dialysis membrane.

All samples were run on mass spectrometers constructed in the Time-of-Flight lab under Dr. K.G. Standing in the Department of Physics at the University of Manitoba. Data was acquired and analysed using Tofma software written in house.

MALDI TOFMS of complete tryptic digests was performed on two different instruments equipped with $N_2$ lasers emitting at 337 nm. TOF II (Tang et al., 1988) was operated in linear mode with an accelerating potential of either 20 or 25 kV. A detector delay (typically in the range of 19000 ns) was used to avoid saturating the detector with low mass ions arising from the matrix, and delayed extraction was used to improve resolution for some samples. Samples run on TOF II were prepared on a non-porous polyurethane membrane (Deerfield PT63105, 0.002 inch thickness) using the method of McComb et al. (1997). Briefly, a small amount (<2 μL) of sample was applied to the membrane and allowed to dry. Two μL of methanol, which is believed to encourage partitioning of sample to the membrane, and salts to the solution phase, was then placed on the sample spot, and also allowed to dry. In many cases the spot was washed with several 20 μL aliquots of nanopure water to remove excess buffer salts. Matrix, typically

sinapinic acid in 2:1 water/acetonitrile, was then applied to the spot. The membrane was glued onto an aluminum disk using a thin layer of 3M Spraymount, trimmed, and placed into a special MALDI probe. Spectra were first calibrated using an external bovine insulin (5733 Da) standard, and then known tryptic peaks were used to calibrate internally.

TOF IV was a modified version of a high resolution hybrid quadrupole (Qq) reflecting TOF mass spectrometer described by Shevchenko et al. (1997). Sample analysis on this instrument was performed with assistance from A.V. Loboda and M. Bromirski. The accelerating voltage was typically 10 kV. Sample volumes of up to 2 μL were mixed with a matrix solution, applied evenly to the entire surface of the target and allowed to air dry. Two matrices were commonly used in TOF IV experiments – a combination of salicylic acid and 2,5-dihydroxybenzoic acid in 1:1 acetonitrile/water, 1 % trifluoroacetic acid (sDHB), and DHB in acetone. Spectra were calibrated externally using a mixture of substance P, mellitin, and enkephalin.

ESI and nanospray ionization TOFMS experiments were performed on TOF III (Verenchikov et al., 1994, Krutchinsky et al., 1995) by A. Krutchinsky and L. Donald. Calibration was external to the sample in all cases, using the +1, +2, and +3 charge states of substance P in 1:1 water/methanol with 1% acetic acid.

The Tofma software was used to generate peaklists, determine peak heights and areas under peaks, and deconvolute the charge envelopes found in ESI MS spectra. Peaks arising from complete and partial proteolytic digests were assigned with the aid of MacBioSpec v. 1.0.1 (PE Sciex, 1991) and BioMultiView PPC, version 1.4b2 (PE Sciex, 1998).

*Capillary Electrophoresis/Mass Spectrometry (CE/MS)*

Online capillary zone electrophoresis (CZE) of complete tryptic digests of photolabeled and unmodified CS coupled to ESI TOFMS was performed by M. McComb as previously described (McComb et al., 1999). Briefly, an interface was created to allow for initial separation of tryptic peptides by CZE, and then analysis of the capillary eluent by the TOF III mass spectrometer as described above. The solvent used for CZE was 1-5% acetic acid. Additions to the Tofma software recorded and stored both the CE total ion electropherogram and the mass spectral data in a more efficient, compressed format (Ens et al., 1999).

# *Results*

# *Allosteric Site Mutants of CS*

### *Preparation of P205A and Y186A*

Table 3 summarizes the purification characteristics of P205A and Y186A in comparison to those of a typical wild type CS preparation. These results are in general agreement with previous preparations of these mutants (Ye and Hacking, unpublished work), except that P205A eluted from the DE-52 salt gradient much sooner than expected. This was also true in comparison to the wild type CS results. The mutant proteins eluted somewhat later from the G 200 column. The specific activities of the mutant proteins were found to be lower than wild type CS.

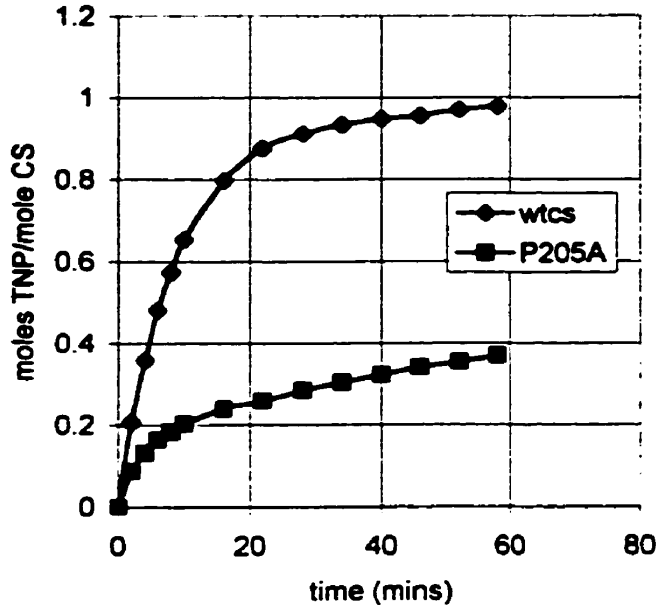
High resolution ESI/TOF MS under denaturing conditions (1:1 methanol/water with acetic acid) was used to confirm that the correct mutant had been purified. The deconvoluted spectrum of P205A has a peak corresponding to a mass of 47 856.4 Da, which is a change of –27.3 Da from the mass of the wild type enzyme measured in the same manner (data not shown). The theoretical mass difference resulting from the P to A mutation is –26 Da. Unfortunately, the mass spectrum acquired for Y186A CS was of poor quality, and a mass comparison with WTCS was not possible.

### *Modification of P205A and Y186A with DTNB and DPDS*

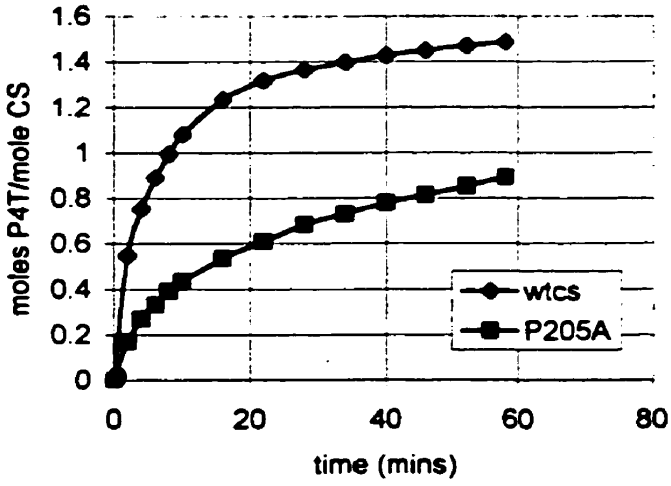
Figures 19 and 20 show the curves resulting from modification of the mutant protein and wild type CS with DTNB and DPDS. Following a rapid initial reaction with the enzyme, the amount of sulfhydryl modification by DTNB and DPDS levels off over the course of an hour, although an extended DTNB reaction with both P205A and wild type

Table 3: Comparison of the purification of wild type, P205A, and Y186A CS

| | DE-52 fraction of highest specific activity | G 200 fraction of highest specific activity | Specific activity of final CS solution |
|---|---|---|---|
| P205A | 25 (17 U/mg) | 61 (28.7 U/mg) | 6.8 U/mg |
| Y186A | 36 (9.1 U/mg) | 61 (15 U/mg) | 16.3 U/mg |
| Wild type CS | 40 (66.4 U/mg) | 50 (86.1 U/mg) | 35-40 U/mg |

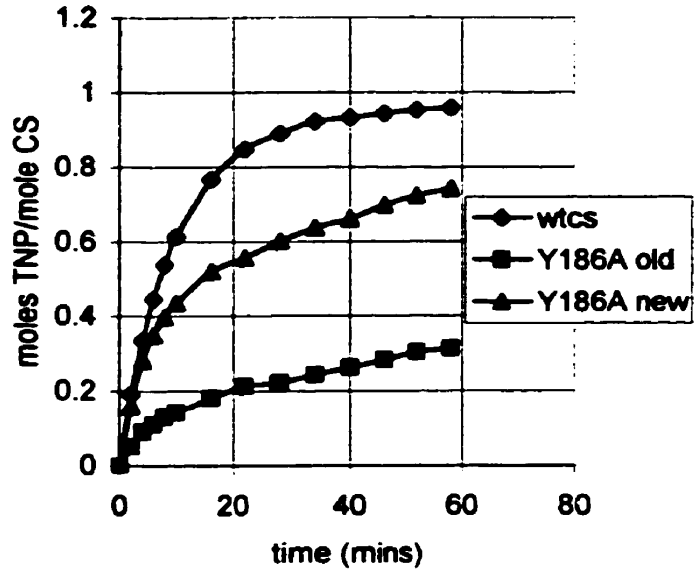



a)

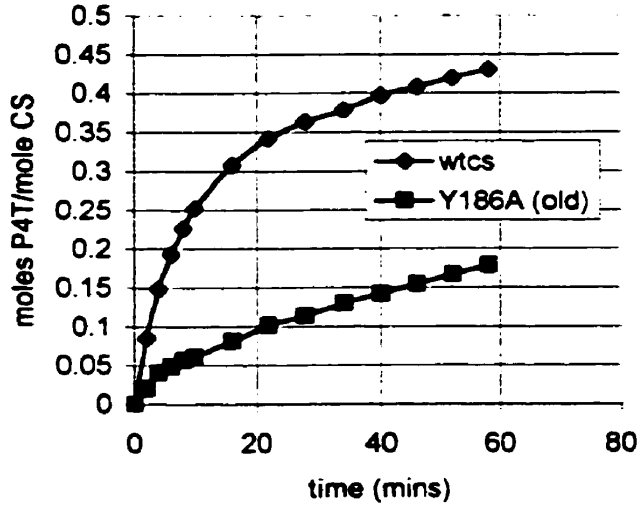



b)

Figure 19: Modification of wild type and P205A CS subunits over the course of one hour with DTNB (a) and DPDS (b).




a)




b)

Figure 20: Modification of wild type and Y186A CS subunits over the course of one hour with DTNB (a) and DPDS (b). The sample denoted as "Y186A old" refers to enzyme prepared by a previous student. "Y186A new" was prepared freshly for these experiments.

CS found that even after 17 hours, levels of TNP (and, one assumes protein modification) were still very slightly increasing (data not shown).

Both mutant enzymes showed decreased binding of the thiol reagents relative to that exhibited by wild type CS. In all cases, DTNB modified the wild type CS controls to the extent of approximately 1 equivalent per subunit as expected (Talgoy et al., 1979), but after one hour, P205A had bound only 0.37 equivalents per subunit (Figure 19). DTNB modification of an older sample of Y186A indicated results similar to those obtained from P205A, with the maximum ratio of bound reagent to enzyme being about 0.3 (Figure 20). However, modification of the Y186A protein freshly prepared for this study shows that ratio to be 0.74, which is considerably higher. The reason for this discrepancy might be the age of the first preparation, as protein degradation through denaturation, limited proteolysis, and oxidation for example, can occur in stored protein solutions over time. In light of this, the DTNB modification results obtained with the newer batch of Y186A seem more acceptable.

DPDS did not react as well with CS as DTNB, possibly a consequence of the poor solubility of DPDS in aqueous buffer solutions. The wild type CS control itself was only partially modified (1.49 and 0.43 equivalents) in both trials, as compared to the 2 modifications per subunit expected (Talgoy et al., 1979). Still, both mutants were modified by DPDS to a lesser extent than the controls (0.89 for P205A and 0.17 for the older batch of Y186A).

## *Mass Spectrometry of CS*

In addition to the dimer and hexamer charge envelopes, the ESI mass spectrum of intact, native CS at a concentration of 10 μM in 5 mM ammonium bicarbonate and acquired on TOF III usually contained several other features (Figure 21). Small amounts of tetramer, thought to arise from non-specific aggregation of dimers, have been previously found in CS spectra of older preparations (Ayed, 1998). The unidentified contaminant of approximately 12 220 Da is believed to be a fragment of CS itself, possibly from limited proteolysis of the enzyme in the ammonium bicarbonate buffer, which lacks EDTA, a buffer component shown to prevent proteolysis of CS (Alex Bell, unpublished work). However, even CS samples prepared for mass spectrometry on the same day as spectral acquisition contained this fragment, and repeated passage through Centricon units with 30 000 MWCO filters failed to remove it.

Sodium (+23) and oxygen (+16) adducts were occasionally seen in spectra of intact CS samples, but the majority contained a persistent adduct of approximately +324 Da. Nicknamed the "Centricon adduct", it is now thought to arise from the plastics of the Centricon units (e.g. phthalates), although in the past it has been attributed to the mass spectrometer itself (Ayed, 1998).

Methanol is often viewed as a protein denaturant, but CS samples prepared in 25 to 50% methanol showed marked improvements in their signal to noise ratio, without any apparent ion mass changes (Figure 22). This effect has also been observed in TOF III spectra of the Lac repressor protein (Donald et al., 1998b).

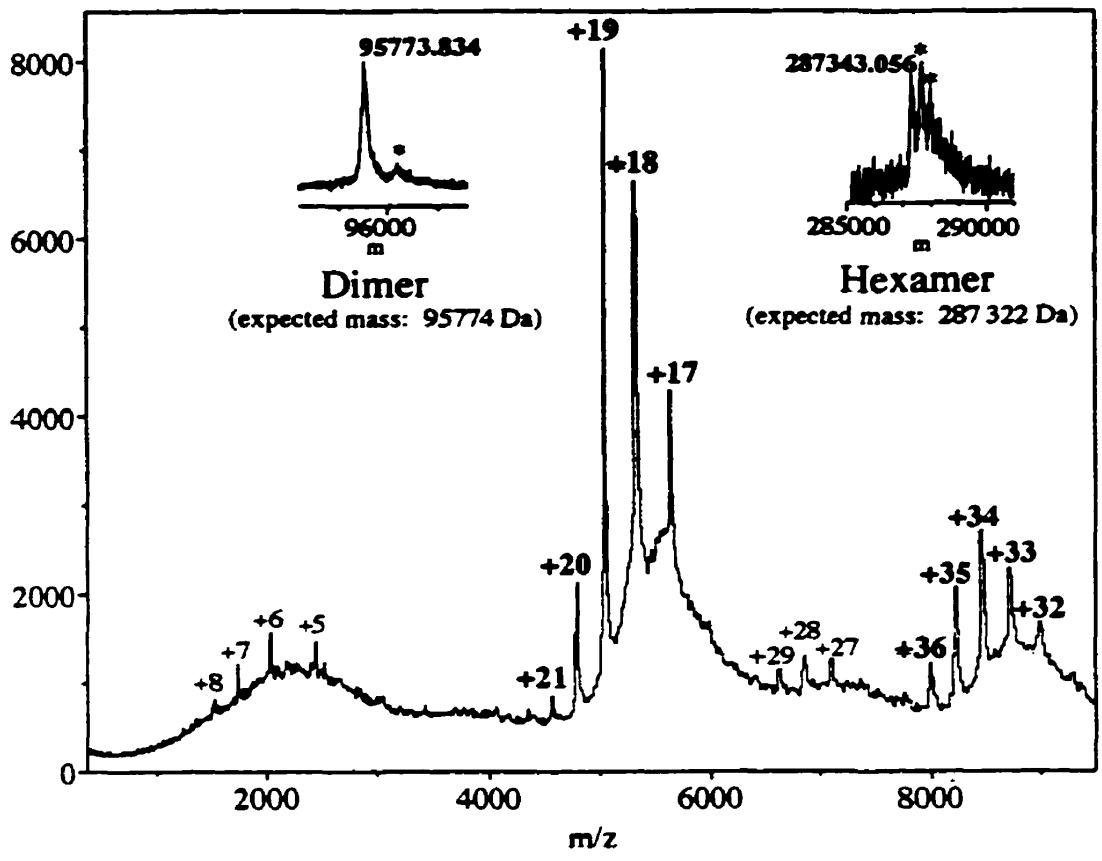



Figure 21: A typical ESI TOF mass spectrum of 10 μM wild type CS in 5 mM ammonium bicarbonate buffer. Deconvolutions of the dimer (+17 to +21) and hexamer (+32 to +36) charge envelopes are shown with the molecular weights indicated. The +324 Da adducts typical of CS prepared with Centricon filtration units are indicated by asterisks. The other two charge envelopes indicated correspond to tetramer (+27 to +29) and an unidentified possible proteolytic CS fragment of 12 211 Da (+5 to +8).

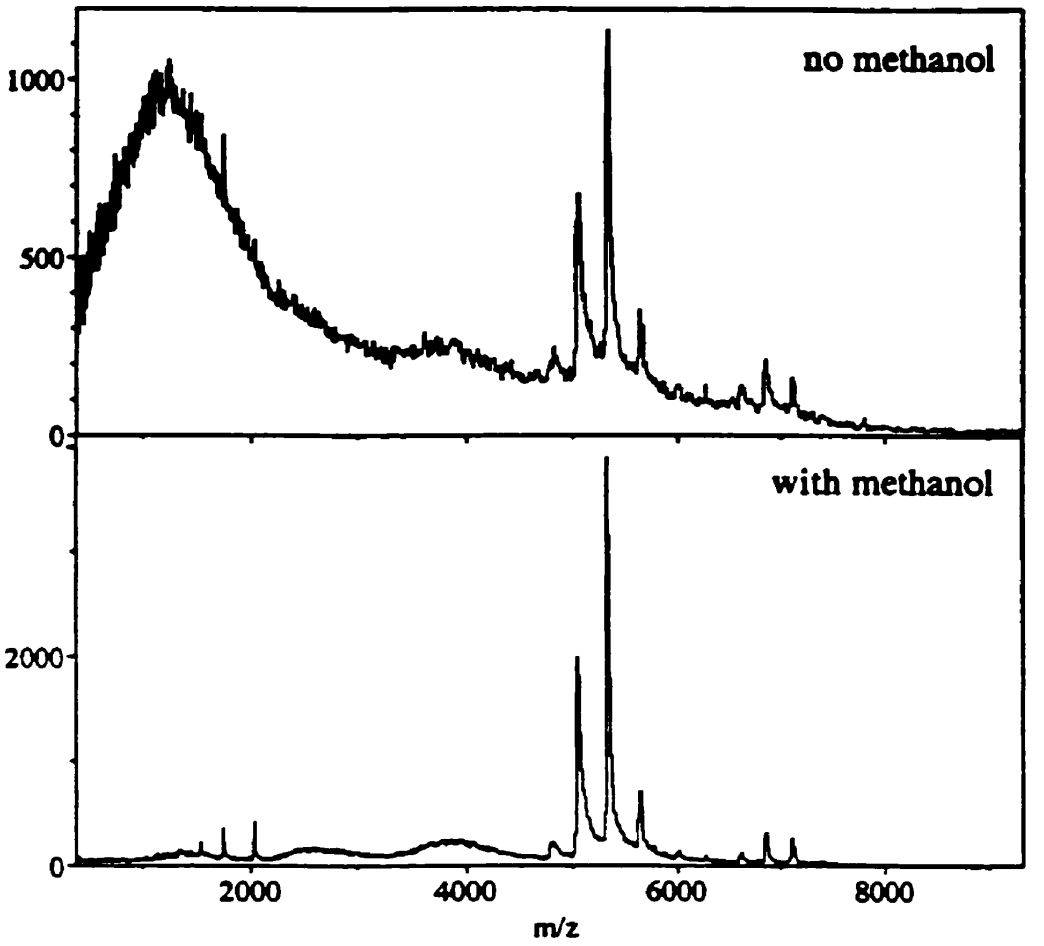



Figure 22: Effect of 30% methanol on the ESI TOF mass spectrum of 10 μM CS in 5 mM ammonium bicarbonate buffer, acquired under identical instrumental conditions.

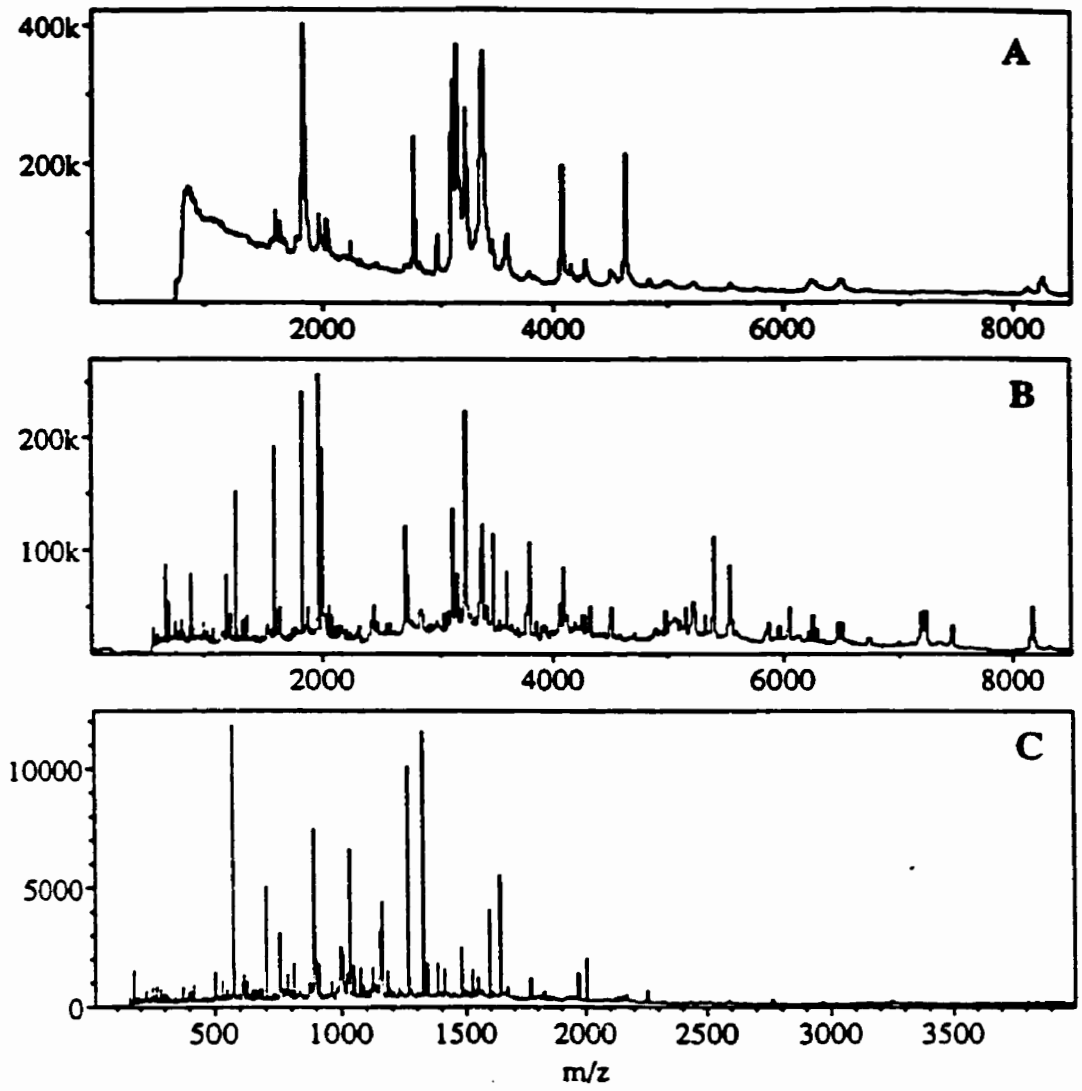



Figure 23: Mass spectra of complete tryptic digests of wild type CS acquired on the following instruments:

A: TOF II (MALDI TOF MS in linear mode with axial sample ion injection)

B: TOF IV (MALDI TOF MS in reflecting TOF mode with orthogonal sample ion injection)

C: TOF III (ESI TOF MS in reflecting TOF mode with orthogonal sample ion injection)

Three different mass spectrometers were used to study tryptic digests of CS, and Figure 23 compares the resulting spectra from TOF II, TOF IV, and TOF III. The m/z ranges of the spectra are different because as explained in the Introduction, MALDI mass spectra tend to contain primarily singly charged ions, while charges states as high as +4 were observed in ESI spectra of complete CS tryptic digests. In all three spectra, the signal is truncated below a certain m/z as a result of the detector delay imposed to prevent strong signals from low molecular weight impurities, buffer, and (for TOF II and TOF IV) matrix ions from saturating the detector and masking signals from the peptide ions.

Another immediately noticeable feature of Figure 23 is the difference in resolution between the three spectra. Peaks from TOF II spectra supplied only average masses with a resolution of approximately 200 for a strong peak such as 2771.2 Da (T3-4, corresponding to residues 7-32). Monoisotopic masses could be obtained from TOF III and TOF IV spectra with resolutions of approximately 7900 and 6600 (respectively) for the first peak of the isotopic series from a strong signal such as m/z = 1265.6 Da (T10, corresponding to residues 110-119). Resolution was calculated in the standard manner by dividing the corresponding m/z of the peak by the peak width at half height. The higher resolution obtainable with both TOF III and TOF IV, as explained in the Introduction, is to be expected from instruments which feature reflecting TOF mass analyzers and orthogonal sample injection via a collisional cooling interface. TOF II, by contrast, has a simple linear TOF analyzer, into which ions are extracted directly from the target.

The extensive list of tryptic fragments of CS observed by mass spectrometry can be found in Appendix 2. Figure 24 compares the sequence coverage provided by peptides


Figure 24: A representation of the CS subunit sequence covered by tryptic fragments indentified in representative mass spectra of CS tryptic digests acquired using the TOF II (top), TOF IV (middle), and TOF III (bottom) instruments. Every tenth amino acid is underlined. Shaded amino acids are those which were found in the tryptic fragments identified in the appropriate mass spectrum. The darker the shading, the greater the number of times that a particular amino acid or sequence of amino acids was found in that digest spectrum.

identified in representative WTCS tryptic digests run on the three mass spectrometers. As can be seen, there is virtually 100% sequence coverage in each case, and the patterns of tryptic fragment occurrence, as indicated by degrees of shading, appear to be quite similar. A large proportion of the peptides arise from the C-terminal region, especially from residue 291 onward. The significance of peptide occurrence and the observed major sites of tryptic cleavage will be discussed more fully in the Results section dealing with Limited Proteolysis of CS.

The majority of peptide mapping in this thesis was performed on unfractionated CS digests, and the assignment of peaks by mass was complicated by several factors. There are several CS tryptic peptides whose masses are essentially the same (e.g. fragment T32 of mass 577.31, and fragment T9 of mass 577.33), so these peaks were assumed to carry a mixture of both ions. As mentioned above, TOF III spectra contained multiply charged peptide ions, whose charge states were assigned on the basis of m/z difference between adjacent peaks in the isotopic series. For example, if the charge on an ion was +2, the mass difference between the adjacent isotopic peaks would be 0.5 U. However, this was only possible for peaks with good signal to noise ratio, rendering identification of low intensity peaks difficult.

Unexpected changes to peptide masses (e.g. oxidation, fragmentation) are another common hazard of peptide mapping. The most interesting example of this was discovered when MS/MS sequencing on TOF IV of an assigned tryptic peptide of 1157.60 Da (T41-42, corresponding to residues 410-418) and a frequently recurring unassigned peak at 1140 Da showed them to have virtually the same sequence. Reduced yields of this peptide during Edman degradation had been previously noted (Bhayana,

1984), and was attributed to the formation of pyroglutamate from the N-terminal glutamine under acidic conditions (in this case likely arising from the acidic matrix).

# *Photoaffinity Labeling of CS*

## *Preliminary Experiments*

### *1. Control of Photoaffinity Labeling*

The initial photoaffinity labeling experiments were designed to demonstrate the specificity and levels of photolabel incorporation using $^{32}P$ as a method of quantitation. The most basic control experiment carried out was to establish that the incorporation of radioactivity into CS was the result of "controlled" photoaffinity labeling (triggered by the application of UV light to a mixture of protein and probe). CS in 20 mM Tris buffer was incubated with 100 μM of either 2- or 8-azidoATP, and either left in the dark, or exposed to UV light for 2 minutes. Autoradiography of the samples on SDS PAGE gels showed virtually no incorporation of probe in the two aliquots which had not been exposed to UV light, while the irradiated samples gave two radioactive bands each, with the 2-azidoATP lane showing the highest level of incorporation. These findings are at least partly confirmed by liquid scintillation counting (Figure 25), which showed around 33% photoincorporation of 2-azidoATP and 8% photoincorporation of 8-azidoATP in the irradiated samples (these percentages represent the ratio of photolabeled CS to total CS added to the reaction, where concentration is with respect to the CS monomer). More problematic was the finding of around 6% incorporation of both probes in the non-irradiated samples. This may have been the result of inaccurate excision of gel bands for scintillation counting. Overall though, these results illustrate firstly that 2-azidoATP interacts better with CS than the 8-azido analogue, and secondly that photolabeling is indeed largely initiated by exposure to UV light.

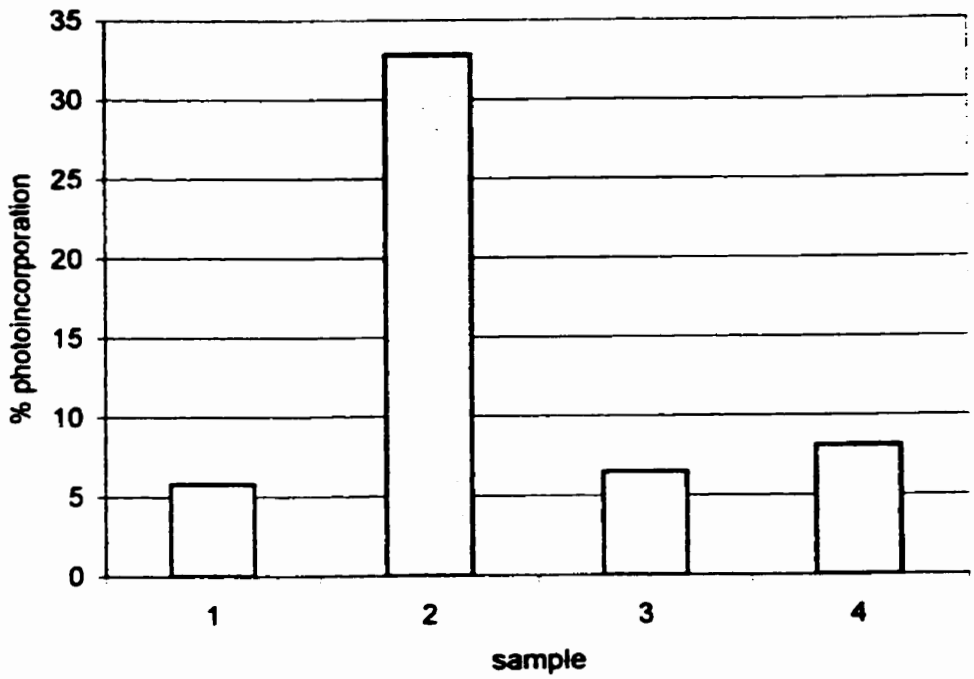



Figure 25: Photolabeling control experiment. Duplicate samples were prepared containing 8.3 μM CS in 20 mM Tris buffer, pH 7.5, and 100 μM of either 2-azidoATP (samples 1 and 2) or 8-azidoATP (samples 3 and 4). Samples 2 and 4 were UV irradiated for 2 minutes, while samples 1 and 3 were left unexposed.

### *2. Saturation and Inhibition Curves of Photolabel Incorporation*

Saturation of the target binding site, and inhibition of photolabel addition by the natural ligand are two methods widely used in photoaffinity labeling studies to demonstrate the specificity of the observed photoincorporation. As shown in Figure 26, both 2- and 8-azidoATP demonstrate saturation of photoincorporation, suggesting that their interaction with CS is specific for a particular binding site. Near saturation occurs at approximately 400 μM 2-azidoATP and 200 μM 8-azidoATP, giving $K_D$s of approximately 100 μM and 75 μM respectively ($K_D$ = dissociation constant for the label-protein complex, defined here as photolabel concentration at the estimated point of half-maximal saturation). These values are somewhat higher that the literature $K_D$ of 46 ± 5 μM for ATP itself (Talgoy & Duckworth, 1979), which is to be expected of a bulkier ligand analogue. Overall, as observed in the control experiment, photoincorporation is higher for 2-azidoATP.

If the labeling were indeed occuring at the allosteric site of CS, NADH would be able to significantly reduce the levels of photoincorporation observed. Both wild type CS and the deletion mutant Δ(264-287)CS were used in this experiment, as ATP is thought to be capable of binding to the active site of CS as well (Harford & Weitzman, 1975). A 24 amino acid deletion in the mutant enzyme effectively destroys the active site, theoretically leaving only the allosteric site available for binding. The mutant Δ(264-287)CS shows essentially the same NADH binding characteristics as wild type CS (Anderson & Duckworth, 1988). As shown in Figure 27, up to 70% of 2-azidoATP photolabeling, and up to 22% of 8-azidoATP photolabeling of Δ(264-287)CS is lost in

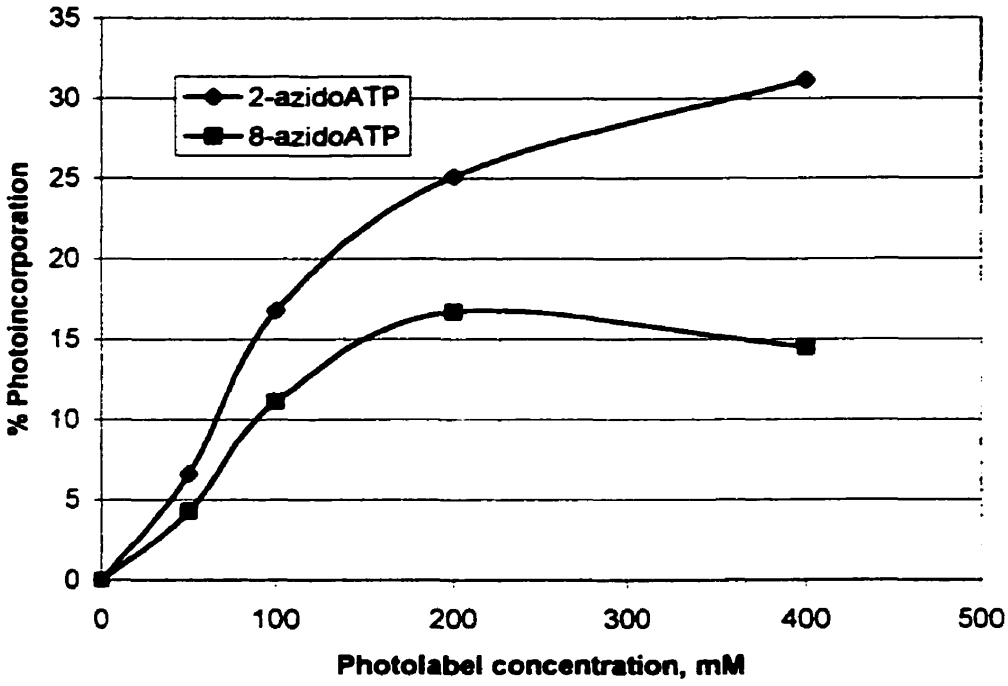



Figure 26: Saturation of wild type CS with increasing concentrations of the 2-azidoATP and 8-azidoATP photolabels.

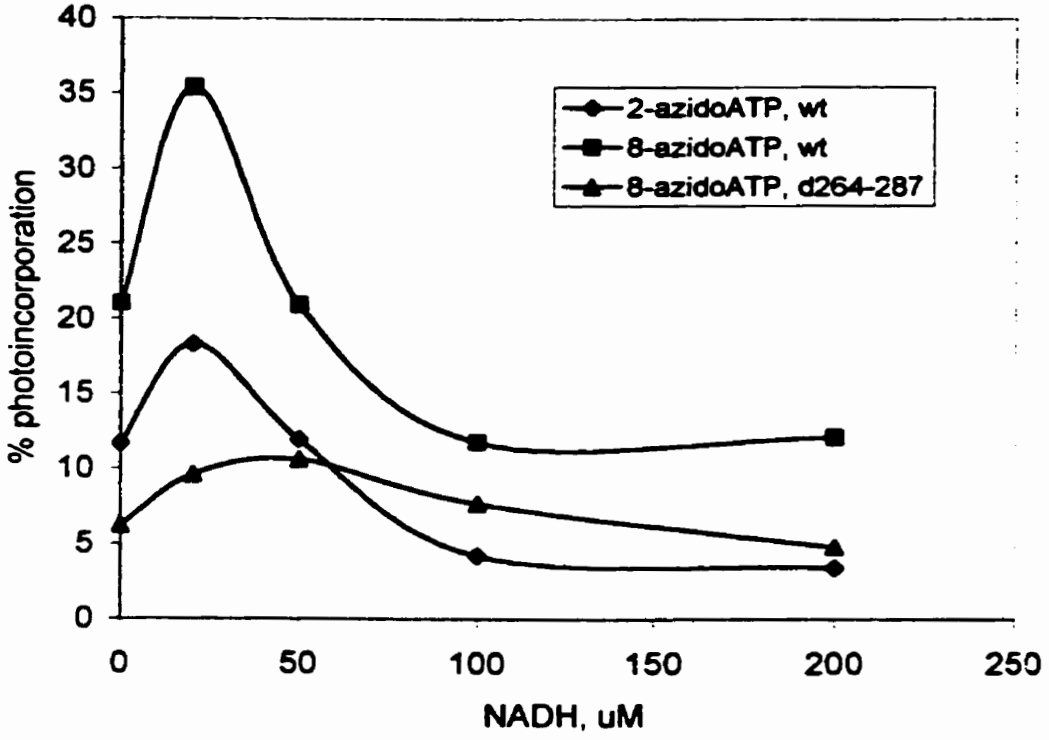



Figure 27: Inhibition of 2- and 8-azidoATP photolabel incorporation into wild type and Δ(264-287) CS by increasing concentrations of NADH. Photolabel concentration is 175 μM. No data are available for inhibition of photoincorporation of 2-azidoATP into Δ(264-287) CS.

the presence of NADH. When wild type CS is substituted for the mutant enzyme, the overall level of 8-azidoATP incorporation increases, indicating that this probe may also be interacting with the active site. However, up to 42% inhibition of photolabeling was observed in this instance.

An unexpected feature of all three inhibition curves is the 1.5-1.7 fold increase in photoincorporation at low (20-50 μM) concentrations of NADH. One possible explanation for this behaviour is that binding of a small amount of NADH to hexameric CS in a dimer-hexamer equilibrium mixture shifts this balance in the latter direction, opening up a large number of allosteric sites for photolabeling. A similar effect has been observed in the allosteric enzyme aspartate transcarbamylase (see Discussion).

### *3. Effect of Multiple Rounds of Photolabeling*

Multiple rounds of photolabeling are suggested as a method for increasing the yield of specific photoincorporation, as simply increasing the concentration of photolabel can lead to non-specific labeling (Bayley & Knowles, 1977; Czarnecki et al., 1979). The process involves successive cycles of photoprobe addition and sample irradiation. A step by step multiple photolabeling experiment was carried out, where incorporation of both 200 and 400 μM 2-azidoATP was determined after each of two-2 minute rounds of UV irradiation. As shown in Figure 28, doubling the concentration of photolabel used almost doubles the amount of incorporation from 12.7% to 22.3%, but two rounds of photolabeling using 200 μM label also increases the level of labeling by a factor of more than two, from 12.7% to 28.1%. Two rounds of labeling using 400 μM label increasing

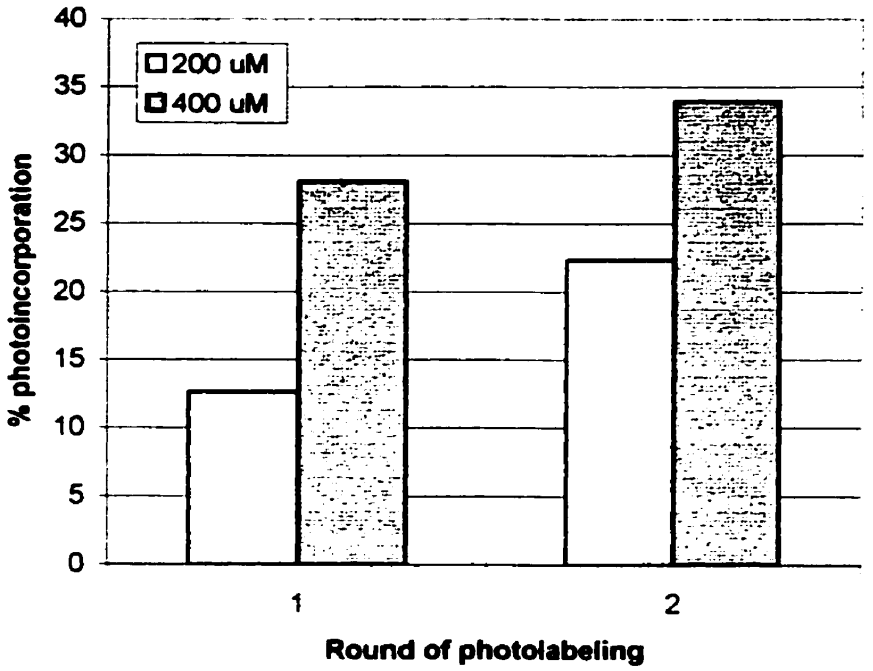



Figure 28: Photoincorporation of 200 μM and 400 μM 2-azidoATP into wild type CS after two rounds of label addition and UV irradiation.

incorporation from 22.3% to 34%. In a separate experiment, multiple photolabeling of CS by three applications of 200 μM of either azidoATP analogue, each followed by two minutes of UV irradiation succeeded in raising the percentage of photoincorporation of CS by 2-azidoATP to 63%, and by 8-azidoATP to 13.5%.

*4. Effect of UV Exposure Time*

Ideally, the UV exposure time for photolabeling experiments should be chosen to maximize label incorporation while minimizing the amount of damage done to the protein by UV light. Figure 29 shows that increasing UV irradiation time will increase the percentage of 2-azidoATP incorporation almost linearly, although to a lesser extent when the UV filter is present. All photoaffinity labeling experiments in this thesis were performed using the UV lamp with the filter removed, as older filters can become less transparent over time due to repeated exposure to high-intensity light of wavelengths less than 260 nm (Potter & Haley, 1983). This problem, called solarization, can be solved by heating the filter in a glassblower's annealing oven, but as the filter for the lamp used in these experiments was glued into its frame, this was out of the question.

The effects of UV exposure on CS were found to be quite rapid and severe. Virtually all CS activity was lost after 10 minutes exposure to the lamp with the filter removed, and even with the filter present, up to 70% of CS activity was lost in the same time period (Figure 30). It also appears that the buffer in which the protein was irradiated (20 mM Tris vs 20 mM ammonium bicarbonate) plays no role in this effect. After assaying for CS activity, the remainders of the irradiated protein samples were run on 8% SDS PAGE gels, which showed increasing levels of high molecular weight aggregates with increased

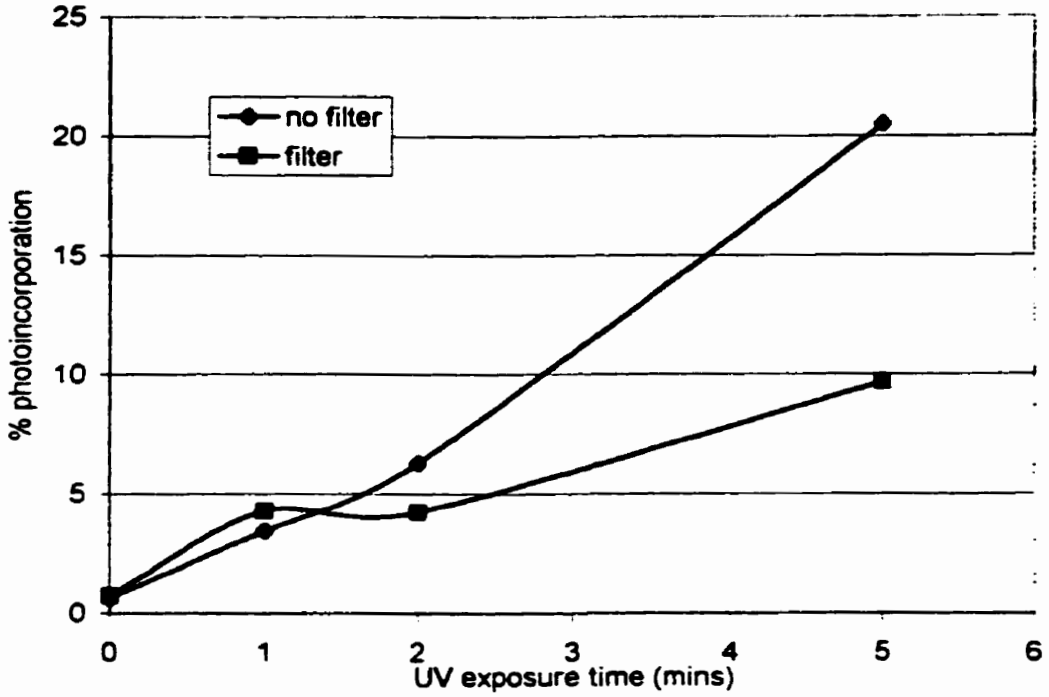



Figure 29: Effect of UV exposure time with and without the lamp filter, on the photoincorporation of 200 μM 2-azidoATP into wild type CS.

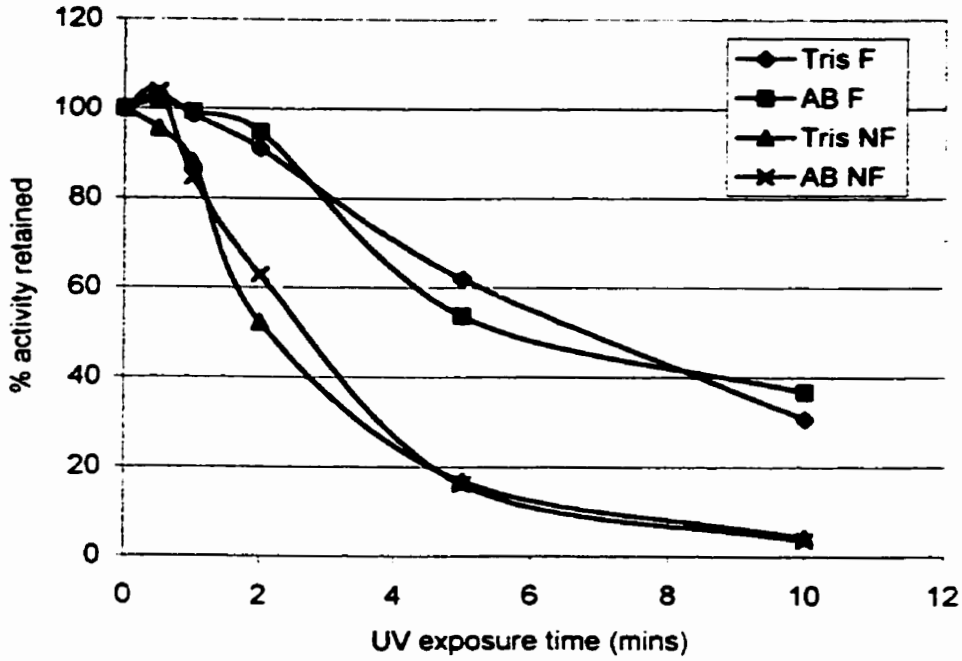



Figure 30: Loss of CS activity after exposure to UV light under the following conditions:
20 mM Tris buffer, lamp filter on (Tris F)
20 mM ammonium bicarbonate buffer, lamp filter on (AB F)
20 mM Tris buffer, lamp filter removed (Tris NF)
20 mM ammonium bicarbonate buffer, lamp filter removed (AB NF)

UV exposure time, although the majority of the protein still ran as expected. These aggregates ran in approximately the same size range as a series of faint photolabeled bands detected by autoradiography of the SDS PAGE gels from the saturation and inhibition of photolabeling experiments previously mentioned. Aggregates induced by dimethylsuberimidate cross-linking of CS were found to be smaller than those observed after UV irradiation, suggesting that the latter bands are the result of non-specific aggregation/crosslinking. ESI TOFMS spectra of both unexposed and UV irradiated CS under non-denaturing conditions (Figure 32a) and b)) are essentially identical.

The exact nature of the damage done to CS by UV light is unclear, but it appears to primarily affect the enzyme's active site. Several amino acids are known to be particularly susceptible to damage by photooxidation, including histidine, tryptophan, tyrosine, methionine, and cysteine (Danson & Weitzman, 1973), so it is possible that several residues either directly or indirectly involved in maintaining the functionality of the active site are damaged when samples are exposed to UV light during photolabeling. The main concern, however was the possibility of damage to the NADH binding site, especially after three rounds of 2 minute photolabeling. A fluorescence binding study of 0.14 mM CS in standard Tris-EDTA buffer exposed to UV light for 6 minutes under the same conditions used in a photolabeling experiment found the $K_D$ of NADH for the UV-exposed CS to be 2.26 μM as compared to 1.96 μM for unexposed CS. The number of molecules of NADH bound per subunit dropped from 0.87 to 0.76. The small difference between both of these sets of values would seem to indicate that the effect of UV light on the NADH binding site of CS is minimal at most. Other options for possibly less damaging light sources capable of producing the appropriate wavelength of light (e.g.

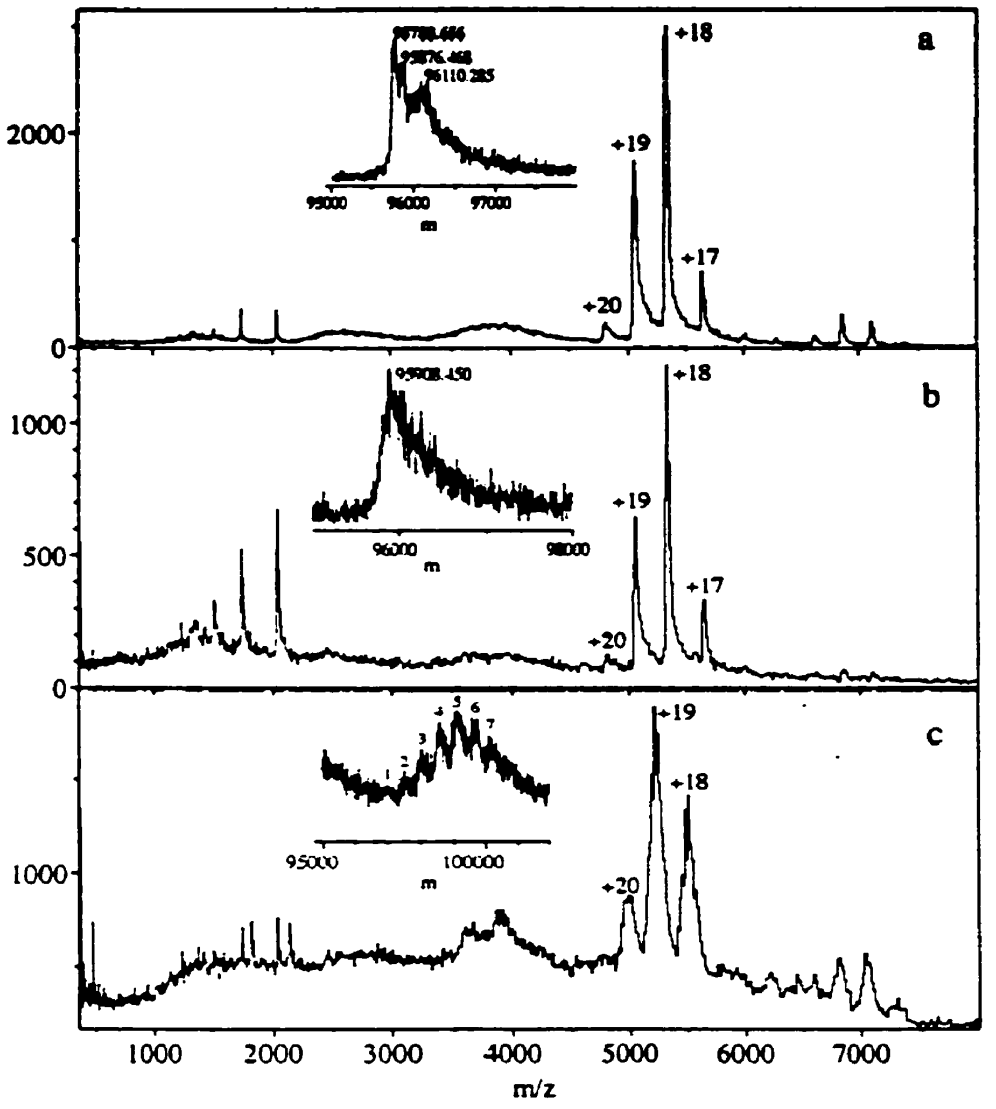



Figure 32: Comparison of the ESI TOF mass spectra of wild type CS (a), UV exposed CS (b), and CS photolabeled with 2-azidoATP (c). All samples were prepared in 20 mM ammonium bicarbonate buffer. Insets show the deconvolution of the dimer charge envelope. In (a) and (b), the main dimer peak is indicated in bold. Peaks 1 through 7 in (c) represent adducts (likely photolabel addition) to the CS dimer.

| Peak in (c) | Mass (Da) |
|---|---|
| 1 | 96 996 (dimer + 2 x 611 Da) |
| 2 | 97 551 (peak 1 + 555 Da) |
| 3 | 98 049 (peak 2 + 498 Da) |
| 4 | 98 583 (peak 3 + 534 Da) |
| 5 | 99 125 (peak 4 + 542 Da) |
| 6 | 99 651 (peak 5 + 526 Da) |
| 7 | 100 147 (peak 6 + 496 Da) |

halogen desk lamps with a specific glass filter, Dr. J. Charlton, personal communication) were explored. However, the maintenance of NADH binding characteristics in UV exposed CS, and the necessity of repeating the preliminary photolabeling experiments with a new light source led to the decision to continue to use the UV lamp and protocol already in place.

# *Mapping Photolabel Incorporation by Mass Spectrometry*

### *1. Identification of a Known Chemical Modification by MALDI TOF MS*

In order to test the use of MALDI TOFMS for mapping photolabel incorporation in unfractionated complete tryptic digests of CS, the alkylation of cysteine 206 by TFBA was successfully identified. A comparison of MALDI spectra from digests of both modified and unmodified CS showed the appearance of a peptide in the TFBA CS digest whose mass was approximately 110 Da greater than that of the tryptic fragment of residues 189-217. This mass shift is demonstrated in the spectral expansions of Figure 31, although both modified and unmodified peaks are present in the TFBA CS spectrum. Thus it appears that TFBA alkylation of C206 is not 100% complete.

The TFBA modified peptide was observed in spectra obtained on both TOF IV and TOF II using the membrane sample preparation method. The latter instrument however, is not able to give monoisotopic resolution of peaks (Figure 31 (II)).

### *2. Attempts to Locate Labelling of CS by 2-azidoATP Using MALDI TOF MS*

Having confirmed through the preliminary experiments that 2-azidoATP specifically labelled the NADH binding site of CS, attempts were made to identify the location of the modification in complete tryptic digests using the peptide mapping method tested on TFBA modified CS. However, unlike the TFBA modification, it was unknown where the photoaffinity label would form an attachment, or if there might be more than one modification per subunit. Thus the resulting spectra of unlabeled and labeled CS digests

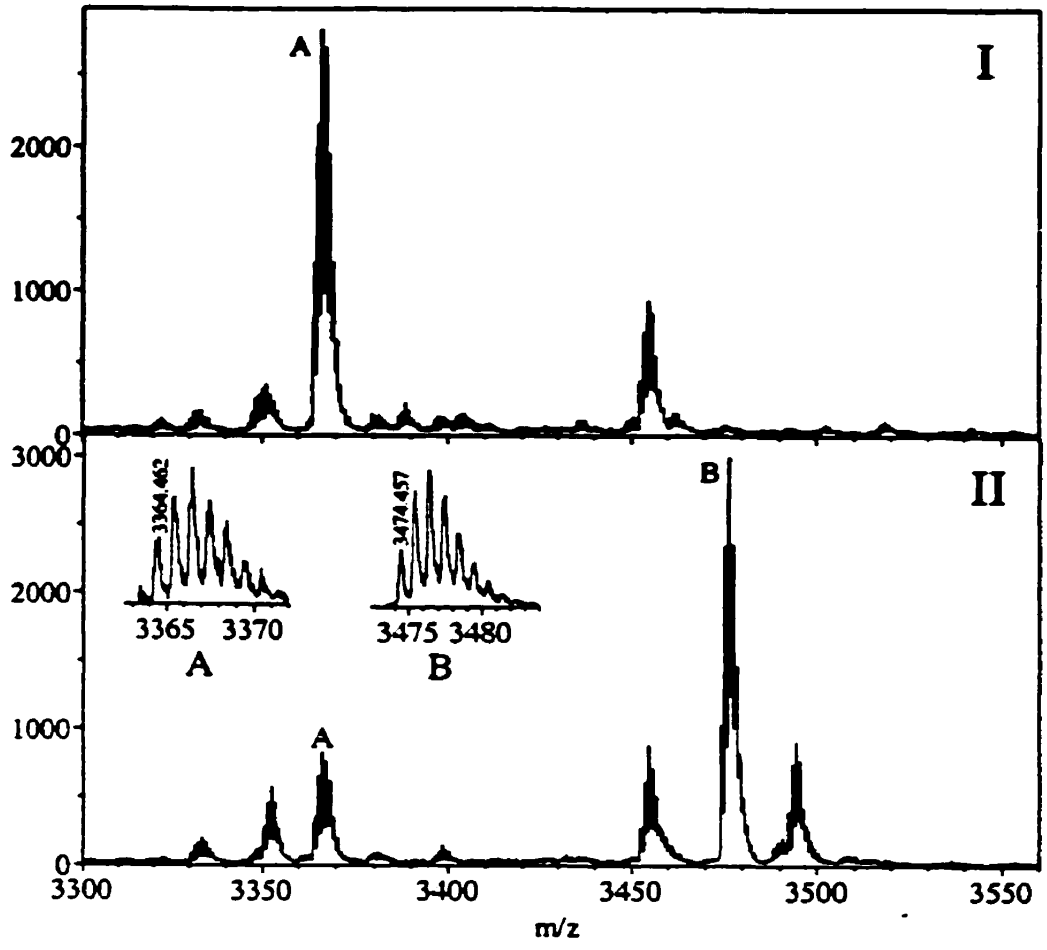



Figure 31: Expansion of the MALDI mass spectra of unfractionated tryptic digests of unmodified (I) and TFBA-alkylated (II) CS. Spectra were acquired on TOF IV using the sDHB matrix. Peak A has an m/z of 3364.462, corresponding to the unmodified tryptic fragment T19 of residues 189-217 (expected mass = 3364.53 Da). The m/z of peak B is 3474.456, corresponding to the mass of fragment T19 + 109.994 Da. The expected mass change associated with TFBA alkylation is 109.91 Da. The insets in (II) demonstrate the monoisotopic resolution typical of TOF IV mass spectra.

had to be analyzed in a thorough and systematic fashion. After the generation of a complete peak list for both spectra, those peaks not assigned to a tryptic CS peptide (based on m/z) were assumed to represent one of four possibilities: Impurities in the sample, fragments of tryptic peptides, tryptic peptides bearing other chemical modifications such as acetylations and oxidations, or photolabeled tryptic peptides. Assuming that the label formed a direct nitrene insertion bond with an amino acid, the mass shift produced by photolabeling with fully protonated 2-azidoATP would be +520 Da. After eliminating unassigned peaks which appeared in both unlabeled and labeled spectra, the remainder were tested to see if a subtraction of 520 Da from their molecular weight resulted in the mass of an assigned tryptic CS peptide. Finally, the relative intensities of the unlabeled and putatively labeled peaks were compared. Generally, peak intensities in mass spectra simply indicate the ease with which a particular analyte was ionized, but within a set of spectra derived under the same instrumental conditions, and from essentially the same sample material, it would be expected that the intensity of the unlabeled tryptic peptide would be lower in the photolabeled digest spectrum, while there would be no trace at all of the putative photolabeled peptide in the unlabeled digest spectrum (based on the preliminary photolabeling experiments, it was known that label incorporation would be less than 100%).

Despite the fact that superficially there often appeared to be differences between the unlabeled and photolabeled digest spectra obtained from both TOF II and TOF IV, once submitted to the above analysis, no satisfactory instances of modification by 2-azidoATP were observed.

### *3. Attempts to Locate Labelling of CS by 2-azidoATP Using ESI TOF MS*

Thinking that the inability to locate the photolabeling of CS by MALDI TOFMS might be linked to the method of ionization or sample preparation, it was decided to instead use ESI TOFMS (TOF III) to generate spectra of the tryptic CS peptides. Not only is ESI MS thought of as a "gentler" method of ionization compared to MALDI (as mentioned in the Introduction), but due to the greater mass range of the TOF III instrument, it could provide mass spectra of intact photolabeled CS.

Figure 32 shows the clear difference between the spectra of unlabeled and labeled CS, including a shift in the position of the dimer charge envelope. Deconvolution of this spectrum showed that the azidoATP modified CS dimer was 1222 Da larger than expected (which could at a stretch account for two photolabel incorporations, one per monomer as expected), with a series of 6 adducts. The masses of these adduct peaks are detailed in Figure 32. The mass differences between them range from 496 Da to 555 Da, and so they could be seen as representing CS dimer with up to 8 photolabel incorporations (four per monomer). It is then likely that some non-specific photolabeling had taken place, as one would expect at most four modifications per dimer (two for the allosteric sites and perhaps two for the active sites). As there would be an excess of unbound, reacted photolabel present in solution after the three rounds of labeling, it is also possible that at least a few of these 8 adducts are non-covalent, non-specific interactions of label with protein as observed, for example, in the titration of CS with NADH (Ayed et al., 1998).

To identify photoaffinity labeled peptides, citrate synthase at 4 mg/mL in 20 mM ammonium bicarbonate was photolabeled with three additions of 200 μM 2-azidoATP.

An unlabeled control sample of the same CS concentration was subjected to UV exposure alongside the labeled sample. After labeling was complete, both control and labeled samples were split into two aliquots, one of which was set aside, and the other of which was washed through a Microcon 30 centrifuge filter unit with 20 mM ammonium bicarbonate in an attempt to remove unbound label and any other impurities associated with the label from the samples prior to mass spectrometry. Both washed and unwashed aliquots were subjected to tryptic digestion at a protein/trypsin ratio of 100:1 overnight at 37°C. Digestion was stopped by freezing the samples, which were then analyzed on TOF III by L. Donald. The unwashed samples were run at 180 and 250 V declustering voltage, while the washed samples were run at 250 V. Thus a series of six spectra (three control, three labeled) were acquired, and systematically analyzed to identify tryptic peptide peaks which disappeared and peaks approximately 520 Da heavier than known tryptic peptides which appeared in the photolabeled spectra.

A total of 99 complete and partial tryptic peptides, covering the entire CS sequence, were identified in the control and labeled sample spectra. Of the many differences between the spectra of labeled and unlabeled CS tryptic digests, the most interesting involved singly charged peptide peaks at m/z 1146 and 1667, in the unlabeled and labeled samples, respectively. The peptide at m/z 1146 could be either T17 (expected mass = 1146.48 Da, corresponding to residues 168-177), or the partial tryptic fragment T15-16 (expected mass = 1147.68, corresponding to residues 158-167). Given that the tryptic digestion was performed overnight, and that peaks representing T15 and T16 alone were present in both the unlabeled and labeled spectra, it is likely that most of the material at m/z 1146 arose from T17. As Figure 33A shows, the peaks at m/z 1146 disappear

completely in the labeled digest spectrum. Among the peaks appearing only in the labeled digest spectra, one of the more prominent is a singly charged peak at m/z 1666.813 (Figure 33B). Although this peak overlaps with a doubly charged peak of approximately the same m/z, the singly charged fragment is seen only in the photolabeled sample. Its mass corresponds to the 2-azidoATP photolabeled derivative of T17 (calculated mass = 1666 Da). It thus appears that a residue of tryptic peptide T17 can be modified by the 2-azidoATP photolabel, although this must be confirmed by sequencing. Figure 34 shows that both T17 (red), and the other potentially photolabeled peptide T15-16 (green) are in the vicinity of C206 (blue), the marker for the NADH binding site. It also shows that parts of both of these peptides are surface accessible.

### *4. Photoaffinity Labeling of Creatine Kinase*

As a control experiment, the photoaffinity labeling of the B subunit of rabbit creatine kinase as performed by Olcott et al. (1994) was studied in order to determine if their results, which were obtained using chromatographic methods, could be reproduced using mass spectrometry. Deconvolution of the ESI TOF mass spectrum under denaturing conditions of intact creatine kinase photolabeled with either 2- or 8-azidoATP showed essentially one main peak of mass 42599 Da (expected mass = 42664) (data not shown). Trailing from this peak were a series of small peaks corresponding to adducts of 30-50 Da each (possibly sodium, potassium, or multiple oxidations), and two low intensity peaks corresponding to adducts of approximately 620 and 1340 Da which were not present in the control spectrum of unlabeled creatine kinase. As with CS, MALDI TOF MS of complete tryptic digests of both labeled and unlabeled creatine kinase on TOF IV

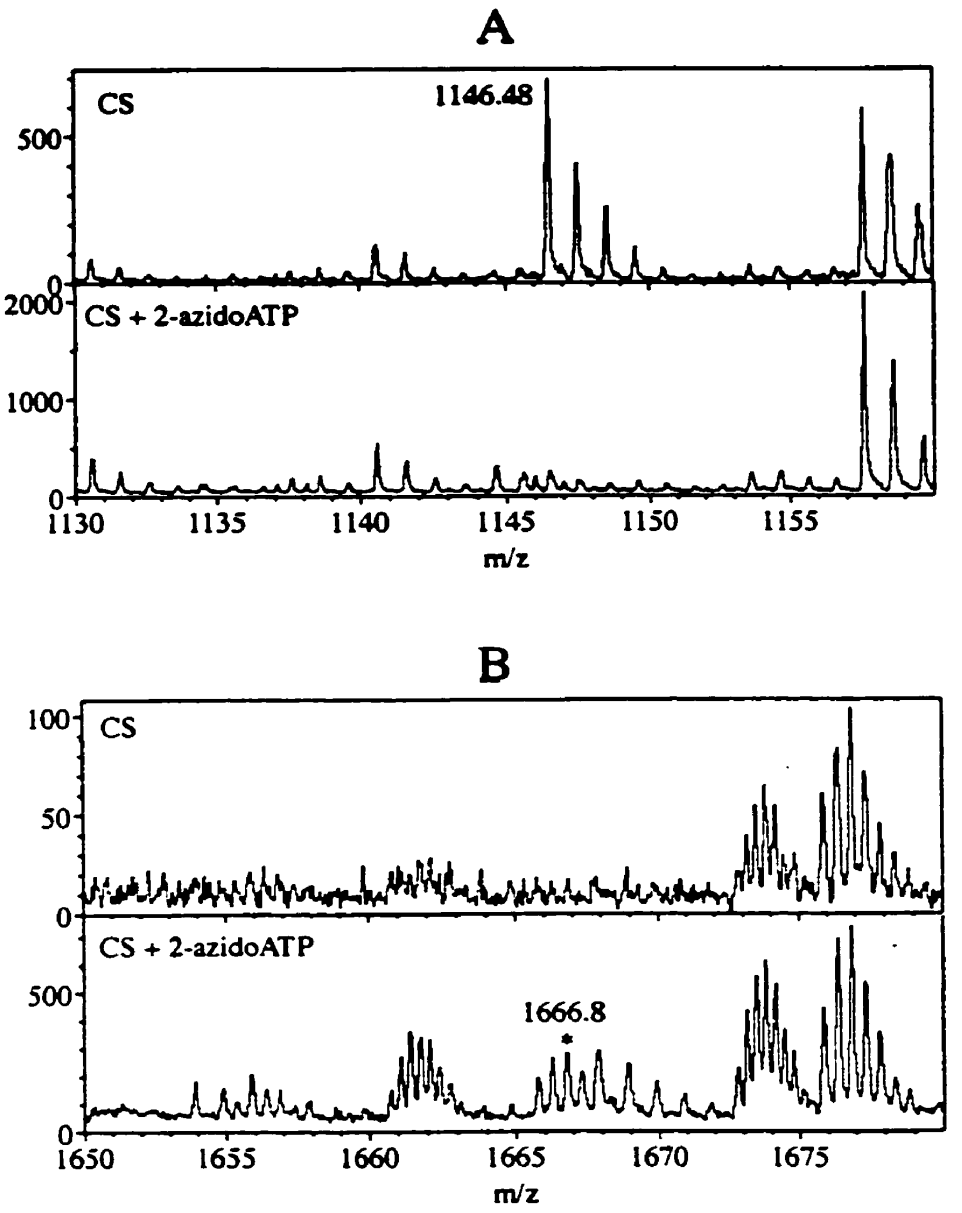



Figure 33: Putative photolabeling of a peptide of mass 1146 Da as detected by ESI TOFMS. The unlabeled peptide (either T17 or T15-16) loses virtually all its intensity in the photolabeled CS digest (panel A), whereas the putatively labeled peptide of mass 1666 Da is nonexistant in the control CS digest, and present in the photolabeled digest (panel B).

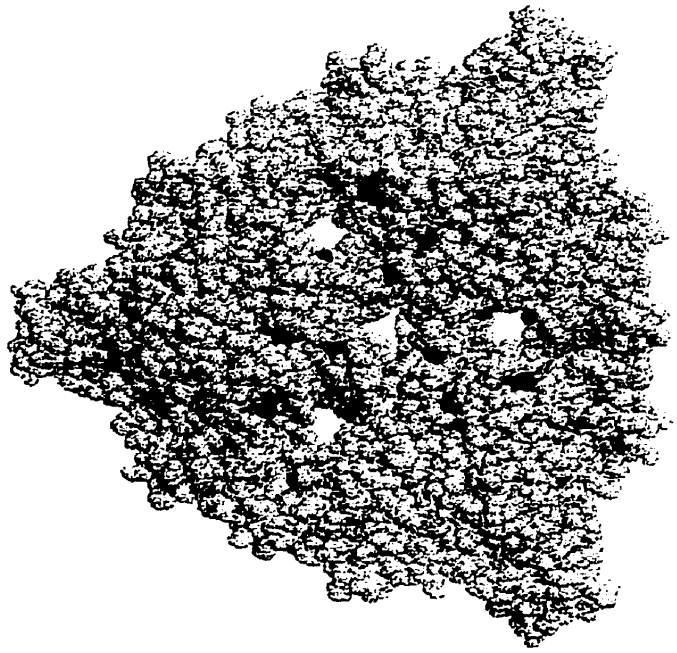


Figure 34: Space-filling model of the crystal structure of *E. coli* CS, showing the relative positions and surface accessibilities of:

- C206 (blue)
- Tryptic peptide T17 of mass 1146.48, corresponding to residues 168-177 (red)
- Tryptic peptide T15-16 of mass 1147.68, corresponding to residues 156-167 (green)

failed to show photolabeling of tryptic fragments, especially peptides 279-291 and 236-241 which constitute part of the creatine kinase B ATP binding domain, and had been previously shown to be labeled by 8-azidoATP (Olcott et al., 1994).

It took three attempts to get useful ESI TOF mass spectra of intact creatine kinase B, and even the best spectra had a number of small adducts. The mass of the main peak was 65 Da less than expected. Although the two larger adducts of approximately 620 and 1340 Da, which appeared only in the photolabeled spectra are in the right general mass region for 1 and 2 additions of azidoATP, no evidence of these additions was found in the tryptic digest spectra. This particular experiment was thus unsuccessful.

### *5. An Attempt to Locate Photoaffinity Labeling Site of CS Using CE/MS*

Difficulties in processing the huge amounts of data generated from ESI mass spectra of unfractionated digests pointed to the need for an additional method for determining peak identities. It was hoped that the use of a second separation technique (in this case, capillary zone electrophoresis, CZE) in concert with mass spectrometry would reduce the problems documented in the preceding sections. A method of interfacing CZE online to TOF III had already been used to produce a tryptic map of the entire CS sequence (McComb et al, 1999).

In CZE, analytes are separated by the electroosmotic flow generated when voltage is applied across a fused silica capillary filled with buffer. The basis of separation is the ratio of a particle's charge to its size (ionic volume). This technique is extremely sensitive, and only a few tens of nanoliters of sample are generally required.

Unfortunately, this sensitivity also applies to any impurities in samples or the running buffer.

The tryptic digest samples analyzed by CE/MS were prepared separately from those analyzed by ESI MS alone. A solution of 4-5 mg/mL CS in 20 mM ammonium bicarbonate was photolabeled by three additions of 200 μM 2-azidoATP. Again, a control sample was also concurrently UV exposed in the absence of label. Tryptic digestion was for one hour at 37°C, and terminated either by freezing, or by the addition of 5% acetic acid. Samples were then split into two aliquots. One was set aside and the other dried down in a SpeedVac to remove the volatile ammonium bicarbonate, and reconstituted in 1% acetic acid. Samples were run by Mark McComb essentially as described for previous unlabeled CS digest samples (McComb et al, 1999).

The CZE electropherograms and mass spectral data from TOF III were collected simultaneously during the sample runs using the Tofma software. Thus, from any selected peak in the total mass spectrum (TMS), the corresponding selected ion electropherogram (SIE) could be generated, and provide a retention time for that peak. Likewise, selection of any portion of the total ion electropherogram (TIE) would give the selected mass spectrum (SMS) corresponding to that particular elution volume. This ability to cross reference between CE and MS data allowed for the elimination of background and artifact peaks, and by comparison of retention times, made it possible to confirm MS peaks thought to represent different charge states of the same tryptic peptide.

A total of 200 "real" peaks were identified in the photolabeled digest CE/MS data. Of these, 97 were assignable to tryptic peptides of CS. Interestingly, the only portion of CS sequence not found corresponded to peptide T17 (residues 168-177, mass of 1146.48

Da), the same tryptic peptide whose presence was found to be greatly reduced in the ESI mass spectra of photoaffinity labeled digests. However, the peak of m/z 1667 representing peptide T17 + photolabel modification was not found in the CE/MS analysis of the photolabeled CS digest. While several unassigned peaks from the ESI TOFMS analysis of the photolabeled digest sample were also found to occur in the CE/MS results, none corresponded to labeling of tryptic peptides from the region of sequence surrounding C206.

## *Studying Other Chemical Modifications of CS by Mass Spectrometry*

As discussed in the Introduction, the sulfhydryl reagents DTNB and DPDS have previously been used to modify the cysteine residues of CS. Both have been found to modify C206, but DPDS also creates a mixed disulfide with one more, unknown cysteine. It has been observed in several cases that modification of surface exposed cysteines (formation of a mixed disulfide in a manner akin to DTNB or DPDS) can occur when a protein is incubated in a buffer containing β-mercaptoethanol, a common reducing agent (Donald et al., 1998; Packman & Berry, 1995). This serendipitous modification can then be detected by mass spectrometery, and is thus a tool for identifying the reactive, surface-exposed sulfhydryl groups of a protein. It was decided to attempt to map this particular modification of one or both of the reactive cysteines of CS in unfractionated tryptic digests using mass spectrometry, and compare the results to those obtained with DTNB and DPDS.

Conflicting results were obtained from the ESI mass spectrum of intact, β-mercaptoethanol-exposed CS (Figure 35). Two adducts of +77.6 Da and +156.8 Da were observed in the deconvolution of the monomer peak envelope (expected mass of the mixed-disulfide is +76 Da), indicating the possible presence of one and two β-mercaptoethanol modifications. Adducts of +72 +151.5 Da were observed in the dimer peak envelope, which could correspond to the modification of one cysteine on one or both subunits of the dimer. It may be that in this case, the second β-mercaptoethanol adduct per subunit might have been found in the baseline noise trailing from the dimer

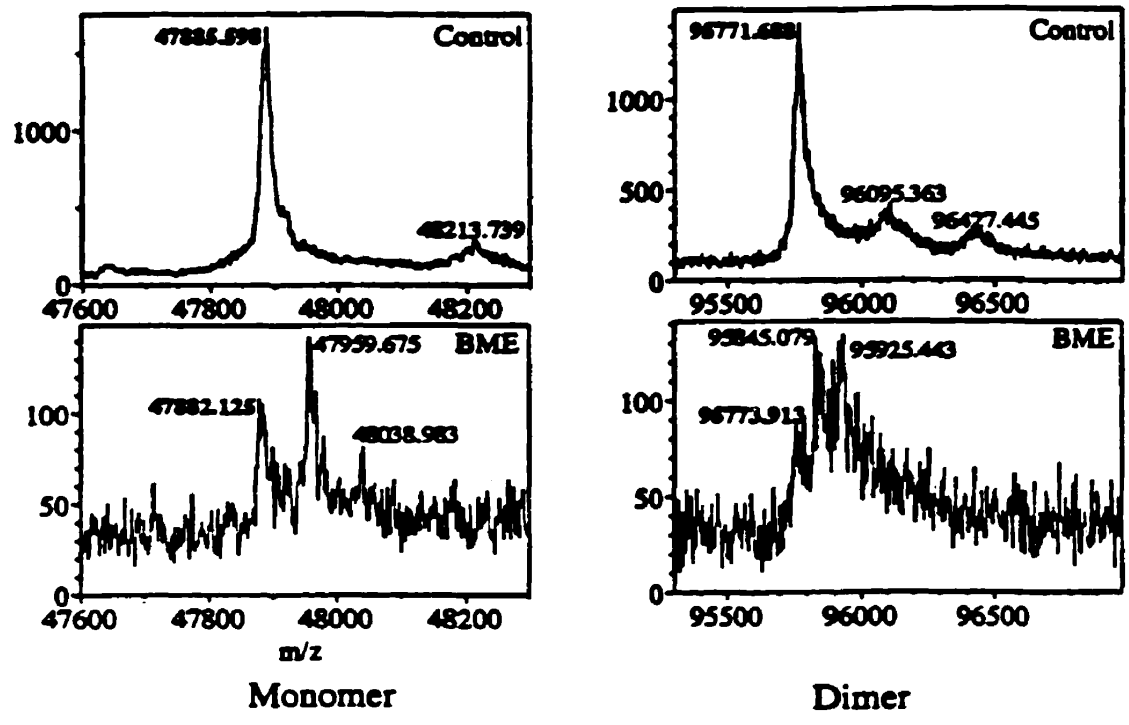



Figure 35: Deconvolutions of the monomer and dimer charge envelopes from ESI TOF mass spectra of unmodified CS ("Control"), and CS modified with β-mercaptoethanol ("BME"). The peak corresponding to unmodified monomer or dimer in each deconvolution is indicated in bold.

peak deconvolution. In any case, based on DTNB and DPDS experiments on CS (Talgoy et al., 1979), one might expect there to be at least one reactive surface cysteine (C206) capable of being modified by β-mercaptoethanol.

In the MALDI mass spectrum of a complete tryptic digest of β-mercaptoethanol-modified CS, a singly charged peak at m/z 3440.482 corresponded to the formation of one β-mercaptoethanol mixed disulfide (+76 Da) for tryptic peptide T19 of residues 189-217 (mass = 3364.53 Da), which contains C206. This result is shown in Figure 36. In addition, β-mercaptoethanol modification of a second, reactive cysteine of CS was also observed. A singly charged peak of m/z 4157.889 corresponds to the formation of one β-mercaptoethanol mixed disulfide for tryptic peptide T8 of residues 70-104 (mass = 4080.94 Da), which contains C86. This result is shown in Figure 37. It could be, then, that this is the second cysteine which typically undergoes DPDS (but not DTNB) modification. The position of C86 in the hexameric CS crystal structure is shown in Figure 38, which demonstrates the surface accessibility of this residue.

Attempts to map both DTNB and DPDS modifications of the reactive cysteines in unfractionated CS tryptic digests were unsuccessful. Despite confirmation of the modification of samples with either DTNB or DPDS by spectrophotometric means, the ESI mass spectra of these samples were generally of poor quality, and did not conclusively show evidence of mixed disulfide adducts. No trace of either DTNB or DPDS modification could be found in the MALDI mass spectra of tryptic digests of these samples. It could be that the unstable mixed disulfides were lost during sample preparation for mass spectrometry.

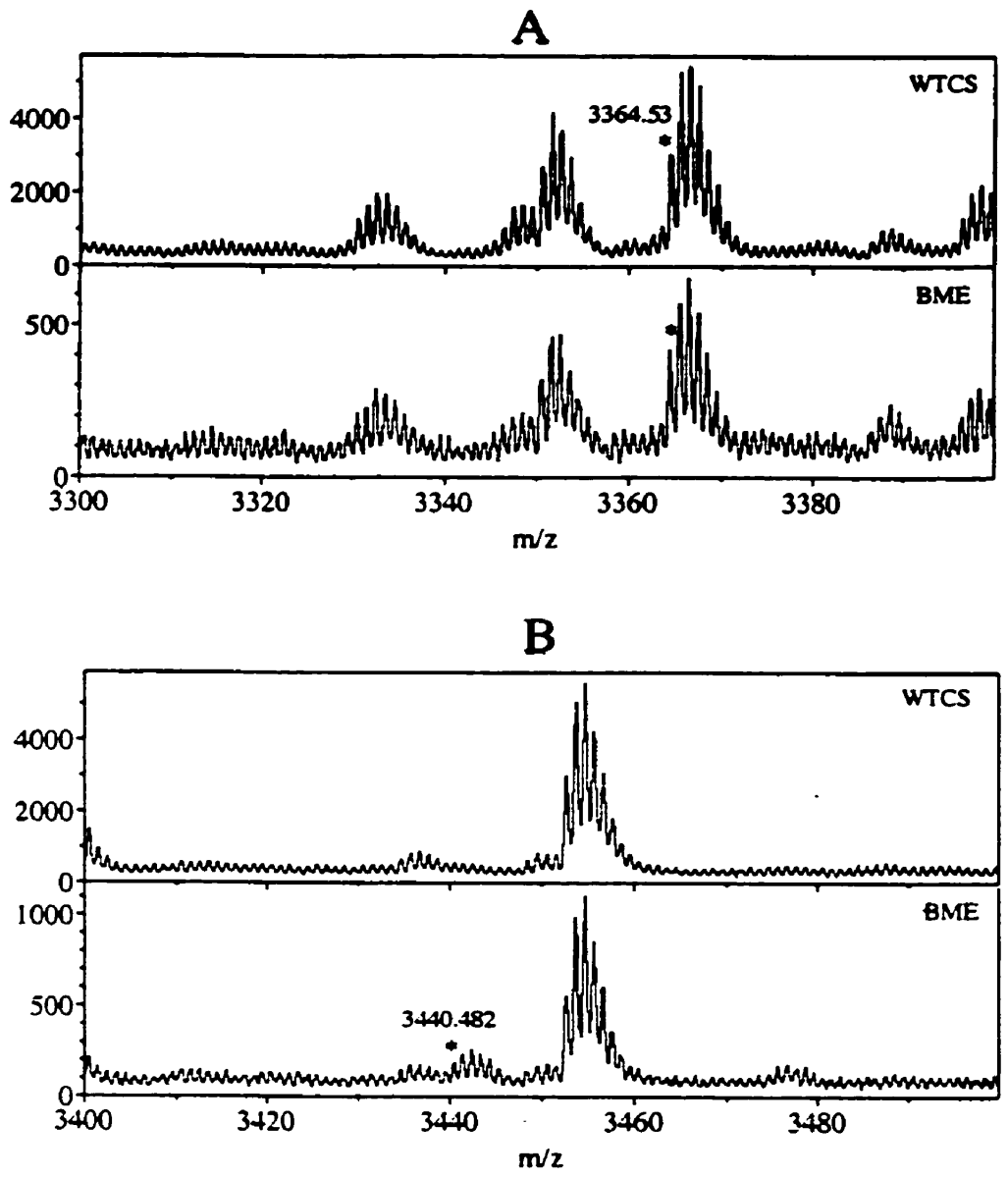



Figure 36: Putative β-mercaptoethanol modification of C206 as detected by MALDI TOFMS. In panel A, the unmodified tryptic peptide bearing C206, T19 (m/z 3364.53, indicated by asterisk) appears in both the control and β-mercaptoethanol labeled digests. In panel B, the putatively labeled peptide of m/z 3440.482 (indicated by asterisk) is nonexistant in the control CS digest.

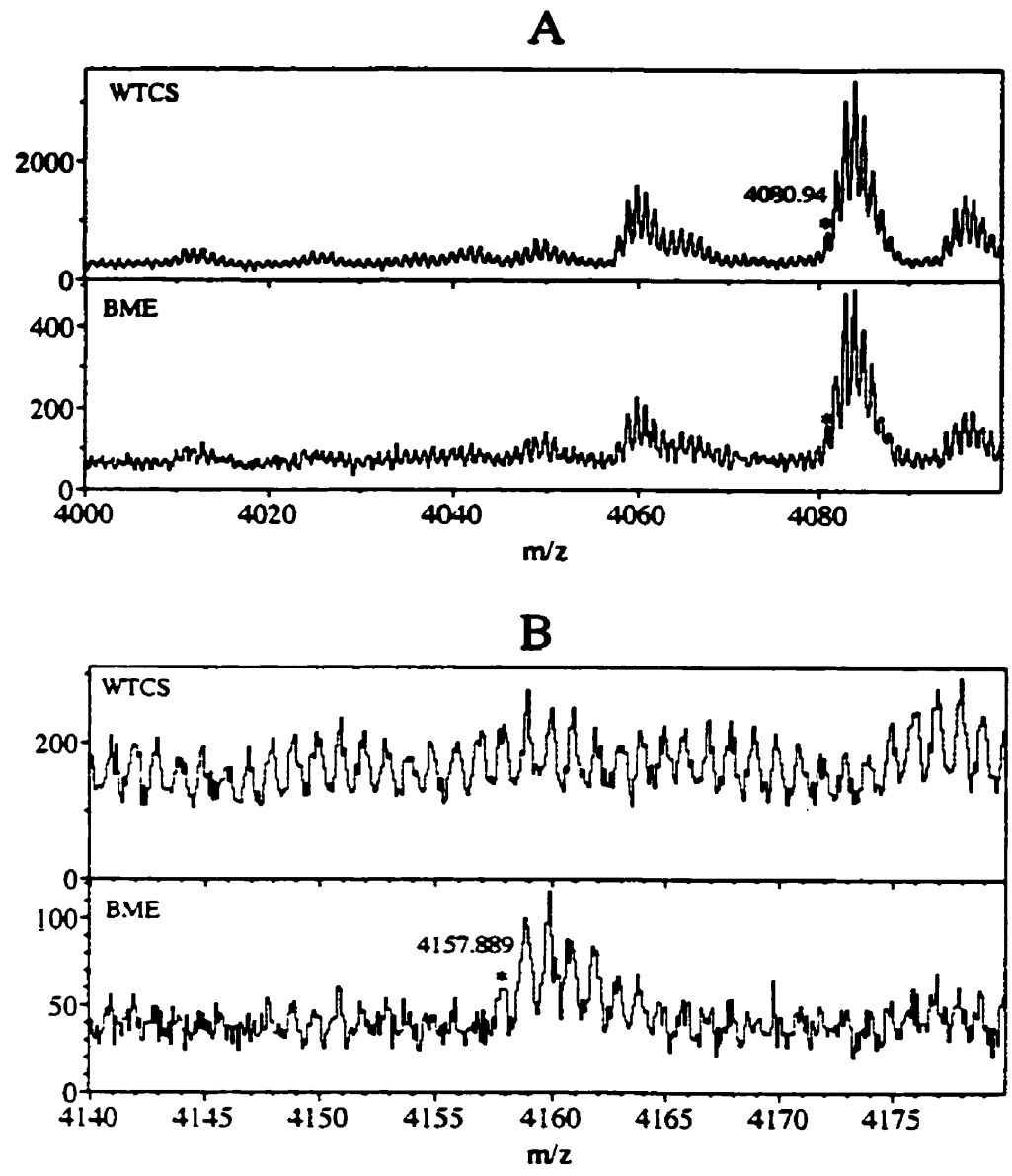



Figure 37: Putative β-mercaptoethanol modification of C86 as detected by MALDI TOFMS. In panel A, the unmodified tryptic peptide bearing C86, T8 (m/z 4080.94, indicated by asterisk) appears in both the control and β-mercaptoethanol labeled digests. In panel B, the putatively labeled peptide of m/z 4157.889 (indicated by asterisk) is not noticeable in the control CS digest.

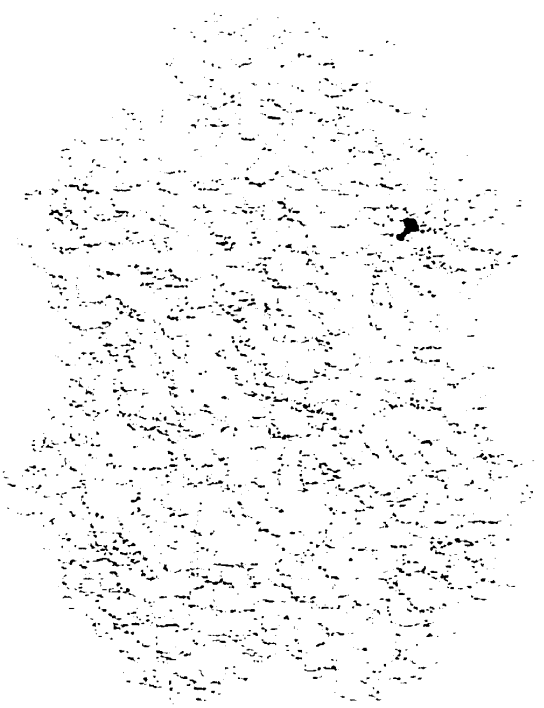


Figure 38: Space-filling model (side-on view) of the crystal structure of *E. coli* CS, showing the position and surface accessibility of C86 (blue).

## *Limited Proteolysis of CS*

The way in which a protein fragments under conditions of limited proteolysis can provide information as to the structural features most exposed and therefore vulnerable to attack by proteolytic enzymes. Changes to the rate and pattern of fragmentation can be used to help reconstruct the conformation of a protein in different solution conditions or in the presence of ligands.

It was hoped that development of rapid mass spectrometry methods for analyzing the limited proteolysis of CS would aid in characterizing conformational changes undergone by the enzyme and its folding intermediate (Ayed & Duckworth, 1999).

### *Individual Timed Digestions*

#### *1. Intact CS*

Citrate synthase was digested at room temperature in 20 mM ammonium bicarbonate by trypsin at a ratio of 1/160 (w/w). As the digestion proceeded, aliquots were withdrawn at 5, 10, 20, and 60 minutes and stopped by the addition of acetic acid to a final concentration of 2.5%.

Figure 39 contrasts the initial undigested CS sample with the progress of tryptic digestion at the chosen time points. Prior to trypsinolysis, the t=0 spectrum shows the presence of intact dimer, and three smaller fragments of masses 12 214, 12 211, and 12 168 Da, whose possible origin has been previously discussed (ref. p.75). These fragments disappeared quickly upon digestion. Both monomer, and to a lesser degree dimer envelopes are visible during the course of the digestion. As shown in Figure 40,

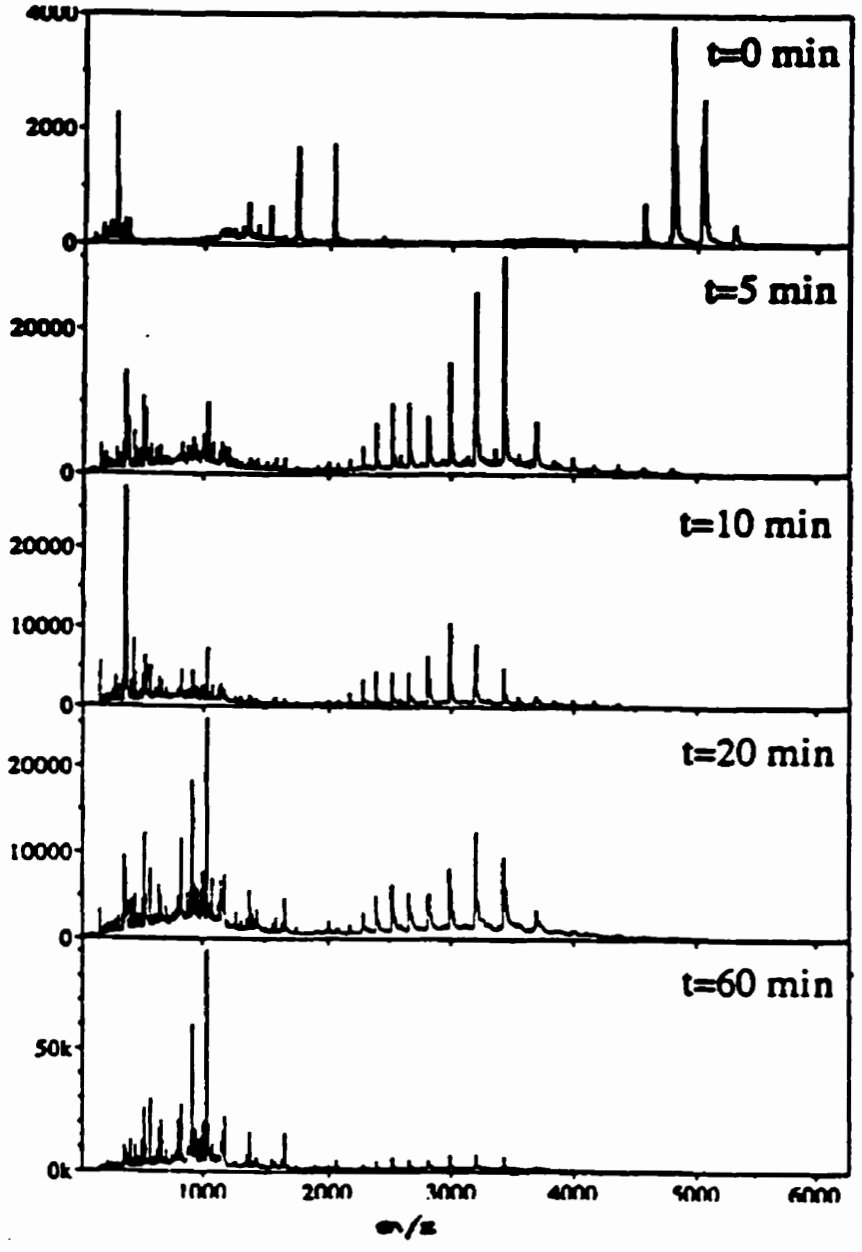



Figure 39: Limited proteolysis timecourse experiment conducted at room temperature. Aliquots of the digest mixture (1:160 w/w trypsin to CS in 20 mM ammonium bicarbonate) were withdrawn at timed intervals and analyzed by ESI TOF MS. The undigested spectrum at t=0 shows the presence of the dimer charge envelope (around 5000 m/z) and two small 12 000 Da fragment charge envelopes (around 2000 m/z). From t=5 minutes onwards, spectra show the monomer charge envelope (around 2000-4000 m/z) and peaks corresponding to small tryptic peptides (around 0-1600 m/z).

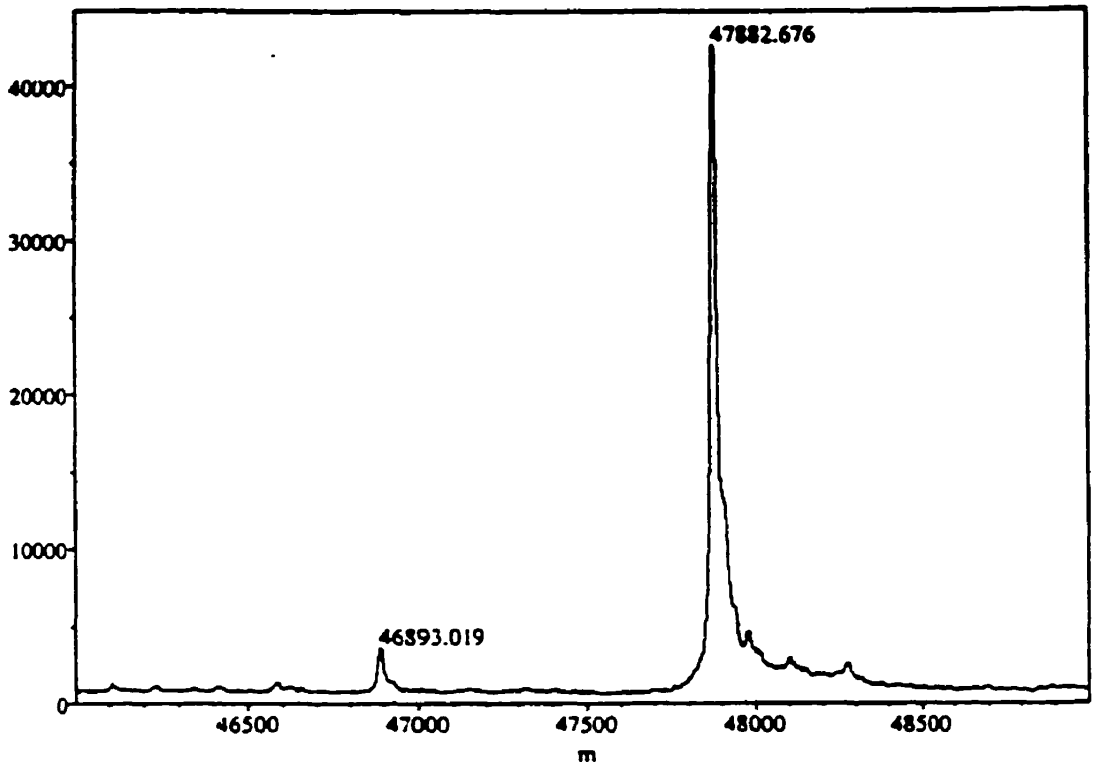



Figure 40: Deconvolution of the 5 minute spectrum from Figure 39. Note the presence of the CS monomer at 47 882.676 (expected mass = 47 887 Da) and a truncated monomer at 46 893.019, corresponding to the loss of 8 C-terminal residues (tryptic fragments T42-T45).

deconvolution of the monomer charge envelope revealed a second precursor peak with a mass of 46 893.4 Da (intact monomer = 47 887 Da), which can be identified as CS which has been truncated by removal of the last eight residues by trypsin (calculated average mass 46 895.8 Da). As the majority of peaks contributing to this truncated monomer's charge are visible only in the 5 minute digest, it is likely that this is a transient intermediate in the digestion process. This CS(1-418) is the only large intermediate observed. The other peaks visible in the digest spectra represent low mass tryptic fragments. Detailed analysis shows that about 92% of the CS sequence is represented, although fragments derived from 70-104 are unaccounted for.

### *2. Proteolysis of Partially Unfolded CS*

Limited proteolysis experiments aimed at characterizing the stable CS folding intermediate of Ayed and Duckworth (1999) were unsuccessful. The original unfolding experiments were conducted using CS in increasing concentrations of up to 9 M urea or guanidine hydrochloride, but both of these substances are unsuitable for mass spectrometry, and even though the first folding transition to the intermediate is said to be irreversible (Ayed, 1998), repeated dialysis would not necessarily remove all of the denaturants. The construction of a two-stage unfolding curve from heat denaturation of the enzyme had also been demonstrated by Ayed (1998), and so it was decided to pursue the intermediate in this way. CS was heated in a water bath at 60°C, the approximate temperature at which the unfolding intermediate was obtained in circular dichroism studies (Ayed, 1998). The protein was then either digested with trypsin while still at 60°C, or removed and immediately placed on ice before trypsinolysis. In either case, ESI

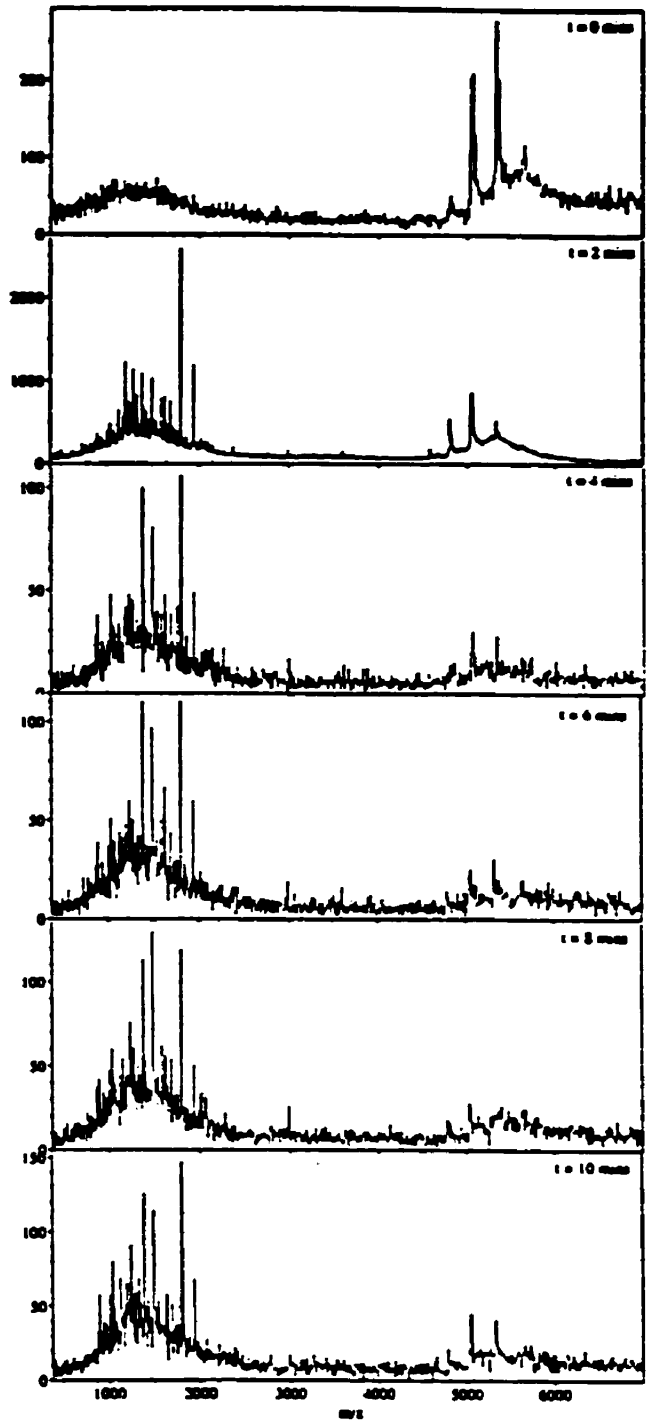


Figure 41: Online limited proteolysis of 10 μM CS by subtilisin BPN at a w/w ratio of 4500:1 in 5 mM ammonium bicarbonate. Spectra are shown at 2 minute intervals up to 10 minutes. Note the disappearance of dimer charge envelope on the right, and accumulation of proteolytic fragments on the left.

MS revealed that even after just 5 minutes of exposure to trypsin at a variety of dilutions, nothing remained of the CS except small tryptic fragments (data not shown). This may well be the result of protein aggregation and precipitation upon cooling, as observed in the past (Ayed, 1998). No evidence was found of a change in the tryptic digestion pattern that could be attributed to the presence of the intermediate form.

***Online Limited Proteolysis***

***1. Digestion by Subtilisin***

Figure 41 shows the timecourse of the limited proteolysis of CS at a ratio of 4500:1 (w/w) with subtilisin BPN. Spectra were recorded at one minute intervals using the Store and Restart feature of the Tofma program. Most of the proteolysis seems to occur within the first two minutes of incubation, although the dimer charge envelope is still visible after 10 minutes. Peaks between 500 to 2500 m/z correspond mostly to small proteolytic fragments of charge +1 to +3, which could not be positively identified due to the wide range of possible subtilisin targets. Also observed in this m/z range throughout the course of the experiment were charge envelopes corresponding to polypeptides of mass 12 211, 12 423, and 37 314 Da.

While these larger mass fragments are in the same approximate range as the 32 000, and 13 500 Da polypeptides observed in the previous limited proteolysis study of *E. coli* CS (Bell et al., 1983), they are not a close match. The 12 211 and 12 423 Da fragments observed here might correspond to the the same unknown polypeptide observed in spectra of intact CS (Figure 21).

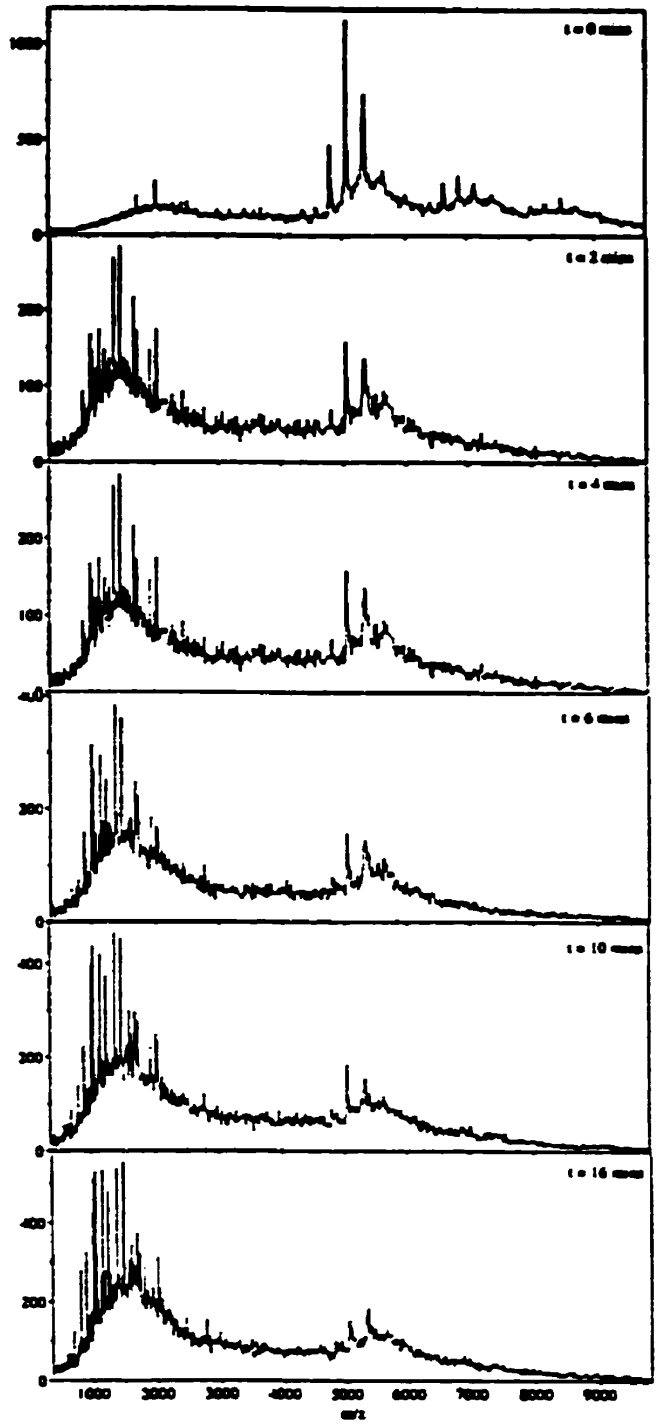


Figure 42: Online limited proteolysis of 10 μM CS by trypsin at a w/w ratio of 20:1 in 5 mM ammonium bicarbonate. Spectra are shown at intervals up to 16 minutes.

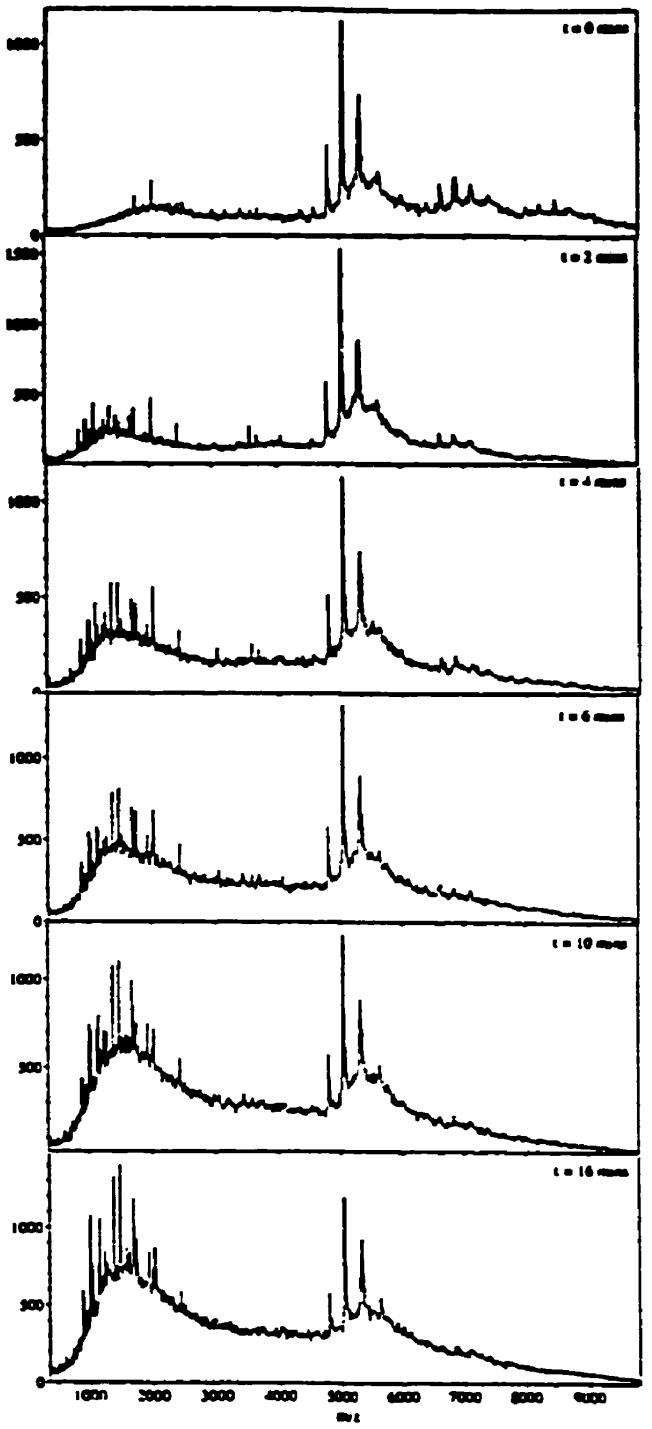


Figure 43: Online limited proteolysis of 10 μM CS by trypsin at a w/w ratio of 100:1 in 5 mM ammonium bicarbonate. Spectra are shown at intervals up to 16 minutes.

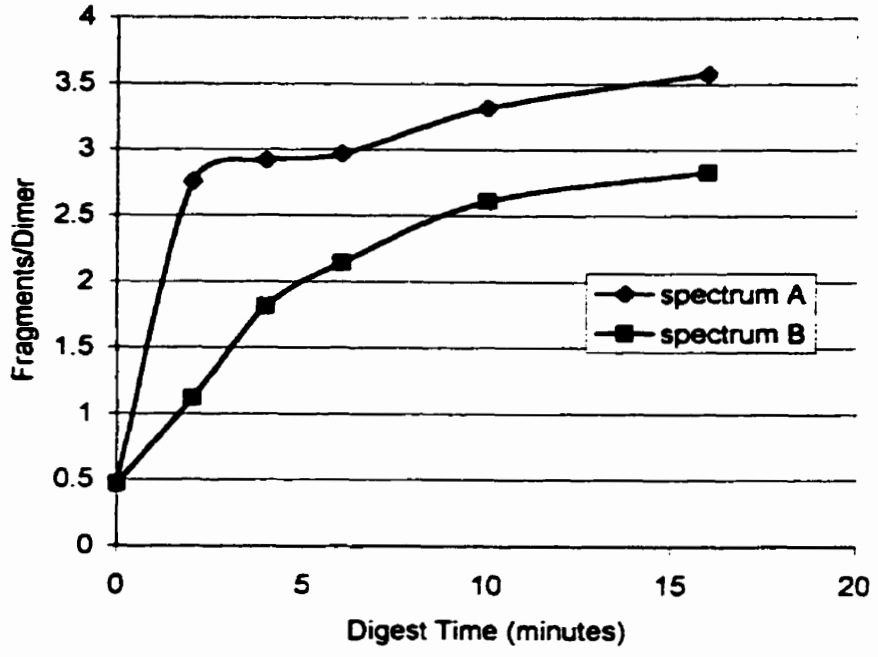



Figure 44: Curves constructed from data shown in Figures 42 and 43, showing the relationship between the tryptic fragment/dimer ratio and the length of the digest for two different trypsin concentrations. Spectrum A corresponds to Figure 42 (20:1 CS/trypsin) and Spectrum B corresponds to Figure 43 (100:1 CS/trypsin).

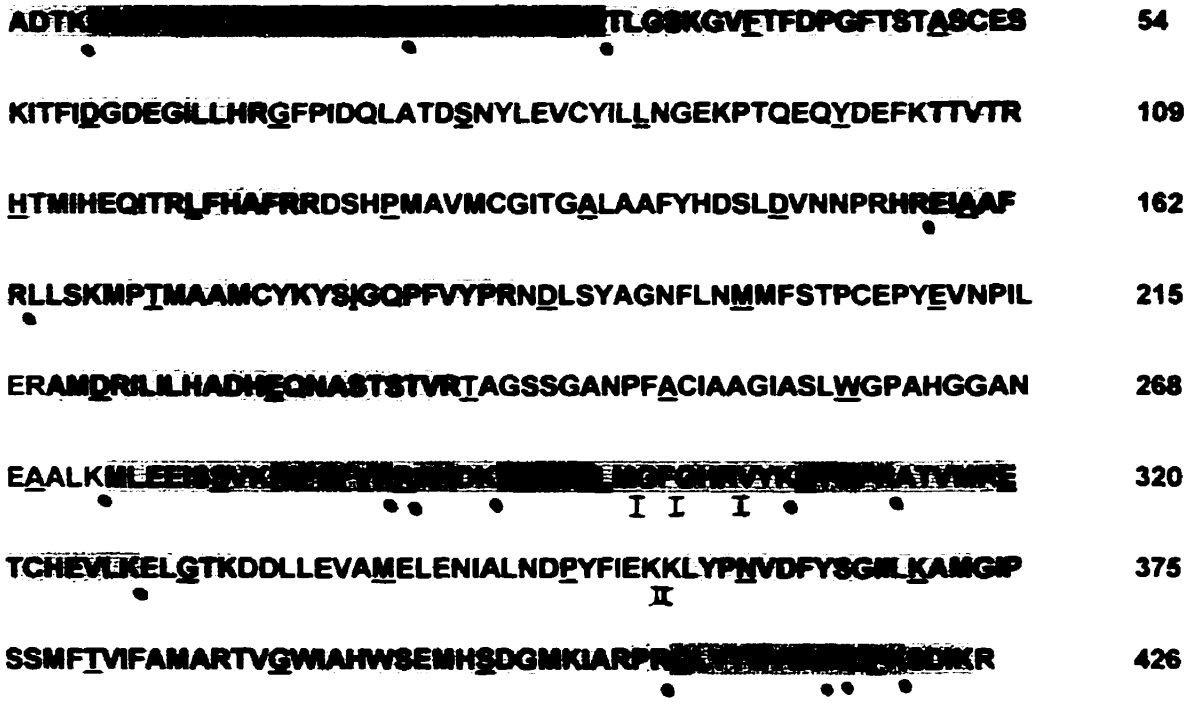



Figure 45: Summary of the tryptic peptides (assigned by mass) found in the 1:20 online digest. Portions of sequence accounted for by these peptides are shaded, where the darker the shading, the greater the number of fragments found with that particular sequence (the darkest shading indicates 5 fragments). Red dots denote the most common sites of cleavage observed. Green numerals denote the primary (I) and secondary (II) subtilisin cleavage sites as determined by Bell et al (1983). Every tenth amino acid of the sequence is underlined.

*2. Digestion by Trypsin*

Figures 42 and 43 show the timecourse of the limited proteolysis of CS at ratios of 20:1 and 100:1 (respectively) (w/w) with trypsin. Once again, fairly rapid conversion of the dimer charge envelope into smaller fragments was observed. It should be noted that these timecourses do not ressemble that of Figure 39, as they were generated under non-denaturing conditions. The timed digest aliquots used to construct the series of spectra in Figure 39 were each treated with 2.5% acetic acid to stop digestion, which could induce denaturation of dimer to monomer.

Figure 44 shows the curves generated by the change in the ratio of total fragment and dimer peak areas as the digests progress. Efforts to fit this data to first order kinetics for the direct disintegration of dimer to small tryptic peptides using the following equation were unsuccessful:

$$[\text{fragments}]_t/[\text{dimer}]_t = (1 - e^{-kt})/(e^{-kt})$$

where t = digest time interval

k = rate constant for conversion of dimer into fragments

Many of the smaller peaks in the online limited trypsinolysis spectra correspond to tryptic peptides of approximately 400 to 4000 Da, which would typically be found in a complete tryptic digest, and account for up to 50% of the CS sequence for the 1:20 digest, and up to 65% of the sequence for the 1:100 digest. Faint charge envelopes corresponding to the 12 200 Da fragment, and intact monomer were also detected in the 2000-3500 m/z range of the 1:100 digest sample spectra at all time intervals recorded.

As seen in Figure 45, the common sites of tryptic cleavage change little over the course of digestion. With reference to the structure of the CS monomer in Figure 5, these are

found in the N- and C-termini, a small section around residue 160 at the N-terminal end of the I helix, and a large region from residue 274 to 332 which encompasses most of the N, O, and P helices of the small domain, and also contains the primary subtilisin cleavage site (Bell et al, 1983) on a non-helical loop between O and P.

### *3. Proteolysis of Ligand Bound CS*

Attempts to characterize the limited proteolysis of ligand bound CS did not meet with much success. Previous limited proteolysis experiments determined that the binding of ligands which act as activators, inhibitors, and substrates of CS altered the rate, but not the site of proteolysis (Bell et al., 1983, Bloxham et al., 1980). Online limited trypsinolysis of CS in the presence of NADH gave spectra which were poorly resolved, and so difficult to analyze (Figure 46). Nonetheless, it was possible to estimate by eye the rate at which hexamer was converted into small tryptic fragments, and note that in comparison to a tryptic timecourse of CS in the absence of NADH performed on the same day, there appears to be very little difference in the rate of proteolysis. Bell et al. (1983) found that NADH increases the rate of CS digestion (as judged by loss of CS activity) by 60%.

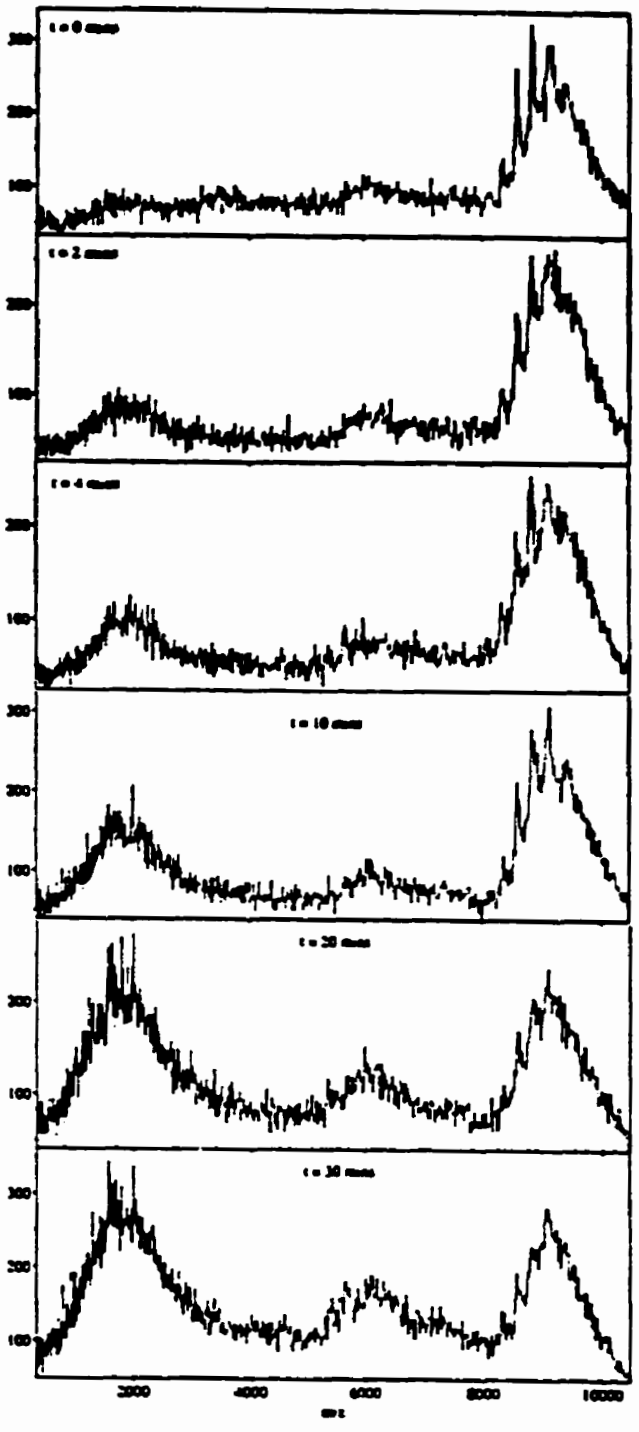


Figure 46: Online limited proteolysis of 10 μM CS by trypsin (100:1, w/w) in 5 mM ammonium bicarbonate and in the presence of 40 μM NADH. Note the gradual conversion of the hexamer charge envelope (far right of the spectra) to tryptic peptides (far left of the spectra). The charge envelope at approximately 6000 m/z is thought to be CS dimer.

## *Allosteric Activation of CS Studied by Mass Spectrometry*

ESI TOF MS has previously been used to demonstrate the hexamerization of CS in the presence of NADH, and even allowed for the determination of the $K_D$ for this inhibitor (Ayed et al., 1998). As mentioned in the Introduction, sedimentation equilibrium studies have found that the shift from dimer to hexamer also occurs upon addition of KCl to the enzyme (Tong & Duckworth, 1975). Fortunately, one of the other monovalent cations able to act as an allosteric activator of CS is the ammonium ion, which is compatible with MS experiments. Thus, CS in 5 to 200 mM ammonium bicarbonate buffer could be examined using nanospray ionization MS, a technique that allows for the use of volatile buffers with concentrations above the standard limit of 5 mM used for ESI MS.

As shown in Figure 47, increasing the concentration of ammonium ion clearly shifts the dimer-hexamer equilibrium to the right. This change mirrors the increase in CS activity observed when the enzyme was assayed in these same ammonium bicarbonate concentrations (Figure 48). The peak areas used to calculate the percentage of hexamer in each spectra (relative to the total hexamer and dimer peak area) were determined by taking the areas under entire charge envelopes as opposed to the individual peaks of the envelope. The peak areas obtained for the dimer envelopes were divided by a factor of 1.3, which accounts for the difference in detector sensitivity to dimer and hexamer (Ayed et al., 1998). The correspondence between the percentage of hexamer and specific activity is not entirely linear, and it would be useful to acquire more spectra of the samples used in this experiment to provide a better data set. Nonetheless, the results

obtained served to confirm a phenomenon first noted over 20 years ago, and demonstrated the usefulness of ESI MS in structure-function studies of enzymes.

Figure 49 presents an interesting question for further study. The nanospray ionization mass spectrum of CS in 100 mM ammonium bicarbonate seems to show not one, but two charge envelopes in the hexamer region (denoted as **a** and **b** in the figure). A similar pattern of two hexamer charge envelopes can barely be seen on the 100 mM spectrum of Figure 47. The charge states of all the hexamer peaks are in fact continuous, and deconvolutions of these envelopes both separately and as one are essentially the same. In addition, there does not appear to be any progression from one form of hexamer envelope to another in the spectral series of Figure 49. A more rigorous study will be required to prove if ESI MS can indeed detect two different forms of the CS hexamer (assumed to be the active and inactive conformations of the enzyme).

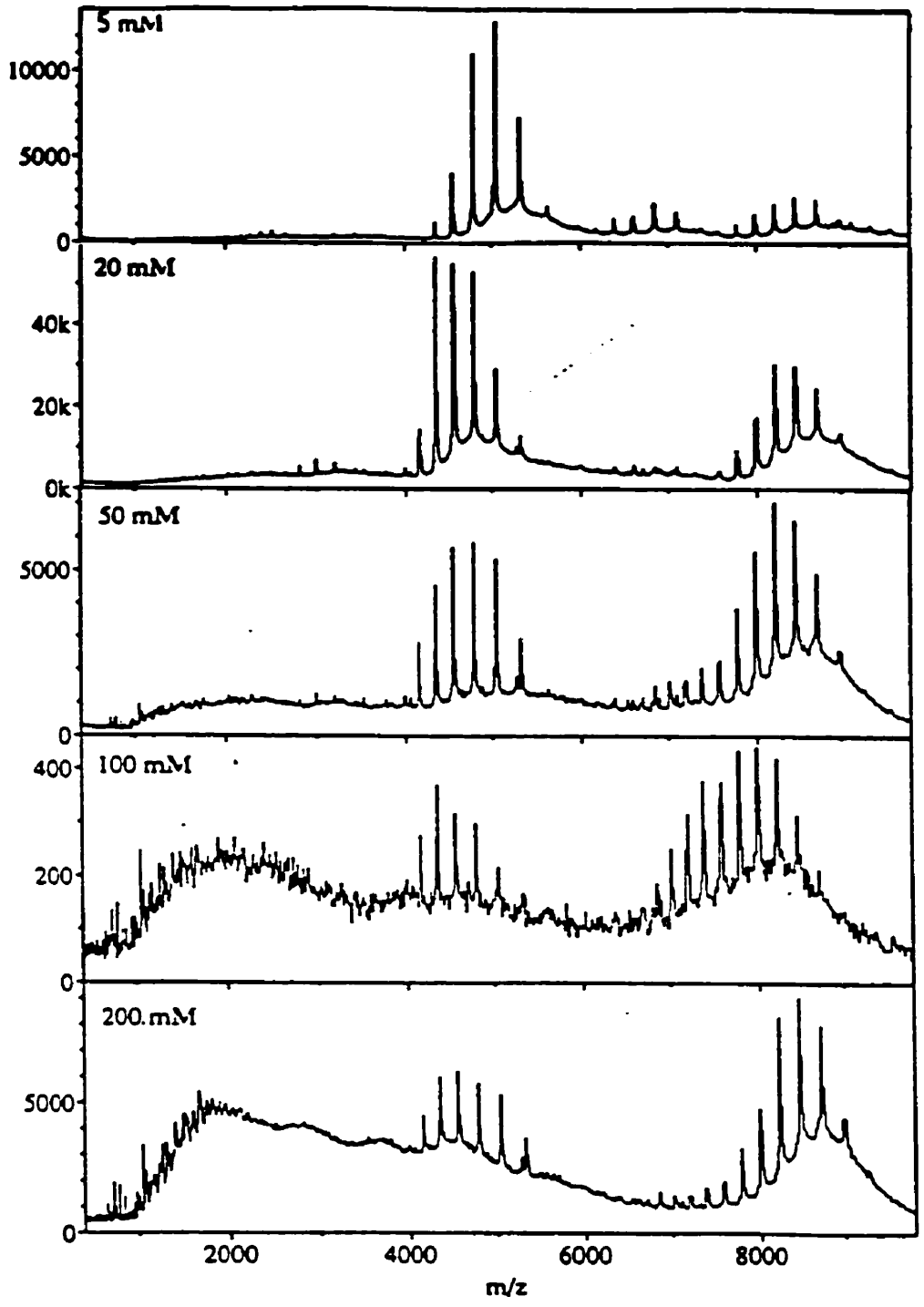



Figure 47: Nanospray ionization mass spectra of 10 μM CS in 5 to 200 mM ammonium bicarbonate buffer. The most prominent charge envelopes belong to dimer and hexamer, as previously observed (refer to Figure 21). Increasing ammonium ion concentration results in a decrease in the dimer charge envelope, and an increase in the hexamer charge envelope.

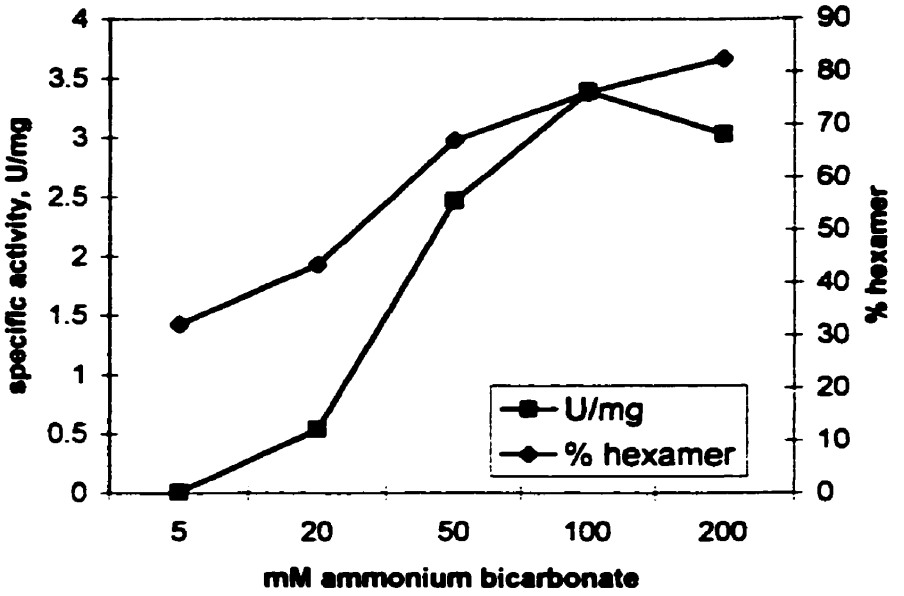



Figure 48 (a): Effect of ammonium bicarbonate on CS specific activity and the percentage of hexamer found in the nanospray ionization mass spectra of Figure 47. The percentage of hexamer was calculated as the percentage of hexamer peak area with respect to total hexamer and dimer peak area. Peak areas were determined by taking the total area under each charge envelope from the baseline.

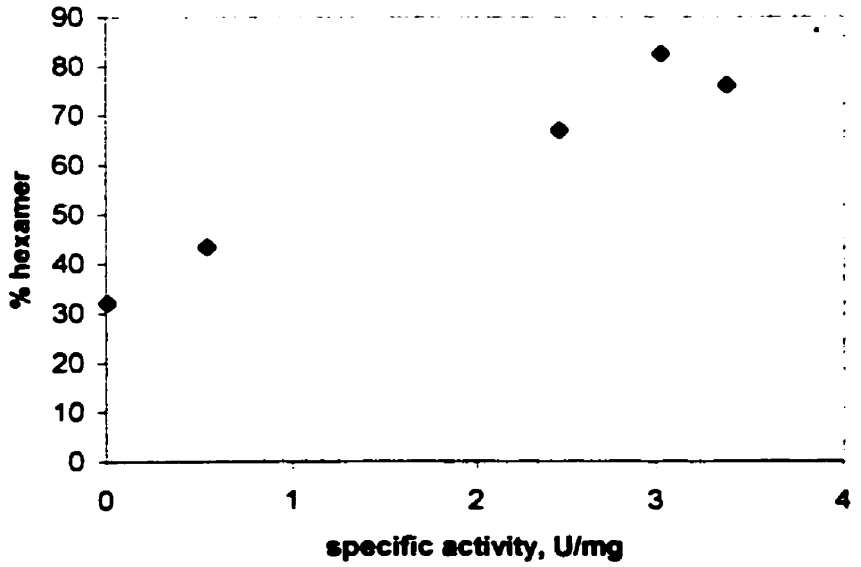



Figure 48 (b): Relationship between the percentage of hexamer and the specific activity of CS in increasing concentrations of ammonium bicarbonate.

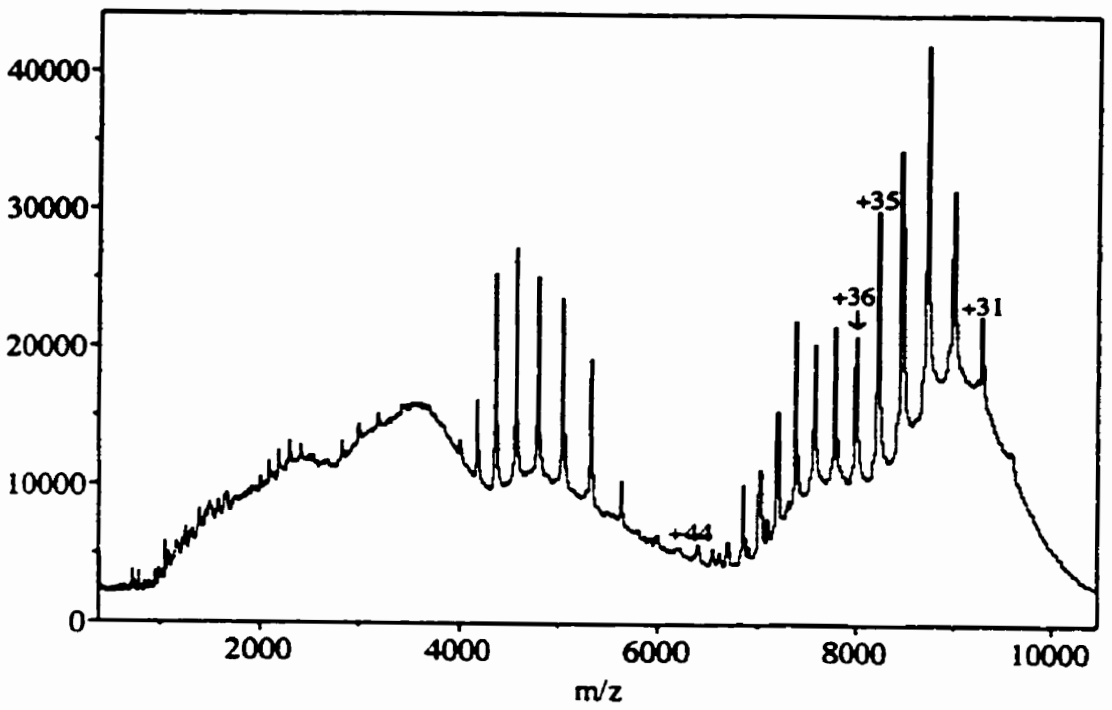



Figure 49: Nanospray ionization mass spectrum of 10 μM CS in 100 mM ammonium bicarbonate. Indicated are the two possible hexamer charge envelopes:

a) +36 to +44
b) +31 to +35

# *Discussion*

## *Photoaffinity Labeling of CS*

## *Preliminary Experiments*

As with any photoaffinity labeling study, a series of preliminary experiments designed to show the strength and specificity of the interaction between CS and the azidoATP analogues were carried out prior to attempts to identify the labeling site with mass spectrometry. The control experiment, which demonstrated the necessity of UV activating light to achieve photolabeling, also gave the first indication that CS bound the 2-azidoATP analogue far better (33% incorporation) than its 8-azidoATP counterpart (8% incorporation). This was confirmed by attempts to saturate the binding site of CS with these analogues. The saturation curves demonstrate that 8-azidoATP reaches near saturation at a level of photoincorporation half that of 2-azidoATP. The azidoATP probes had higher estimated $K_D$ values when compared to the $K_D$ of ATP itself of 46 ± 5 μM (Talgoy & Duckworth, 1979), as expected for the substituted version of the natural ligand. The $K_D$ values were 100 μM and 75 μM for 8-azidoATP and 2-azidoATP, respectively, confirming that the binding of the latter label to CS is most favorable.

The specificity of labeling using these photoprobes was further proven by the notable drop in label incorporation in the presence of increasing concentrations of the natural inhibitor of CS, NADH. Up to 70% of 2-azidoATP binding and up to 22% of 8-azidoATP binding is lost as the NADH concentration in the labeling mixture is raised to 200 μM. The results of a parallel experiment using the active site deletion mutant of CS, Δ(264-287)CS, also confirmed suspicions that at least the 8-azidoATP label, like ATP itself, was capable of binding to a small degree to the active site as well as the allosteric

site. However, at the lower concentrations of NADH (20-50 μM), photoincorporation is actually increased. A possible explanation for this unexpected result is that low levels of NADH binding to T state CS molecules pulls the allosteric equilibrium towards a T state which more fully exposes the allosteric site, thus rendering the binding of adenylates (in this case, the azidoATP photolabels) more favorable. At higher concentrations of NADH, the natural inhibitor successfully competes for the available allosteric sites. This theory is derived from a similar phenomenon observed in aspartate transcarbamylase, another allosteric enzyme (Gerhart & Pardee, 1963). At low concentrations of its allosteric activator aspartate, its isosteric inhibitors maleate and succinate can pull the allosteric equilibrium from T to R state by binding to one active site on an R state molecule. This creates a surplus of active sites available for aspartate to bind, aiding activation of the enzyme. However, at higher concentrations, the isosteric inhibitors will begin to seriously compete with aspartate for the active sites, and inhibit the enzyme.

There is some cause for concern when comparing the relatively inconsistent levels of photoincorporation achieved in the saturation, and labeling inhibition experiments as shown in Table 4. In addition, a background of 6% photoincorporation was observed in the absence of UV light during the control experiment. All of these preliminary experiments were performed using the same preparation of CS, and at the same CS concentration. It is likely that these discrepancies resulted from the method used to quantitate photoincorporation, that is, excision and scintillation counting of bands from SDS PAGE gels. The photolabeled CS band was never directly observable on the gels due to the low level of protein present, but was instead excised solely on the basis of the

Table 4: Levels of incorporation of photoaffinity labels at a label concentration of 175 μM. Data are taken from Figures 26 and 27.

| | % photoincorporation, saturation curve | % photoincorporation, inhibition curve |
|---|---|---|
| **2-azidoATP** | 23 | 12 |
| **8-azidoATP** | 16 | 21 |

position of a visible CS band from samples loaded at either end of the gel. Although precautions were taken, such as running samples in alternating lanes, and excising a portion of the gel surrounding the suspected region of the photolabeled CS band, inaccurate removal of these bands may have occurred. In addition, the presence of a small amount of what was later determined to be nonspecific high molecular weight aggregates of CS was observed in the autoradiograms of the SDS PAGE gels from the preliminary experiments. As these bands were never excised and counted along with the expected CS gel bands it is possible that some of the variability in levels of photoincorporation shown in Table 4 may have arisen from variability in the amounts of these aggregates from sample to sample.

Large amounts of radioactivity were found in the fix and Coomassie stain used on the gels. The length of time required for fixing, staining, and destaining the gels was limited as much as possible, but loss of radioactive label by hydrolysis in these acidic solutions cannot be ruled out, nor could it be stringently controlled.

The 6% incorporation of photolabel into CS in the absence of activating light may be accounted for by the SDS PAGE method of quantitation. Nonetheless, despite assurances from photolabel suppliers and the precautions outlined in the Materials and Methods section of this thesis for avoiding sample exposure to light, it might be that ambient room light caused some photolysis of the azidoATP labels prior to UV exposure.

Efforts were made to increase levels of photoincorporation by experimenting with higher concentrations of photolabel in the reaction mixture, multiple rounds of photolabeling, and increased UV exposure time. The best result was obtained when three rounds of photolabeling with 200 μM 2-azidoATP produced photoincorporation of 63%.

Given that consistently higher levels of photoincorporation were obtainable using the 2-azidoATP probe, it alone was selected for use in the peptide mapping experiments.

While increasing the length of UV irradiation of the photolabeling mixture increased levels of photoincorporation almost linearly, it caused the loss of up to 70% CS activity (as determined by enzyme assay) within 10 minutes. This loss of activity was concurrent with the appearance of higher molecular weight aggregates of CS as evidenced by SDS PAGE of CS samples subjected to timed intervals of UV exposure (data not shown). These aggregates were found in approximately the same location on the gels as those observed in the autoradiograms of gels containing photolabeled CS samples, and did not match the bands produced by dimethylsuberimidate cross-linking of CS (specific aggregation). These aggregates were thus assumed to be non-specific.

The ESI TOF mass spectra of both UV irradiated and unexposed CS were virtually the same (Figure 32), and only a 13% drop was observed in the $K_D$ of NADH for CS, and in the number of NADH bound per CS subunit upon UV exposure for the same duration (3 x 2 minutes) as in a typical photoaffinity labeling experiment. Both of these factors suggest that despite the production of aggregates, the UV damage to CS is limited to the active site. The damage may arise from photooxidation of residues such as histidine (Danson & Weitzman, 1973), of which there are three residues (264, 229, and 305) thought to be involved in the active site of CS.

## *Mapping Photolabel Incorporation by Mass Spectrometry*

Mass spectrometry of unlabeled and photolabeled unfractionated digests of CS to map the position of 2-azidoATP photoincorporation proved more challenging than expected, largely due to the volume of data contained in the mass spectrum of an unfractionated CS digest. No photolabel incorporation could be detected by MALDI TOF MS, but ESI TOF MS identified a peak of mass 1667, which appears only in the photolabeled CS digest, as a putatively photolabeled variant of the tryptic peptides 1146.48 Da (T17, corresponding to residues 168-177) and/or 1147.68 (T15-16, corresponding to residues 158-167), whose intensity decreases significantly in the photolabeled CS digest (Figure 33). It is virtually impossible to determine if the peak at 1146 Da is T17 or T15-16 by ESI MS alone, as their masses are just over 1 mass unit apart, the same difference observed between peaks in the monoisotopic series of a singly-charged peak. However, it stands to reason that T17 would be the dominant component of this peak, as T15 and T16 were individually identified in the digest spectra. Both T17 and T15-16 are found in the vicinity of C206, and demonstrate some degree of surface exposure of their residues in the new crystal structure of CS (Figure 34). It should be borne in mind when using this structure to evaluate the positions and surface exposure of putative allosteric site residues, that the crystals were obtained in a high salt environment, which would favor the R-state hexamer, and not the T-state hexamer which ably binds NADH. It must also be remembered that only one residue on each of these peptides is putatively photolabeled, not the entire peptide.

The tryptic fragment T17 was also found to be absent in the CE/MS analysis of the photolabeled CS digest. However, no trace of the putatively photolabeled peak of mass 1667 Da could be found in these experiments. Several other peaks were identified as possibly bearing photolabel incorporations, and a few of these peaks are also found in the ESI MS results. However, none of these peaks are truly in the vicinity of C206. This raises the possibility of nonspecific photoaffinity labeling, something which fits with the number of photolabel-sized adducts observed in the mass spectrum of intact photolabeled CS (Figure 32).

There are a number of cautions to be expressed about the above results. The assignments of both the tryptic peptides and their putatively photolabeled counterparts are based solely on mass, and on the assumed mass change of +520 Da after labeling with 2-azidoATP. The CE/MS data do allow for confirmation that an observed peak arises from the protein sample (not background noise), and has a specific retention time. Nonetheless, none of the above-mentioned peptides has undergone any kind of sequencing to confirm their identities. In addition, there may have been further photolabel incorporations which went undetected. In particular, phosphates are known to interfere with the formation of ions in mass spectrometry processes (M. McComb, personal communication). This may have contributed to the inability to see the putatively photolabeled 1667 Da peptide in the CE/MS results. If photolabel modification of a particular peptide were of low efficiency (perhaps as a result of steric hindrance) it also might have been overlooked. The systematic method of elimination used to sift through the mass spectral data did not take into account the possibility that there could be more than one photolabel incorporation per peptide, or that an assigned tryptic peptide could

have the same mass as the photolabeled variant of another peptide. Finally, the negatively charged phosphates of the label may have picked up a number of sodium adducts while in solution.

The fact that label incorporation may have occurred on more than one potential allosteric site peptide (T17 and/or T15-16) is not cause for alarm, but is expected of a chemical modification system whose targeting method (the ATP moiety of 2-azidoATP) is specific but whose modification method (the azido moiety) is not. In fact, multiple sites of incorporation can help to more fully define a particular binding site. For this reason, different analogues of a photolabel (e.g. the 2- and 8- forms of an azidoadenylate label) are also employed in labeling studies. In this study, the level of 8-azidoATP photoincorporation into CS was judged to be insufficient for detection in an unfractionated CS digest. The appearance of more than one potentially photolabeled CS peptide might also suggest a certain amount of variability in the way that 2-azidoATP interacts with the allosteric site, a possibility not inconsistent with the higher $K_D$ values estimated for the azidoATP analogues in the preliminary round of photolabeling experiments.

A survey of the literature concerning the use of mass spectrometry in photoaffinity labeling studies suggested that mass spectrometry of unfractionated proteolytic digests has not previously been used to map photolabel incorporation. Those few groups who did have access to mass spectrometry facilities used them either directly online with an HPLC system to identify the peptides of fractionated digests, or as a method of confirming that a radioactively-tagged photolabeled peptide isolated by HPLC was of the

correct mass (e.g. Branchini et al., 1995; Cho et al., 1997; Rousselot et al., 1997; Shapiro et al., 1997).

Both the conventional HPLC-based methods used for isolating and identifying photolabeled peptides, and the method described in this thesis, which involves mass spectrometry alone, have their advantages and disadvantages. The mass spectrometry method does require very pure samples for ESI MS, but is more economical with respect to the amounts of both the target protein and the photolabel required per experiment. Providing one has access to a high-resolution ESI mass spectrometer, spectra of both intact and digested labeled and unlabeled protein can be rapidly acquired, whereas HPLC-based separations may take several hours to run to completion. Based on the observations in this thesis, the photoinsertion bond is at least partially stable to ESI MS conditions. Loss of label has been reported in some HPLC separations (Chavan, 1996), but these difficulties have largely been solved by optimizing flow rates, and by the use of immobilized metal affinity chromatography.

The single greatest drawback to mapping a modification in an unfractionated digest by mass spectrometry is the amount of tedious analysis involved in catalogueing and identifying the many peaks found in the control and labeled protein spectra. This was particularly the case for CS, which produces a large number of complete and incomplete tryptic peptides (see Appendix 2). In the ESI MS results, some higher m/z fragments (particularly beyond 3000 m/z) proved difficult to characterize, as their resolution was too poor to establish their charge state. Analysis of CE/MS results, which provides confirmation as to which peaks are real and which are simply background and thus decreases the number of peaks to work with, is even more lengthy, as both CE and MS

data must be examined and cross-referenced. By contrast, once an HPLC method for separation of proteolytic peptides is in place, identification of photoincorporation by a radioactive or fluorescent tag on the photolabel should be relatively straightforward. The isolated peptide can then be subjected to further tests (e.g. sequencing, mass analysis) to determine its identity and the site of photoincorporation.

The complexity of the data, and the stringency of sample preparation are reduced when MALDI MS is used instead of ESI MS, as multiple charge states are few, and it was the original intent of this project to run digest samples on a MALDI mass spectrometer. However, the inability to identify any photolabeling in both CS and creatine kinase led to the decision to use ESI MS instead.

## *Studying Other Chemical Modifications of CS by Mass Spectrometry*

In addition to the 2-azidoATP photolabel, mass spectrometry was applied to the mapping of four other chemical modifications which target the reactive cysteines of CS, particularly C206. The alkylation of C206 by TFBA was the first of these to be attempted, as a method of demonstrating that it was possible to identify a chemical modification of the protein in an unfractionated CS digest using MALDI MS. As shown in Figure 31, the modified peptide appears in the spectra as a peak of mass 110 Da greater than the unmodified peptide T19, which contains residues 189-217. Thus modification of C206 is confirmed.

The TFBA alkylation of CS is a known modification, and gives high yields. This would not necessarily be the case with DTNB, DPDS, and β-mercaptoethanol modifications of the enzyme, although DTNB and DPDS are known to react with C206 (Talgoy et al., 1979). The second cysteine residue which reacts with DPDS has never been identified, and it was unknown which cysteines would react, if at all, with the β-mercaptoethanol.

ESI TOFMS on TOF III was used to observe the characteristics of the intact modified CS, while unfractionated tryptic digests of these samples were analyzed by MALDI TOFMS on TOF IV. Unfortunately, only the spectra of CS incubated with β-mercaptoethanol were able to give clear evidence of sulfydryl group modification. As the formation of mixed sulfhydryls in CS samples incubated with DTNB and DPDS was demonstrated spectrophotometrically, it is then likely that these modifications were lost

during the cleanup of samples in preparation for mass spectrometry. Repeating this experiment, with immediate transfer of freshly modified samples to the mass spectrometer could provide better results. Mass spectra of intact β-mercaptoethanol modified protein showed the presence of up to two β-mercaptoethanol additions per CS monomer, based on deconvolution of the monomer charge envelope. Deconvolution of the dimer envelope indicated only one addition per monomer, but further adduct peaks might have been hidden in the trailing edge of the peak.

The method used to sort through the unfractionated digest data from TOF IV was virtually the same as that used on the tryptic digest data of photolabeled CS. It showed two potential modifications of CS peptides by β-mercaptoethanol. One of these is peptide T19, which contains C206. A second peptide of residues 70 to 104 (unmodified peptide = T8, mass of 4080.94 Da) also showed β-mercaptoethanol modification. This latter fragment contains C86, which is partially surface exposed and a possible candidate for the second cysteine modified by DPDS. As with the photolabeling study, it must be stressed that these assignments are based on mass alone, and require sequencing in order to conclusively prove their identity.

## *Limited Proteolysis of CS*

In a previous limited proteolysis study of CS, it was found that subtilisin initially cleaved the enzyme into a larger N-terminal fragment or approximately 32 000 Da, and a smaller 13 500 Da C-terminal fragment (Bell et al., 1983). Polypeptides of 7 500 and 24 000 Da appeared later in the digest. Their results indicated that trypsin, chymotrypsin, and thermolysin also produced similar large and small CS fragments, with the smaller of the two being less stable to further proteolysis than that generated by subtilisin. Proteolytic enzymes were used at a ratio of 1:200 (w/w, with CS), and digestion was at room temperature (with the exception of thermolysin, at 60°C). The above observations mirrored those of limited proteolysis studies of pig heart CS (Wiegand et al., 1979; Bloxham et al., 1980; Bayer et al., 1981; Bloxham et al., 1981), and presented an argument for structural similarity between pig heart and *E. coli* CS.

From the limited proteolyis results presented in this thesis, it would initially seem that proteolysis of CS is an "all or nothing" process, as only a few large proteolytic peptides were observed during the course of both the conventional and online limited proteolysis experiments. One, at 46 893 Da appears to be a simple truncation of the CS sequence at residue 418, and could have arisen if this C-terminal "tail" were in any way freely exposed in solution. This truncated species was observed in the conventional limited proteolysis experiment, and virtually disappears after five minutes of tryptic digestion. Of the large fragments observed throughout the 10 minute online subtilisin digest of CS, the largest (37 314 Da) could conceivably correspond to the large N-terminal fragment

observed by Bell et al. Coincidentally, the others (12 211, 12 423 Da) are of the same mass as those found in CS preparations in the absence of specific proteolytic enzymes.

The apparent resistance of the native forms of CS to stepwise disintegration as the digest progresses is not so puzzling in light of observations from previous studies. Limited proteolysis of CS results in the gradual inactivation of the enzyme (Bell et al., 1983; Bloxham et al., 1980; Bayer et al., 1981; Else et al., 1988). The large proteolytic fragments, however, are generally isolated only under denaturing conditions, and gel filtration chromatography of a partial subtilisin digest of *E. coli* CS produced one large peak whose fractions showed decreased enzyme activity, but virtually the same NADH binding characteristics as the native enzyme (Bell et al., 1983). Thus it is likely that what appears to be native dimer or hexamer in the timecourse spectra presented here are in fact aggregations of limited proteolysis products which remain inseparable due to the ability of ESI MS to preserve non-covalent interactions. There may be then some threshold of nicks which the subunit assembly can tolerate before completely disintegrating into small proteolytic fragments.

While the use of mass spectrometry speeds up the identification of limited proteolysis products, results are entirely dependent on the quality of spectra obtained. This is especially apparent in the online experiments, where data are being collected and saved over very brief intervals. As found with the photoaffinity labeling experiments, mass spectral data of unfractionated protein digests can be overwhelmingly difficult to analyze, particularly for ESI MS, where the identification of charge states is key to assigning peak identities.

It was somewhat disappointing that these limited proteolysis experiments could not clearly document a stepwise pattern of proteolysis for CS, and the effect of ligand binding and partial denaturation on proteolysis, which might have provided further clues as to this enzyme's structure. Further mass spectrometry experiments conducted under carefully controlled, denaturing conditions may be successful in this goal. Nonetheless, online limited proteolysis was shown to greatly simplify the basic timecourse of digestion experiment. This technique could be of particular use in the preliminary trial and error optimization of the ratio of protein to digestive enzyme.

## *Allosteric Activation of CS Studied by Mass Spectrometry*

Nanospray ionization mass spectrometry was successfully used to demonstrate the changes in the dimer-hexamer equilibrium of CS in increasing concentrations of the ammonium ion, an allosteric activator of the enzyme. Although absolute quantitation of this effect was not the goal of this experiment, a correlation between the shift from dimer to hexamer and activation of the enzyme as determined by the standard assay was found, although it is not entirely linear. Repeating the experiment several times under more rigorous conditions, and finding a more reproducible method of determining the areas under charge envelope peaks of spectra in Tofma would allow for a more quantitative assessment of this phenomenon.

While this experiment confirmed results previously reported in the literature concerning the allosteric activation and subunit associations of CS, it also posed an entirely new question. Figure 49 appears to show two charge envelopes for the CS hexamer in 100 mM ammonium bicarbonate. Although the charge states of both putative envelopes are continous from one to the other, and deconvolute to the same single peak with the mass of the hexamer, an intriguing question arises – are the two charge states an artifact of the nanospray process, or do they give a hint that there are two different forms of the hexamer? The latter is likely, as the binding of both allosteric inhibitors (NADH) and activators (monovalent cations) has been demonstrated to shift the enzyme from dimer to hexamer, and yet NADH cannot bind to the activated form of CS, strongly suggesting that an additional conformational change has taken place. While a more rigorous investigation of this question by nanospray ionization mass spectrometry may

serve to confirm the true existence of two hexamer charge envelopes, it is more likely that a definitive answer will involve x-ray crystal structures of CS under a variety of free and ligand-bound conditions.

## *Allosteric Site Mutants Y186A and P205A*

Both DTNB and DPDS modification experiments seem to indicate that the mutation of Y186 or P205 to alanine have somehow partially blocked the reactive cysteines from reacting with DTNB and DPDS. One possible theory to explain this is that the changes in amino acids from the bulkier proline and tyrosine residues to the simplicity of alanine, which has a single methyl group as a side chain, may have induced some sort of structural collapse (a "cave-in") around C206, thereby lowering but not completely impeding access to the reactive sulfhydryl groups. This could also account for the greatly reduced NADH inhibition observed for both of these mutants by Ye and Hacking (unpublished results).

The new crystal structure of hexameric *E. coli* CS will allow cysteine 206 and all of the various mutant CS created over the years to be placed in context to each other, and structural features such as the active site, and domain and intersubunit contacts will be located. With such a detailed structural map of *E. coli* CS available, further site-directed mutagenesis studies will doubtless continue in the search for the allosteric site of this enzyme.

## *Concluding Remarks*

The work presented in this thesis demonstrates the usage of mass spectrometry in the characterization of certain properties of *E. coli* CS, and the mapping of specific chemical modifications of this enzyme directed at the allosteric site. Several points arising from these experiments can be made.

In theory, mapping chemical modifications of a protein in unfractionated proteolytic digests sounds easier than conventional HPLC-based methods which typically involve larger sample requirements and radioactively tagged ligands. In practice, the analysis of unfractionated CS tryptic digests was tedious and time consuming due to the large amount of data to be analyzed. The introduction of a second, tandem separation method (in this case, CE) did improve the confidence of MS peak assignments. The identification of 2-azidoATP photolabel incorporation was hampered by other factors, including the potential for several different label masses, and decreased ion production due to phosphate groups on the label. If possible, mapping of chemical modifications in unfractionated digests is more easily accomplished using MALDI MS (as for the β-mercaptoethanol labeling of CS) than ESI MS, as fewer multiply charged ions are produced.

Several of the experiments in this thesis made use of mass spectrometry to repeat previously published studies of *E. coli* CS. Mapping TFBA and β-mercaptoethanol modifications of CS by mass spectrometry confirmed that these reagents target C206, and suggested, via the β-mercaptoethanol results, a possible candidate for the second, as yet unknown reactive cysteine of CS. The hexamerization of the enzyme upon titration with

ammonium ion, and concurrent activation of CS activity, was also shown. Limited proteolysis experiments provided the opportunity to study the limited proteolysis of CS to greater mass accuracy under non-denaturing conditions and within the first few minutes of digestion.

Overall, *E. coli* citrate synthase is a good candidate for mass spectrometric studies. It gives excellent spectra, and is robust enough to survive the rigors of extensive sample cleanup prior to MS. Moreover, its allosteric nature provides for a number of interesting structure and function questions which still need answering.

# *Appendices*

## *Appendix 1*

Sequence alignment of representative hexameric and dimeric CS from the following organisms:

*Escherichia coli* (Gram negative bacteria, hexamer
*Pseudomonas aeruginosa* (Gram negative bacteria, hexamer)
*Acinetobacter acidophilum* (Gram negative bacteria, hexamer)
The second (methyl) CS of *E. coli* (Gram negative bacteria, dimer)
*Thermoplasma acidophilum* (Archaea, dimer)
*Bacillus subtilis* (CSII) (Gram positive bacteria, dimer)
*Sus scrofa* (Pig, dimer)
*Gallus gallus* (Chicken, dimer)

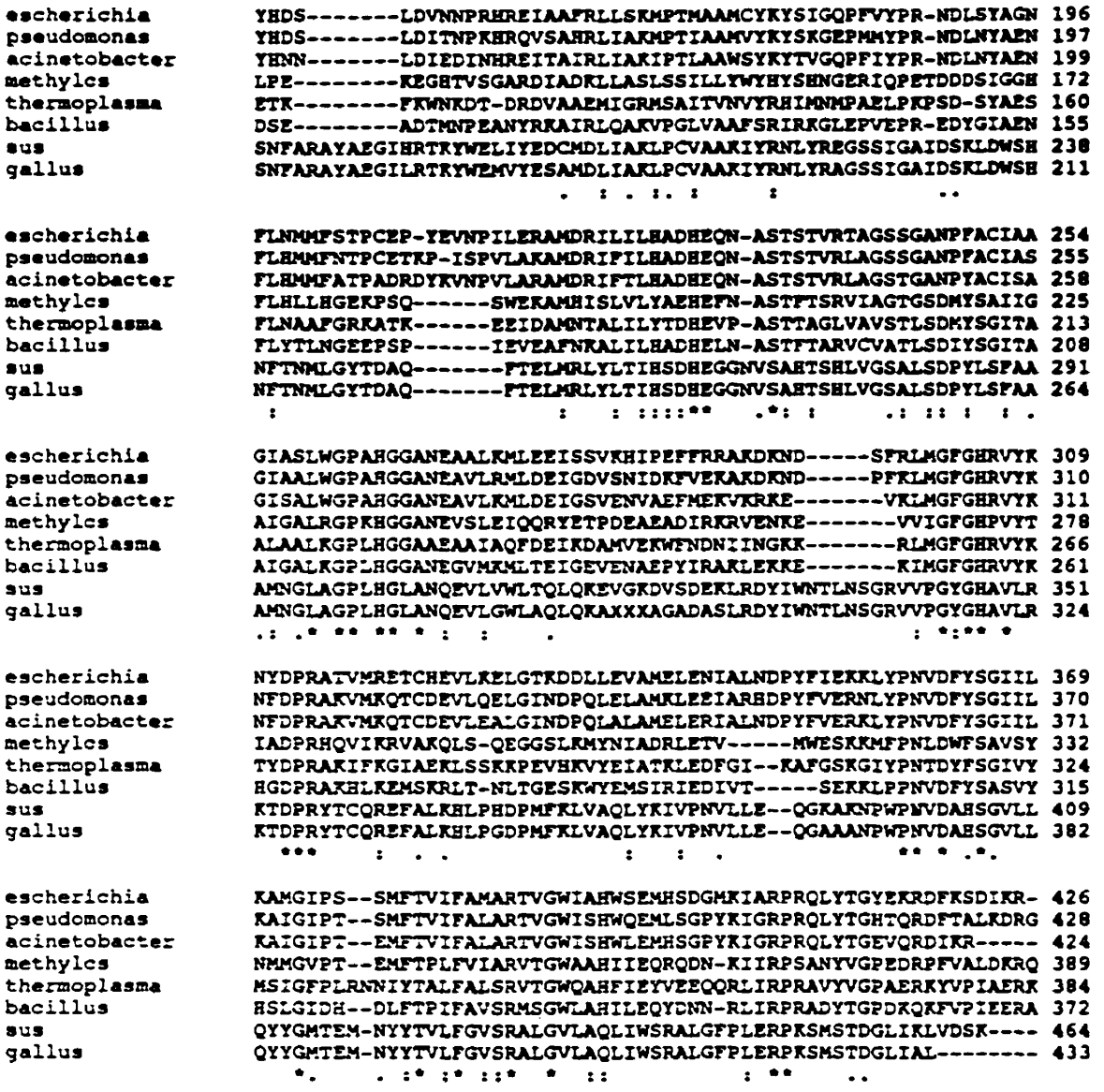
*Alignment constructed using ClustalW (default settings) (Thompson et al., 1994)*

* *indicates complete conservation of the residue at that position*
: *indicates good conservation of the residue*
. *indicates fair conservation of the residue*

```
escherichia         ---------------------------ADTKAKLTLNGDTAVELDVLKGTLGQDVIDIRTL 34
pseudomonas         ---------------------------MADKKAQLIIEGSAPVELPVLSGTMGPDVVDVRGL 35
acinetobacter       -----------------------MSEATGKKAVLHLDGKEIELPIYSGTLGPDVIDVKDV 37
methylcs            ---------------------------------------------------MSDTTILQNS 10
thermoplasma        ------------------------------------------------------------
bacillus            ------------------------------------------------------------
sus                 MALLTAAARLFGAKNASCLVLAARHASASSTNLKDILADLIPKEQARIKTFRQQHGNTVV 60
gallus              ---------------------------ASSTNLKDVLAALIPKEQARIKTFRQQHGGTAL 33
                                                                              .

escherichia         GSKGVFTFDPGFTSTASCESKITFIDGDEGILLHRGFPIDQLATD---------SNYLEV 85
pseudomonas         TATGHFTFDPGFMSTASCESKITYIDGDKGVLLHRGYPIEQLAEK---------SDYLET 86
acinetobacter       LASGHFTFDPGFMATASCESKITFIDGDKGILLHRGYPIDQLATQ---------ADYLET 88
methylcs            THVIKPKKSVALSGVPAGNTALCTVGKSGNDLHYRGYDILDLAKH---------CEFEEV 61
thermoplasma        -MPETEEISKGLEDVNIKWTRLTTIDGNKGILRYGGYSVEDIIASG---------AQDEEI 51
bacillus            -----MTATRGLEGVVATTSSVSSIIDDT--LTYVGYDIDDLTEN---------ASFEEI 44
sus                 GQITVDMMYGGMRGMKGLVYETSVLDPDEG-IRFRGYSIPECQKMLPKAKGGEEPLPEGL 119
gallus              GQITVDMSYGGMRGMKGLVYETSVLDPDEG-IRFRGFSIPECQKLLPKGGKGGEPLPEGL 92
                              .:          :  .   : . *: : :

escherichia         CYILLNGEKPTQEQYDEFKTTVTRHTMIHEQITRLFHAFRRD-SHPMAVMCGITGALAAF 144
pseudomonas         CYLLLNGELPTAAQKEQFVGTIKNHTMVHEQLKTFFNGFRRD-AHPMAVMCGVIGALSAF 145
acinetobacter       CYLLLNGELPTAEQKVEFDAKVRAHTMVHDQVSRFFNGFRRD-AHPMAIMVGVVGALSAF 147
methylcs            AHLLIHGKLPTRDELAAYKTKLKALRGLPANVRTVLEALPAA-SHPMDVMRTGVSALGCT 120
thermoplasma        QYLFLYGNLPTEQELRKYKETVQKGYKIPDFVINAIRQLPRE-SDAVAMQMAAVAAMAAS 110
bacillus            IYLLWHLRLPNKKELEELKQQLAKEAAVPQEIIEHFKSYSLENVHPMAALRTAISLLGLL 104
sus                 FWLLVTGQIPTEEQVSWLSKEWAKRAALPSHVVTMLDNFPTN-LHPMSQLSAAITALNSE 178
gallus              FWLLVTGQIPTGAQVSWLSKEWAKRAALPSHVVTMLDNFPTN-LHPMSQLSAAITALNSE 151
                      ::   . *.  :              :   :   :        ..:         :
```


escherichia YHDS-------LDVNNPRHREIAAFRLLSKMPTMAAMCYKYSIGQPFVYPR-NDLSYAGN 196
pseudomonas YHDS-------LDITNPKHRQVSAHRLIAKMPTIAAMVYKYSKGEPMMYPR-NDLNYAEN 197
acinetobacter YHNN-------LDIEDINHREITAIRLIAKIPTLAAWSYKYTVGQPFIYPR-NDLNYAEN 199
methylcs LPE--------KEGHTVSGARDIADKLLASLSSILLYWYHYSHNGERIQPETDDDSIGGH 172
thermoplasma ETK--------FKWNKDT-DRDVAAEMIGRMSAITVNVYRHIMNMPAELPKPSD-SYAES 160
bacillus DSE--------ADTMNPEANYRKAIRLQAKVPGLVAAFSRIRKGLEPVEPR-EDYGIAEN 155
sus SNFARAYAEGIHRTKYWELIYEDCMDLIAKLPCVAAKIYRNLYREGSSIGAIDSKLDWSH 238
gallus SNFARAYAEGILRTKYWEMVYESAMDLIAKLPCVAAKIYRNLYRAGSSIGAIDSKLDWSH 211
escherichia FLNMMFSTPCEP-YEVNPILERAMDRILILHADHEQN-ASTSTVRTAGSSGANPFACIAA 254
pseudomonas FLHMMFNTPCETKP-ISPVLAKAMDRIFILHADHEQN-ASTSTVRLAGSSGANPFACIAS 255
acinetobacter FLHMMFATPADRDYKVNPVLARAMDRIFTLHADHEQN-ASTSTVRLAGSTGANPYACISA 258
methylcs FLHLLHGEKPSQ------SWEKAMHISLVLYAEHEFN-ASTFTSRVIAGTGSDMYSAIIG 225
thermoplasma FLNAAFGRKATK------EEIDAMNTALILYTDHEVP-ASTTAGLVAVSTLSDMYSGITA 213
bacillus FLYTLNGEEPSP------IEVEAFNKALILHADHELN-ASTFTARVCVATLSDIYSGITA 208
sus NFTNMLGYTDAQ-------FTELMRLYLTIHSDHEGGNVSAHTSHLVGSALSDPYLSFAA 291
gallus NFTNMLGYTDAQ-------FTELMRLYLTIHSDHEGGNVSAHTSHLVGSALSDPYLSFAA 264
escherichia GIASLWGPAHGGANEAALKMLEEISSVKHIPEFFRRAKDKND-----SFRLMGFGHRVYK 309
pseudomonas GIAALWGPAHGGANEAVLRMLDEIGDVSNIDKFVEKAKDKND-----PFKLMGFGHRVYK 310
acinetobacter GISALWGPAHGGANEAVLKMLDEIGSVENVAEFMEKVKRKE-------VKLMGFGHRVYK 311
methylcs AIGALRGPKHGGANEVSLEIQQRYETPDEAEADIRKRVENKE-------VVIGFGHPVYT 278
thermoplasma ALAALKGPLHGGAAEAAIAQFDEIKDAMVEKWFNDNIINGKK-------RLMGFGHRVYK 266
bacillus AIGALKGPLHGGANEGVMKMLTEIGEVENAEPYIRAKLEKKE-------KIMGFGHRVYK 261
sus AMNGLAGPLHGLANQEVLVWLTQLQKEVGKDVSDEKLRDYIWNTLNSGRVVPGYGHAVLR 351
gallus AMNGLAGPLHGLANQEVLGWLAQLQKAXXXAGADASLRDYIWNTLNSGRVVPGYGHAVLR 324
escherichia NYDPRATVMRETCHEVLKELGTKDDLLEVAMELENIALNDPYFIEKKLYPNVDFYSGIIL 369
pseudomonas NFDPRAKVMKQTCDEVLQELGINDPQLELAMKLEEIARHDPYFVERNLYPNVDFYSGIIL 370
acinetobacter NFDPRAKVMKQTCDEVLEALGINDPQLALAMELERIALNDPYFVERKLYPNVDFYSGIIL 371
methylcs IADPRHQVIKRVAKQLS-QEGGSLKMYNIADRLETV-----MWESKKMFPNLDWFSAVSY 332
thermoplasma TYDPRAKIFKGIAEKLSSKKPEVHKVYEIATKLEDFGI--KAFGSKGIYPNTDYFSGIVY 324
bacillus HGDPRAKHLKEMSKRLT-NLTGESKWYEMSIRIEDIVT-----SEKKLPPNVDFYSASVY 315
sus KTDPRYTCQREFALKHLPHDPMFKLVAQLYKIVPNVLLE--QGKAKNPWPNVDAHSGVLL 409
gallus KTDPRYTCQREFALKHLPGDPMFKLVAQLYKIVPNVLLE--QGAAANPWPNVDAHSGVLL 382
escherichia KAMGIPS--SMFTVIFAMARTVGWIAHWSEMHSDGMKIARPRQLYTGYEKRDFKSDIKR- 426
pseudomonas KAIGIPT--SMFTVIFALARTVGWISHWQEMLSGPYKIGRPRQLYTGHTQRDFTALKDRG 428
acinetobacter KAIGIPT--EMFTVIFALARTVGWISHWLEMHSGPYKIGRPRQLYTGEVQRDIKR----- 424
methylcs NMMGVPT--EMFTPLFVIARVTGWAAHIIEQRQDN-KIIRPSANYVGPEDRPFVALDKRQ 389
thermoplasma MSIGFPLRNNIYTALFALSRVTGWQAHFIEYVEEQQRLIRPRAVYVGPAERKYVPIAERK 384
bacillus HSLGIDH--DLFTPIFAVSRMSGWLAHILEQYDNN-RLIRPRADYTGPDKQKFVPIEERA 372
sus QYYGMTEM-NYYTVLFGVSRALGVLAQLIWSRALGFPLERPKSMSTDGLIKLVDSK---- 464
gallus QYYGMTEM-NYYTVLFGVSRALGVLAQLIWSRALGFPLERPKSMSTDGLIAL-------- 433

*Appendix 2*

Compilation of all complete CS tryptic fragments, and those fragments observed in mass spectra (MALDI TOF and ESI TOF) of unfractionated CS digests. Masses given are theoretical (calculated), and monoisotopic.

| Fragment | Sequence | Mass | Fragment | Sequence | Mass |
|---|---|---|---|---|---|
| T36 | 356 to 356 | 147.11 | T18 | 178 to 188 | 1326.68 |
| T12 | 126 to 126 | 175.12 | T28, T29 | 295 to 306 | 1436.69 |
| T45 | 426 to 426 | 175.12 | T14-16 | 156 to 167 | 1440.84 |
| T25 | 290 to 290 | 175.12 | T33, T34 | 320 to 332 | 1486.76 |
| T42 | 418 to 418 | 175.12 | T32, T33 | 315 to 327 | 1516.76 |
| T2 | 5 to 6 | 218.15 | T41-43 | 410 to 421 | 1547.79 |
| T26 | 291 to 292 | 218.15 | T16, T17 | 164 to 177 | 1587.78 |
| T27 | 293 to 294 | 262.14 | T40, T41 | 405 to 417 | 1594.87 |
| T14 | 156 to 157 | 312.18 | T7 | 56 to 69 | 1598.85 |
| T25, T26 | 290 to 292 | 374.25 | T3 | 7 to 21 | 1601.86 |
| T43 | 419 to 421 | 409.21 | T37 | 357 to 370 | 1641.89 |
| T30 | 307 to 309 | 409.24 | T4, T5 | 22 to 37 | 1672.92 |
| T1 | 1 to 4 | 434.22 | T36, T37 | 356 to 370 | 1769.98 |
| T16 | 164 to 167 | 460.31 | T2, T3 | 5 to 21 | 1800.99 |
| T26, T27 | 291 to 294 | 461.27 | T9, T10 | 105 to 119 | 1823.96 |
| T44 | 422 to 425 | 462.26 | T28-30 | 295 to 309 | 1826.91 |
| T20 | 218 to 221 | 492.22 | T38 | 371 to 387 | 1829.91 |
| T5 | 33 to 37 | 505.3 | T29-31 | 300 to 314 | 1852.93 |
| T34 | 328 to 332 | 547.31 | T6 | 38 to 55 | 1880.84 |
| T42, T43 | 418 to 421 | 565.31 | T23, T24 | 274 to 289 | 1914.02 |
| T32 | 315 to 319 | 577.31 | T39 | 388 to 404 | 1971.86 |
| T9 | 105 to 109 | 577.33 | T41-44 | 410 to 425 | 1991.02 |
| T40 | 405 to 409 | 612.39 | T21 | 222 to 239 | 2005.05 |
| T44, T45 | 422 to 426 | 618.36 | T10, T11 | 110 to 125 | 2037.06 |
| T1, T2 | 1 to 6 | 633.36 | T23-25 | 274 to 290 | 2070.12 |
| T28 | 295 to 299 | 638.29 | T27-30 | 293 to 309 | 2070.03 |
| T31 | 310 to 314 | 664.31 | T41-45 | 410 to 426 | 2147.12 |
| T15 | 158 to 163 | 706.39 | T31-33 | 310 to 327 | 2162.05 |
| T11 | 120 to 125 | 790.44 | T10-12 | 110 to 126 | 2193.16 |
| T29 | 300 to 306 | 817.41 | T1-3 | 1 to 21 | 2216.2 |
| T43, T44 | 419 to 425 | 852.45 | T5, T6 | 33 to 55 | 2367.12 |
| T27, T28 | 293 to 299 | 881.41 | T25-30 | 290 to 309 | 2425.27 |
| T24 | 283 to 289 | 897.49 | T17, T18 | 168 to 188 | 2454.15 |
| T11, T12 | 120 to 126 | 946.54 | T20, T21 | 218 to 239 | 2478.25 |
| T33 | 320 to 327 | 958.47 | T30-33 | 307 to 327 | 2552.28 |
| T14, T15 | 156 to 163 | 999.55 | T39, T40 | 388 to 409 | 2565.26 |
| T41 | 410 to 417 | 1001.49 | T35 | 333 to 355 | 2694.32 |
| T42-44 | 418 to 425 | 1008.55 | T27-31 | 293 to 314 | 2715.32 |
| T23 | 274 to 282 | 1035.54 | T3, T4 | 7 to 32 | 2769.49 |
| T24, T25 | 283 to 290 | 1053.6 | T35, T36 | 333 to 356 | 2822.42 |
| T30, T31 | 307 to 314 | 1054.53 | T16-18 | 164 to 188 | 2895.44 |
| T26-28 | 291 to 299 | 1080.54 | T2-4 | 5 to 32 | 2968.62 |
| T17 | 168 to 177 | 1146.48 | T25-31 | 290 to 314 | 3070.55 |
| T15, T16 | 158 to 167 | 1147.68 | T13 | 127 to 155 | 3102.42 |
| T41, T42 | 410 to 418 | 1157.6 | T22 | 240 to 273 | 3138.54 |
| T4 | 22 to 32 | 1186.64 | T34, T35 | 328 to 355 | 3222.61 |
| T29, T30 | 300 to 309 | 1207.64 | T12, T13 | 126 to 155 | 3258.52 |
| T31, T32 | 310 to 319 | 1222.6 | T29-33 | 300 to 327 | 3350.67 |
| T10 | 110 to 119 | 1265.64 | T34-36 | 328 to 356 | 3350.71 |

| Fragment | Sequence | Mass |
|---|---|---|
| T19 | 189 to 217 | 3364.53 |
| T13, T14 | 127 to 157 | 3395.58 |
| T37, T38 | 357 to 387 | 3452.78 |
| T6, T7 | 38 to 69 | 3460.67 |
| T36-38 | 356 to 387 | 3580.88 |
| T38, T39 | 371 to 404 | 3782.78 |
| T19, T20 | 189 to 221 | 3837.74 |
| T28-33 | 295 to 327 | 3969.94 |
| T8 | 70 to 104 | 4080.94 |
| T39-43 | 388 to 421 | 4094.03 |
| T22, T23 | 240 to 282 | 4155.06 |
| T23-30 | 274 to 309 | 4320.26 |
| T24-32 | 283 to 319 | 4507.32 |
| T8, T9 | 70 to 109 | 4639.25 |
| T18, T19 | 178 to 217 | 4672.2 |
| T32-36 | 315 to 356 | 4848.45 |
| T26-34 | 291 to 332 | 4940.49 |
| T23-31 | 274 to 314 | 4965.55 |
| T21, T22 | 222 to 273 | 5124.57 |
| T22-26 | 240 to 292 | 5388.77 |
| T37-39 | 357 to 404 | 5405.65 |
| T31-36 | 310 to 356 | 5493.74 |
| T38-42 | 371 to 418 | 5514.73 |
| T36-39 | 356 to 404 | 5533.74 |
| T7, T8 | 56 to 104 | 5660.78 |
| T19-21 | 189 to 239 | 5823.77 |
| T7-9 | 56 to 109 | 6219.09 |
| T17-20 | 168 to 221 | 6272.87 |
| T23-33 | 274 to 327 | 6463.3 |
| T20-23 | 218 to 282 | 6614.3 |

# *References*


Anderson, D.H. (1988) Ph.D. Thesis, University of Manitoba.

Anderson, D.H., Donald, L.J., Jacob, M.V., & Duckworth, H.W. (1991) *Biochem. Cell Biol. 69*, 232-238.

Anderson, D.H., & Duckworth, H.W. (1988) *J. Biol. Chem. 263*, 2163-2169.

Ayed, A. (1998) Ph.D. Thesis, University of Manitoba.

Ayed, A., Krutchinsky, A.N., Ens, W., Standing, K.G., & Duckworth, H.W. (1998) *Rapid Commun. Mass Spectrom. 12*, 339-344.

Ayed, A., & Duckworth, H.W. (1999) *Protein Science 8*, 1116-1126.

Bayer, E., Bauer, B., & Eggerer, H. (1981) *FEBS Letters 123*, 258-260.

Bayley, H., & Knowles, J.R. (1977) *Methods in Enzymology 46*, 69-114.

Bayley, H., & Staros, J.V. (1984) in Azides and Nitrenes – Reactivity and Utility. (Scriven, E.F.V., Ed.) Academic Press, 434-490.

Bell, A.W., Bhayana, V., & Duckworth, H.W. (1983) *Biochemistry 22*, 3400-3405.

Bhayana, V. (1984) Ph.D. Thesis, University of Manitoba.

Bhayana, V., & Duckworth, H.W. (1984) *Biochemistry 23*, 2900-2905.

Biesecker, G., & Wonacott, A.J. (1977) *Biochem. Soc. Trans. 5*, 647-652.

Birktoft, J.J., Fernley, R.T., Bradshaw, R.A., & Banaszak, L.J. (1982) *Proc. Nat. Acad. Sci. USA 79*, 6166-6170.

Bloxham, D.P., Ericsson, L.H., Titani, K., Walsh, K.A., & Neurath, H. (1980) *Biochemistry 19*, 3979-3985.

Bloxham, D.P., Parmelee, D.C., Kumar, S., Wade, R.D., Ericsson, L.H., Neurath, H., Walsh, K.A., & Titani, K. (1981) *Proc. Nat. Acad. Sci. USA 78*, 5381-5385.

Branchini, B.R., Murtiashaw, M.H., Eckman, E.A., Egan, L.A., Alfano, C.V., & Stroh, J.G. (1995) *Arch. Biochem. Biophys. 318*, 221-230.

Chavan, A.J. (1996) Photoaffinity Labeling: Methods and Applications; Volume I:

Isolation of Active Site Peptides. Research Products International Corp. USA.
Cho, H., Park, H., Zhang, X., Riba, I., Gaskell, S.J., Widger, W.R., & Kohn, H. (1997) *J. Org. Chem. 62*, 5432-5440.

Chowdhry, V., & Westheimer, F.H. (1979) *Ann. Rev. Biochem. 48*, 293-325.

Czarnecki, J., Geahlen, R., & Haley, B.E. (1979) *Methods in Enzymology 56*, 643-653.

Czarnecki, J.J. (1984) *Biochim. Biophys. Acta 800*, 41-51.

Danson, M.J., & Weitzman, P.D.J. (1973) *Biochem. J. 135*, 513-524.

Danson, M.J., Harford, S., & Weitzman, P.D.J. (1979) *Eur. J. Biochem. 101*, 515-521.

Davies, G.E., & Stark, G.R. (1970) *Proc. Nat. Acad. Sci. USA 66*, 651-656.

Donald, L.J., Crane, B.R., Anderson, D.H., & Duckworth, H.W. (1991) *J. Biol. Chem. 266*, 20709-20713.

Donald, L.J., Hosfield, D.J., Cuvelier, S., & Duckworth, H.W. (1998) *41st Annual Meeting of the Canadian Federation of Biological Societies*, Edmonton, AB.

Donald, L.J., Loboda, A.V., Potier, N., Swint-Kruse, L., Ens, W., Matthews, K.S., Duckworth, H.W., & Standing, K.G. (1998b) *41st Annual Meeting of the Canadian Federation of Biological Societies*, Edmonton, AB.

Duckworth, H., Anderson, D., Bell, A., Donald, L., Chu, A., & Brayer, G. (1987) in Biochemical Society Symposia (Kay, J., & Weitzman, P., Eds.), 83-92.

Duckworth, H.W., & Bell, A.W. (1982) *Can. J. Biochem 60*, 1143-1147.

Duckworth, H.W., & Tong, E.K. (1976) *Biochemistry 15*, 108-114.

Eklund, H., Branden, C.I., & Jornvall, H. (1976) *J. Mol. Biol. 102*, 61-73.

Else, A.J., Danson, M.J., & Weitzman, P.D.J. (1988) *Biochem. J. 254*, 437-442.

Ens, W., Krutchinsky, A., Loboda, A., Lock, C., McComb, M., & Standing, K.G. (1999) *Proceedings of the 47th ASMS Conference on Mass Spectrometry and Allied Topics.*

Ens, W., & Standing, K.G. (1999) *Analytical Chemistry News and Features, 4*, 452A-461A.

Faloona, G.R., & Srere, P.A. (1969) *Biochemistry 8*, 4497-4503.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F, & Whitehouse, C.M. (1989) *Science 246*, 64-71.

Fleet, G.W.J., Porter, R.R., & Knowles, J.R. (1969) *Nature 224*, 511-512.

Gerhart, J.C., & Pardee, A.B. (1963) *Cold Spring Harbor Symp. Quant. Biol. 28*, 491.

Gerike, U., Hough, D.W., Russell, N.J., Dyall-Smith, M.L., & Danson, M. (1998) *Microbiology 144*, 929-935.

Gruic-Solvulj, I., Ludemann, H., Hillenkamp, F., Weygand-Durasevic, I., Kucan, Z., & Peter-Katalinic, J. (1997) *J. Biol. Chem. 272*, 32084-32091.

Guillory, R.J., & Jeng, S.J. (1983) *FASEB Proceedings 42*, 2826-2830.

Haley, B.E. (1977) *Methods in Enzymology 46*, 339-346.

Harford, S., & Weitzman, P.D.J. (1975) *Biochem. J. 151*, 455-458.

Harford, S., & Weitzman, P.D.J. (1978) *Biochem. Soc. Trans. 6*, 433-435.

Henneke, C.M., Danson, M.J., Hough, D.W., & Osguthorpe, D.J. (1989) *Protein Eng. 2*, 597-604.

Jangaard, N.O., Unkeless, J., & Atkinson, D.E. (1968) *Biochim. Biophys. Acta 151*, 225-235.

Jin, S., & Sonenshein, A.L. (1996) *J. Bacteriology 178*, 3658-3660.

Jungblut, P., & Thiede, B. (1997) *Mass Spec. Reviews, 16*, 145-162.

Karas, M., & Hillenkamp, F. (1988) *Anal. Chem. 60*, 2299-2301.

Karpusas, M., Holland, D., & Remington, S.J. (1991) *Biochemistry 30*, 6024-6031.

Kim, H., & Haley, B.E. (1990) *J. Biol. Chem. 265*, 3636-3641.

King, S.M., Kim, H., & Haley, B.E. (1991) *Methods in Enzymology 196*, 449-466.

Kroeger, M., & Wahl, R. (1997) *Nucleic Acids Res. 25*, 39-42.

Krutchinsky, A.N., Chernusevich, I.V., Spicer, V., Ens, W., & Standing, K.G. (1995) *Proceedings of the 43rd ASMS Conference on Mass Spectrometry and Allied Topics.* Atlanta, GA. 126.

Lehninger, A.L., Nelson, D.L., & Cox, M.M. (1993) Principles of Biochemistry, 2nd Edition. Worth Publishers, Inc, NY.

Liao, D.I., Karpusas, M., & Remington, S.J. (1991) *Biochemistry 30*, 6031-6036.

McComb, M.E., Donald, L.J., & Perreault, H. (1999) *Can. J. Chem 77*, 1752-1760.

McComb, M.E., Oleschuk, R.D., Manley, D.M., Donald, L., Chow, A., O'Neil, J.D.J., Ens, W., Standing, K.G., & Perreault, H. (1997) *Rapid Commun. Mass Spec. 11*, 1716-1722.

Mitchell, C.G. (1996) *Biochem. J. 313*, 769-774.

Mitchell, C.G., & Anderson, S.C.K. (1996) *Biochem. Soc. Trans. 24*, 46S.

Mitchell, C.G., Anderson, S.C.K., & El-Mansi, E.M.T. (1995) *Biochem. J. 309*, 507-511.

Mitchell, C.G., & Weitzman P.D.J. (1983) *J. Gen. Microbiol. 132*, 737-742.

Molgat, G.F., Donald, L.J., & Duckworth, H.W. (1992) *Arch. Biochem. Biophys. 298*, 238-246.

Monod, J., Wyman, J., & Changeux, J-P. (1965) *J. Mol. Biol. 12*, 88-118.

Ner, S.S., Bhayana, V., Bell, A.W., Giles, I., Duckworth, H.W., & Bloxham, D. (1983) *Biochemistry 22*, 5243-5248.

Olcott, M.C., Bradley, M.L., & Haley, B.E. (1994) *Biochemistry 33*, 11935-11941.

Orr, A., Ivanova, V.S., & Bonner, W.M. (1995) *Biotechniques 19*, 204-206.

Ouchterlony, O (1953) *Acta Path. Microbiol. Scand. 32*, 231-236.

Packman, L.C., & Berry, A. (1995) *Eur. J. Biochem. 227*, 510-515.

Patton, A.J., Hough, D.W., Towner, P., & Danson, M.J. (1993) *Eur. J. Biochem. 214*, 75-81.

Pereira, D.S., Donald, L.J., Hosfield, D.J., & Duckworth, H.W. (1994) *J. Biol. Chem. 269*, 412-417.

Potter, R.L., & Haley, B.E. (1983) *Methods in Enzymology, 91*, 613-633.

Remington, S., Wiegand, G., & Huber, R. (1982) *J. Mol. Biol. 158*, 111-152.

Robinson, C. (1996) in Protein and Peptide Analysis by Mass Spectrometry (Chapman, J.R., Ed.) Humana Press, N.J.

Rosenkrantz, M.T., Alam, K., Kim, K., Clark, B.J., Srere, P.A., & Guarente, L.P. (1986) *Mol. Cell. Biol. 6*, 4509-4515.

Rousselot, P., Mappus, E., Blachere, T., Rolland de Ravel, M., Grenot, C., Tonnelle, C., & Cuilleron, C.Y. (1997) *Biochemistry 36*, 7860-7868.

Rowe, A.J., & Weitzman, P.D.J. (1969) *J. Mol. Biol. 43*, 345-349.

Russell, R.J.M., Ferguson, J.M.C., Hough, D.W., Danson, M.J., & Taylor, G.L. (1997) *Biochemistry 36*, 9983-9994.

Russell, R.J.M., Hough, D.W., Danson, M.J., & Taylor, G.L. (1994) *Structure 2*, 1157-1167.

Russell, R.J.M., Gerike, U. Danson, M.J., Hough, D.W., & Taylor, G.L. (1998) *Structure 6*, 351-361.

Salvucci, M.E., Chavan, A.J., Klein, R.R., Rajagopolan, K., & Haley, B.E. (1994) *Biochemistry 33*, 14879-14886.

Shapiro, R.E., Specht, C.D., Collins, B.E., Woods, A.S., Cotter, R.J., & Schnaar, R.L. (1997) *J. Biol. Chem. 272*, 30380-30386.

Shevchenko, A., Chernushevich, I., Ens, W., Standing, K.G., Thomson, B., Wilm, M., & Mann, M. (1997) *Rapid Commun. Mass Spectr. 11*, 1015-1024.

Singh, A., Thornton, E.R., & Westheimer, F.H. (1962) *J. Biol. Chem. 237*, PC3006-PC3008.

Srere, P.A., Brazil, H., & Gonen, L. (1963) *Acta Chem. Scand. 17*, S129-S134.

Sternberg, M.J.E., & Taylor, W.R. (1984) *FEBS Lett. 175*, 387-392.

Talgoy, M.M., Bell, A.W., & Duckworth, H.W. (1979) *Can. J. Biochem. 57*, 822-833.

Talgoy, M.M., & Duckworth, H.W. (1979) *Can. J. Biochem. 57*, 385-395.

Tang, X., Beavis, R.C., Ens, W., Lafortune, F., Schueler, B., & Standing, K.G. (1988) *Int. J. Mass Spec. Ion Proc. 85*, 43.

Textor, S., Wendisch, V.F., DeGraaf, A.A., Muller, U., Linder, M.I., Linder, D., & Buckel, W. (1997) *Arch. Microbiol. 168*, 428-436.

Tong, E.K., & Duckworth, H.W. (1975) *Biochemistry 14*, 235-241.

Verenchikov, A.N., Ens, W., & Standing, K.G. (1994) *Analytical Chemistry 66*, 126-133.

Weber, P.J.A., & Beck-Sickinger, A.G. (1997) *Journal of Peptide Research 49*, 375-383.

Weitzman, P.D.J. (1966a) *Biochim. Biophys. Acta 128*, 213-215.

Weitzman, P.D.J. (1966b) *Biochem. J. 101*, 44-45c.

Weitzman, P.D.J. (1967) *Biochim. Biophys. Acta 139*, 526-528.

Weitzman, PDJ, & Jones, D. (1968) *Nature 219*, 270-272.

Weitzman, P.D.J., Ward, B.A., & Rann, D.L. (1974) *FEBS Lett. 43*, 97-100.

Weitzman, P.D.J., Kinghorn, H.A., Beecroft, L.J., & Harford, S. (1978) *Biochem. Soc. Trans. 6*, 436-438.

Wiegand, G., Kukla, D., Scholze, H., Jones, A., & Huber, R. (1979) *Eur. J. Biochem. 93*, 41-50.

Wiegand, G., Remington, S., Deisenhofer, J., & Huber, R. (1984) *J. Mol. Biol. 174*, 205-219.

Wierenga, R.K, & Hol, W.G.J. (1983) *Nature 302*, 842-844.

Willard, H.H., Merritt, L.L. Jr., Dean, J.A., & Settle, F.A., Jr. (1988) Instrumental Methods of Analysis, 7th Edition. Wadsworth Publishing Company, Ca.

Williamson, M.P. (1994) *Biochem. J. 297*, 249-260.

Wilm, M.S., & Mann, M. (1994) *Int. J. Mass Spectrom. Ion Proc., 136*, 167-180.

Wood, D.O., Williamson, L.R., Winkler, H.H. & Krause, D.C. (1987) *J. Bacteriol. 169*, 3564-3572.

Wright, J.A., Maeba, P., & Sanwal, B.D. (1967) *Biochem. Biophys. Res. Commun. 29*, 34-38.

Zhi, W., Srere, P.A., & Evans, C.T. (1991) *Biochemistry 30*, 9281-9286.