This is to certify that the

dissertation entitled

Location and Structure of the Substrate Hexose Binding Site of Rat Brain Hexokinase

presented by

Douglas M. Schirch

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Biochemistry

John E. Wilson
Major professor

Date November 11, 1986

LOCATION AND STRUCTURE OF THE SUBSTRATE HEXOSE BINDING SITE OF RAT BRAIN HEXOKINASE

by

Douglas M. Schirch

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry

1986

ABSTRACT

LOCATION AND STRUCTURE OF THE SUBSTRATE HEXOSE BINDING SITE OF RAT BRAIN HEXOKINASE

by

Douglas M. Schirch

A glucose analog, N-(bromoacetyl)-D-glucosamine (GlcNBrAc), previously used to label the glucose binding sites of rat muscle Type II and bovine brain Type I hexokinases, also inactivates rat brain hexokinase (ATP: D-hexose 6-phosphotransferase, EC 2.7.1.1) with pseudo-first order kinetics. Several methods of kinetic analysis indicate that the affinity label binds reversibly to the glucose binding site prior to inactivating the enzyme. However, nonspecific (i.e., without prior complex formation) inactivation also occurs, and equations are derived to describe this behavior. Inactivation is dependent on the deprotonation of a residue with an alkaline pK_a consistent with the modified residue being a sulfhydryl group as reported to be the case with the bovine brain enzyme.

The affinity label modifies three residues at indistinguishable rates. Amino acid analysis of the labeled protein indicates that each of these residues is a cysteine. However, only one of these cysteines appears to be critical

for activity. Peptide mapping techniques have permitted localization of the critical residue, and thus the glucose binding site, in a 40 kDa domain at the C-terminus of the enzyme. This is the same domain recently shown to include the ATP binding site. Thus, catalytic function is assigned to the C-terminal domain of rat brain hexokinase.

Peptide mappping of hexokinase modified by radiolabeled affinity label yields three labeled peptides, two of which are protected by inclusion of Glc or GlcNAc during reaction with GlcNBrAc. The sequences of these two peptides show considerable homology to portions of the sequences of yeast hexokinase isozymes A and B. One of these peptides is homologous to a region of yeast hexokinase which is clearly in a position to be labeled by the bromoacetyl arm on the 2-carbon of GlcNBrAc. An essential serine residue in yeast hexokinase, implicated in glucose binding, is also conserved in rat brain hexokinase. This provides strong evidence that the 40 kDa region of rat brain hexokinase Type I shares a common ancestry with yeast hexokinase.

This is dedicated to my parents,
for giving me the freedom and encouragement
to pursue my own goals,
and for being such good parents that
their goals have become my goals.

ACKNOWLEDGMENTS

Above all I would like to thank John Wilson for serving as my advisor, as well as being patient with my graduate school plans when they changed from time to time. Asking him to serve as my advisor is probably the only decision I made in graduate school that I never second guessed. Much thanks also go to Al Smith, who, either as a technician or a graduate student, taught me almost every technique I mention in this dissertation. For the major portion of my time in the lab he also filled my insatiable appetite for hexokinase, for which he reminded me more than once. I am indebted to Dave Schwab also, for writing computer programs for me when I needed them, as well as teaching me a thing or two about pool. A final thanks to Paul Polakis, Ken Finney, Mohsen Nemat-Gorgani, Tracey White, Guochen Xie, Annette Thelen and Laura Chiarantini for making room 301 the most enjoyable lab in the building.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DTNB	5,5-dithiobis-(2-nitrobenzoate)
GalNBrAc	N-(bromoacetyl)-D-galactosamine
GlcNAc	N-acetylglucosamine
GlcBrAc	N-(bromoacetyl)-D-glucosamine
GlcNHexBr	N-(N-bromoacetyl)-6-amino-hexanoyl-glucosamine
HPLC	high performance liquid chromatography
IAM	iodoacetamide
PAGE	polyacrylamide gel electrophoresis
PMSF	phenylmethylsulfonyl flouride
PTC	phenyl thiocarbamyl derivatives of amino acids
PTH	phenylthiohydantoin derivatives of amino acids
SDS	sodium dodecyl sulfate
TFA	trifluoroacetic acid
TNB	thionitrobenzoate anion
tm	trypsin modified
TPCK	L-1-tosylamide-2-phenylethyl chloromethyl ketone

INTRODUCTION

Hexokinase catalyzes the conversion of Glc and Mg-ATP to Glc-6-P and Mg-ADP. This phosphorylation of Glc represents the first step in a number of metabolic pathways, including glycolysis and the hexose monophosphate shunt. In mammalian brain tissue, where Glc represents the principal energy substrate (1), hexokinase assumes a major role in metabolic regulation (2).

Rat Type I hexokinase (ATP:D-hexose 6-phosphotransferase, EC 2.7.1.1), the predominant isozyme in mammalian brain, is composed of a single polypeptide chain with M_r 100,000 (3). Studies with limited tryptic digestion of the enzyme suggest that the polypeptide chain is folded into three major structural domains (4). The approximate molecular masses of these domains are, from the N-terminus, 10 kDa, 50 kDa, and 40 kDa. Investigations of other enzymes have shown that discrete structural domains can often be associated with specific functions (e.g.,5,6), and so it is with rat brain hexokinase. Thus, this enzyme specifically and reversibly binds to the outer mitochondrial membrane (7), and the mitochondrial binding function has been localized to the 10 kDa N-terminal domain (8). More recently, the binding site for the nucleotide substrate,

ATP, has been shown to be contained within the 40 kDa C-terminal domain (9).

A glucose analog affinity label, N-(bromoacetyl)-D-glucosamine (GlcNBrAc), has been used to alkylate a sulfhydryl at the glucose binding sites of rat muscle Type II hexokinase (10) and the Type I isozyme from bovine brain (11); in both cases, derivatization with GlcNBrAc caused a stoichiometric irreversible loss of enzyme activity. A similar alkylating reagent, GalNBrAc, was also reported to function as an affinity label for yeast hexokinase isozyme B, again reacting with a critical sulfhydryl (12,13). In the present work, GlcNBrAc has been used to label the Glc binding site of rat brain hexokinase. Though three sulfhydryls were modified, only one appeared to be critical for activity. Subsequent peptide mapping of the enzyme, derivatized with radiolabeled GlcNBrAc, permitted localization of the glucose binding site in the C-terminal 40 kDa domain previously shown to include the ATP binding site (9). This conjunction of the binding sites for both substrates in the 40 kDa domain makes it reasonable to assign catalytic function to the C-terminal domain of Type I hexokinase from brain.

In the present work, three tryptic peptides derived from the rat brain enzyme and containing the GlcNBrAc modified sulfhydryls have been isolated and partially sequenced. Two of these peptides are homologous to regions of the yeast hexokinase sequence which, based on X-ray

crystallographic studies (14-16), are located near the glucose binding site of this enzyme. The demonstration of homology in a functional region of these two hexokinases supports earlier suggestions (17-23) that mammalian hexokinase isozymes I-III (M_r 100,000) may have evolved by duplication and fusion of an ancestral gene similar to that coding for yeast, wheat germ, and mammalian Type IV hexokinases (M_r 50,000).

The homology in the active site regions of these hexokinases provides a reasonable basis for suggesting that similarities in the secondary and tertiary structure may also exist, and it will be shown that the properties of Type I mammalian hexokinase can be reasonably interpreted in terms of the X-ray structure of the yeast enzyme.

LITERATURE REVIEW

Yeast Hexokinase.

Originally it was believed that there were three hexokinase isozymes in yeast, designated A, B and C in order of elution from a DEAE-cellulose column (24). It was later determined that hexokinase C is a native enzyme which, without proper precautions, is modified to Type B during purification (25). The two isozymes are now referred to as A and B. Some authors refer to isozymes A and B as PI and PII, respectively (26).

Hexokinase A is composed of two identical alpha subunits. Likewise, two beta subunits comprise the B isozyme (27). The molecular weight of each subunit is approximately 50 kDa. The subunits are dissociated by increased Glc concentrations, ionic strength, pH or temperature (26). Some data had suggested that the enzyme is active in the dimer form (28,29). However, more recent evidence from measurements of the molecular weight of isozyme B under reacting conditions (30) demonstrates that the active enzyme is monomeric.

Studies of isozyme B mutants indicate that it is involved in Glc repression (31), a system regulating the repression of various enzymes involved in gluconeogenesis and other catabolic pathways. The mechanism for involvement of the B isozyme in Glc repression is not yet understood, although mutants have been prepared in which the catalytic

activity of the enzyme is unaffected, but the ability to regulate Glc repression has been lost (31).

The structures of the A and B isozymes have been refined from X-ray crystallography to resolutions of 3.0 Å (32) and 2.5 Å (14), respectively. A high degree of structural similarities between the two isozymes is evident (33). The predominant structural feature of yeast hexokinases is a cleft which separates the enzyme into two lobes (34), a property shared with several other kinases (35). The locations of the Glc (15,16) and ATP (33) binding sites inside the cleft has also been determined. In the B isozyme several residues have been identified as essential for binding Glc (15,16). The ATP binding site has been less accurately elucidated (33). The structure of the dimer has also been determined, showing the subunits in a nonsymmetrical arrangement (36).

The amino acid sequences for both isozymes, as determined from the nucleotide sequence of the cloned cDNA, has recently been completed (37,38). Both isozymes contain 485 amino acids, of which 76% are homologous, a figure lower than isozymes for many other glycolytic enzymes (38). These nonhomologous positions are located in discrete clusters of the polypeptide chain (37,38). A refined three dimensional structure of yeast hexokinase, integrating the recently determined sequence data with the X-ray crystallographic structures, is in progress (37).

Mammalian Hexokinases.

Four different hexokinase isozymes are found in mammalian tissues (39,40). The isozymes are designated I-IV in order of increased electrophoretic mobility toward the anode of a starch gel (41). Alternatively, the isozymes are also referred to as A-D (39), corresponding to the same order as I-IV.

There are several properties which make isozymes I-III distinct from the Type IV isozyme. The K_G for Glc for Types I-III is much lower (10^{-4} to 10^{-6} M) than that observed with Type IV (10^{-2} M) (39-41). Each of the isozymes is composed of a single polypeptide chain, but the molecular weight of Type IV, approximately 50 kDa, is half that of Types I-III (reviewed in (7)). The isozymes also differ in their tissue distribution (7). Type I is the predominant isozyme in the brain, erythrocytes, kidney, lung, spleen and heart tissues. Skeletal muscles contain predominantly the Type II isozyme, while Type III hexokinase exists only in minor amounts in several tissues. Type IV is the primary hexokinase in the liver.

Amino acid composition comparisons between hexokinases from yeast and mammalian sources indicate it is likely that they share a common evolutionary origin and are structurally similar (21,42). Recent demonstrations of antigenic cross reactivities between all four isozymes in mammals (42,43) also indicate structural homologies.

Hexokinase isozymes I-III have four major functions: catalysis, regulation of catalysis by Glc-6-P, reversible binding to the outer mitochondrial membrane, and regulation of mitochondrial binding by Glc-6-P. Although several aspects of the mechanism of catalysis are unknown, a ternary complex between the enzyme, hexose and Mg-ATP is generally agreed upon (7). Glc-6-P is also a potent inhibitor of activity (17). This inhibition is believed to be mediated through an allosteric regulatory site (44), although some authors do not rule out the possibility that Glc-6-P exerts its influence by competition between its phosphate and the gamma-phosphate of ATP for the same binding site (45).

Reversible binding to the outer mitochondrial membrane, presumably at porin (46), the pore-forming protein, gives hexokinase preferential access to ATP generated in the mitochondria (47,48). Glc-6-P solubilizes the enzyme from the mitochondria (49), which decreases the affinity of the enzyme for ATP (50,51) and increases its affinity for Glc-6-P (51). Thus, high levels of Glc-6-P induce a less active enzyme species, as well as directly inhibiting the enzyme. This is believed to be a major mechanism of hexokinase regulation in vivo (52). Because the characteristics of Glc-6-P inhibition and the ability of Glc-6-P to solubilize the enzyme from the mitochondria are similar in many respects (7), a single Glc-6-P binding site for both functions seems likely.

Hexokinase can exist in several discrete conformations which are induced by the binding of various ligands (53-55). These conformations have been roughly compared by determining the protection afforded by each conformation against inactivation of the enzyme by chymotrypsin, glutaraldehyde, heat and DTNB. These comparisons indicate that the suitability of carbohydrates to serve as substrates or inhibitors depends on their ability to induce necessary conformational changes (53). For example, mannose induces the same conformation as glucose and it is also phosphorylated by the enzyme when Mg-ATP is present. GlcNAc, on the other hand, is neither phosphorylated nor able to induce the same conformation as Glc, although it does act as a competitive inhibitor of Glc. Likewise, for a hexose-6-phosphate to inhibit the enzyme it must not only bind to the Glc-6-P regulatory site, but also induce the necessary conformational change.

Postulated Evolution of Hexokinases.

Because mammalian hexokinase isozymes I-III, M_r 100 kDa, are twice the size of yeast hexokinase and mammalian Type IV hexokinase, it has often been suggested (17-23) that all hexokinases evolved from the same ancestral gene coding for a 50 kDa enzyme. Evolution of mammalian hexokinase isozymes I-III could have occurred after duplication and fusion of a hexokinase gene coding for a 50 kDa enzyme. This would lead to two catalytic sites in the "fused"

enzyme, one of which could have evolved into the Glc-6-P allosteric site evident in mammalian isozymes I-III, but not present in yeast hexokinase or mammalian isozyme IV.

Circumstantial evidence for this evolutionary relationship is provided by the amino acid composition comparisons discussed earlier (21,42). Analysis of hexokinases using a peptide mapping technique for peptides containing tyrosine also suggested structural homology between yeast hexokinase and rat hexokinase isozymes II and IV. The first direct evidence of homology between yeast and mammalian hexokinases was provided recently by Marcus and Ureta (56), who determined the amino acid sequence of seven peptides from rat Type III hexokinase, several of which showed significant homology to yeast hexokinase.

Gene fusion is believed to have occurred in the evolutionary development of several other proteins. A series of fusions between genes coding for smaller, distinct enzymes is believed to have led to the variety of fatty acid synthases (57) apparent in several organisms, as suggested by the partial sequences of these enzymes. The entire amino acid sequence for yeast DNA topoisomerase II has been determined and subsequently its N- and C-terminal halves shown to be homologous to the B and A subunits of bacterial gyrase (58).

Similar to the proposed evolution for hexokinases, the amino acid sequence of rabbit muscle phosphorylase, compared with that of a phosphorylase in Escherichia coli, suggests

that the rabbit enzyme arose after gene duplication and fusion, with one of the phosphoryl binding sites developing into an allosteric regulatory site (59). Yet another example of this type of protein evolution is found in rabbit muscle phosphofructokinase (60), which has two halves that are both highly homologous to the smaller phosphofructokinase in Bacillus stearothermophilus. This homology has allowed a three-dimensional model of the rabbit enzyme to be constructed from detailed information of the prokaryotic enzyme. Like hexokinase, the mammalian enzyme is regulated by effectors which do not alter the activity of the prokaryotic enzyme. It has been possible to predict the locations of these effector sites in the rabbit enzyme by identifying amino acid substitutions at key positions between the rabbit and the Bacillus sequences.

Sulfhydryl Groups in Hexokinases.

Reaction of yeast hexokinase B with the alkylating reagent IAM indicated the presence of four sulfydryls (61), which has since been confirmed by sequencing of the cloned cDNA (37,38). Of the four sulfhydryls, two react rapidly with IAM, only above 30° C, and are accompanied by a loss of enzyme activity (61). Glc gave strong protection against inactivation, while Mg-ATP gave less significant protection.

Reaction with GalNBrAc, a galactose derived alkylating reagent, labeled the same two sulfhydryls labeled by IAM (12,13). In the presence of Glc, one sulfhydryl, described

as critical for activity, could be blocked from reaction with GlNBrAc and enzyme activity was retained. GlcNBrAc, structurally more similar to the enzymes's natural substrate that GalNBrAc, was not an affinity label.

The critical sulfhydryl of the B isozyme was identified at a position 20% of the polypeptide length from the C-terminus. Anderson et al. (15) identified two cysteines in this region by X-ray crystallography, although in the three dimensional structure of the protein both cysteines were located about 20 Å from the Glc binding site. From the X-ray data it was proposed (15) that a cysteine at position 243, of a sequence tentatively defined by X-ray crystallography (14), was in a position to react with GalNBrAc and may be critical for enzyme activity. However, from the recently published amino acid sequences for yeast hexokinase, as determined from the cDNA (37,38), it is now apparent that the amino acid at this position is actually an aspartate. The X-ray sequence (14) and the cDNA sequences (37,38) both show a cysteine at the adjacent position 244 (of the X-ray sequence), but the sulfur atom of this cysteine is located more than 10 Å from the bound Glc molecule (15). Therefore, the location of the critical cysteine in yeast hexokinase remains unclear.

Several mammalian hexokinases are also inactivated by sulfhydryl reagents. DTNB inactivates Type I hexokinases from rat (62) and bovine brains (63,64), as well as rat hexokinase Type IV (65,66). Glc and Glc-6-P protect either

the rat brain or bovine brain hexokinases from inactivation. Glc, but not Glc-6-P, blocks the inactivation of the rat Type IV hexokinase (65). The ability of Glc to protect these hexokinases from inactivation by DTNB suggests a conserved sulfhydryl at their Glc binding sites.

Fourteen sulfhydryl groups are present in rat brain hexokinase, all of which are reactive toward DTNB in the protein's native conformation (62). Saturating levels of Glc, Glc-6-P, ATP and P_i each protect a different number of otherwise reactive sulfhydryls. Because the protection afforded by some ligands is substantial (Glc-6-P protects 10 sulfhydryls), it is highly unlikely that the ligands are physically covering the protected cysteines. The protective effects are probably mediated by the conformational changes induced by the ligands. The reaction rates of sulfhydryls with DTNB fall into three distinct categories, which are designated as fast (F), intermediate (I) or slow (S). For enzyme devoid of any protecting ligand, the breakdown of sulfhydryls into the F, I and S categories is 2, 3 and 9, respectively, The rate constant for inactivation is the same as the rate constant for the reaction of the I-type sulfhydryls (62), indicating that it is one of the I-type sulfhydryls (67) which is critical for activity.

Similar sulfhydryl reactivities with DTNB are seen in bovine brain hexokinase (63,64). Of the enzyme's 11-13 sulfhydryls, 10-11 are reactive in the unfolded protein. The reactivities of the sulfhydryls were classified into

four categories. Only one sulfhydryl was implicated as crytical for activity. Using stoichiometric amounts of DTNB it was also possible to show that the critical sulfhydryl could be oxidized to a second, vicinal sulfhydryl, forming an intramolecular disulfide bridge. Two such pairs of vicinal sulfhydryls were demonstrated (64).

Sulfhydryl oxidation has also been linked to inactivation of rabbit red blood cell hexokinase (68,69) and hexokinase II in ascites tumor cells (70). Magnani et al. (68,69) demonstrated that the oxidation of sulfhydryls in rabbit red blood cell hexokinase may regulate enzyme activity in vivo.

GlcNBrAc has been sucessfully used as an affinity label for bovine brain (11) and rat muscle Type II hexokinases (10). A single residue, most likely a sulfhydryl, is labeled in the muscle enzyme (10), while two sulfhydryls are labeled by GlcNBrAc in bovine brain hexokinase (11). The presence of Glc protected both enzymes from inactivation. In the bovine brain enzyme Glc protected only the sulfhydryl which was critical for activity, while the reactivity of the second sulfhydryl was unaffected. GlcNBrAc, although it binds to the rat Type IV hexokinase, does not inactivate the enzyme (71).

A similar Glc derived affinity label, GlcNBrHex, with an extended alkylating "arm", did inactivate the rat Type IV isozyme (71). Two groups, one presumably a carboxyl and the other undetermined, were labeled. Glc selectively protected

against the reaction of the critical residue, assumed to be a carboxyl. GlcNBrHex is also an affinity label for rat Type II hexokinase (71). However, two groups, suggested to be a sulfhydryl and a carboxyl, appeared to be critical. They were the only two residues labeled.

A table summarizing the reactions of hexose derived alkylating reagents with several hexokinases is included in the discussion section of Chapter 1. Also included are the results of GlcNBrAc inactivation of rat brain hexokinase.

An ATP analog affinity label, 6-mercapto-9- -D-ribofuranosylpurine 5'-triphosphate, has also been used to inactivate bovine brain hexokinase (72). Although it appeared that an enzyme sulfhydryl was labeled, no oxidized enzyme-affinity label product could be demonstrated. Glc, and to a slightly lesser extent ATP, protected the enzyme against inactivation. It was proposed that the affinity label catalyzed the oxidation of one of the same pairs of vicinal sulfhydryls, containing the critical residue, which was oxidized by DTNB.

Sulfhydryl Groups at Enzyme Active Sites.

Critical or "essential" sulfhydryls at binding sites of enzymes are commonly reported for many enzymes, as is indicated by a sampling of such reports during the first half of 1986 alone (73-76). It would seem that the presence of sulfhydryls at enzyme binding sites is a common occurrence in a wide variety of enzymes. Not as common, but

still a frequent occurrence, is the identification of vicinal sulfhydryls at the binding sites of various enzymes (77-82).

On the basis of their proximity to enzyme binding sites, sulfhydryls are often designated as "essential", implying that they play a direct role in substrate binding or catalysis. However, the location of a sulfhydryl at an active site should not be considered proof that the sulfhydryl is essential, as there have been cases were active site sulfhydryls, previously called "essential", have been modified with small blocking groups which leave enzyme activity intact (77,83). It seems likely that the labeling of sulfhydryls at the active sites of many other proteins may cause inactivation by introducing a steric bulk which the catalytic site cannot accommodate, rather than by chemically blocking the activity of an essential residue. Therefore, residues mentioned in this text whose chemical modification leads to enzyme inactivation will be termed "critical", unless there is evidence that the residue participates directly in catalysis and can be termed "essential".

Some enzymes, such as papain or thiotransferases, have active site sulfhydryls whose roles are unambiguous. However, it would seem that the body of enzymes reported to have sulfhydryls at their active sites is much larger than the number of enzymes for which the roles of the sulfhydryls are understood. Given the susceptability of sulfhydryls to

oxidation, it would seem that their presence at enzyme active sites would be detrimental to enzyme activity. This would be especially true for enzymes possessing vicinal sulfhydryls at their catalytic sites. That sulfhydryls do, nevertheless, often occur at catalytic sites suggests that there may be a common reason, not yet understood, for their presence. However, the evidence of functional roles for these sulfhydryls is limited. It has been proposed (84) that the oxidation or reduction of the sulfhydryls may regulate the activity of some enzymes, such as phosphofructokinase (84) and 1,6-diphosphatase (85), both of which are key enzymes of glycolysis and gluconeogenesis. Sulfhydryl oxidation has also been proposed to regulate the activity of other enzymes (86-88).

MATERIALS AND METHODS

Materials

TPCK-treated trypsin and D(+)-glucosamine hydrochloride were purchased from Sigma Chemical Co. (St. Louis, MO). Staphylococcal aureus V8 protease was obtained from Miles Research Laboratories (Elkhart, IN). Bromoacetic anhydride was a product of Pfaltz and Bauer (Stanford, CT). D-Glucosamine hydrochloride [1-^{14}C] and iodoacetamide [1-^{14}C] were purchased from ICN Pharmaceuticals, Inc. (Irvine, CA). Safety-Solve liquid scintillation cocktail was a product of Research Products International Corp. (Mount Prospect, IL). Centricon 30 microconcentrators were obtained from Amicon Corp. (Danvers, MA). Molecular weight standards used for SDS-PAGE were from Bio-Rad Laboratories (Rockville Centre, NY). Ultra pure urea was a product of Schwarz/Mann (Cambridge, MA). HPLC grade ammonium acetate and Scinti Verse LC liquid scintillation cocktail were obtained from Fisher Scientific Co. (Fairlawn, NJ), and DC-Plastikfolien cellulose plates were purchased from American Scientific Products (Romulus, MI). Model 1750 sample cups were a product of Isco, Inc. (Lincoln, NE).

Methods

Purification of rat brain hexokinase.

Rat brain hexokinase was prepared according to Polakis and Wilson (4).

Protein and hexokinase assays.

A molar extinction coefficient of 5.1 x 10^4 cm^{-1} M^{-1} at 280 nm was used to determine hexokinase concentrations (3). Hexokinase activity was coupled to Glc-6-P dehydrogenase activity and measured spectrophotometrically as described by Polakis and Wilson (89). Trypsin solutions were freshly prepared in 1 mM HCl and the concentration determined from absorbance at 280 nm based on an absorbance of 1.43 for a 1 mg/ml solution (90).

Synthesis of N-(bromoacetyl)-D-glucosamine.

GlcNBrAc was prepared from D-glucosamine and bromoacetic anhydride as described by Otieno *et al*. (12). Analysis by thin layer chromatography on silica gel G, with methanol-acetone (1:10, v/v) as solvent, gave a single symmetrical component at R_f 0.5. The melting point was 141-145° C with decomposition. [^{14}C]-GlcNBrAc was prepared from [1-^{14}C]-D-glucosamine and had a specific activity of 0.53 mCi/mmole.

Reaction of GlcNBrAc with hexokinase.

Purified hexokinase was chromatographed on a Sephadex G-25 column equilibrated with 50 mM Tris-HCl, 0.5 mM EDTA, pH 8.85, freeing the enzyme from Glc, P_i, and thioglycerol

that were used to maintain enzyme activity during storage. When different buffers were used, this is indicated in the text. Prior to column equilibration, the buffer was degassed and purged with N_2 to remove oxygen. After trypsin modification of native hexokinase (see below), a similarly equilibrated Sephadex G-75 column was used for removal of the protective agents, noted above; this also effected the removal of trypsin, which was retarded by the G-75 column. When necessary, samples were concentrated using a Centricon 30 microconcentrator.

Inactivations were done at room temperature. GlcNBrAc was used at 1mM unless noted otherwise. Other ligands were used, where indicated, at the following concentrations: 10mM Glc, 10 mM GlcNAc, 1.0 mM Glc-6-P, 5.0 mM ATP and 5.0 mM Mg-ATP; these concentrations are at least 10-fold higher than the K_d values estimated for these ligands (7). When other ligands were used, their concentration is indicated in the text. Control samples (no GlcNBrAc) lost no more than 5% of the original activity during the time course of the experiments. Reactions were quenched either by addition of 10 mM 2-mercaptoethanol or by adding an aliquot of the reaction to a larger volume of reaction buffer which included 10 mM 2-mercaptoethanol, followed by incubation for at least 15 min to ensure complete reaction of residual GlcNBrAc.

Measurement of incorporation of [^{14}C]-GlcNBrAc.

At various stages of inactivation, aliquots were removed from the reaction mixture and quenched prior to addition to bovine serum albumin-coated polypropylene centrifuge tubes (91) containing 0.5 ml cold 10% (w/v) trichloroacetic acid. Bovine serum albumin (50 ug) was added as a carrier, and the samples centrifuged for 5 minutes in a microcentrifuge. The resulting pellet was washed twice with 0.5 ml cold 10% trichloroacetic acid and redissolved in 0.3 ml formic acid. This solution was added to 5.0 ml Safety-Solve scintillant for scintillation spectrometry. Samples were counted in a Packard 300 CD scintillation counter using an external standard and a quench curve to determine efficiencies.

Trypsin modification of native hexokinase.

This was done as previously described (4). Hexokinase was incubated with trypsin in 0.1 M sodium phosphate, 0.1 M Glc, 0.01 M thioglycerol and 0.5 mM EDTA, pH 7.0, at room temperature. The protease:substrate ratio was 1:5 (w/w) unless noted otherwise. Proteolysis was terminated by addition of 1.0 mM PMSF, followed by incubation for at least 15 minutes at room temperature. This trypsin modified native hexokinase is referred to as "tm-hexokinase" throughout this paper.

Trypsin modification of [^{14}C]-GlcNBrAc labeled hexokinase.

After reaction with GlcNBrAc, hexokinase samples were diluted 40 fold with the same pH 7.0 buffer used for trypsin

modification of native hexokinase. Samples were concentrated in a Centricon 30 microconcentrator to 50-100 ul. Digestion by trypsin was done as described above.

[^{14}C]-IAM labeling of denatured tm-Hexokinase.

Tm-hexokinase in 50 mM Tris, 0.5 mM EDTA, 1% SDS, pH 8.0, was incubated for 30 min with 5 mM [^{14}C]-IAM. The reaction was quenched by addition of 10 mM 2-mercaptoethanol. The sample was subsequently run on an SDS-PAGE gel as described below.

SDS-PAGE and fluorography.

Both one and two dimensional separations were performed as described by Polakis and Wilson (4). Gels to be fluorographed were soaked in water for 30 minutes, followed by 30 minutes in 1.0 M sodium salicylate prior to drying the gels (92). Densitometric scanning of stained gels or fluorographs was done using a Gelman ACD-18 densitometer, which also integrated peak areas.

Molecular weight standards (see Fig. 9) were lysozyme (14,400), soybean trypsin inhibitor (21,500), carbonic anhydrase (31,000), ovalbumin (45,000), bovine serum albumin (66,200) and phosphorylase B (92,500). Based on these standards, the M_r of native hexokinase is approximately 115,000, somewhat larger than the value of 98,000 repeatedly found in previous work from this laboratory (e.g. 3,4,93) using both SDS-PAGE as well as other methods. As noted in the Introduction, and reviewed in (7), many other investigators have also reported molecular weights of about

100,000 for mammalian hexokinases. We have no explanation for this recently encountered discrepancy. However, for the sake of continuity, the molecular mass values for the 57 kDa, 67 kDa, and 33 kDa bands in Fig. 10, lane D, were calculated using the bands in lane C as standards with molecular weights as previously reported (4).

Immunoblotting.

The previously published method (4) was followed with the exception that the secondary antibody was used at a 1:3,000 dilution. Also, the blots were developed in the dark with a solution prepared immediately before use by adding 10 ml of ice-cold 0.3% (w/v) 4-chloro-1-naphthol in methanol to 50 ml of 0.015% (v/v) H_2O_2 in 20 mM Tris-HCl, 0.5 M NaCl, pH 7.5. The peroxidase reaction was quenched after 30-45 min by washing the filter in 0.02% (w/v) sodium azide.

Nonlinear regression analysis.

Data were analyzed by a Gauss-Newton algorithm (94) using a computer program written and kindly made available by Dr. S. Brooks (Biochemistry Dept., Michigan State University).

Reaction of DTNB with hexokinase.

Hexokinase was reacted with equimolar amounts of DTNB in 50 mM Tris, 0.5 mM EDTA, pH 8.5. DTNB stock solutions were prepared in the same buffer and their concentration determined by reacting an aliquot of the solution with an excess of 2-mercaptoethanol in the pH 8.5 buffer. The

concentration of the TNB anion was measured using an extinction coefficient of 13600 M^{-1} cm^{-1} at an absorbance of 412 nm (95). Hexokinase oxidized by DTNB was assayed as described above, with the exception that thioglycerol was excluded from the assay solution.

Reduction and carboxymethylation of [^{14}C]-GlcNBrAc labeled hexokinase.

This was performed essentially as described by Crestfield et al. (96). Additional 2-mercaptoethanol was added to the quenched reaction solution to provide a final concentration that was 100-fold in excess of the hexokinase sulfhydryl content. Solid urea and a stock solution containing Tris-HCl and EDTA were added to yield final concentrations of 8.0 M urea, 0.2 M Tris-HCl, 5.0 mM EDTA, pH 8.5. After incubation for 30 minutes at 37^{o} C, iodoacetamide was added in an amount stoichiometrically equivalent to the 2-mercaptoethanol, and incubation continued for an additional 15 min.

Exhaustive tryptic digestion of [^{14}C]-GlcNBrAc-labeled hexokinase.

The [^{14}C]-GlcNBrAc-labeled and carboxymethylated enzyme was extensively dialyzed against 1% ammonium bicarbonate, pH 7.8. Precipitated protein was redissolved by addition of solid urea to a final concentration of 8.0 M. The solution was diluted with a three fold excess of 1% ammonium bicarbonate and 10 mM glycylglycine, pH 7.8; the protein did not reprecipitate after dilution to 2.0 M urea. Trypsin was

added immediately at a trypsin to hexokinase ratio of 1:5 (w/w), and the sample incubated with stirring at 37° C for 45 minutes. Proteolysis was terminated by adding two drops of glacial acetic acid for each ml of digestion volume.

Cyanate formed spontaneously in urea solutions (97) can modify the side chains of several amino acids, including lysine (98). Therefore, glycylglycine was added as a scavenger of cyanate (98), and a high trypsin:hexokinase ratio was used to keep the digestion time, and therefore possible cyanate modification, to a minimum.

Primary HPLC of tryptic peptides.

The digestion mixture was dried on a rotary evaporater with careful attention to prevent sudden degassing of CO_2. The dried residue was redissolved with 0.1% TFA to a final urea concentration between 4.0 and 8.0 M. The solution was centrifuged for 5 minutes at 38,000 x g prior to HPLC. The peptides were separated on an Altex Ultrasphere-ODS reverse phase column (4.6 mm I.D. x 25 cm), monitoring absorbance at 214 nm with a Beckman Model 160 detector. Solvent A was 0.10% TFA in water and solvent B was 0.09% TFA in acetonitrile. The column was equilibrated in 5% solvent B and the peptides were eluted with gradients as indicated in Fig. 1 at a flow rate of 1.0 ml/min. Radioactivity in the effluent, after exiting the absorbance monitor, was detected by a Flo-One Model HS radioactive flow detector (Radiomatic Instruments and Chemical Co., Inc., Tampa, FL) using Scinti

Verse LC liquid scintillation cocktail at a flow rate of 3.0 ml/min.

Secondary HPLC of [^{14}C]-GlcNBrAc-derivatized peptides.

Peptides were collected in silanized polypropylene centrifuge tubes immediately after exiting the UV detector, and lyophilized. Peptides were further purified by HPLC with the following changes from the above procedure. Solvent A was 5.0 mM ammonium acetate in water, pH 6.0, and solvent B was acetonitrile. Peptides I, II and III (see Fig. 1) were redissolved in 5%, 15% and 30% solvent B in solvent A, respectively, and loaded onto a column equilibrated with the same solvent mixture. An Altex Ultrasphere-I.P. reverse phase column (4.6 mm I.D. x 25 cm) was used for isolation of Peptides I and II, while the Altex Ultrasphere ODS column was used for Peptide III. Peptides were eluted with gradients as indicated in Fig. 2.

Peptide mapping of the [^{14}C]-GlcNBrAc labeled 40 kDa domain.

Tm-hexokinase (0.24 mg) was labeled with [^{14}C]-GlcNBrAc as described above. For subsequent tryptic digestion and separation of the peptides by HPLC it was necessary to reduce and carboxymethylate the protein at this point, also as described above, prior to separating the fragments by SDS-PAGE. The carboxymethylated protein was layered across the top of a SDS-polyacrylamide gel for preparative separation of the protein fragments. After electrophoreisis the protein bands were visualized by immersing the gel in cold 0.2 M KCl (99). The 40 kDa band was excised. A

representative vertical slice of the gel was stained in the usual manner to verify that the 40 kDa band had been correctly identified. The excised 40 kDa band was electroeluted from the gel slice using a Model 1750 sample cup. The buffer chambers were filled with the same Tris-glycine buffer used for SDS-PAGE, as referenced above. Electroelution was performed for four hrs at three watts. The 40 kDa protein was lyophilized and redissolved in 75 ul of water. SDS was removed by an ion pairing technique (100) using acetone:triethylamine:acetic acid (90:5:5). The precipitated protein was redissolved with 8.0 M urea and an aliquot counted for ^{14}C content to determine the amount of hexokinase present for subsequent digestion with trypsin. After dialysis against 1% ammonium carbonate, the protein was digested by trypsin in 2.0 M urea and the peptides separated by HPLC as before.

Amino acid analyses and peptide sequencing.

Twenty-four hour hydrolysis and analysis of PTC-derivatized amino acids were performed by Dr. Y. M. Lee of the Macromolecular Structure Facility, Michigan State University, East Lansing, MI. A similar hydrolysis and amino acid analysis of underivatized amino acids was performed by Mr. Alan Smith of the Protein Structure Laboratory, University of California, Davis. Automated Edman degradation of peptides were performed by both facilities. The facility which performed a given analysis

or sequence determination are as indicated in the appropriate tables.

Mass spectrometry of peptides.

The molecular weight of Peptide III was estimated by fast atom bombardment mass spectrometry. Mass spectral data were obtained from the Michigan State University Mass Spectrometry Facility supported by a grant (RR-00480) from the Biotechnology Resources Branch, Division of Research Resources, NIH. Peptide III, 2-3 nmoles, was dissolved in a 5:1 dithiothreitol:dithioerythritol matrix. We are grateful to Dr. John Stults for assistance in this part of the work.

Preparation of N-carboxymethyllysine.

Polylysine was alkylated by iodoacetate as described by Gundlach et al. (101) with the exception that the pH was maintained manually at 10.0 rather than with a pH stat. The reaction time was extended from 3.0 to 3.5 hours. The carboxymethylated polylysine was hydrolyzed for 24 hours and derivatized to the PTC amino acid prior to amino acid analysis. A single novel peak, corresponding to N-carboxymethyllysine, eluted at a position between the PTC derivatives of glutamate and S-carboxymethylcysteine.

Preparation of S-[^{14}C]-GlcNAc-cysteine.

[^{14}C]-GlcNBrAc (50 nmoles) and cysteine (100 nmoles) in 15 ul of 66 mM Tris-HCl, pH 8.5, was allowed to react at room temperature for 2 hours. The reaction product was purified by thin layer chromatography on DC-Plastikfolien cellulose plates with ethanol-water (7:3, v/v) as solvent.

[^{14}C]-GlcNBrAc had an R_f of 0.6, as determined by scanning the plate for radioactivity, while cysteine was detected by ninhydrin at an R_f of 0.4. A single ninhydrin positive and radioactive peak was detected in the reaction mixture at R_f 0.2. A scan of the radioactivity on the plate indicated a 90% conversion of [^{14}C]-GlcNBrAc to S-[^{14}C]-GlcNAc-cysteine. The product was eluted from the the plate with 200 ul of water for a final yield of 60%.

Preparation of S-[^{14}C]-GlcNAc-glutathione and S-[^{14}C]-acetamido-glutathione.

Forty nmoles of glutathione were reacted with either 20 nmoles of [^{14}C]-IAM or [^{14}C]-GlcNBrAc. Both reagents were 0.53 uCi/umole. Reactions were in 20 ul of 20 mM Tris, pH 8.5 for a duration of 20 min. The reaction mixtures were separated by HPLC using an Altex Ultrasphere-Octyl reverse phase column (4.6 mm I.D. x 15 cm). The column was equilibrated with 0.10% TFA in water at a flow rate of 1.0 ml/min. The reactants, glutathione, iodoacetamide and GlcNBrAc, eluted 5.5, 4.5 and 3.0 min after injection, respectively. Tris was not retained by the column and eluted at 2.0 min. Oxidized glutathione did not elute from the column unless an acetonitrile gradient was used. The reaction products of glutathione with [^{14}C]-iodoacetamide or [^{14}C]-GlcNBrAc eluted at 4.5 and 3.5 min, respectively. Both products were hydrolyzed, derivatized to the PTC amino acid and subjected to amino acid analysis as described above.

CHAPTER I

LOCATION OF THE GLUCOSE BINDING SITE

GlcNBrAc was evaluated as an affinity label for rat brain hexokinase by several methods of kinetic analysis. The results were consistent with the label binding reversibly to the substrate Glc site prior to inactivating the enzyme. The number and general locations of sulfhydryls labeled by GlcNBrAc were also determined.

RESULTS

GlcNBrAc Inactivation of Hexokinase in the Presence of Various Ligands.

Reaction of rat brain hexokinase with GlcNBrAc causes pseudo-first order loss of enzyme activity (Fig. 1). Saturating amounts of Glc, Glc-6-P and GlcNAc give differing degrees of protection against inactivation, while ATP and MgATP slightly increase the rate of inactivation (Fig. 1).

It was expected that Glc would give excellent protection against inactivation by GlcNBrAc. It is unclear why Glc-6-P, which is believed to bind to an allosteric regulatory site and not the Glc substrate site, would also give excellent protection. Yet protection by Glc-6-P against GlcNBrAc inactivation was also seen with the bovine Type I (11) and the rat Type II (10) hexokinases. It is unlikely, based on studies of substrate and inhibitor specificities (53,44,102) that GlcNBrAc would bind to the Glc-6-P regulatory site. It is also very unlikely that Glc-

Figure 1.

Protection by ligands against inactivation of hexokinase by GlcNBrAc. Activity was monitored as a function of time after addition of 1 mM GlcNBrAc, and with no added ligand or with addition of the various ligands indicated beside the corresponding lines. Concentrations of added ligands are given in Methods.

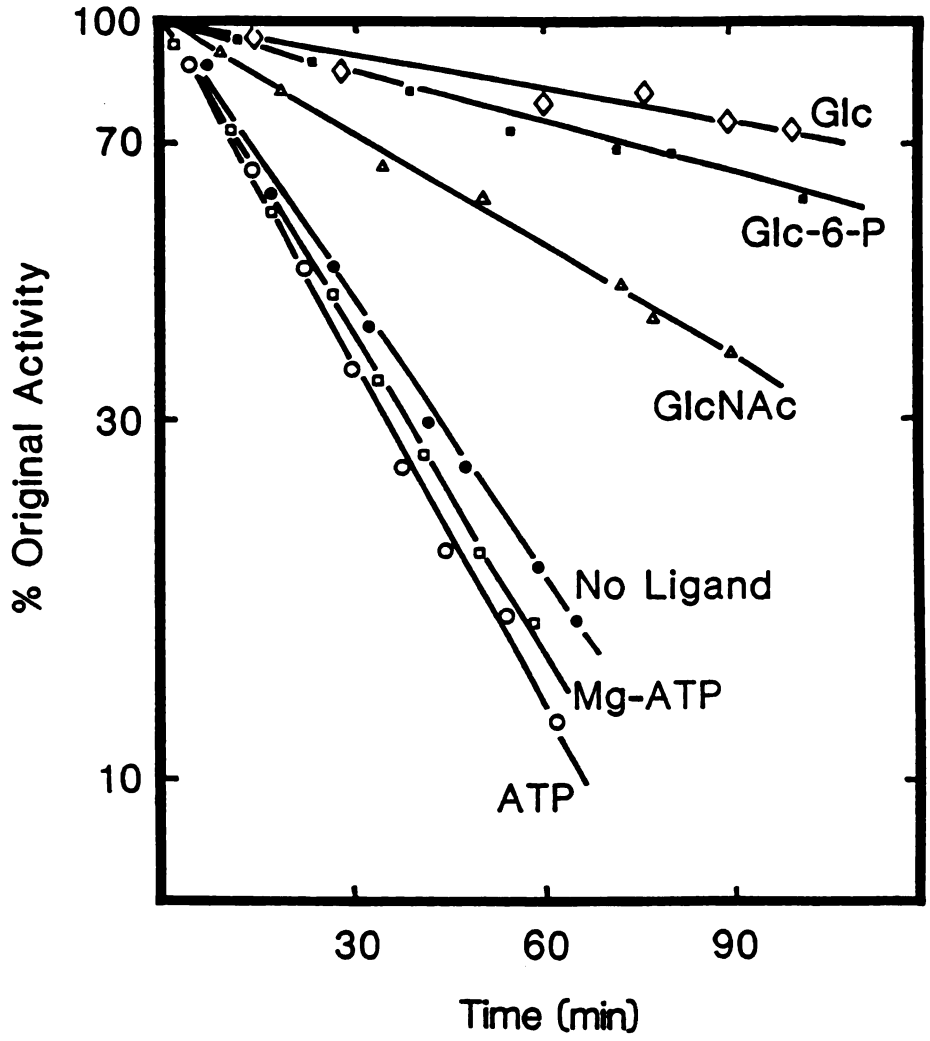

100
70
30
10
% Original Activity
Glc
Glc-6-P
GlcNAc
No Ligand
Mg-ATP
ATP
30
60
90
Time (min)

Figure 1

6-P competes with GlcNBrAc for the Glc substrate site, as Glc-6-P is a noncompetitive inhibitor vs. Glc (103). It is more likely that Glc-6-P protects against inactivation by virtue of the marked conformational change it induces (53,54) resulting in exclusion of ligands bearing bulky substituents at the 2-position from the hexose site. This would be consistent with previous results demonstrating synergistic binding of Glc-6-P and hexoses such as Glc, but no such effects with Glc-6-P and GlcNAc (53,54).

That Glc-6-P protects against inactivation by virtue of the conformational change it induces, rather than by binding at the site of GlcNBrAc mediated inactivation, was tested by using other hexose-6-phosphates in conjunction with GlcNBrAc. Three hexose-6-phosphates, Gal-6-P, Man-6-P and 2-deoxyglucose-6-P, which bind to the same allosteric site as Glc-6-P, but do not induce the same conformation or inhibit hexokinase activity, were tested for their ability to protect against GlcNBrAc inactivation. Also tested was 1,5-anhydroglucitol-6-P, which is similar to Glc-6-P in respect to inhibition and ability to induce the same conformational change. Thus, if Glc-6-P protection is mediatated through the conformational change it induces, and not through its binding to the enzyme, then 1,5-anhydroglucitol-6-P would be predicted to also protect against inactivation by GlcNBrAc, while Gal-6-P, Man-6-P and 2-deoxyglucose would have little or no protective effects. This was seen to be the case (Fig. 2). Therefore, it is

Figure 2.

Protection by various hexose-6-phosphates against inactivation by GlcNBrAc. This was performed as in Fig. 1, except that 50 mM glycylglycine, 0.5 mM EDTA, pH 8.5 was used as the buffer. Glc-6-P, 1,5-anhydroglucitol-6-P, and 2-deoxyglucose-6-P are indicated by the open circles, squares and triangles, respectively. The closed circles, squares and triangles represent no ligand used, mannose-6-P, and galactose-6-P. All ligand concentrations were 1 mM, as used previously in this laboratory (53).

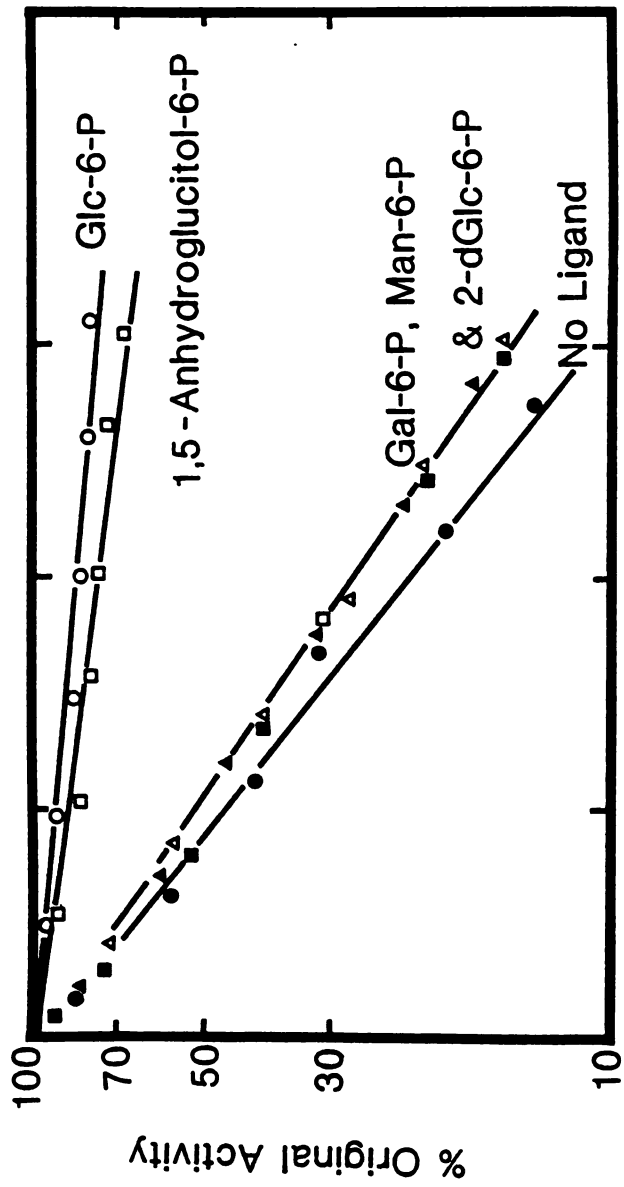

100
70
50
30
10
% Original Activity
40
80
120
Time(min)
Glc-6-P
1,5-Anhydroglucitol-6-P
Gal-6-P, Man-6-P
& 2-dGlc-6-P
No Ligand

Figure 2

possible to conclude that GlcNBrAc and Glc-6-P do not compete for the same binding site on rat brain hexokinase.

A similar experiment was also conducted with hexoses other than Glc. Here it might be expected that hexoses which bind to hexokinase, regardless of whether they induce the same conformational change as Glc and can serve as substrates, would all give equal protection. However, this was not the case (Fig. 3). GlcNAc (also in Fig. 1) and fructose, which are competitive inhibitors vs. Glc, but do not induce the same conformational change, did not perform as well as Glc. While GlcNAc still provides a considerable degree of protection, the protection afforded by fructose is only slightly better than that given by Gal-6-P, Man-6-P and 2-deoxyglucose-6-P (Fig. 2). The concentration of GlcNAc used, 10 mM, was shown to be saturating by repeating the experiment with 100 mM GlcNAc, which did not increase protection against inactivation (results not shown). Glc and 2-deoxyglucose, which induce the same conformational change as Glc, give excellent protection against inactivation, as expected.

The experiment in Fig. 3 is therefore a poor demonstration that GlcNBrAc inactivation is mediated through the Glc substrate site. Strong evidence that the Glc site is indeed targeted by GlcNBrAc will be presented later in this chapter. The results from Fig. 3 do indicate, however, that the excellent protection afforded by Glc is due not only to competition with GlcNBrAc for the same binding site,

Figure 3.

Protection by various hexoses against inactivation by GlcNBrAc. This was performed as in Fig. 1, except that 50 mM glycylglycine, 0.5 mM EDTA, pH 8.5 was used as the buffer. Glc (10 mM), 2-deoxyglucose (100 mM) and fructose (100 mM) are indicated by the open circles, triangles and squares, respectively. The closed circles, triangles and squares represent no ligand used, GlcNAc (10 mM) and mannose (10). All ligand concentrations are at least 10 fold higher than the K_m values estimated for these ligands (102).

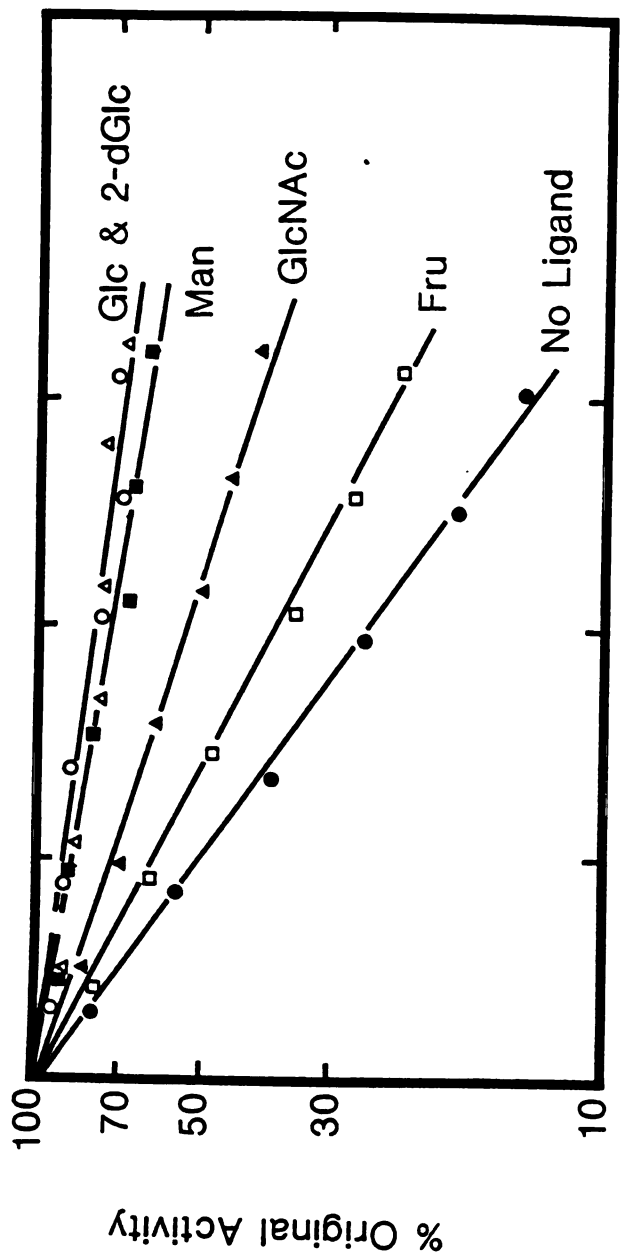

% Original Activity
100
70
50
30
10
40
80
120
Time (min)
Glc & 2-dGlc
Man
GlcNAc
Fru
No Ligand

Figure 3

but also because of the specific conformational change which Glc induces. An explanation of this behavior will become apparent later in this chapter.

Kinetics of GlcNBrAc Inactivation of hexokinase.

Classical affinity labels form reversible complexes with the enzyme prior to covalent modification and formation of an inactive enzyme species:

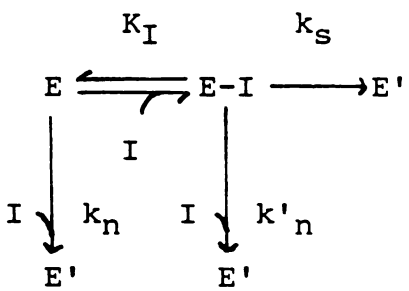
$$E + I \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E\text{-}I \xrightarrow{k_2} E' \qquad K_I = \frac{k_{-1}}{k_1}$$

Kitz and Wilson (104) have derived an equation which relates the pseudo-first order rate constant for inactivation, k_{app}, to other parameters. Equation 1 allows the determination of the values of K_I and k_2 when the inhibitor concentration is varied:

$$(1) \quad \frac{[I]}{k_{app}} = \frac{[I]}{k_2} + \frac{K_I}{k_2}$$

Inactivation dependent on formation of a reversible complex will result in a linear plot which intersects the abscissa at $-K_I$ when $[I]/k_{app}$ is plotted *vs.* [I]. In contrast, a horizontal line will be defined by data arising from purely nonspecific inactivation (i.e., in the case of a classical second order reaction proceeding without formation of a

complex). At low concentrations of GlcNBrAc, inactivation of hexokinase follows kinetics consistent with reversible complex formation, but at higher inhibitor concentrations it is obvious that nonspecific inactivation is also occurring (Fig. 4); this was also seen in the previous work with the bovine brain enzyme (11). An equation allowing for both specific and nonspecific inactivation was derived based on the following scheme:



The rate constants for "nonspecific" inactivation of free enzyme, E, and the enzyme-inhibitor complex, E-I, are k_n and k'_n, respectively, while k_s is the pseudo-first order rate constant for the "specific" inactivation caused by GlcNBrAc complexed with the enzyme. E' designates inactive GlcNBrAc-labeled enzyme, regardless of how that inactivation arises. Equation 2 defines the apparent pseudo-first order rate constant for inactivation of hexokinase proceeding as indicated in the above scheme:

$$(2) \quad k_{app} = \frac{k_n K_I [I] + k'_n [I]^2 + k_s [I]}{[I] + K_I}$$

Figure 4.

Inactivation at varying concentrations of GlcNBrAc. Pseudo-first order rate constants were determined from semilogarithmic plots of the type shown in Fig. 1. The line was determined by regression analysis of the data according to equation (3), given in the text. The values ±SD for K_I, k'_n and k_s determined for this set of data were 0.13 ±0.02 mM, 0.0081 ±0.0005 mM^{-1} min^{-1} and 0.027 ±0.002 min^{-1}.

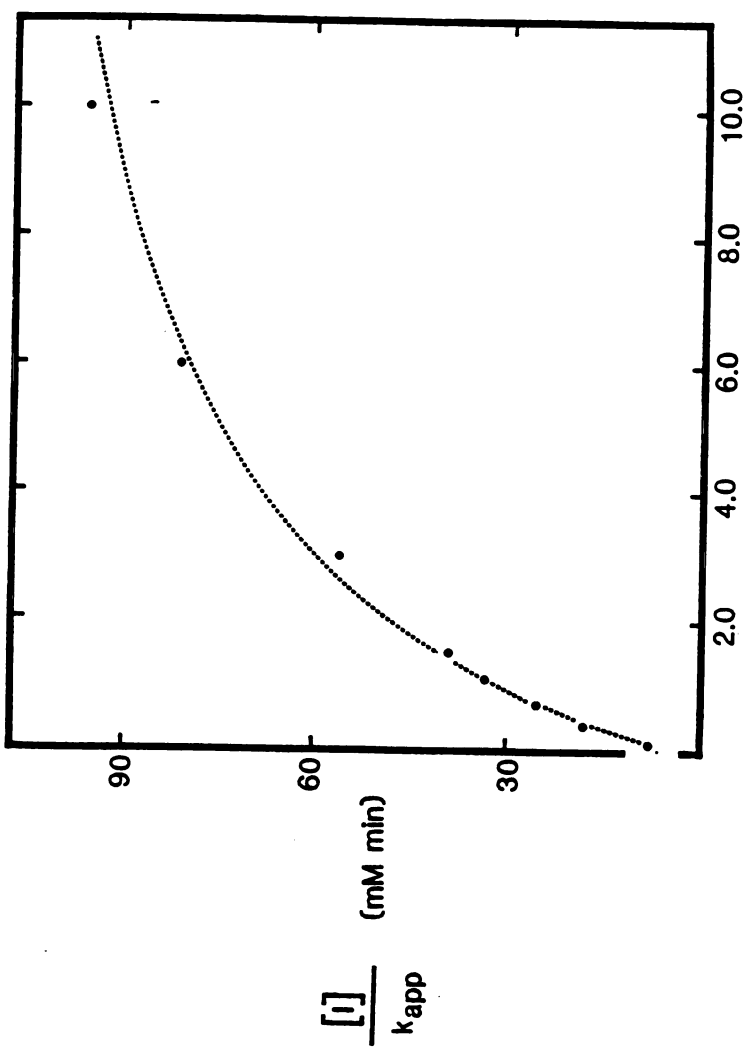


Figure 4

Rearranging this equation to the same form as equation 1 gives:

$$(3) \quad \frac{[I]}{k_{app}} = \frac{[I] + K_I}{k_n K_I + k'_n[I] + k_s}$$

Experimental data were fitted by nonlinear regression to equation 3, yielding values for K_I, k'_n and k_s, with k_n assumed to be equal to k'_n.[1] The latter is a reasonable assumption since the factor most likely to influence k'_n would be a conformational change, induced by the binding of GlcNBrAc and affecting the accessibility of the nonspecifically modified residue(s). Although binding of Glc itself is known to induce a marked conformational change, binding of GlcNAc, to which GlcNBrAc is structurally very similar, does not (53). The means ±SD from three experiments were K_I = 0.27 ±0.12 mM, k'_n = 0.0092 ±0.0012 min^{-1} mM^{-1}, and k_s = 0.037 ±0.011 min^{-1}.

GlcNBrAc as a Reversible Competitive Inhibitor.

At pH 8.5, irreversible inactivation is sufficiently slow (see below) that GlcNBrAc can be evaluated as a reversible competitive inhibitor of hexokinase with respect to Glc. That is, negligible irreversible inhibition occurs during the brief time required for measurement of initial rates; analogous experiments were done with the bovine brain enzyme by Swarup and Kenkare (11). Results, plotted

according to Dixon (105), indicate that inhibition by GlcNBrAc is competitive vs Glc (Fig. 5). From four determinations, the mean K_I ±SD was 0.28 ±0.04 mM, a value in excellent agreement with the dissociation constant for the enzyme-GlcNBrAc complex (0.27 ±0.12 mM) deduced from the inactivation experiments described above.

Protection by Glc.

If GlcNBrAc inactivation is mediated through the Glc binding site, it should also be possible to treat Glc as a competitive inhibitor of inactivation. If nonspecific inactivation were not occurring, it would be appropriate to treat the data according to the equation used by Swarup and Kenkare (11):

$$(4)\quad \frac{1}{k_{app}} = \frac{K_I}{k_2[I]} + \frac{1}{k_2} + \frac{K_I[Glc]}{k_2[I]K_G}$$

where k_2 is defined as above, and K_G is the dissociation constant for the enzyme-Glc complex. However, it is clear that in the present case, nonspecific inactivation must also be taken into account. This is represented in the following scheme:

Figure 5.

Dixon plot for GlcNBrAc as a reversible inhibitor of hexokinase. Glc concentrations were 0.05 mM (circles) and 0.10 mM (triangles). The horizontal line corresponds to the maximum velocity, as determined by using 10 mM Glc. The assay conditions were as described in Methods with the exceptions that 50 mM glycylglycine was used in place of Tris, Glc concentrations were altered as indicated and thioglycerol was deleted from the assay solution. During the time required for rate measurements, less than 5% of enzyme activity would have been lost to irreversible inactivation.

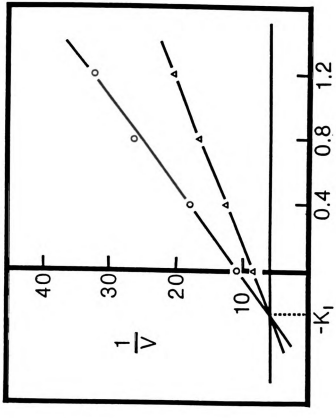
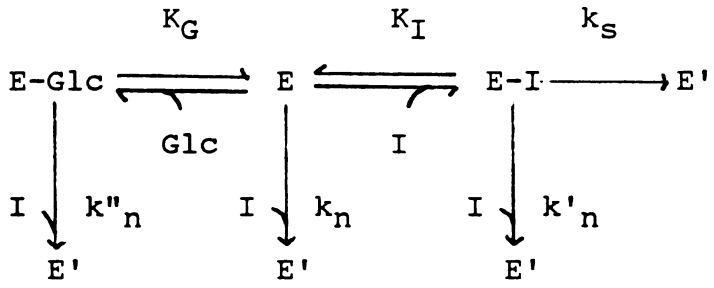


Figure 5



Based on the latter scheme, the pseudo-first order rate constant for inactivation by GlcNBrAc in the presence of Glc is given by the following equation, analogous to equation 2:

$$(5)\quad k_{app} = \frac{k''_n K_I [I][Glc]/K_G + k_n K_I [I] + k'_n [I]^2 + k_s [I]}{[I] + K_I(1 + [Glc]/K_G)}$$

where k''_n is the second-order rate constant for nonspecific inactivation of the enzyme-Glc complex, and all other constants are as defined earlier. Inactivation of the enzyme in the presence of varying concentrations of Glc was analyzed, fitting the data by nonlinear regression to the reciprocal form of equation 5 (Fig. 6). The values of K_I, k_n, k'_n and k_s were those obtained from inactivation in the absence of Glc and using equation 2, as indicated above. Best fit curves from three determinations gave mean values $\pm$SD of 0.14 $\pm$0.05 mM for K_G and 0.0016 $\pm$0.0001 min^{-1} mM^{-1} for k''_n.

The value for K_G obtained in the present work is similar to K_G values reported previously, which are in the

Figure 6.

Protection by varying concentrations of Glc against inactivation by GlcNBrAc. Hexokinase was inactivated by 0.5 mM GlcNBrAc. The line was determined by regression analysis of the data using the reciprocal of equation (5), given in the text. The values of K_I, k_n, k'_n and k_s used in the equation are those determined from the plot in Fig. 4. The values $\pm$SD for K_G and k''_n determined from this set of data were 0.078 $\pm$0.004 mM and 0.0015 $\pm$0.0002 mM^{-1} min^{-1}.

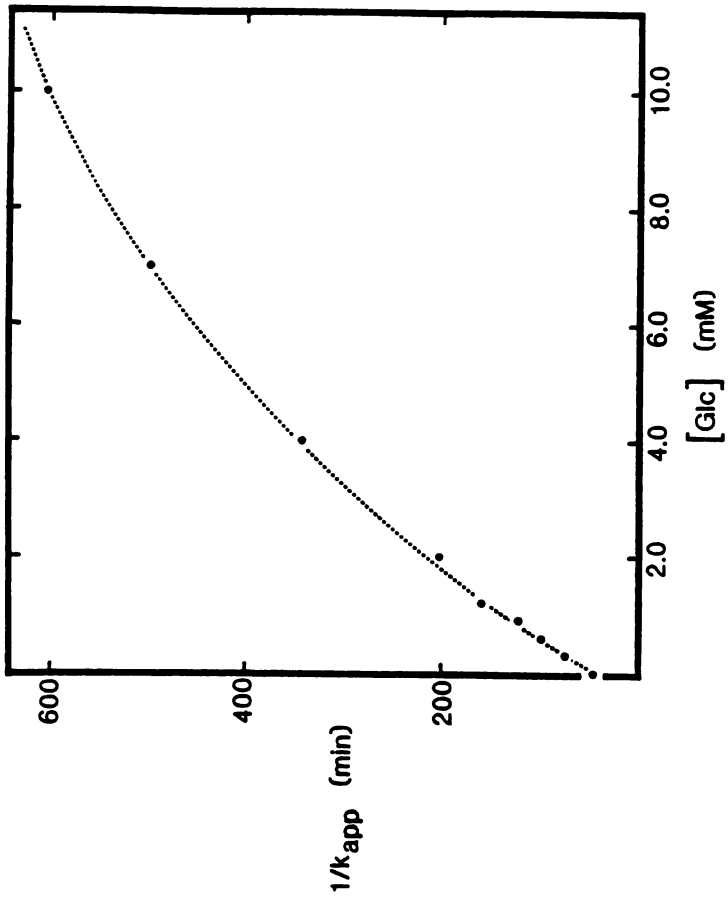


Figure 6

range of 0.04-0.25 mM (54,106); the variation in the reported values may reflect the different methods used for determination of K_G. It should also be noted that the present results were obtained at pH 8.85 rather than at pH 7-7.5 as in the earlier work. In this range pH appears to have little influence on the binding of Glc; Solheim and Fromm (107) have reported that the kinetic parameters of bovine brain hexokinase, including the K_m for Glc, are not influenced by pH in this range.

The value of k''_n is considerably less than values of k_n and k'_n, given above. Therefore, binding of Glc appears to offer appreciable protection against nonspecific inactivation of the enzyme. It is likely that the inability of GlcNAc to fully protect against inactivation is because it can only protect against one "route" of inactivation, the specific route. Thus, binding of GlcNAc blocks GlcNBrAc from binding to the enzyme and inactivating the enzyme via the specific route, while it does not protect against nonspecific inactivation. If this were the case, then it would be predicted that the rate of inactivation in the presence of GlcNAc would be the same as that determined to arise from nonspecific inactivation by GlcNBrAc in the absence of any ligand. The apparent pseudo-first order rate constant for inactivation in the presence of GlcNBrAc had a mean value, from three determinations, of 0.013 $\pm$0.002 min^{-1}. This is in close agreement with the apparent pseudo-

first order rate constant for nonspecific inactivation, 0.0092 ±0.0012, determined from k'_n at 1.0 mM GlcNBrAc.

This ability of Glc to induce a conformation change which "buries" a critical residue is in accord with previous results, which demonstrated that binding of glucose consistently provides protection against a variety of inactivating agents, including DTNB (53,54,62).

pH Dependence of Inactivation.

When the inactivation of an enzyme depends on the deprotonation of a single group, the relation between $[H^+]$ and the rate of inactivation are described by the following equation (108):

$$(6) \quad \frac{1}{k_{app}} = \frac{1}{k_{max}} + \frac{[H^+]}{K_a\, k_{max}}$$

K_a is the dissociation constant for the critical group involved, and k_{max} is the rate constant for reaction of the dissociated group. Thus, a plot of $1/k_{app}$ against $[H^+]$ is expected to be linear, with extrapolation to the abscissa providing a value for K_a.

Results obtained with the Type II isozyme from rat skeletal muscle (10) and with the Type I isozyme from bovine brain (11) conform to this expectation, and yield 9.1 and 8.9, respectively, as the pK_a values for the groups reactive with GlcNBrAc in these two enzymes. However, the results obtained with rat brain hexokinase were not in accord with

those predicted by equation 6; although the data did indeed define a line, the line appeared to extrapolate to the origin, or slightly below. The meaning of this is unclear, but it obviously precludes direct estimation of the pK_a by this method. Nonetheless, a plot of the rate of inactivation as a function of pH (inset, Fig. 7) indicates that inactivation of rat brain hexokinase is dependent on deprotonation of a group with an alkaline pK_a. The only three sidechains which are reactive with haloacetyl groups and have alkaline pK_a's are those of cysteine, lysine and tyrosine (109,110). Of these, the sulfhydryl group of cysteine is the most reactive (111). Thus, a cysteine is likely to be the critical residue labeled by GlcNBrAc.

Amino Acid Analysis of GlcNBrAc Labeled Hexokinase.

Hexokinase was subjected to acid hydrolysis and amino acid analysis before and after labeling with GlcNBrAc (to 90% inactivation) to determine the nature of the derivitized residue(s). The expected hydrolysis product of cysteine derivitized with GlcNBrAc is S-carboxymethylcysteine (11). Amino acid analyses of two independently prepared GlcNBrAc labeled hexokinase samples indicated the presence of 3 ± 1 moles of S-carboxymethylcysteine per mole of protein. The levels of amino acids susceptible to alkylation by haloacetyl groups, including lysine, tyrosine, aspartate, glutamate, and histidine, were not significantly altered by alklyation with GlcNBrAc. The methionine content of the labeled and unlabeled samples were substantially lower than

Figure 7.

Effect of pH on the rate of inactivation of hexokinase by GlcNBrAc. In the larger graph, the data are plotted in the format dictated by equation (6), given in the text. The inset shows the pseudo-first order rate constant for inactivation as a function of pH.

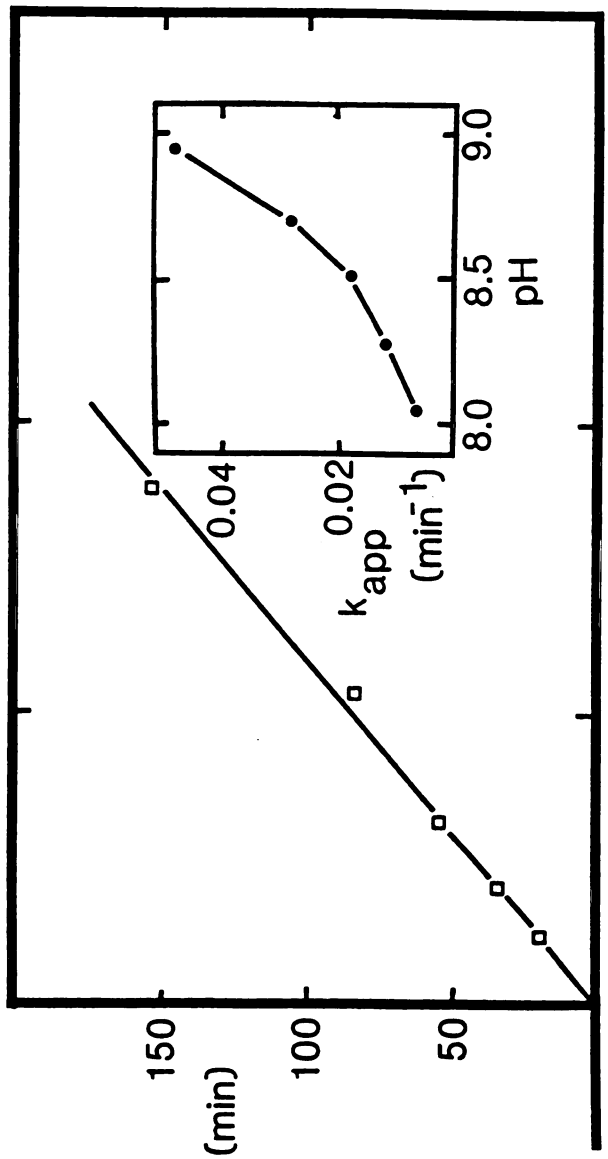


Figure 7

Table 1. Amino Acid Analysis of GlcNBrAc labeled Rat Brain Hexokinase.

	Relative moles amino acid[a]				
		Analysis 1[b]		Analysis 2[b]	
Amino Acid	Previously Reported[c]	Before GlcNBrAc	After GlcNBrAc	Before GlcNBrAc	After GlcNBrAc
ASX	5.36 + 0.11	5.55	5.58	5.65	5.69
THR	2.91 + 0.10	3.04	3.04	3.11	3.13
SER	2.84 + 0.05	3.27	3.20	2.88	2.88
GLX	5.09 + 0.11	5.26	5.25	5.29	5.22
PRO	1.44 + 0.03	1.37	1.35	1.27	1.28
GLY	4.41 + 0.08	4.90	4.81	4.57	4.61
ALA	2.84 + 0.08	3.01	3.03	2.71	2.69
VAL	3.49 + 0.06	3.39	3.58	3.61	3.59
MET	1.96 + 0.06	0.98	0.61	0.73	0.84
ILE	2.70 + 0.05	2.60	2.69	2.83	2.81
LEU	4.86 + 0.11	4.64	4.67	4.93	4.88
TYR	1.15 + 0.03	1.12	1.18	1.11	1.13
PHE	2.31 + 0.41	2.18	2.20	2.30	2.30
HIS	(1.00)	1.25	1.25	1.20	1.24
LYS	3.47 + 0.05	3.40	3.44	3.52	3.47
ARG	3.08 + 0.07	2.97	3.01	3.20	3.18
cmCYS and $METSO_2$[d]		0.33	0.56	0.42	0.56
TOTAL[a]	48.9	48.9	48.9	48.9	48.9

[a]Results were previously reported as moles amino acid per mole histidine (3,112). Based on the known molecular weight of hexokinase, the presence of 14 moles of cysteine and 4 moles of tryptophan per mole of hexokinase, it was determined that the total number of moles amino acid per mole of histidine (48.9) represented one-eighteenth the total amino acid composition of the enzyme. However, the levels of histidine detected in the current study were considerably higher than those of previous analyses, making it misleading to represent the current results relative to the histidine content. So that the current results can be accurately compared to those of the previous studies, the amino acid compositions presented here are instead relative to one-eighteenth the total amino acid composition.

[b]Two samples were independently labeled, hydrolyzed and analyzed for amino acid composition. Both were performed by the University of Davis facility.

[c]Ref (112).

Table 1 (con't).

[d]Methionine sulfoxide and S-carboxymethylcysteine both eluted at the same time, immediately before aspartate. The S-carboxymethylcysteine content of the labeled sample was therefore determined by subtracting the amount of methionine sulfoxide background detected in the unlabeled sample.

that detected previously (3,112) and demonstrated considerable variance. This is likely caused by the instability of methionine to acid hydrolysis. Nonetheless, because no unusual peaks were detected, included the positions of elution for N-carboxymethyllysine (101), O-carboxymethyltyrosine (113), and carboxymethylhomocysteine (the breakdown product of the carboxymethyl sulfonium salt of methionine (111)), the residues labeled by GlcNBrAc must therefore be only cysteines.

Stoichiometric Relationship Between Incorporation of GlcNBrAc and Inactivation.

In Fig. 8, residual activity is plotted as a function of the number of residues modified by reaction with radiolabeled GlcNBrAc. A linear relationship is evident throughout most of the inactivation. Extrapolation of the lines obtained in three such experiments gave a value of 3.4 $\pm$0.1 GlcNBrAc incorporated at zero activity. This is in good agreement with the number of cysteines modified per hexokinase, as detected by amino acid analysis. Slight additional incorporation into the inactivated enzyme is also suggested by the departure from linearity at low activities. This is not unexpected since previous work (62) has shown that all 14 of the sulfhydryl groups in the enzyme are reactive with the general sulfhydryl reagent, DTNB, and thus may also be expected to be susceptible to alkylation by the bromoacetyl moiety of GlcNBrAc.

Figure 8.

Stoichiometric relationship between incorporation of radioactivity from [^{14}C]-GlcNBrAc and residual activity. The protein-bound radioactivity at various times during inactivation was determined as described in Methods. Fractional activity remaining (a) is plotted against the number of moles of [^{14}C]-GlcNBrAc reacted per mole of enzyme. The line shows an extrapolation of the linear portion of the data to complete inactivation. The dotted line describes the curvature expected if the specific and nonspecific modes of inactivation arose from the labeling of two different residues, as described in the text.

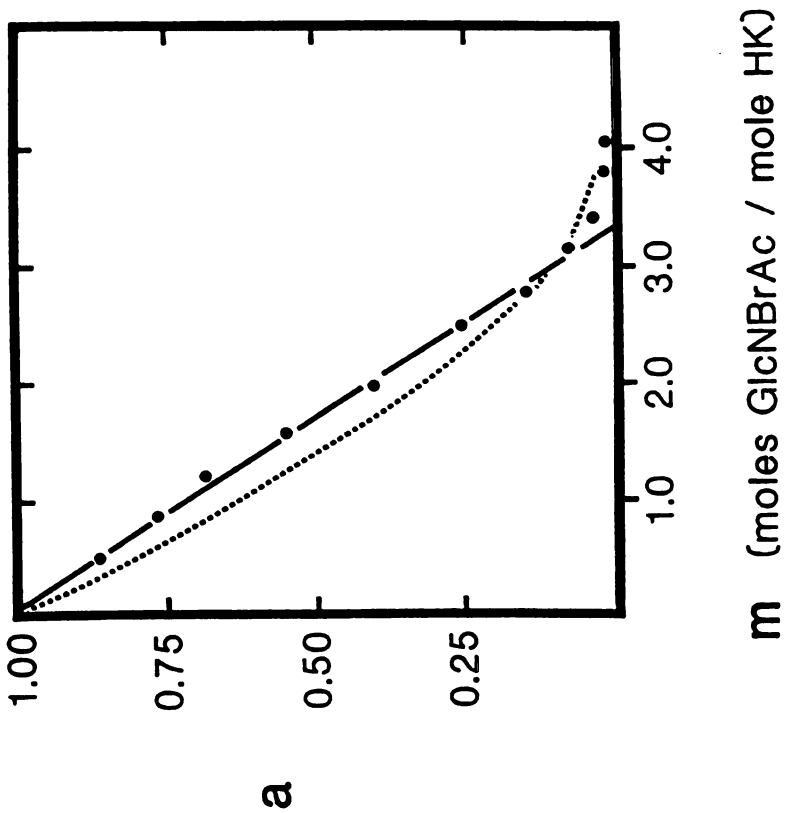


Figure 8

According to Tsou (114), a linear relationship between residual activity and the number of modified residues indicates that only one of the modified residues is critical for catalytic activity. However, an assumption in the treatment of Tsou (114) is that the critical groups, if more than one, are reacting at similar rates. In the present case, it is evident that both nonspecific and specific inactivation are occurring at appreciably different rates. It was conceivable that if <u>different</u> critical sulfhydryls were modified by the specific and nonspecific paths, curvature in the plot might not be evident because of the substantial difference in their reaction rates. Therefore, the following equation was derived, by a treatment analogous to that used by Tsou (114), describing the modification of two classes of residues reacting at appreciably different rates, and with the possibility of one or more critical groups in each class:

$$(7) \qquad m = n - pa^{\frac{1}{i+\alpha j}} - (n - p)a^{\frac{\alpha}{i+\alpha j}}$$

where m is the number of groups modified (per mole of enzyme) at any given time and a is the fraction of the original activity remaining at this same time, n is the total number of reactive groups, p is the number of groups in the first class (and thus n-p is the number of groups in the second class), i and j are the number of critical groups in the first and second classes, respectively, and α is the

ratio of the apparent rate constants for reaction of the second class of residues relative to that of the first.[2] In the present case, the first class corresponds to sulfhydryls modified with the same k_{app} as the critical sulfhydryl reacting with the bound inhibitor (i.e., the specific path), while the second class corresponds to sulfhydryls reacting at the slower rate of the nonspecific path. The value for α is calculated using equation 2 by setting either $k_n = k'_n = 0$, to give k_{app} for the the first class, or $k_s = 0$ to give k_{app} for the second class. Using the values for k_n, k'_n, k_s, and K_I given above, and at 1 mM GlcNBrAc, $\alpha = 0.32$. If there were a single critical residue in each class (i.e., if there were two distinct critical sulfhydryls in the enzyme), i=j=1, and with n=4 and p=3 (estimated from Fig. 8), a theoretical plot of a vs. m (Fig. 8) shows curvature that is easily distinguishable from the experimental data. Therefore, k_n, k'_n and k_s must pertain to the reaction of the same critical sulfhydryl and the treatment of Tsou (114), which makes no assumptions concerning the reaction path (i.e., specific or nonspecific), is applicable.

Thus, of the 3-4 sulfhydryls reacting with GlcNBrAc (during the time period of the present experiments), one can be considered critical. It should be noted that a single critical sulfhydryl residue was indicated by previous work with rat brain hexokinase (62) in which the rate of inactivation was found to be identical to the rate at which a group of three sulfhydryl groups reacted with DTNB;

according to the treatment of Ray and Koshland (67), this coincidence of rates indicates that only one of the three sulfhydryls is critical for activity. Therefore, the single critical sulfhydryl detected in the earlier work (62) must corresponds to the critical sulfhydryl modified by GlcNBrAc.

Reaction of the noncritcal sulfhydryls is unlikely to be the result of binding GlcNBrAc at sites other than the substrate binding site since hexokinase contains a single binding site for Glc (106) and thus, presumably, for the Glc analog, GlcNBrAc. Therefore, it is reasonable to expect that the noncritical sulfhydryls reacting with GlcNBrAc are not at the catalytic site, and one would anticipate that the latter sulfhydryls would not be (directly) protected by binding of ligands at that catalytic site. Or, in other words, one might be able to selectively prevent labeling of the critical sulfhydryl with ligands such as Glc or GlcNAc, while having relatively little effect on the reactivity of the noncritical sulfhydryls. Plots of residual activity *vs.* labeling by [^{14}C]-GlcNBrAc in the presence of Glc or GlcNAc are shown in Fig. 9. Due to the complexity of this situation, interpretation of the results is not immediately obvious from the plots themselves. However, equation 7 is again applicable to analysis of this situation, with the first class of residues consisting of a single critical sulfhydryl (p=i=1), now reacting at a decreased rate due to protection by the ligand, and the second class being the two

Figure 9.

Stoichiometric relationship between incorporation of radioactivity from [^{14}C]-GlcNBrAc and residual activity in the presence of added Glc or GlcNAc. These experiments were conducted and results plotted as described in the legend to Fig. 8, with the exception that inactivation was conducted in the presence of added Glc (left panel) or GlcNAc (right panel). The lines drawn were calculated by nonlinear regression onto equation 7 and with other conditions as given in the text.

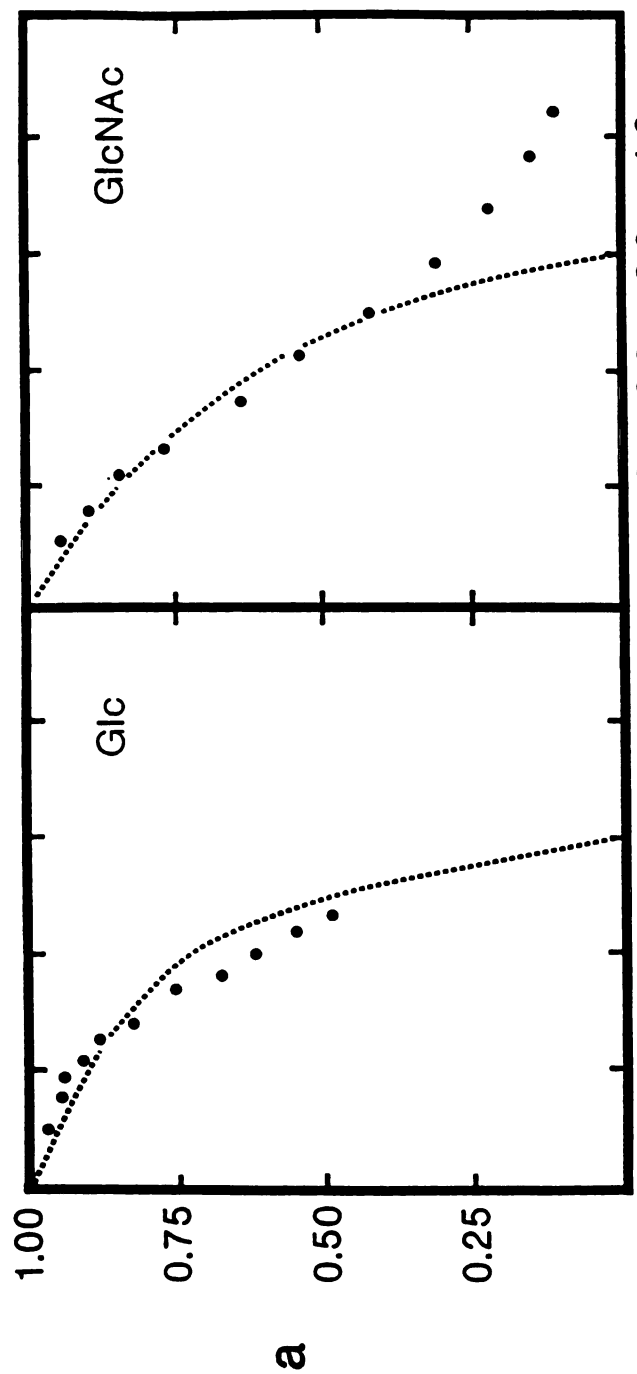

GlcNAc
Glc
1.00
0.75
0.50
0.25
a
1.0
2.0
3.0
4.0
m (moles GlcNBrAc / mole HK)

Figure 9

remaining noncritical sulfhydryls that are assumed to react at the rate seen in the absence of ligand (n-p=2, j=0).

In the presence of Glc, the calculated value for α was 6 $\pm$ 1. This is in agreement with the experimentally determined value, which corresponds to the ratio of the pseudo-first order rate constant for inactivation in the absence of Glc to that seen in its presence (Fig. 1); in four such experiments, the mean $\pm$SD for α was 7 $\pm$ 1. Due to the extremely slow rate of inactivation in the presence of saturating levels of Glc, it was not practical to follow the inactivation to near completion. Nevertheless, over the range examined, the data provide a reasonable fit to the theoretical curve. However, it is evident that some nonrandom deviation is occurring. Binding of Glc, but not GlcNAc, is known to induce a conformational change that affects reactivity of sulfhydryl groups in brain hexokinase (53,54,62). It is not unexpected that this should result in significant deviation from the curve calculated from equation 7, since the assumption that the ligand affects reactivity only of the critical sulfhydryl is unlikely to be totally valid.

In the presence of GlcNAc, the data also gave reasonable conformity with equation 7, though the obvious deviation from the theoretical curve at prolonged incubation times (low residual activity values) again indicated that appreciable nonspecific incorporation occurred, as noted above with the experiments conducted in the absence of added

ligand (Fig. 8). The value of α, calculated by nonlinear regression of the data onto equation 7 and using only the results where $a \geq 0.4$, was 3.2 ± 0.3. This calculated value is again in reasonable agreement with the experimentally determined ratio of the rates of inactivation in the absence and presence of GlcNAc, which was, from three experiments, 2.5 ± 0.2 (Fig. 1).

The above results demonstrate reasonable agreement between experimental data and behavior predicted on the assumption that Glc and GlcNAc protect a single sulfhydryl, i.e., with i=p=1 in equation 7. Theoretical curves calculated with the assumption that these ligands protected more than one sulfhydryl (e.g., i=1, but p=2) gave calculated α values greatly in excess of experimentally observed values. For example, in the case of the inactivation in the presence of Glc, regression analysis with p=2 gave a calculated α value of more than 30. However, as will become apparent in Chapter 2, GlcNAc and Glc clearly protected two sulfhydryls from alkylation by GlcNBrAc. It is unclear why this was not detected by the experiment conducted in Fig. 9.

Tryptic cleaveage of GlcNBrAc modified hexokinase.

As noted above, limited tryptic cleavage of rat brain hexokinase results in generation of three major fragments whose disposition in the overall structure has been determined (4). Thus, tryptic cleavage of [^{14}C]-GlcNBrAc modified hexokinase into the three major fragments would

seem to be the most straightforward method for identifying those regions containing sulfhydryls reacting with GlcNBrAc. However, modification of the enzyme with GlcNBrAc results in an altered tryptic cleavage pattern (Fig. 10). Although, initially, this complicated the analysis, it turned out to be advantageous, as will become evident below.

The GlcNBrAc-modified hexokinase is far more susceptible to proteolysis by trypsin. Besides the major cleavage fragments seen with the unmodified enzyme (Fig. 10, lane C), additional major fragments with molecular masses of approximately 67 kDa, 57 kDa, and 33 kDa are generated while production of the 40 kDa fragment is nearly abolished. These additional fragments are further degraded with digestion times longer than 10 minutes. Fluorography demonstrates that the 33 kDa peptide is the major labeled species, with only minor amounts of radioactivity seen in the 57 kDa and 67 kDa peptides (Fig. 10, lane D'').

The altered cleavage pattern can be explained by an increased susceptibility at a tryptic cleavage site within the 40 kDa C-terminal domain of the modified enzyme, designated T_Z in the tryptic cleavage map represented in Fig. 11. Thus, cleavage at 33% of the distance from the C-terminus would explain the appearance of new 33 kDa and 67 kDa fragments, as well as the disappearance of the 40 kDa fragment. The 57 kDa fragment would be produced by cleavage at T_1 and T_Z.

Figure 10.

Limited tryptic digestion of [^{14}C]-GlcNBrAc labeled hexokinase. SDS-PAGE of the following samples: A) Molecular weight standards (described in Methods). B) Trypsin alone, incubated under the same conditions as samples C and D. C) Native hexokinase (25 ug) digested by trypsin for 10 minutes under the conditions described in Methods. D) Hexokinase (25 ug), 90% inactivated with [^{14}C]-GlcNBrAc and digested under the same conditions as sample C. C') Immunoblotting with monoclonal antibody 5A of a lane identical to lane C. D') Similar immunoblotting of a lane identical to lane D. D") Fluorograph of lane D.

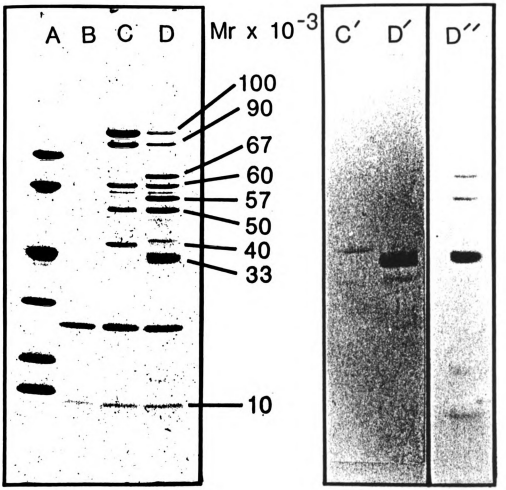


Figure 10

Figure 11.

Schematic representation of relevant proteolytic cleavage sites in rat brain hexokinase. Numbers above or below the line representing the polypeptide chain are M_r values (x 10^{-3}) of major cleavage products. T_1 and T_2 are the tryptic cleavage sites in native hexokinase, as previously designated by Polakis and Wilson (4). T_z is the additional tryptic cleavage site produced in GlcNBrAc-modified hexokinase (see text). S12, a fragment produced by further digestion with S. aureus V8 protease (4), is positioned above the region of the 40K domain from which it is derived. The cross-hatched area, marking the overlap of S12 with the 33 kDa peptide resulting from cleavage at T_z, denotes the region containing the cysteines labeled by reaction with [^{14}C]-GlcNBrAc.

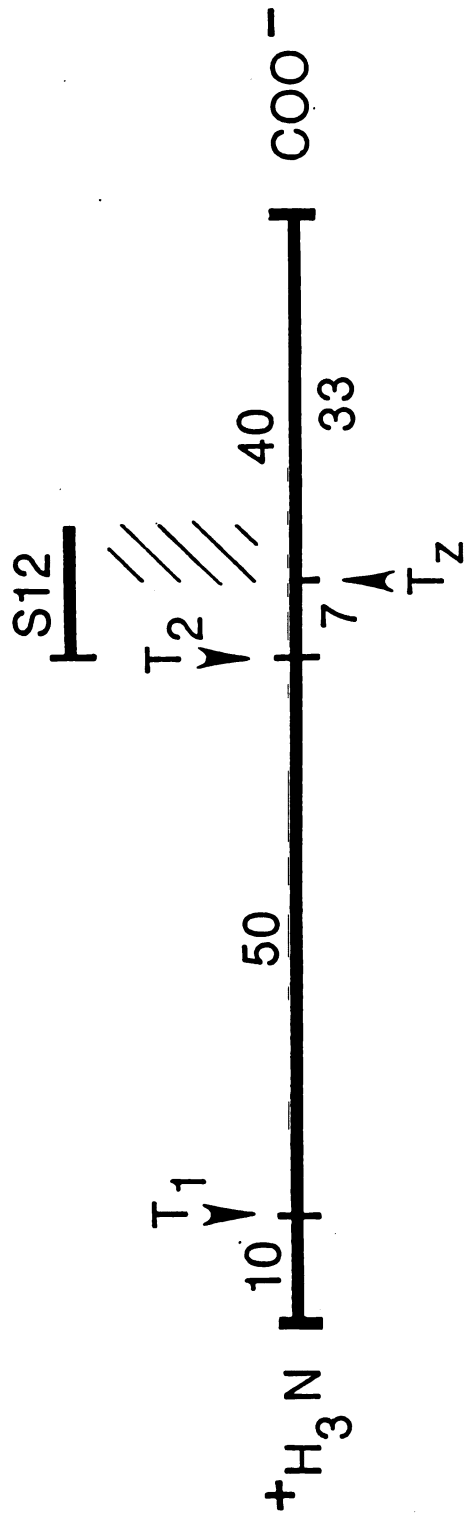


Figure 11

That the 33 kDa fragment did indeed originate from the C-terminal region was confirmed by immunoblotting with a monoclonal antibody whose epitope was previously mapped to a region within the 40 kDa domain (4). Monoclonal antibody 5A gave a strong positive reaction against the 33 kDa species (Fig. 10, lane D'). Thus, the C-terminal domain, and more specifically, a 33 kDa fragment derived from the C-terminal region of this domain, is the site of GlcNBrAc modification.

[^{14}C]-GlcNBrAc labeling of tm-hexokinase.

An alternative approach to locating the sites of GlcNBrAc modification is to react the affinity label with hexokinase after the enzyme has been proteolyzed. It has previously been demonstrated (4) that, even after tryptic cleavage, the domains remain associated via noncovalent interactions, with retention of catalytic activity, little change in the kinetic parameters for Glc, Glc-6-P or ATP, and continued ability to bind in a Glc-6-P-sensitive manner to the outer mitochondrial membrane. In other words, in terms of the functions that have been examined, tm-hexokinase is virtually indistinguishable from the intact enzyme. Thus, not unexpectedly, GlcNBrAc was also found to inactivate tm-hexokinase with kinetic parameters and protective effects of various ligands virtually identical to those seen with the intact enzyme (Fig. 1). The relation between inactivation and incorporation of [^{14}C]-GlcNBrAc was also the same as that shown in Fig. 8 for the native enzyme. SDS-PAGE and fluorography of tm-hexokinase labeled by [^{14}C]-

GlcNBrAc confirmed the 40 kDa C-terminal domain as the site of labeling by GlcNBrAc (Fig. 12).

[^{14}C]-IAM labeling of denatured tm-hexokinase.

By an alteration of the above procedure it was possible to estimate the total number of sulfhydryl groups in each of the domains of rat brain hexokinase. After cleaving the enzyme into its respective domains with trypsin, as above, the enzyme was denatured with SDS, to expose all buried sulfhydryls, and [^{14}C]-IAM was added. After SDS-PAGE and fluorography of the fragments it was possible to determine the relative [^{14}C]-carboxamidocysteine in each domain by densitometric scanning of the fluorograph (Fig. 13). No label was apparent at all in the 10 kDa domain, meaning the 40 kDa and 50 kDa domains must contain all 14 sulfhydryl groups (62) in rat brain hexokinase. Because cleavage at T_2 yields the 40 kDa and either 50 kDa or 60 kDa fragments (see Figure 11), the amount of 40 kDa fragments must be equimolar to the sum of the amounts of 50 kDa and 60 kDa fragments present in the gel. Thus, the ratio of the number of sulfhydryls between the 40 kDa and 50 kDa domains can be determined from the ratio of label detected in the 40 kDa band to that detected in the 50 kDa band plus that detected in the 60 kDa band. The labeling of the 100 kDa and 90 kDa proteins was ignored as it provides no information regarding the relative labeling of the 40 kDa and 50 kDa domains. At a total of 14 sulfhydryls (62) per hexokinase molecule, the 40 kDa and 50 kDa domains were determined to have 8.16 $\pm$0.04

Figure 12.

Labeling of tm-hexokinase with [^{14}C]-GlcNBrAc. Lane A is an SDS-PAGE gel pattern, stained with Coomassie Blue, showing the cleavage fragments present in a preparation of tm-hexokinase, 90% inactivated with [^{14}C]-GlcNBrAc. Lane B is a fluorograph of lane A. The tm-hexokinase was produced by digestion with trypsin for 60 min as described in Methods.

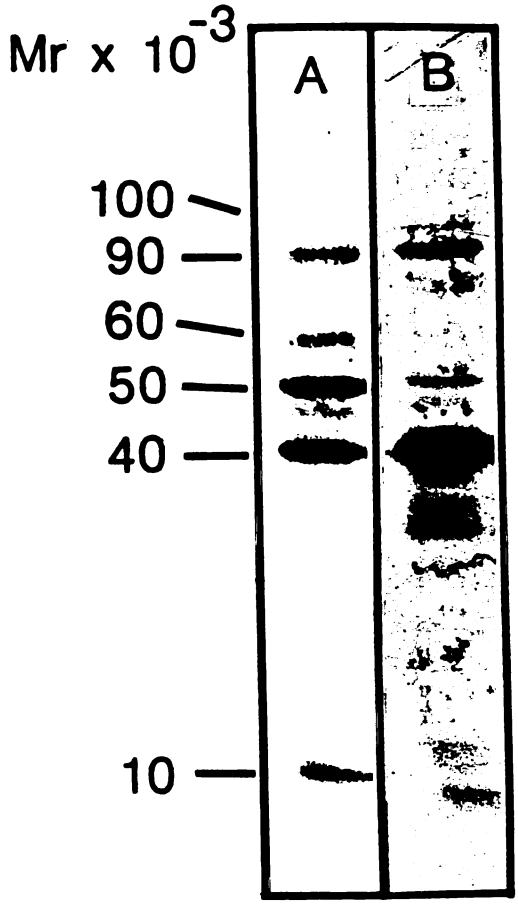


Figure 12

Figure 13.

Quantitation of sulfhydryls in rat brain hexokinase domains. After labeling of tm-hexokinase with [^{14}C]-IAM in the presence of 1% SDS the protein was analyzed by SDS-PAGE and fluorography of the resulting gel. The relative amount of ^{14}C label in each domain was determined by densitometric scanning of the fluorograph, which is shown here. The positions of each of the major protein fragments in the gel are indicated.

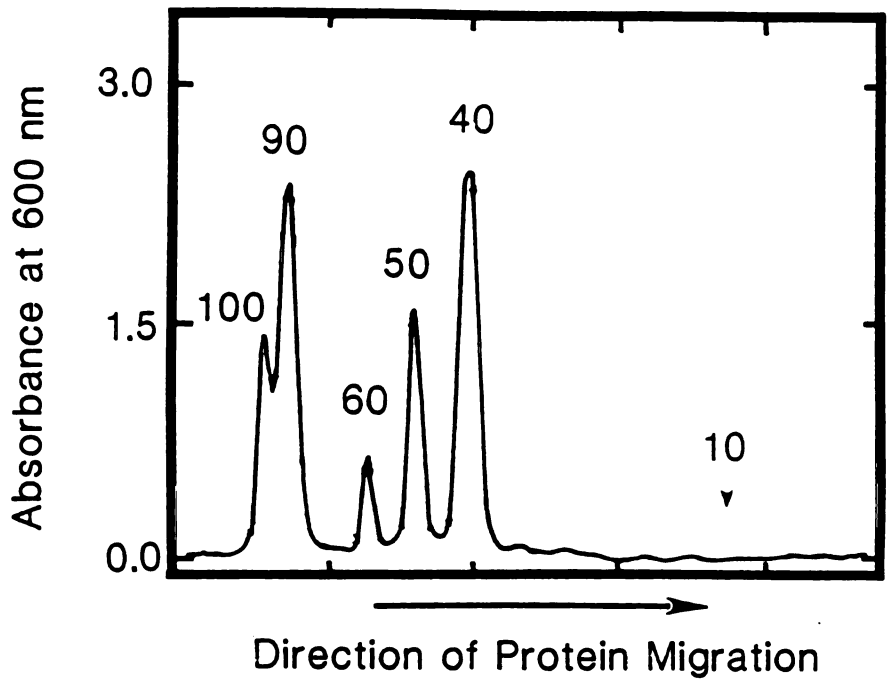

Absorbance at 600 nm
3.0
1.5
0.0
100
90
60
50
40
10
Direction of Protein Migration

Figure 13

and 5.84 ±0.03 cysteines, respectively, as determined from two such experiments. Therefore, after rounding to the nearest integers, the 50 kDa, 40 kDa and 10 kDa domains of rat brain hexokinase are expected to contain 6, 8 and 0 sulfhydryls, respectively.

Two-dimensional peptide mapping of [^{14}C]-GlcNBrAc labeled tm-hexokinase.

By the method of Polakis and Wilson (4) it is possible to digest tm-hexokinase with S. aureus V8 protease into further subregions of the enzyme. This two-dimensional method of peptide mapping was used for identical samples of [^{14}C]-GlcNBrAc-modified tm-hexokinase, one of which had been protected by GlcNAc during inactivation. Both inactivations were quenched when the unprotected sample was 60% inactivated, at which time the GlcNAc-protected sample still retained 75% of its original activity. From the fluorographs (Fig. 14), it is apparent that the peptide designated "S12" by Polakis and Wilson and known to be derived from the N-terminal region of the 40 kDa domain (4), is the major labeled species. The relative specific activity of the S12 peptide from the GlcNAc-protected sample was 56% of the specific activity of the S12 peptide from the unprotected sample, as determined by densitometric scans of the stained gels and the fluorographs. Comparable protection by GlcNAc was seen in each of three experiments.

The S12 fragment includes that region of the hexokinase molecule in which the new tryptic cleavage site, T_Z, is

Figure 14.

Fluorographs after two-dimensional peptide mapping of tm-hexokinase labeled with [^{14}C]-GlcNBrAc. Identical samples of tm-hexokinase were incubated with [^{14}C]-GlcNBrAc with no ligand or with addition of GlcNAc, as described in the text. Subsequently, 2-dimensional peptide mapping was done according to Polakis and Wilson (4); resulting peptide maps were as shown in previous publications from this laboratory (4,9). Migration in the first dimension was from left to right, and the position of the 40 kDa tryptic fragment is indicated by the arrows. Subsequent further proteolysis with S. aureus V8 protease was followed by electrophoresis in the second dimension (in the direction of the arrows). Fluorographs obtained with these maps are shown above. Left panel, no added ligand; right panel, added GlcNAc. The prominently labeled peptide was "S12", as designated by Polakis and Wilson (4).

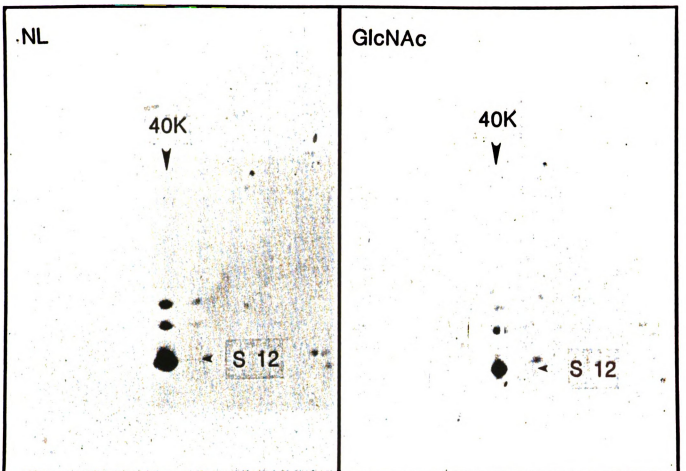


Figure 14

produced after modification with GlcNBrAc (see Fig. 11). Since both S12, derived from the N-terminus of the 40 kDa domain, as well as the 33 kDa species, derived from the C-terminus of the 40 kDa domain by tryptic cleavage at T_Z, contain the GlcNBrAc-modified sulfhydryls, it is apparent that the location of these residues must be confined to the approximately 5 kDa region of overlap between these fragments (crosshatched region in Fig. 11). The protective effect of GlcNAc against [^{14}C]-GlcNBrAc incorporation into this region indicates that it includes the critical sulfhydryl at the Glc binding site. Moreover, the lack of significant labeling elsewhere in the molecule indicates that the additional (noncritical) modified sulfhydryls may also be confined to this region. Indeed, it may well be the case that the particular reactivity of these noncritical sulfhydryls results from their proximity to high local concentrations of GlcNBrAc that may exist in the vicinity of the binding site for this ligand.

Vicinal Sulfhydryls in Rat Brain Hexokinase.

As mentioned in the literature review, bovine brain hexokinase has two pairs of vicinal sulfhydryls, the slower reacting pair contains the critical sulfhydryl believed to be at the Glc binding site. It may therefore be expected that rat brain hexokinase also contains vicinal sulfhydryls, one of which may be the sulfhydryl labeled by GlcNBrAc.

Rat brain hexokinase incubated with stoichiometric levels of DTNB revealed the existence of two pair of vicinal

sulfhydryls (Table 2). Thus, reaction of equimolar amounts of DTNB and hexokinase produces two TNB anions. As has been seen with other enzymes (63,76-81), this indicates the oxidation of two enzyme sulfhydryl groups as a disulfide. Because rat brain hexokinase is monomeric (3), it is unlikely that the vicinal sulfhydryls are intermolecular, and can therefore be expected to be intramolecular. Unlike the bovine enzyme, however, the enzyme retains considerable activity after formation of two disulfide bonds. However, if the critical sulfhydryl labeled by GlcNBrAc inactivates the enzyme due to the steric bulk of the attached GlcNAc moiety, then it is possible that its oxidation to an adjacent sulfhydryl, a modification which does not introduce a bulky group into the active site, would not inactivate the enzyme. If this were the case, the critical sulfhydryl would be resistant to alkylation by GlcNBrAc and the oxidized enzyme would not be further inactivated. This was tested by incubating enzyme oxidized with one or two equivalents of DTNB with GlcNBrAc and comparing the results to a control (Fig. 15). The enzyme species with one or two pairs of oxidized sulfhydryls showed only a slight decrease in the rate of inactivation by GlcNBrAc, indicating that the critical sulfhydryl labeled by GlcNBrAc can not be one of the vicinal sulfhydryls.

Table 2. Oxidation of vicinal sulfhydryls in rat brain hexokinase.

Moles DTNB added/ mole hexokinase	Moles TNB released/ mole hexokinase	Activity remaining
1.0	2.1 $\pm$0.2[a]	73 $\pm$2%
2.0	4.1 $\pm$0.4[b]	62 $\pm$10%

Standard deviations were determined from three experiments.

[a]10–15 min was required for complete reaction.

[b]20–30 min was required for complete reaction.

Figure 15.

Effect of vicinal sulfhydryls on inactivation by GlcNBrAc. Enzyme was incubated with either one or two equivalents of DTNB, forming one or two disulfide bridges, as described in the text. After allowing sufficient time for the completion of the reaction, 1.0 mM GlcNBrAc was added to each sample, as well as to a control (no DTNB added). The reactions with DTNB and GlcNBrAc were done in 50 mM Tris, 0.5 mM EDTA, pH 8.5. The activity of the control sample prior to reaction with GlcNBrAc represents 100% activity.

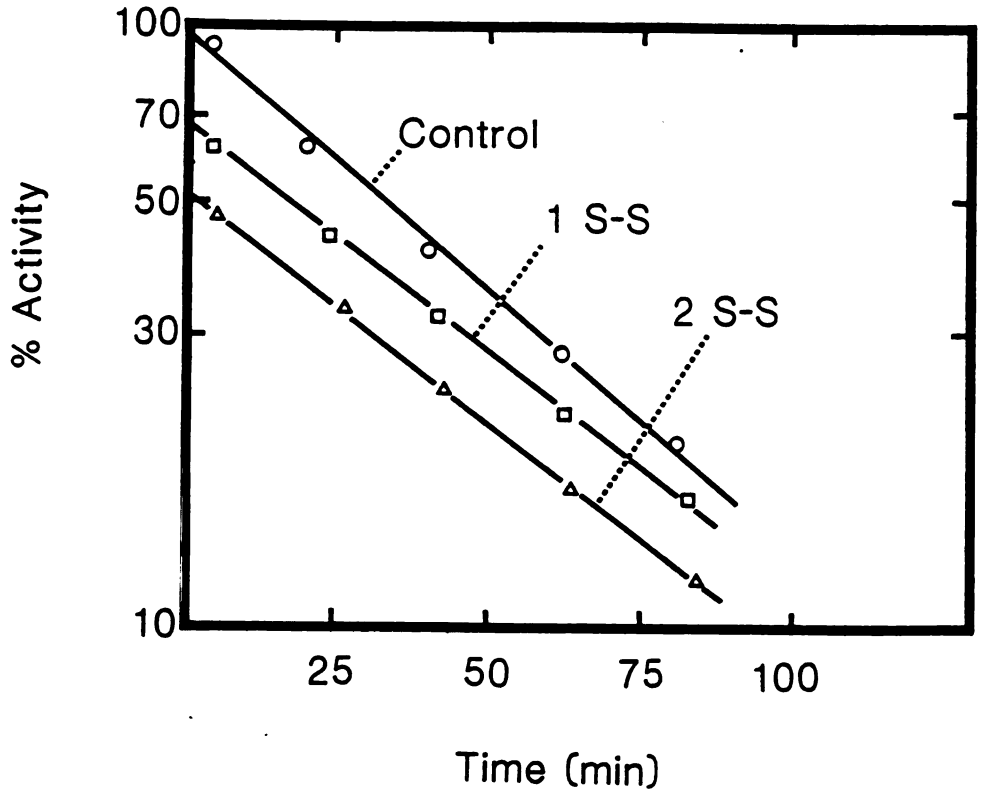


Figure 15

DISCUSSION

The results of this chapter make possible a comparison of GlcNBrAc inactivation of rat brain hexokinase with the inactivation of other hexokinases by the same or related affinity labels (Table 3). Rat brain hexokinase is clearly similar to bovine brain and rat Type II hexokinases with respect to inactivation by GlcNBrAc. The situation becomes more complicated with respect to the number and identity of residues labeled in each enzyme. In the case of the Type II isozyme of hexokinase from rat skeletal muscle (10), a single residue, presumed to be a sulfhydryl group, was modified by reaction with radiolabeled GlcNBrAc with resulting complete inactivation of the enzyme. With the Type I isozyme of hexokinase from bovine brain (11), the situation was slightly more complicated in that two sulfhydryl residues were reactive with GlcNBrAc; however, by suitable kinetic analysis and demonstration of selective protection by Glc, only one of these sulfhydryls was shown to be critical for activity, with labeling characteristics consistent with its location at the Glc binding site. In the present case of the Type I isozyme from rat brain, the situation becomes still more complex in that appreciable labeling of three residues is observed. Despite this complication, kinetic analysis indicates that, as in the case with the rat skeletal muscle (10) and bovine brain (11)

Table 3. Summary of reactions of various hexokinases with glucose- and galactose-derived alkylating reagents.

		Active Site Directed		Inactivation			
Reagent	HK[a]	?[b]	K_i or label[c]	?	k_2[d] (pH)	Residue(s) Modified	Critical Residue(s)
GlcNBrAc	Bovine[e] Type I	Y	2.0 (0.45)	Y	0.03 (8.5)	2 Cys	1 Cys
	Rat Type I	Y	0.27 (0.28)	Y	0.04[f] (8.85)	1 Cys (?) 2 unknown	1 Cys (?)
	Rat[g] Type II	Y	0.57	Y	0.03 (8.5)	1 Cys	1 Cys
	Rat[h] Type IV	Y		N			
	Yeast B[h]	Y		N			
GlcNBrHex	Rat[h] Type II	Y		Y	0.11 (8.5) 0.01 (7.0)	1 Cys (?) 1 COOH (?)	1 Cys (?) 1 COOH (?)
	Rat[h] Type IV	Y		Y	0.01 (7.0)	1 COOH (?) 1 unknown	1 COOH (?)
	Yeast B[h]	Y		N			
GalNBrAc	Bovine[h] Type I	N					
	Rat[h] Type II	N					
	Rat[h] Type IV	N	ca. 100				
	Yeast B[i]	Y	14.0	Y	0.06 (8.5)	2 Cys	1 Cys

[a]Hexokinase. [b]Yes (Y) or no (N). [c]Determined from inactivation kinetics. Values in parentheses were determined by treating the label as a reversible inhibitor. [d]As defined by Kitz and Wilson (104) and mentioned in the text. [e]Results from Swarup and Kenkare (11). [f]k_s in the present work. [g]Results from Connoly and Trayer (10). [h]Results from Darby et al. (71). [i]Results from Otieno et al. (12,13).

enzymes, modification of a single critical sulfhydryl accounts for the inactivation of the enzyme.

Moreover, the present results are consistent with this critical sulfhydryl being located at the Glc binding site of the enzyme. Thus, there is excellent agreement between the kinetic parameters for the action of GlcNBrAc as a reversible inhibitor, competitive vs. Glc, and as an irreversible inhibitor of the enzyme. Protection by Glc and GlcNAc are obviously in accord with the effects expected if inactivation by GlcNBrAc is mediated through binding at the Glc binding site. Although the present study has indicated that a significant amount of "nonspecific" inactivation can occur without formation of a complex between hexokinase and this inhibitor, an effect also noted by Swarup and Kenkare (11), this does not preclude the use of this compound, with appropriate precautions, for selectively labeling the Glc binding site.

Although results presented here and previous studies (10,11) demonstrate the presence of a cysteine located near the Glc binding site of mammalian hexokinases, we do not believe that this group is involved in either substrate binding or the catalytic event itself. The basis for this deduction is that dissociation of this group, and hence its reactivity with GlcNBrAc, is markedly affected by pH in the range of 7-9 (10,11, and the present work). However, Solheim and Fromm (107) have shown that the kinetic parameters of bovine brain hexokinase are totally

independent of pH in this region. It is highly unlikely that a residue directly involved in either substrate binding or catalysis could dissociate without noticeable effect on kinetic parameters. Thus, although it is apparent that reaction of this residue with GlcNBrAc results in inactivation - and it is on this basis that it is designated to be "critical" - it is likely that the inactivation results from steric obstruction of the glucose binding site rather than from modification of a sulfhydryl directly involved in catalysis or binding.

The present results, together with other recent work from this laboratory (9), have demonstrated that the binding sites for both substrates, Glc and ATP, are located within the 40 kDa C-terminal domain. It is therefore reasonable to consider this domain as serving the catalytic function. What then, can be made of the rest of this rather large molecule? It hardly seems likely that nearly half of this 100 kDa protein is superfluous, with no function of physiological significance. The location of catalytic function in the C-terminal half of the molecule is consistent with the suggestion (17-23) that present-day mammalian hexokinases Types I-III, with M_r 100 kDa may have evolved by duplication and fusion of an ancestral gene similar to that coding for yeast, wheat germ, and mammalian Type IV hexokinases, all of which have M_r 50 kDa. According to this view, one of the catalytic sites in the ancestral "fused" protein evolved into the regulatory site for Glc-6-P

present in mammalian hexokinases Types I-III but absent in the hexokinases with M_r 50 kDa. Thus it is predicted that the binding site for Glc-6-P will be found to be associated with the N-terminal half of the molecule, and that the functions associated with this region of the molecule will be regulatory in nature. In this context, it is worth noting that the Glc-6-P-sensitive binding of hexokinase to the outer mitochondrial membrane, a property suggested to have major regulatory significance (7), has been shown to be dependent on structural features in the N-terminal half of the molecule (8,115).

DERIVATIONS

Equation 2 was derived as follows:

$$\frac{d(\%\ \text{Activity})}{dt} = \frac{1}{[E_T]} \frac{d[E + EI]}{dt} = \frac{-k_{app}\ [E + EI]}{[E_T]}$$

$$k_{app} = \frac{k_n\ [I]\ [E]\ +\ (k'_n\ [I] + k_s)\ [EI]}{[E\ +\ EI]}$$

$$k_{app} = \frac{k_n\ [I][E] + (k'_n\ [I] + k_s)[E][I]/K_I}{[E] + [E][I]/K_I}$$

$$k_{app} = \frac{k_n\ [I] + k'_n\ [I]^2/K_I + k_s[I]/K_I}{1 + [I]/K_I}$$

$$k_{app} = \frac{k_n\ K_I\ [I] + k'_n\ [I]^2 + k_s\ [I]}{K_I + [I]}$$

Equation 5 was derived in a similar manner, but starting from the following equation:

$$\frac{d(\%\ \text{Activity})}{dt} = \frac{1}{[E_T]} \frac{d[E + EI + EG]}{dt} = \frac{-k_{app}\ [E + EI + EG]}{[E_T]}$$

Equation 7 was derived as follows, assuming two classes of groups, F and S, reacting according to two different first order rate constants, k_1 and αk_1. The fraction of F and S classes, x_f and x_s, remaining at any given time, t, will be

$$x_f = e^{k_1 t} \quad \text{and} \quad x_s = e^{\alpha k_1 t} = x_f^{\alpha}$$

The fractional activity, a, at any given time will be

$$a = (x_f^{i}) * (x_s^{j})$$
$$= (x_f^{i}) * (x_f^{\alpha j})$$
$$= x_f^{i+\alpha j} \quad \text{or} \quad x_f = a^{\frac{1}{i+\alpha j}}$$

(I) $$\text{and} \quad x_s = a^{\frac{\alpha}{i+\alpha j}}$$

The fraction, x, of total reactive groups, n, at any given time will be

$$nx = px_f + (n-p)x_s$$

where p is the number of groups in the F class.

Substituting from equation I,

(II) $$nx = pa^{\frac{1}{i+\alpha j}} + (n-p)a^{\frac{\alpha}{i+\alpha j}}$$

The number of labeled groups per protein, m, at any given time can be determined from

$$x = \frac{n - m}{n} \quad \text{or} \quad xn = n - m$$

Substituting with equation II,

$$n - m = pa^{\frac{1}{i+\alpha j}} + (n-p)a^{\frac{\alpha}{i+\alpha j}}$$

or

$$m = n - pa^{\frac{1}{i+\alpha j}} - (n-p)a^{\frac{\alpha}{i+\alpha j}}$$

FOOTNOTES

[1]In principle, k_n could be determined directly by observing the rate of inactivation at very low concentrations of GlcNBrAc, at which inactivation *via* the "specific" pathway, involving formation of the E-I complex was negligible. However, the very low rates of inactivation under such conditions make this impractical.

[2]When $j = 0$, equation 7 simplifies to a form equivalent to equation 5 of Tsou (114).

CHAPTER II

STRUCTURE OF THE GLUCOSE BINDING SITE

Hexokinase labeled with [^{14}C]-GlcNBrAc was exhaustively digested with trypsin and the peptides separated by HPLC. Three major radioactive peptides were found. The sequences of two of these peptides are highly homologous to amino acid sequences of yeast hexokinase which are near the Glc binding site. A serine residue in yeast hexokinase, identified as essential for a major conformational change necessary for catalysis, is also conserved in rat brain hexokinase. Several other structural parallels are made between yeast and rat brain hexokinases.

RESULTS

Peptide Mapping of [^{14}C]-GlcNBrAc-Labeled Hexokinase.

HPLC-peptide mapping of tryptic fragments generated from hexokinase labeled with [^{14}C]-GlcNBrAc showed three major radioactive peptides (Fig. 16). Additional digestion with trypsin did not increase the yield of peptides I-III or produce new radioactive peaks.

Peptides I and II were both found to have coeluted with one other major component in the primary HPLC separation, and were further purified by HPLC under modified conditions as described in Methods (Fig. 17). In contrast, further HPLC analysis of Peptide III (Fig. 17) showed no substantial contamination and it was subsequently prepared for amino

Figure 16.

Primary HPLC fractionation of a tryptic digest of [^{14}C]-GlcNBrAc labeled hexokinase. Top Panel: Profile of peptides eluted from HPLC column. The dashed line indicates the programmed elution gradient. The elution positions for radioactive Peptides I-III are indicated. Bottom panel: Profile of [^{14}C]-GlcNBrAc labeled peptides eluted from HPLC column.

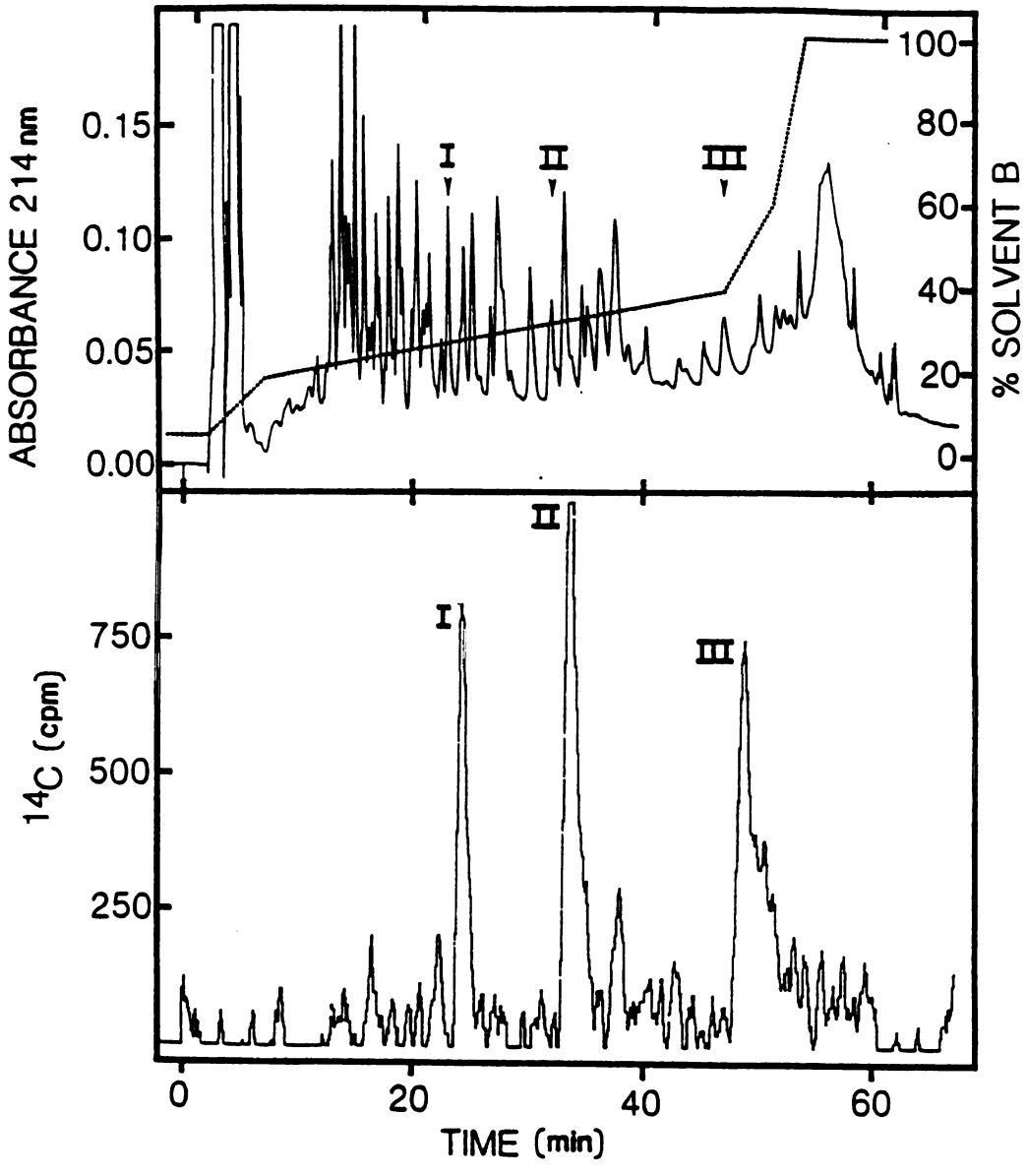

ABSORBANCE 214 nm
0.15
0.10
0.05
0.00
I
II
III
100
80
60
40
20
0
% SOLVENT B
^{14}C (cpm)
750
500
250
0
20
40
60
TIME (min)

Figure 16

Figure 17.

Secondary HPLC fractionation of Peptides I-III. The dashed line in each figure indicates the elution gradient. Arrows designate the only radioactive peak detected in each separation.

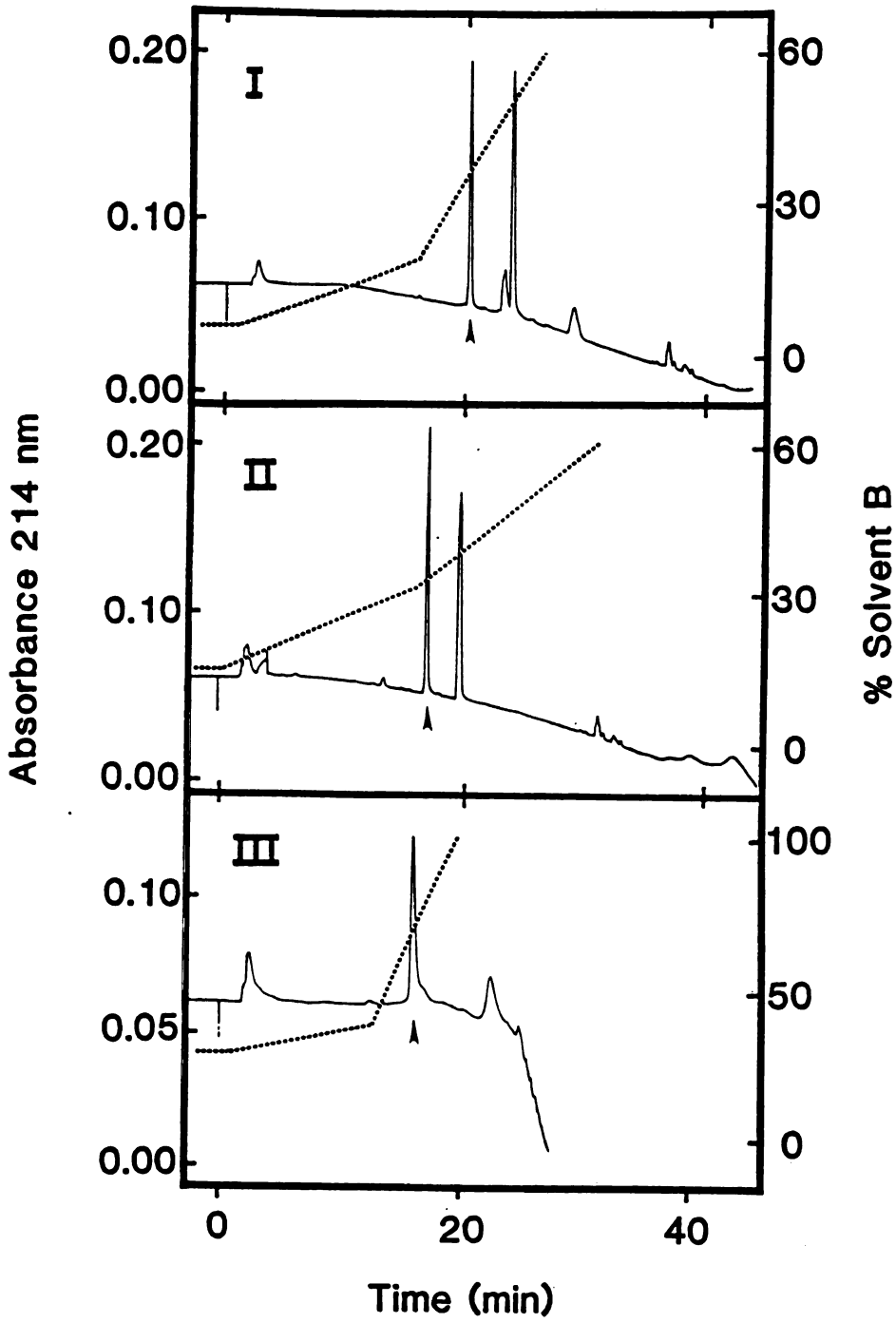

I
II
III
0.20
0.10
0.00
0.10
0.05
0.00
60
30
0
100
50
0
0
20
40
Time (min)
Absorbance 214 nm
% Solvent B

Figure 17

acid analysis and sequencing without secondary HPLC purification.

The correspondence between the number of labeled peptides seen and the number of sulfhydryls modified by GlcNBrAc (Chapter I), together with evidence that Peptides I-III were indeed discrete species (see below), indicated that each of these peptides contained one of the sulfhydryls susceptible to facile reaction with GlcNBrAc.

The results in Chapter I indicated that all three labeled sulfhydryls were in the 40 kDa domain. Part of the proof of their location came from the labeling of tm-hexokinase, assuming that the same sulfhydryls are labeled as in the native enzyme. This was tested by peptide mapping, as in Fig. 16, the labeled 40 kDa fragment. Hexokinase was proteolyzed by trypsin, labeled with [^{14}C]-GlcNBrAc and the fragments separated by SDS-PAGE as in Fig. 12. The 40 kDa fragment was excised from the gel, electroeluted, the SDS removed, and then digested with trypsin for peptide mapping by HPLC. Analysis of the digestion mixture by HPLC (Fig. 18), showed radioactive components eluting at the same retention times as seen for Peptides I-III in Fig. 16. Peptides I-III were also visible in the HPLC absorbance profile (results not shown) which conclusively demonstrates the location of all three labeled sulfhydryls within the 40 kDa domain.

Figure 18.

HPLC fractionation of a tryptic digest of [^{14}C]-GlcNBrAc labeled 40 kDa domain. The labeled 40 kDa domain was isolated as described in the text. The protein was subjected to peptide mapping in the same manner as done in Fig. 16 for the entire 100 kDa protein. Only the ^{14}C profile is shown. The elution times of Peptides I-III are indicated.

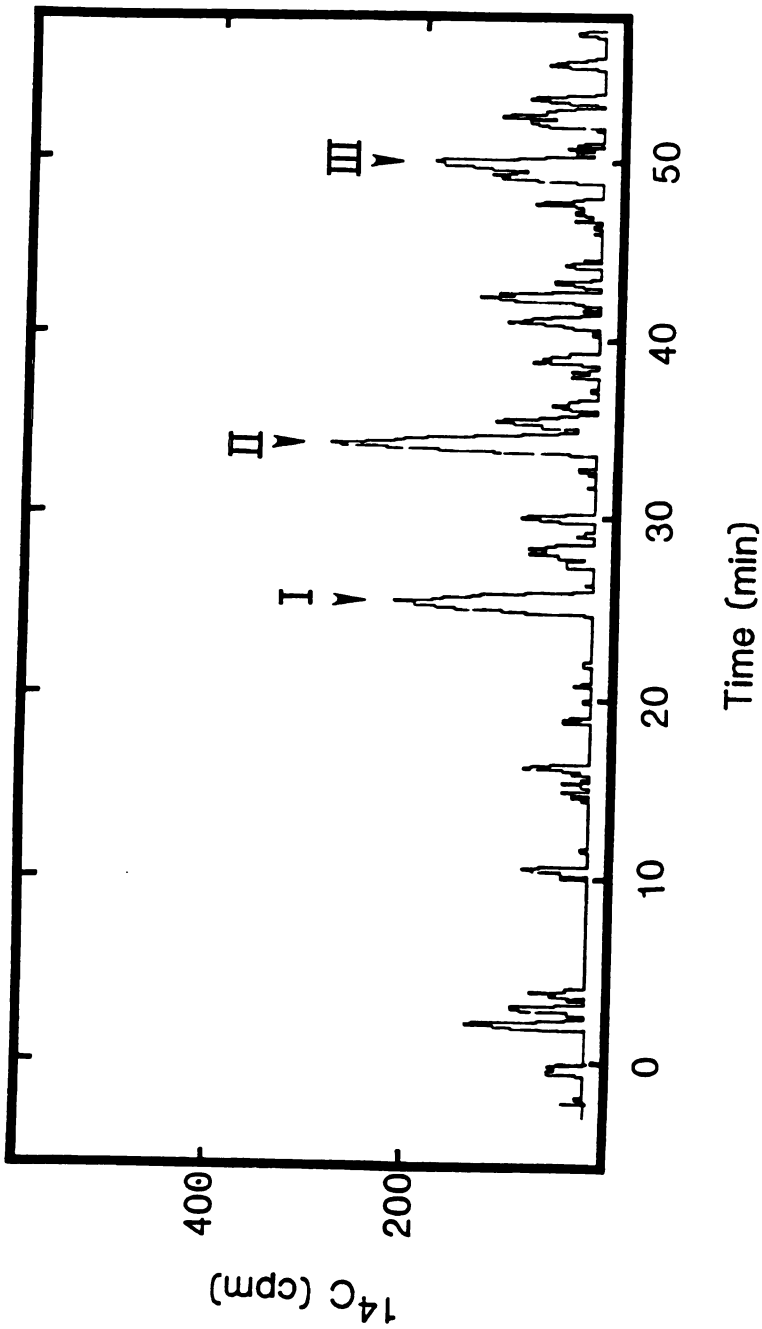


Figure 18

Effects of Glc and GlcNAc on Alkylation of Peptides I-III.

Of the three sulfhydryls labeled by GlcNBrAc, only one was believed to be critical and to be selectively protected by Glc or GlcNAc. To determine which peptide contained the critical sulfhydryl, enzyme was labeled with [^{14}C]-GlcNBrAc in the presence of either Glc or GlcNAc and subjected to peptide mapping as in Fig. 16. Results were compared to a control sample which had been labeled by [^{14}C]-GlcNBrAc in the absence of a protective ligand. The expectation was that the labeling of one of the peptides, containing the critical sulfhydryl, would be selectively decreased. However, the data from three experiments did not show such unambiguous results. The results from one such experiment are represented in Fig. 19. It is apparent that both Peptides I and III are protected by the inclusion of either Glc or GlcNAc during incubation with [^{14}C]-GlcNBrAc. Labeling of Peptide II was quite insensitive to the addition of protective ligands, clearly indicating that this peptide can not include the critical sulfhydryl within its sequence. These results, as well as those from two other similar experiments, are diagramed in Fig. 20. The protection of both Peptides I and III approximates the level of protection expected if either was to contain the critical sulfhydryl. In every case, however, protection of Peptide III exceeded that of Peptide I.

Figure 19.

Effects of Glc or GlcNAc on the labeling of Peptides I-III by [^{14}C]-GlcNBrAc. Enzyme was incubated with [^{14}C]-GlcNBrAc in the presence of Glc, GlcNAc or no protecting ligand at all, which served as a control. The reactions were terminated when the control sample was 90% inactivated. Each sample was subjected to peptide mapping in the same manner as the sample in Fig. 16. The control, GlcNAc and Glc protected samples are in the top, middle and bottom panels, respectively. The positions of Peptides I-III are indicated in each panel.

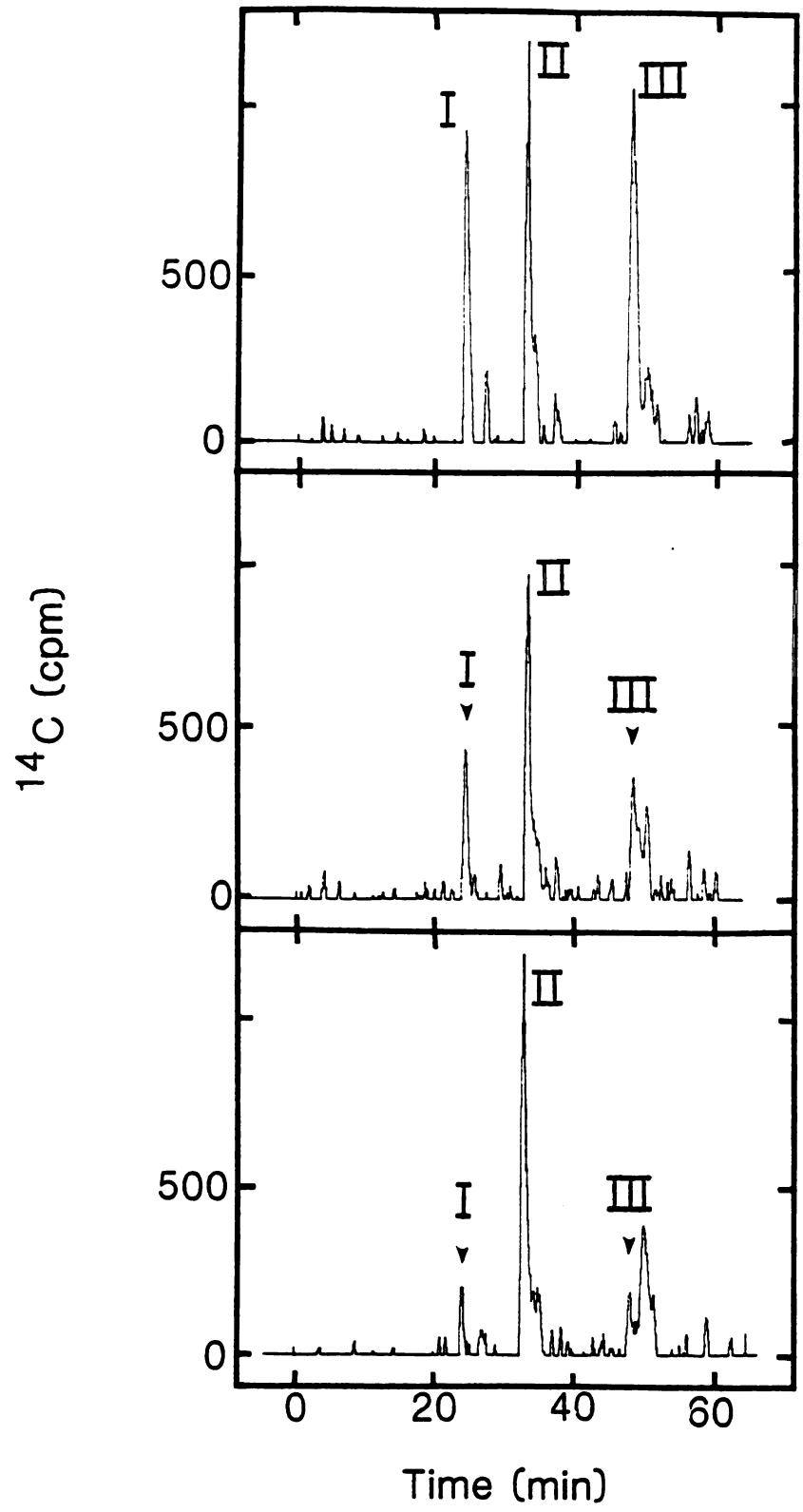


Figure 19

Figure 20.

Correlation between protection by Glc and GlcNAc against [^{14}C]-GlcNBrAc inactivation and the labeling of Peptides I-III. The results displayed in Fig. 19, as well as those for two other such experiments, are represented in graphical form to compare the alkylation of each peptide to the extent of enzyme inactivation. Thus, in three experiments, enzyme was inactivated by [^{14}C]-GlcNBrAc in the presence of Glc, GlcNAc, or no protecting ligand. The latter sample served as the basis for comparison, with labeling of Peptides I-III and extent of inactivation expressed relative to that seen in the sample with no protective ligands. The reactions were initiated at the same time and terminated when the control sample was approximately 90% inactivated. The [^{14}C]-GlcNBrAc labeled samples were subjected to peptide mapping as in Fig. 16. The relative labeling of Peptides I-III was determined by comparing the area of the radioactive peak for each peptide to its corresponding peak in the control sample. Peak areas were determined by cutting and weighing. Each data point represents a radioactive peak from a sample protected by either Glc or GlcNAc, as indicated in the lower right hand side of the figure. The diagonal line indicates the 1:1 correlation expected between inactivation of the enzyme and labeling of a critical residue in that peptide.

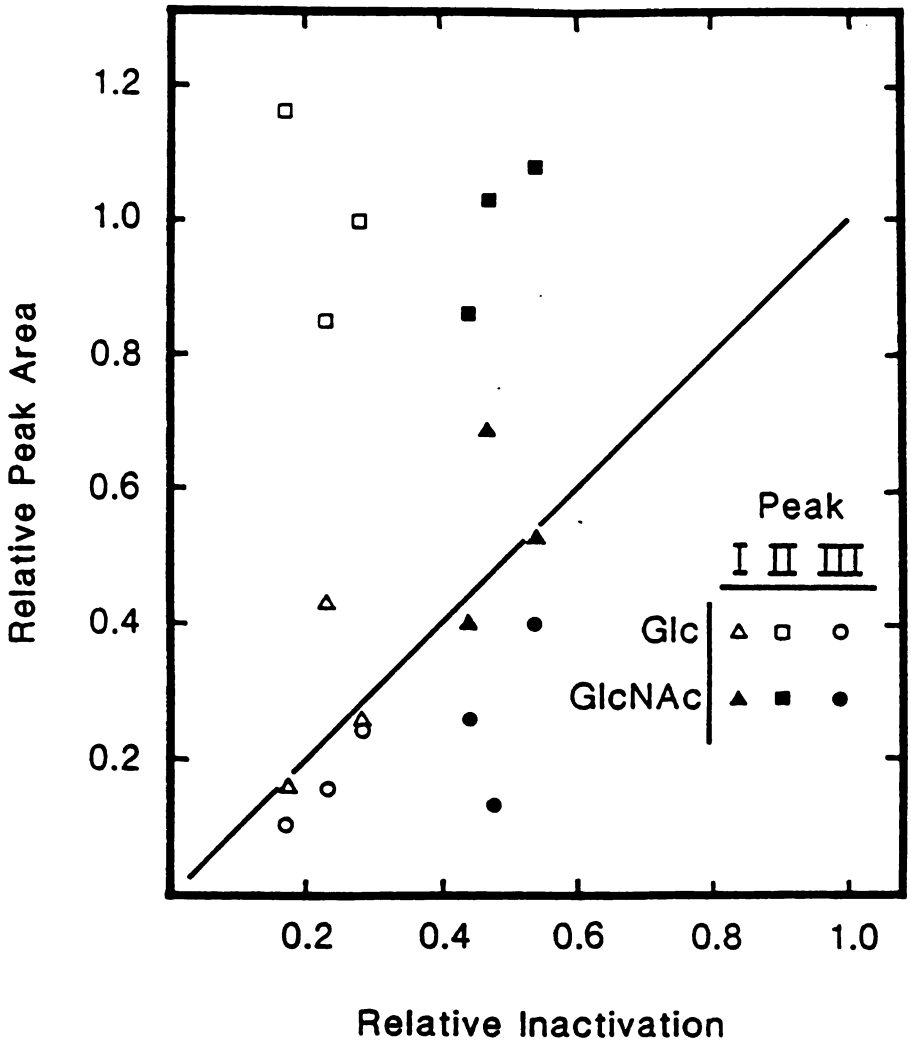

1.2
1.0
0.8
0.6
0.4
0.2
Relative Peak Area
0.2
0.4
0.6
0.8
1.0
Relative Inactivation
Peak
I II III
Glc
GlcNAc

Figure 20

Amino Acid Compositions of Peptides I-III.

Results of amino acid analyses of Peptides I-III are shown in Table 4. There are no particularly notable features such as relatively high content of acidic, basic, or hydrophobic amino acids. A single basic amino acid, lysine or arginine, would be expected in each peptide if tryptic cleavage were complete. This was the case with Peptide I but not with the larger Peptides II and III. Since, as mentioned above, continued digestion with trypsin did not result in further cleavage of these latter peptides, it seems that they may contain a potential cleavage site that is resistant to tryptic attack in the derivatized (and perhaps in the underivatized) protein. Swarup and Kenkare (11) apparently met similar resistance with the bovine brain hexokinase, using prolonged digestion times and still finding evidence of incomplete tryptic cleavage. Since haloacetyl derivatives can react with lysyl sidechains (109,110), modification of a lysine by GlcNBrAc is a possible explanation, although no derivitized lysine was detected after hydrolysis of the labeled protein (Chapter I). Furthermore, none of the peptides gave rise to a hydrolysis product that corresponded to N-carboxymethllysine. Other possible causes of the resistance to tryptic cleavage might include adjacent lysyl or arginyl residues (116), or steric hindrance of the cleavage site resulting from proximity to a sulfhydryl modified by GlcNBrAc. The latter being a causative factor is consistent

Table 4. Amino acid compositions of Peptides I-III[a].

	Peptide					
	I		II		III	
Asx	2.2	(2)[b]	4.5	(5)	3.8	(4)
Glx	1.5	(2)	3.4	(3)	2.7	(3)
cmCys[c]	0.2	(1)	0.2	(1)	0.6	(1)
Ser	1.6	(2)	3.1	(3)	2.9	(3)
Gly	1.6	(2)	3.4	(3)	6.3	(6)
His	1.1	(1)	1.2	(1)	1.0	(1)
Arg	[1.0]	(1)	[1.0]	(1)	[1.0]	(1)
Thr	0.9	(1)	2.0	(2)	2.8	(3)
Ala	1.9	(2)	3.5	(4)	1.0	(1)
Pro	0.4	(0)	1.5	(2)	2.1	(2)
Tyr	0.1	(0)	0.5	(1)	0.3	(0)
Val	1.3	(1)	2.9	(3)	1.1	(1)
Met	0.1	(0)	0.3	(0)	0.8	(1)
Ile	0.2	(0)	2.2	(2)	2.1	(2)
Leu	1.9	(2)	5.7	(6)	3.1	(3)
Phe	0.3	(0)	1.0	(1)	2.7	(3)
Lys	0.3	(0)	1.9	(2)	0.9	(1)
Trp	ND[d]		ND		ND	
Estimated amino acids/ peptide	17		40		36	

[a]Expressed relative to arginine, which was found in all three peptides in amounts stoichiometrically related to other amino acids present in significant amount.

[b]Numbers in parentheses are nearest integer values.

[c]S-carboxymethylcysteine. Although the cm-Cys contents are low, they are interpreted as indicating the presence of cysteine in the peptide (see text).

[d]Not determined.

with the observation that the amount of Peptide III, which is well resolved from other digestion products seen in the 214 nm absorbance profile (Fig. 16), was markedly reduced in digests of enzyme that had been reacted with GlcNBrAc in the presence of a protective ligand, Glc or GlcNAc. This suggests that Peptide III had undergone further digestion in the unmodified enzyme, the mere absence of a GlcNAc moiety being considered unlikely to radically alter the retention time of such a large peptide (i.e., the apparent decrease was <u>not</u> simply due to altered retention of the corresponding peptide derived from the unmodified enzyme); if this is indeed the case, then it obviously implies that such digestion was hindered in the derivatized protein.

As discussed in Chapter I, GlcNBrAc labels three cysteines in the enzyme. Thus, it is expected that each of the GlcNBrAc-labeled peptides must contain (minimally) one alkylated cysteine. Though S-carboxymethylcysteine, the expected hydrolysis product from cysteine derivatized with GlcNBrAc (11), was found in all three peptides, the yield was much less than this. Since S-carboxymethylcysteine is not a naturally occurring amino acid, its presence in each peptide clearly suggested that cysteine was, in fact, present. When samples of glutathione, labeled with either [^{14}C]-GlcNBrAc or [^{14}C]-iodoacetamide, were subjected to similar amino acid analysis, both derivatized peptides yielded the expected amounts of glycine and glutamate, based on ^{14}C content of the samples prior to hydrolysis, but only

0-30% of the expected carboxymethylcysteine (results not shown). Thus the low amounts of S-carboxymethylcysteine detected in these analyses (Table 4) reflect the instability of this compound under the conditions used. Analysis of another preparation of Peptide III using conventional ion exchange methods, i.e., without derivatization to the PTC amino acid, yielded 0.8 moles of S-carboxymethylcysteine per mole of arginine; similar analyses were not performed on Peptides I or II.

Mass Spectrometry of Peptide III.

A more accurate analysis of the size of Peptide III was determined by Fast Atom Bombardment Mass Spectroscopy. A protonated molecular ion was detected at 3394.4, as well as a diprotonated molecular ion at 1697.8 (1697.2 is half the mass/charge ratio of 3394.4). A fragment ion detected at 3176 corresponds well with loss of a GlcNAc moiety (M_r 220) from the peptide. A peptide of M_r 3176, based on an average molecular weight of 105 per amino acid residue, calculated from the composition shown in Table 4, corresponds to a peptide 30 amino acid residues in length, slightly smaller than that estimated from amino acid analysis (Table 4).

Amino Acid Sequences of Peptides I-III.

Partial amino acid sequences of Peptides I-III are presented in Tables 5-7. As will be discussed below, Peptides I and III show considerable homology to sequences in yeast hexokinase isozymes A and B (37,38) and mammalian Type III hexokinase (56). No significant homology was

Table 5. Sequence analysis of Peptide I.[a]

Cycle	Residue	Yield[b]
1	Ala (Ser, Gly)	175 (-[c],38)
2	Thr	-
3	Asp	108
4	-	-
5	Glu	-
6	Gly	-
7	-	-
8	Asp	50
9	Val	36
10	Ala	-
11[d]	Ser	-
12[d]	Leu	24

[a]The contents of each cycle were analyzed by two non-identical HPLC systems, unless otherwise noted. This analysis was performed by the Michigan State University facility.

[b]Yields, when determined, are indicated in pmoles.

[c]The Ser was not quantitated, but it was a minor component.

[d]Cycles 11 and 12 were analyzed by only one HPLC system.

Table 6. Sequence analysis of Peptide II[a].

Cycle	Residue	Yield[b]
1	Ala	506
2	Ile	367
3	Leu	434
4	Gln	194
5	Gln	146
6	Leu	349
7	Gly	154
8	Leu	321
9	Asn	147
10	Ser	-
11	Thr	-
12	-	-
11[c]	Asp	-
14[c]	Asp	-
15[c]	Ser	-

[a]The contents of each cycle were analyzed by two non-identical HPLC systems, unless noted otherwise. The analysis listed here was performed by the Michigan State University facility. Cycles 1 - 9 of an identical sample were also sequenced by the University of Davis facility and the same residues were detected at each step.

[b]Yields, when determined, are indicated in pmoles.

[c]Cycles 13-15 were analyzed by only one HPLC system.

Table 7. Sequence analysis of Peptide III[a].

Analysis:	1		2		3		Consensus[d]
Cycle	AA[b]	Yield[c]	AA	Yield	AA	Yield	
1	Ile	90	Gly (Met)	548 (-)	Met (Gly)	607 (107)	X
2	Pro	-	Pro	72	Pro	108	Pro
3	Leu	60	Leu (Val)	403 (-)	Leu (Val)	504 (90)	Leu
4	Gly	-	Gly	260	Tyr (Gly)	333 (52)	Gly[e]
5			Phe	267	Phe	464	Phe
6			Thr	-	-	-	Thr
7			Phe	232	Phe	447	Phe
8			Ser	-	Ser	-	Ser
9			Phe	187	Phe	378	Phe
10			Pro	151	-	-	Pro
11			X	-	-	-	X
12			His	-	-	-	His
13			Gln	-	Glc[f]	-	Gln
14			X	-	-	-	X
15			Asn	-	Asn[f]	-	Asn
16			Leu	130	Leu[f]	-	Leu
17			Asp	-			Asp
18			X	-			X
19			Gly	-			Gly
20			X	-			X
21			Leu	95			Leu

[a]The contents of each cycle were analyzed by two non-identical HPLC systems, unless noted otherise. Analyses 1 and 2 were performed by the University of Davis facility. Analysis 3 was performed by the Michigan State University facility.

[b]Amino acid.

[c]Yields, when determined, are indicated in pmoles.

[d]Consensus for each cycle.

[e]Although tyrosine was determined to be the major amino acid at this position of analysis 3, it was concluded to be a glycine on the basis of analyses 1 and 2, where no tyrosine was detected at this position, but only glycine.

[f]These cylces were analyzed by only one HPLC system.

detected between the sequence of Peptide II and those reported for other hexokinases (37,38,56).

In principle, the position of the modified sulfhydryl might be determined by monitoring radioactivity released in each cycle during sequencing of peptides prepared from enzyme labeled with [^{14}C]-GlcNBrAc. However, no significant radioactivity was detected during sequencing of Peptide II. Nor was significant radioactivity left in the cup containing the unsequenced portion of this peptide. Thus it was apparent that the Glcnbrac labeled cysteine was unstable to the rather harsh conditions encountered during the sequencing operation, with the radioactive moiety being lost during one of the wash phases of the cycle. Previous workers (117) have encountered similar problems of stability in sequencing peptides containing cysteine derivatized with a related alkylating agent, 2-bromo-4'-nitroacetophenone. Moreover, when cysteine that had been derivatized by reaction with [^{14}C]-GlcNBrAc was subjected to 5 reaction cycles in the sequencer, no detectable PTH amino acid and less than 5 % of the radioactivity was released as a PTH derivative in any cycle, nor was there significant radioactivity left in the cup. In view of these results, this approach was not pursued with peptides I and III. It is obvious, however, that the position of the derivatized sulfhydryl must correspond to one of the unidentified residues or to unsequenced portions of these peptides.

DISCUSSION

Location of Peptides I and III within the Overall Sequence of Brain Hexokinase and Comparison with Location of Homologous Peptides in Yeast Hexokinase.

In Chapter I it was demonstrated that the sulfhydryl(s) labeled by GlcNBrAc and protected by GlcNAc are located within a 5 kDa region located near the N-terminus of the 40 kDa C-terminal domain of rat brain hexokinase. The present work shows that these sulfhydryls are contained within Peptides I and III. Since, together, Peptides I and III account for a sequence containing approximately 50 residues, it is apparent that essentially the entire sequence in this 5 kDa region must be represented by these two peptides.

That Peptides I and III are derived from the same 5 kDa region of rat brain hexokinase is supported by their homology to amino acid sequences in an approximately 5 kDa region of yeast hexokinase (Fig. 21). If Peptides I and III are from the same region of rat brain hexokinase, it raises the possibility that Peptide I may be included in the larger Peptide III. Based on their homology to the yeast enzyme, Peptide I would represent the C-terminal portion of Peptide III. This would explain the similar protection of Peptides I and III against labeling by GlcNBrAc (Figs. 19 and 20). However, the results in Table I do not support this possibility since the alanine content of Peptide I exceeds that of Peptide III and hence the former cannot be contained within the latter. Additional evidence that Peptides I and

Figure 21.

Amino acid sequence comparisons between yeast hexokinase and selected peptides from rat hexokinase isozymes I and III. The sequences are as follows: RI(I) and RI(III) are rat hexokinase isozyme I, Peptides I and III, respectively; RIII(4), RIII(5) and RIII(6) are rat hexokinase isozyme III, peptides 4, 5 and 6 (56); YA and YB are yeast hexokinase isozymes A and B, as determined by sequencing the cloned gene (37,38); YB(X) is the sequence of yeast hexokinase isozyme B, as tentatively determined by X-ray crystallography (14); γ, Δ and ∂ designate amino acids thought to have two, three or four atoms in their side chain. Where the the crystallographic sequence agrees with the actual sequence, or where the correct number of atoms in the amino acid side chain were identified, the residues are underlined. Homologous regions between the yeast and rat isozymes are enclosed in boxes. An X denotes an unidentified residue in the sequence.

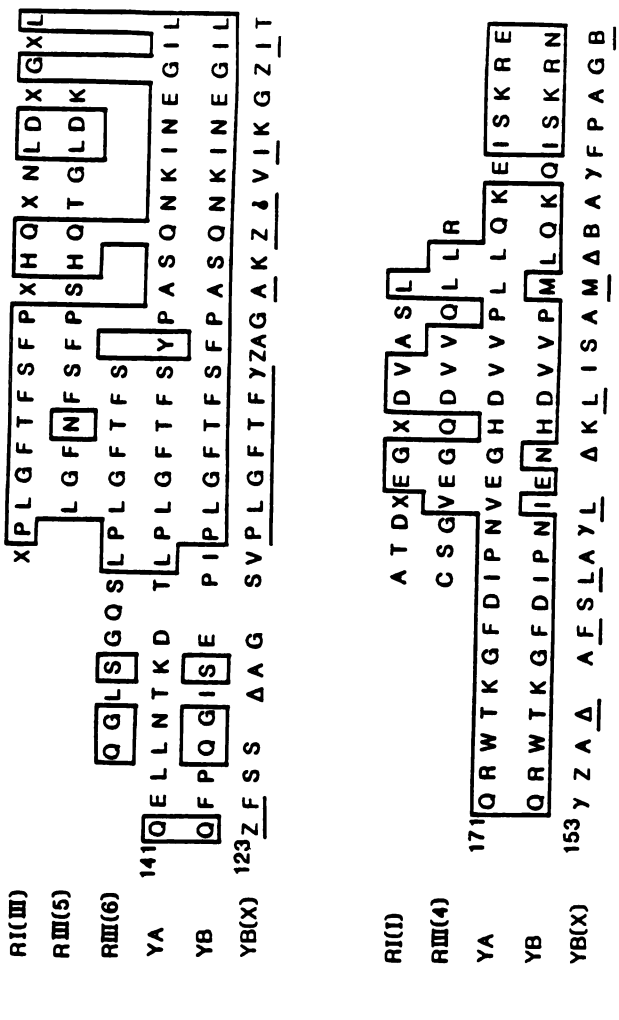


Figure 21

III were discrete nonoverlapping species came from sequencing. For example, 3 leucines were found in the partial sequence determined for Peptide III, accounting for the 3 leucines seen in the amino acid analysis. Yet amino acid analysis of Peptide I also indicated the presence of 2 leucines which, based on sequencing, were clearly distinct from those of Peptide III. Thus, the sequence of Peptide I cannot be contained within that of Peptide III.

Similar considerations also permit the conclusion that there is no overlap of these peptides. Thus, if the C-terminal segment of Peptide III overlapped with the N-terminal segment of Peptide I, it is apparent that the latter segment must contain a basic amino acid, susceptible to tryptic cleavage, that would correspond to the C-terminal residue of Peptide III. Thus, Peptide I would have to contain a lysine or arginine in addition to the basic residue that represents its own C-terminus. This is not the case (Table 4). Therefore, based on their homology with known sequences in the yeast enzymes (37,38), it can be deduced that Peptide III represents the N-terminal region and Peptide I the C-terminal region of this 5 kDa segment of rat brain hexokinase.

Examination of the amino acid sequence of the yeast enzymes (37,38) indicates that there are 29 residues from the start of the segment homologous to Peptide III to the start of the segment homologous to Peptide I. This is virtually identical to the 30 residues in Peptide III,

estimated from its molecular weight determined by mass spectrometry. It is clear that the overall length of the region represented by Peptides I and III in the rat brain enzyme must be virtually identical to that of the corresponding region in yeast hexokinase. Thus, in rat brain hexokinase, the nonoverlapping Peptides I and III must be contiguous, or nearly so, with no appreciable insertions or deletions in the sequence determined for yeast hexokinase.

The placement of Peptides I and III within the overall sequence of the 40 kDa domain of rat brain hexokinase is shown in Fig. 22, and compared to the disposition of the homologous sequences within the overall sequence of yeast hexokinase. The similarity is striking, and in accord with the postulated evolutionary relationship of these enzymes (17-23) in which the 40 kDa catalytic domain (9, and the present work) of rat brain hexokinase can be considered to have evolved by a process of gene duplication and fusion from an ancestral hexokinase similar to the present-day yeast enzyme. The representation in this figure also suggests that deletion of approximately 80 amino acid residues, corresponding to the N-terminal region of the yeast enzyme, occurred in the course of evolution of the 40 kDa domain from the ancestral form. Because estimation of the number of residues deleted from the N-terminus depends on molecular weights determined by SDS-PAGE, a technique with a 5-10% margin of error (118,119), the figure of 80

Figure 22.

Locations of Peptides I and III within the 40 kDa domain of rat brain hexokinase compared to the locations of homologous sequences in the yeast isozymes. The horizontal lines represent the relative length of the yeast enzyme and the 40 kDa C-terminal catalytic domain of the rat brain enzyme. N-termini are oriented on the left. The darkened regions represent the homologous sequences between the two proteins, positions 151-195 from yeast hexokinase and Peptides I and III from the rat brain hexokinase 40 kDa domain. Homologous regions are aligned.

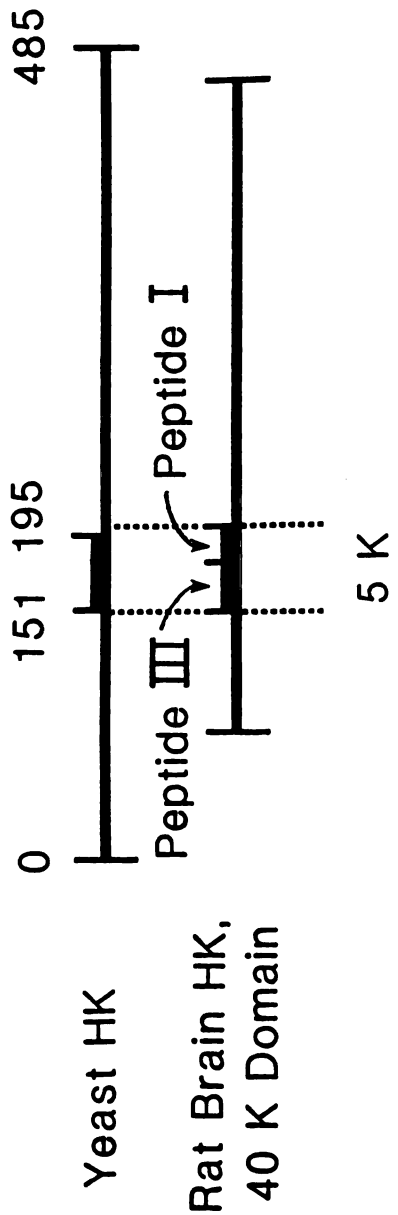

485
195
151
0
Peptide I
Peptide III
5 K
Yeast HK
Rat Brain HK,
40 K Domain

Figure 22

residues is only an approximation. As will be discussed below, this can readily be interpreted in light of the postulated evolutionary relationship and current knowledge of the structure of yeast hexokinase.

Proposed Locations of Peptides I and III within the Tertiary Structure of the 40 kDa Catalytic Domain of Rat Brain Hexokinase.

Prior to deduction of the yeast hexokinase amino acid sequence based on the nucleotide sequence of the cloned cDNA (37,38), Steitz and coworkers had attempted to determine this sequence using X-ray crystallographic methods (14). In subsequent discussions of their structural studies of yeast hexokinase, the Steitz group understandably referred to this proposed sequence, then the only one available. However, comparison of the sequence proposed on the basis of X-ray diffraction measurements (14) with that now known from nucleotide sequencing (37,38) discloses numerous discrepancies, some of which may be seen from the alignments shown in Fig. 21; in aligning these sequences, we have made insertions or deletions deemed appropriate for increasing agreement of the sequence proposed from X-ray data with the sequence determined from nucleotide sequencing. It will be apparent from Fig. 21 that the position of specific residues, expressed conventionally beginning from the N-terminus, will be different in the "X-ray" sequence and in the sequence deduced from the nucleotide sequence. This can obviously lead to confusion. In an attempt to minimize that

confusion in the following discussion, residue positions in the "X-ray" sequence will be given in quotes, whereas a position numbered without this indication may be considered to correspond to a residue as given in the sequence deduced from the cloned cDNA.

The three dimensional structure of yeast hexokinase has been determined with considerable resolution (14-16,33,36,120). Thus the location of those regions homologous to Peptides I and III can be identified within the yeast enzyme's tertiary structure. This is depicted in Fig. 23. It is immediately obvious that these peptides are in close proximity to the binding site for Glc, as proposed by Anderson et al. (14-16). That this is true for Peptides I and III themselves (in brain hexokinase) seems likely, considering the identification of these peptides based on their reactivity with a glucose analog. This close similarity in both sequence and function can reasonably be taken to indicate similarity in tertiary structure (121,122). Thus, the structure of the glucose binding site of the 40 kDa domain of rat brain hexokinase is very likely to closely approximate that of yeast hexokinase, as will be discussed later in more detail.

Location and Nature of "Critical" Sulfhydryls in Brain Hexokinase.

Results presented in Chapter I indicated that there was a single critical sulfhydryl in rat brain hexokinase that was modified by GlcNBrAc and which was protected against

Figure 23.

Region of yeast hexokinase active site to which Peptides I and III from rat brain hexokinase are homologous. The structure of the active site was redrawn from Fig. 6 of Anderson et. al. (120). The cleft, which runs from the center of the figure to the upper right hand corner, is in the open conformation. The regions which are homologous to Peptides III and I are the solid and cross-hatched regions, respectively. Although the entire length of Peptides I and III were not sequenced, the length of each peptide, as estimated by amino acid analysis or mass spectrometry, is indicated in the figure. The locations of asn-"188", asp-"189", asn-"215", asn-"245", and gln-"277" are indicated. These residues have been identified as playing essential roles in the binding of hexoses (16). The position of ser-"138", a highly conserved residue playing a critical role in substrate binding and conformational changes required for catalysis (see text), is shown, as is the position of residue "155" (R-155), predicted to be in a position to be labeled by an appropriate glucosamine derivative (17). GlcNAc is oriented in the Glc binding site in the same orientation as that determined for Glc (120). The ATP binding site is located at the far right hand side of the figure (33).

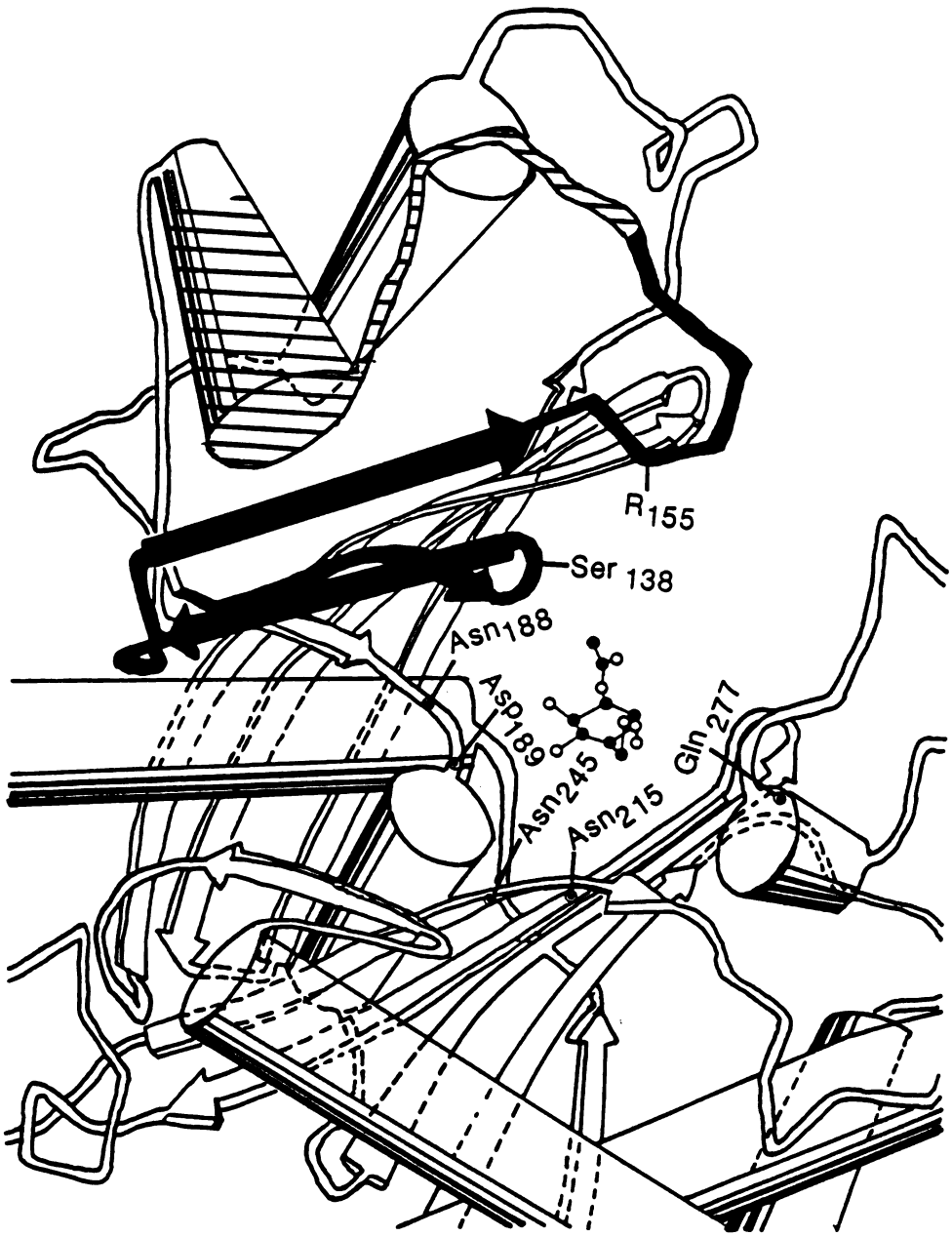


Figure 23

modification in the presence of Glc or GlcNAc, as had been found with the bovine Type I (11) and rat Type II (10) isozymes. It is evident from the comments above that the location of this sulfhydryl has not been defined by the present studies, and indeed even the existence of a single critical sulfhydryl may need to be questioned. Based on its proximity to the Glc binding site, it would seem that modification of a sulfhydryl within the sequence defined by Peptide III would be most likely to markedly affect catalytic activity.

Indeed, Bennett and Steitz (16) predicted that an appropriate glucosamine derivative could be used to label the residue at position "155", which is included in the region corresponding to the unsequenced portion of Peptide III. Residue "155" is located on a loop of the polypeptide backbone (Fig. 23) identified as a region of high mobility in yeast hexokinase (16). When the cleft is in the closed position, this loop has moved several Å closer in the direction of the bound Glc molecule. With the cleft closed, the β carbon of residue "155" is located within 5 Å of the 2-hydroxyl of the substrate hexose. Given that the bromoacetyl group of GlcNBrAc can rotate on a 4 Å arm (123), it is also possible that one of the residues adjacent to position "155" could be labeled. Since this region was predicted to be labeled by a glucosamine derivative, and such a glucosamine derivative labeled a homologous peptide in rat brain hexokinase which contains this region, it is

very likely that the derivatized sulfhydryl of Peptide III is located within this region.

An explanation of the kinetics of GlcNBrAc inactivation, as discussed in Chapter I, is possible if it assumed that the rat brain hexokinase Glc binding site is as depicted in Fig. 23 and the alkylated sulfhydryl corresponds to position "155" (or possibly an adjacent position). Alkylation of a sulfhydryl in this region of the molecule would block cleft closure, a step considered necessary for catalysis (124), and therefore inactivate the enzyme.

Furthermore, the dynamics of the cleft closure have the characteristics described in Chapter I of the conformational change induced by Glc, but not GlcNAc, which makes the critical sulfhydryl inaccessible to alkylation. As mentioned above, binding of Glc closes the cleft and residue "155" is buried next to the bound Glc molecule. This action would protect residue "155" from alkylation by GlcNAc, either by a specific or nonspecific mechanism. The X-ray crystallography data indicated, however, that Glc derivatives with bulky substituents at the number 2 position, such as GlcNAc, would impede cleft closure, and thus block catalysis (16,124). This observation supported previous demonstrations that such Glc derivatives were competitive inhibitors, but not substrates, of yeast hexokinase (125). The same is true for mammalian Type I hexokinases (54,102). With GlcNAc bound at the active site, and the cleft remaining in the open position, it appears

that residue "155" would remain open to the solvent and GlcNAc bound to the enzyme would not be in a position to block nonspecific alkylation.

Therefore, it seems likely, based on the X-ray crystallography data, the predicted structural homology between yeast and rat brain hexokinase, and the kinetics of GlcNBrAc inactivation, that the critical sulfhydryl in rat brain hexokinase is located at a position analogous to residue "155" in the yeast enzyme. Work is underway in this laboratory to deduce the rat brain hexokinase amino acid sequence based on the nucleotide sequence of the cloned cDNA. Completion of this project would determine the nature of the amino acid at the predicted position for the reactive sulfhydryl in Peptide III, and whether it indeed is the only cysteine in Peptide III.

Based on the structure shown in Fig. 23, it is unclear why a sulfhydryl in Peptide I might also be labeled by this reactive substrate analog, with protection by competing ligands such as Glc. In the open cleft form of the yeast enzyme (Fig. 23), this peptide segment is located at an appreciable distance from the postulated Glc binding site, and this remains the case even in the closed form of the enzyme (16). However, this is not unheard of, as there have clearly been other cases when affinity labeling has occurred at amino acids far removed from the targeted binding site (110,126). Alternatively, it may be that there is some variance in the structure in this region of the molecule,

and that in the rat brain enzyme the segment represented by Peptide I is in closer proximity to the Glc binding site. If this were the case, then the protective effects of Glc and GlcNAc on labeling of Peptide I would be understandable, and it might be expected that modification of the sulfhydryl in this peptide segment, as in Peptide III, could lead to decreased catalytic activity.

If the sulfhydryl in Peptide I is critical for activity, the possibility must be raised that GlcNBrAc labeled two sulfhydryls critical for activity. This suggests that the reasons for concluding there could only be a single critical sulfhydryl should be reevaluated. The previous conclusion that there was a single critical sulfhydryl was based on the assumption that modification of that sulfhydryl totally inactivated the enzyme. If the modified sulfhydryls in Peptides I and III were not directly involved in catalysis, then it is conceivable that their modification might result in only partial loss of catalytic activity, e.g., by hindering but not totally preventing access to the active site or closure of the cleft. Hence, the conclusion that there was a single critical sulfhydryl might have been incorrect because the assumption on which it was based was incorrect, the relationship between modification and loss of activity being more complex than assumed.

From the present work it is apparent that the critical sulfhydryl(s) in rat brain hexokinase can not be the same

critical sulfhydryl which is observed in yeast hexokinase. The yeast hexokinase sequence which corresponds to peptides I and III are totally devoid of cysteine. It also explains the inability of GlcNBrAc to inactivate the yeast enzyme (12). In view of the evidence that these hexokinases are fundamentally similar in their structure and catalytic mechanism, it is obvious that this absence of cysteines precludes their direct function in catalysis. This conclusion was reached earlier by Anderson et al. (15), who, on the basis of X-ray crystallography data, stated that sulfhydryls are unlikely to play a role in substrate binding or catalysis by the yeast enzyme.

Structural Homology Between Yeast and Rat Brain Hexokinases.

Despite the evidence that there is no conserved, essential cysteine in hexokinases, it is possible to make several comparisons between yeast hexokinase and the rat brain hexokinase 40 kDa domain and see that several structural features must be conserved between the two enzymes. The strongest evidence for this comes from the extensive homology between the N-terminal portion of Peptide III and the sequence around serine "138" (serine 157 in the corrected sequence) of the yeast enzyme (Fig. 21). X-ray studies (16) have indicated that Ser-"138" is involved in hydrogen bonding to the 6-hydroxyl group of Glc, and that this interaction is critical for the Glc-induced closure of a cleft necessary for catalysis (124). This serine is present in all hexokinases examined thus far (Fig. 21), both

yeast and mammalian, clearly implying functional importance. This is evidence that not only is the secondary and tertiary structure of the immediate region likely to be conserved, but also the predominant structural features of yeast hexokinase: a cleft formed by two lobes of the enzyme connected by a flexible "hinge" region.

If indeed it is the predicted C-terminal region of Peptide III which is labeled by GlcNBrAc, then this region of the cleft must also be similar between the two enzymes, as well as the orientation at which Glc is bound in the cleft.

Evidence that the ATP binding site in the 40 kDa domain of rat brain hexokinase is also structurally similar to that of the yeast enzyme can be inferred from recent work with the ATP analog affinity label, 8-azido-ATP (9). In yeast hexokinase, the 8-position of the bound ATP is in close proximity to the residue at position "393" (33), i.e., a residue located toward the C-terminal region of the yeast enzyme and thus expected to occupy a comparable position in the 40 kDa domain of the brain enzyme (Fig. 22). As expected, labeling by 8-azido-ATP was shown to occur in this region of the 40 kDa domain (9). Binding of ATP does not involve residues in the regions defined by Peptides I and III (33) and hence would be predicted not to interfere with inactivation by GlcNBrAc, and this has been found to be the case, as discussed in Chapter I.

Structural Relationships in Monomeric Mammalian Hexokinases Compared with the Dimeric Yeast Hexokinase.

That mammalian hexokinase evolved from an ancestral gene shared with that in yeast is evident from the present study. Marcus and Ureta (56) also found such evidence, as well as proof of gene duplication from the existence of two highly homologous peptides from the Type III isozyme (peptides 5 and 6, shown in Fig. 21). Based on results with the Type I isozyme in the present work and comparison with the yeast hexokinase, it is now recognized that these peptides include the highly conserved Ser-"138" that functions in binding of hexose moieties and closure of the cleft (16,124). The conservation of this residue in two homologous but nonidentical peptides implies that mammalian hexokinases have two structurally related functional clefts whose opening or closing fulfill essential catalytic or regulatory roles. The role of this residue in catalytic function, now associated with the C-terminal domain of mammalian hexokinase, has been discussed above.

According to the proposals for hexokinase evolution (17-23), duplication and fusion of an ancestral gene for hexokinase initially resulted in a fusion protein possessing duplicate catalytic sites. It is further proposed that one was retained for catalytic function while the other site evolved to serve a regulatory function, providing the basis for the potent allosteric regulation by Glc-6-P seen with the mammalian hexokinases but not with yeast hexokinase.

Identification of the 40 kDa domain at the C-terminus of brain hexokinase as serving catalytic function obviously implies that the modified catalytic site, now serving a regulatory function, must be associated with the N-terminal region of the enzyme, as discussed in Chapter I. According to this evolutionary scenario, the duplicate cleft can be expected in the N-terminal region, probably in the 50 kDa domain of brain hexokinase (4), and to be the site of binding for the allosteric effector, Glc-6-P. Binding of this ligand has been shown to induce closing of the cleft with the yeast enzyme (127).

Given this evidence that rat brain hexokinase evolved by gene duplication and fusion of a gene similar to that in yeast, it can be postulated that the subunit structure of the yeast dimer is conserved in the binding between the 40 kDa and 50 kDa domains of rat brain hexokinase. In the X-ray crystallographic structure of the yeast dimer (36), reproduced in Fig. 24, there are no close contacts between the C- and N-termini of either subunit, indicating that structural alterations would be required to join the two subunits as a single polypeptide chain. The present work provides evidence of such an alteration leading to a "fused dimer" structure in mammalian hexokinase.

As mentioned above and depicted in Fig. 22, alignment of homologous sequences from the region of the Glc binding sites in yeast hexokinase and in the 40 kDa catalytic domain of brain hexokinase suggests that approximately 80 amino

Figure 24.

Structure of yeast hexokinase dimer, proposed to approximate the structure of the 40 kDa and 50 kDa domains of rat brain hexokinase. This is a duplication of Fig. 13 of Steitz et al. (36). The subunits bind to each other nonsymmetrically. The "down" and "up" subunits are indicated. Both subunits have their N-termini at the top of the figure and their C-termini at the bottom. The region of the down subunit which is homologous to Peptides I and III is darkened. The numbering used in the figure was approximated from the total number of amino acids in the protein, and does not correspond to the sequence as determined by either X-ray crystallography (14) or sequencing of the cloned cDNA (37,38), as this figure was published prior to any of those sequence determinations. As discussed in the text, it is proposed that the rat brain hexokinase 40 kDa domain is analogous to the "down" subunit, minus the A and B helices, and the 50 kDa domain is analogous to the "up" subunit. The point at which the 40 kDa and 50 kDa domains are believed to be fused is indicated by a dotted line.

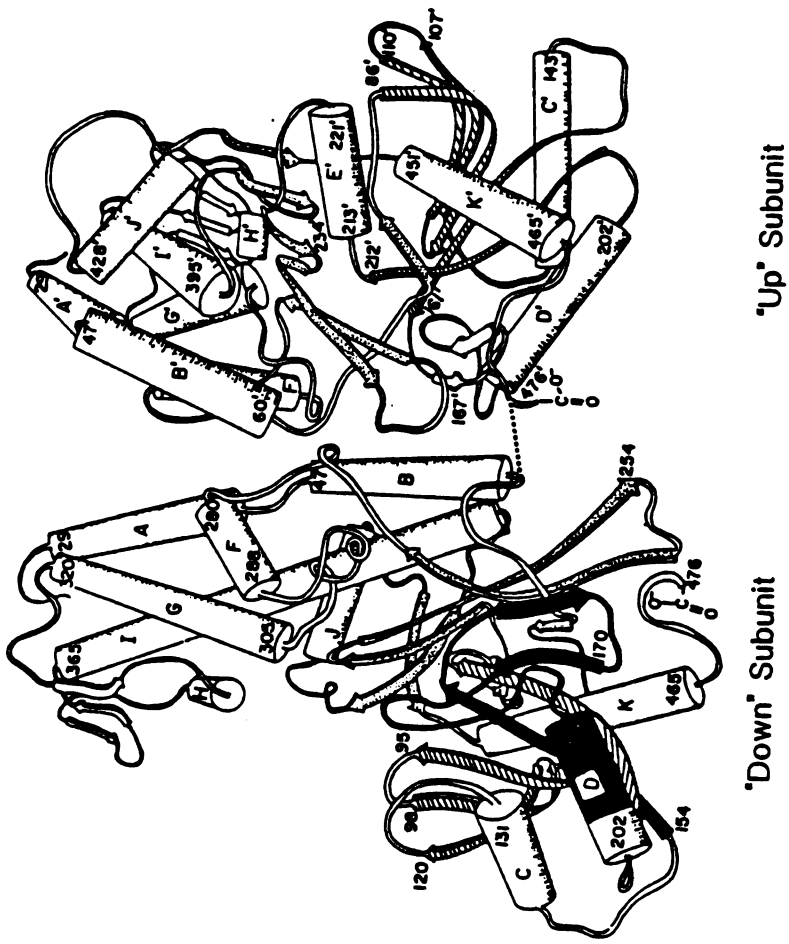


Figure 24

acid residues have been deleted in the course of the evolution of the 40 kDa domain from its ancestral 50 kDa precursor. The initial portion (approx. 25 residues) at the N-terminal region of yeast hexokinase could not be localized by X-ray diffraction, possibly due to localized disorder (36). Taking these residues into consideration, deletion of approximately 80 residues would correspond to loss of helices A and B, as defined in Fig. 13 of Steitz et al. (36), and reproduced here in Fig. 24. After loss of these residues, the resulting N-terminus in the "down" subunit is in the same region as the C-terminus of the "up" subunit and it is evident that fusion at this point could occur with little disruption of structure. This proposed site of fusion would also correspond to a "surface loop", generally considered as favorable sites for such events, and according to the analysis of Craik et al. (128) it is predicted that this may correspond to an intron-exon junction in the mammalian hexokinase gene. Though helix B provides some of the interactions maintaining the dimeric structure of the yeast enzyme (36), its loss is unlikely to be catastrophic since other interactions exist and might even be enhanced by loss of helix B.

Thus, the structure of dimeric yeast hexokinase may represent a very close approximation to the structure of mammalian hexokinase, with the "down" subunit of the yeast enzyme corresponding to the C-terminal catalytic domain, and the evolutionary descendant of the "up" subunit being the 50

kDa domain associated with the N-terminal half of the brain enzyme and serving a regulatory function in binding of the allosteric effector, Glc-6-P.

Binding of Glc-6-P to the cleft in the "regulatory" domain in the 50 kDa domain of mammalian hexokinases must reasonably be expected to have a significant influence on the structure and function of the C-terminal catalytic domain if regulation is to be mediated by Glc-6-P. There is already direct evidence that these regions of the molecule are linked by intense noncovalent interactions (4).

Previous data on the conformational states induced by Glc and Glc-6-P in conjunction with data from this work also allow predictions of the nature of these conformational changes. Of the 14 sulfhydryl groups in rat brain hexokinase, Glc and Glc-6-P protect six and ten sulfhydryls, respectively, from reaction with DTNB (62). Since the present study has shown that six sulfhydryls are located in the putative 50 kDa regulatory domain, binding of Glc-6-P to the 50 kDa domain must therefore alter the reactivity of <u>at least</u> half of the eight sulfhydryls in the catalytic 40 kDa domain. This clearly suggests a large conformational change in the 40 kDa domain. The most obvious conformational change which can take place in the 40 kDa domain is the closing of the cleft. This would concur with the ability of Glc-6-P to protect the critical sulfhydryl in the 40 kDa domain from labeling by GlcNBrAc by inducing a conformational change with the same effect as that induced

by Glc. Other evidence (53,54) has also indicated that the conformational changes induced by Glc and Glc-6-P are very similar. The ability of either Glc or Glc-6-P to induce a cleft closure in both domains would clearly explain the observed synergistic binding between the two ligands (53).

The gene fusion scheme, which is believed to have led to the "fused dimer" structure of Fig. 24, does not provide any information about the evolutionary background of the 10 kDa domain at the N-terminus of the 50 kDa domain. It is possible that the 10 kDa N-terminal domain originated from another protein which must also bind to porin on the outer mitochondrial membrane, such as has been suggested for mitochondrial glycerol kinase (129).

A description of the general rat brain hexokinase structure proposed in this study is diagramed in Fig. 25. Evidence presented here indicates that the 40 kDa domain contains the catalytic function, while the mitochondrial binding function has previously been assigned to the 10 kDa domain. It is apparent that the 50 kDa domain, similar in structure to the 40 kDa but proposed to contain the Glc-6-P binding site, is in contact with both domains whose activities it regulates, catalysis and mitochondrial binding. Hopefully, proof of the location of the Glc-6-P regulatory site within the 50 kDa domain will be forthcoming.

Figure 25. Generalized diagram of rat brain hexokinase structure. The three major domains and their molecular weights are indicated. The binding sites for Glc, ATP (P-P-P-Ad) and Glc-6-P are indicated. A second, low affinity ATP or ADP binding site is indicated by broken lines, which may represent the residual ATP-binding site described by Gregoriou et al. (22).

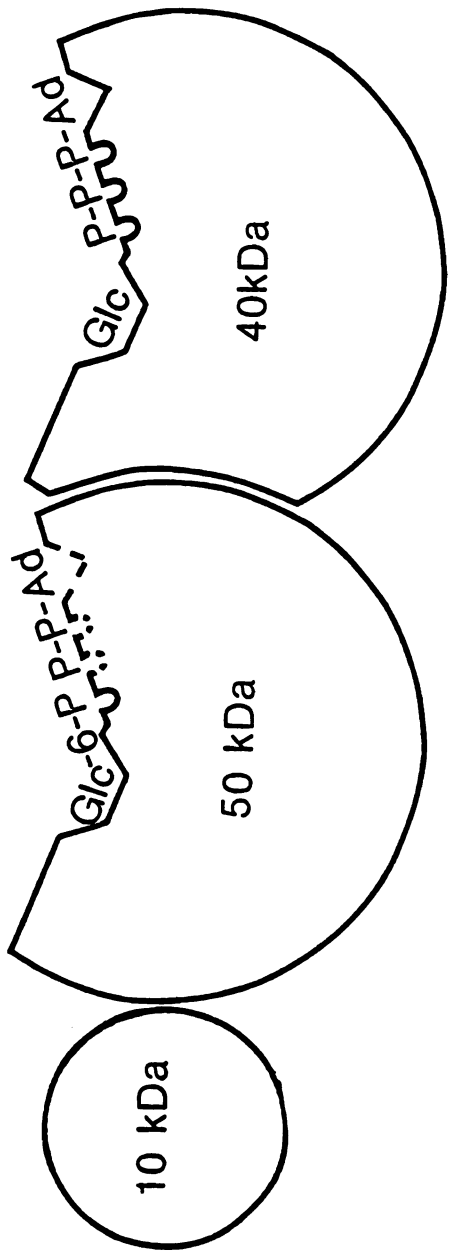


Figure 25

LIST OF REFERENCES

1. Sokoloff, L., (1977) in "Nutrition and the Brain" (R. J. Wurtman and J. J. Wurtman, eds.), Vol. 1, pp. 87-139. Raven, New York.

2. Bachelard, H. S., (1970) in "Handbook of Neurochemistry" (A. Lajtha, ed.), Vol. 4, pp. 1-12.

3. Chou, A. C., and Wilson, J. E. (1972) Arch. Biochem. Biophys. 151, 48-55.

4. Polakis, P. G., and Wilson, J. E. (1984) Arch. Biochem. Biophys. 234, 341-352.

5. Tucker, M. M., Robinson, J. B., Jr., and Stellwagen, E. (1981) J. Biol. Chem. 256, 9051-9058.

6. Gervais, M., Risler, Y., and Corazzin, S. (1983) Eur. J. Biochem. 130, 253-259.

7. Wilson, J. E., (1984) in Regulation of Carbohydrate Metabolism (Beitner, R., ed.) pp.45-85, CRC Press, Boca Raton, Florida.

8. Polakis, P. G., and Wilson, J. E. (1985) Arch. Biochem. Biophys. 236, 328-337.

9. Nemat-Gorgani, M., and Wilson, J. E. (1986) Arch. Biochem. Biophys. in press.

10. Connolly, B. A., and Trayer, I. P. (1979) Eur. J. Biochem. 93, 375-385.

11. Swarup, G., and Kenkare, U. W. (1980) Biochemistry 19, 4058-4064.

12. Otieno, S., Bhargava, A. K., Barnard, E. A., and Ramel, A. H. (1975) Biochemistry 14, 2403-2410.

13. Otieno, S., Bhargava, A. K., Serelis, D., and Barnard, E. A. (1977) Biochem. 16, 4249-4255.

14. Anderson, C. M., Stenkamp, R. E., and Steitz, T. A. (1978) J. Mol. Biol. 123, 15-33.

15. Anderson, C. A., Stenkamp, R. E., McDonald, R. C., and Steitz, T. A. (1978) J. Mol. Biol. 123, 207-219.

16. Bennett, Jr., W. S., and Steitz, T. A. (1980) J. Mol. Biol. 140, 211-230.

17. Colowick, S. P. (1973) in The Enzymes (P. D. Boyer, ed) 3rd ed., Vol. 9, pp. 1-48, Academic Press, New York.

18. Easterby, J. S., and O'Brien, M. J. (1973) Eur. J. Biochem. 38, 201-211.

19. Rose, I. A., Warms, J. V. B., and Kosow, D. P. (1974) Arch. Biochem. Biophys. 164, 729-735.

20. Holroyde, M. J., and Trayer, I. P. (1976) FEBS Lett. 62, 215-219.

21. Ureta, T. (1982) Comp. Biochem. Physiol. 71B, 549-555.

22. Gregoriou, M., Trayer, I. P., and Cornish-Bowden, A. (1983) Eur. J. Biochem. 134, 283-288.

23. Manning, T. A., and Wilson, J. E., (1984) Biochem. Biophys. Res. Commun. 118, 90-96.

24. Rustum, Y. M., Ramel, A. H., and Barnard, E. A. (1971) Prep. Biochem. 1, 309.

25. Ramel, A. H., Rustum, Y. M., Jones, J. G., and Barnard, E. A. (1971) Biochemistry 10, 3499-3508.

26. Schulze, I. T., and Colowick, S. P., (1969) J. Biol. Chem. 244, 2306-2316.

27. Rustum, Y. M., Massaro, E. J., and Barnard, E. A. (1971) Biochemistry 10, 3509-3516.

28. Shill, J. P. and Neet, K. E. (1975) J. Biol. Chem. 250, 2259-2268.

29. Steitz, T. A., Anderson, W. F., Fetterick, R. J., and Anderson, C. M., (1977) J. Biol. Chem. 252, 4494-4500.

30. Furman, T. C. and Neet, K. E. (1983) J. Biol. Chem. 258, 4930-4936.

31. Entian, K.-D. and Frohlich, K. U. (1984) J. Bacteriol. 158, 29-35.

32. Bennett, W. S., Jr., and Steitz, T. A. (1978) Proc. Natl. Acad. Sci. USA 75, 4848-4852.

33. Shoham, M., and Steitz, T. A. (1980) J. Mol. Biol. 140, 1-14.

34. Steitz, T. A., Fletterick, R. J., and Hwang, K. J. (1973) J. Mol. Biol. 78, 551-561.

35. Anderson, C. M., Zucker, F. H., and Steitz, T. A. (1979) Science 204, 375-380.

36. Steitz, T. A., Fletterick, R. J., Anderson, W. F., and Anderson, C. M. (1976) J. Mol. Biol. 104, 197-222.

37. Stachelek, C., Stachelek, J., Swan, J., Botstein, D., and Konigsberg, W. (1986) Nucl. Acids Res. 14, 945-963.

38. Kopetzki, E., Entian, K.-D., and Mecke, D. (1985) Gene 39, 95-102.

39. Gonzalez, C., Ureta, T., Babul, J., Rabajille, E., and Niemeyer, H. (1967) Biochemistry 6, 460-468.

40. Katzen, H. M., Soderman, D. D., and Nitowsky, H. M. (1965) Biochem. Biophys. Res. Commun. 19, 377-382.

41. Katzen, H. M. and Schimke, R. T. (1965) Proc. Natl. Acad. Sci. USA 54, 1218-1225.

42. Lawrence, G. M. and Trayer, I. P. (1984) Comp. Biochem. Physiol. 79B, 233-238.

43. Lawrence, G. M., Walker, D. G., and Trayer, I. P. (1983) Biochim. Biophys. Acta 743, 219-225.

44. Crane, R. K., and Sols, A. (1954) J. Biol. Chem. 210, 597-606.

45. Fromm, H. J., (1981) in "The Regulation of Carbohydrate Formation and Utilization in Mammals", (C. M., Veneziale, ed.), University Park Press, Baltimore, p.45.

46. Felgner, P. L., Messer, J. L., and Wilson, J. E. (1979) J. Biol. Chem. 254, 4946-4949.

47. Gots, R. E. and Bessman, S. P. (1972) Biochem. Biophys. Res. Commun. 49, 1249-1255.

48. Inui, M. and Ishibashi, S. (1979) J. Biochem. 85, 1151-1156.

49. Rose, I. A. and Warms, J. V. B. (1967) J. Biol. Chem. 242, 1635-1645.

50. Kurokawa M., Tokuoka, S., Oda, S., Tsubotani, E., Ishibashi, S. (1981) Biochem. Int. 2, 645-650.

51. Wilson, J. E. (1968) J. Biol. Chem. 243, 3640-3647.

52. Wilson, J. E. (1978) Trends Biochem. Sci. 3, 124-125.

53. Wilson, J. E. (1979) Arch. Biochem. Biophys. 196, 79-87.

54. Wilson, J. E. (1978) Arch. Biochem. Biophys. 185, 88-99.

55. Wilson, J. E. (1973) Arch. Biochem. Biophys. 159, 543-549.

56. Marcus, F., and Ureta, T. (1986) Biochem. Biophys. Res. Commun. 139, 714-719.

57. McCarthy, A. P. and Hardie, D. G. (1984) Trends Biochem. Sci. 9, 60-63.

58. Lynn, R., Giaever, G., Swanberg, S. L., and Wang, J. C. (1986) Science 233, 647-649.

59. Palm, D., Goerl, R., and Burger, K. J. (1985) Nature 313, 500-502.

60. Poorman, R. A., Randolph, A., Kemp, R. G., and Heinrikson, R. L. (1984) Nature 31, 467-469.

61. Jones, J. G., Otieno, S., Barnard, E. A. and Bhargava, A. K. (1975) Biochemistry 14, 2396-2403.

62. Chou, A. C., and Wilson, J. E. (1974) Arch. Biochem. Biophys. 163, 191-199.

63. Redkar, V. D. and Kenkare, U. W. (1972) J. Biol. Chem. 247, 7576-7584.

64. Redkar, V. D. and Kenkare, U. W. (1975) Biochemistry 14, 4704-4712.

65. Monasterio, O., Orellana, O., Heberlein, U., Niemeyer, H. (1981) Arch. Biol. Med. Exp. 14, 139-142.

66. Niemeyer, H., Heberlein, U., Monasterio, O. (1981) Arch. Biol. Med. Exp. 14, 177-182.

67. Ray, W. J., and Koshland, D. E., Jr. (1961) J. Biol. Chem. 236, 1973-1979.

68. Magnani, M., Stocchi, V., Ninfali, P., Dacha, M., Fornaini, G. (1980) Biochim. Biophys. Acta 615, 113-120.

69. Magnani, M., Stocchi, V., Dacha, M., Fornaini, G. (1984) Biochim. Biophys. Acta 804, 145-153.

70. Rose, I. A. and Warms, J. V. B. (1982) Arch. Biochem. Biophys. 213, 625-634.

71. Darby, M. K., Connolly, B. A., and Trayer, I. P. (1981) Arch. Biol. Med. Exp. 14, 163-170.

72. Subbarao, B. and Kenkare, U. W. (1977) Arch. Biochem. Biophys. 181, 19-25.

73. Batra, S. P. and Colman, R. F. (1986) Biochemistry 25, 3508-3515.

74. King, G. C., Martin, C. T., Pham, T. T., and Coleman, J. E. (1986), Biochemistry 25, 36-40.

75. Gonzalez, D. H., Iglesias, A. A., and Andreo, C. S. (1986) Arch. Biochem. Biophys. 245, 179-186.

76. Grubmeyer, C. T. and Cray, W. R. (1986) Biochemistry 25, 4778-4784.

77. Alter, S. C. and DeLuca, M. (1986) Biochemistry 25, 1599-1605.

78. Marshall, M. and Fahien, L. A. (1985) Arch. Biochem. Biophys. 241, 200-214.

79. Carlson, G. M., Colombo, G., and Lardy, H. A. (1971) Biochemistry 17, 5329-5338.

80. Ray, M. and Bhaduri, A. (1980) J. Biol. Chem. 255, 10777-10781.

81. Gavilares, F., Peterson, D., Bullis, B., and Schirch, L. (1983) J. Biol. Chem. 258, 13155-13159.

82. Olson, S. T. and Massey, V. (1980) Biochemistry 19, 3137-3144.

83. Smith, D. J. and Kenyon, G. L. (1974) J. Biol. Chem. 249, 3317-3318.

84. Gilbert, H. F. (1982) J. Biol. Chem. 257, 12086-12091.

85. Nakashima, K., Pontremoli, S., and Horecker, B. L. (1969) Proc. Natl. Acad. Sci. USA 64, 947-951.

86. Tschesche, H. and MaCartney, H. W. (1981) Eur. J. Biochem 120, 183-190.

87. Pontis, H. G., Babio, J. R., and Salerno, G. (1981) Proc. Natl. Acad. Sci. USA 78, 6667-6669.

88. Doehlert, D. C. and Huber, S. C. (1985) Biochim. Biophys. Acta 830, 267-273.

89. Polakis, P. G., and Wilson, J. E. (1982) Biochem. Biophys. Res. Commun. 107, 937-943.

90. Decker, L. A., ed. (1977) Worthington Enzyme Manual, p.221, Worthington Biochemical Corp., Freehold, NJ.

91. Felgner, P. L., and Wilson, J. E. (1976) Anal. Biochem. 74, 631-635.

92. Chamberlain, J. P. (1979) Anal. Biochem. 98, 132-135.

93. Wilson, J. E. (1973) Arch. Biochem. Biophys. 154, 332-340.

94. Fox, J. (1984) Linear Statistical Models and Related Models, pp. 204-213, John Wiley & Sons, New York.

95. Ellman, G. L. (1959) Arch. Biochem. Biophys. 82, 70-77.

96. Crestfield, A. M., Moore, S., and Stein, W. H. (1963) J. Biol. Chem. 238, 622-627.

97. Dirnhuber P., and Schutz, F. (1948) Biochem. J. 42, 628-632.

98. Stark, G. R. (1967) Methods. Enzym. 11, 590-594.

99. Hagar, D. A. and Burgess, R. R. (1980) Anal. Biochem. 109, 76-86.

100. Henderson, L. E., Oroszlan, S., and Konigsberg, W. (1979) Anal. Biochem. 93, 153-157.

101. Gundlach, H. G., Stein, W. H., and Moore, S. (1959) J. Biol. Chem. 234, 1754-1760.

102. Sols, A., and Crane, R. K. (1954) J. Biol. Chem. 210, 581-595.

103. Grossbard, L., and Schimke, R. T. (1966) J. Biol. Chem. 241, 3546-3560.

104. Kitz, R., and Wilson, I. B. (1962) J. Biol. Chem. 237, 3245-3249.

105. Dixon, M. (1953) Biochem. J. 55, 170-171.

106. Chou, A. C., and Wilson, J. E. (1974) Arch. Biochem. Biophys. 165, 628-633.

107. Solheim, L. P., and Fromm, H. J. (1980) Biochemistry 19, 6074-6080.

108. Lindley, H. (1960) Biochem. J. 74, 577-584.

109. Wilchek, M. and Givol, D. (1977) Methods. Enzym. 46, 153-157.

110. Plapp, B. V. (1982) Methods Enzym. 87, 469-499.

111. Gurd, F. R. N. (1972) Methods Enzym. 25. 424-438.

112. Wilson, J. E. (1982) Arch. Biochem. Biophys. 218, 254-262.

113. Plummer, T. H. and Lawson, W. B. (1966) J. Biol. Chem. 241, 1648-1650.

114. Tsou, C. -L. (1962) Sci. Sin. 11, 1535-1558.

115. Wilson, J. E., and Smith, A. D. (1985) J. Biol. Chem. 260, 12838-12843.

116. Kasper, C. B. (1975) in Protein Sequence Determination (S. B. Needleman, ed.), p. 132, Springer-Verlag, New York.

117. Parkinson, A., Ryan, D. E., Thomas, P. E., Jerina, D. M., Sayer, J. M., van Bladeren, P. J., Haniu, M., Shively, J. E., and Levin, W. (1986) J. Biol. Chem. 261, 11478-11486.

118. Weber, K. and Osborn, M. (1969) J. Biol. Chem. 244, 4406-4412.

119. Dunker, A. K. and Rueckert, R. R. (1969) J. Biol. Chem. 244, 5074-5080.

120. Anderson, C. M., McDonald, R. C, and Steitz, T. A. (1978) J. Mol. Biol. 123, 1-13.

121. Weber, I. T., Takio, K., Titani, K., and Steitz, T. A. (1982) Proc. Natl. Acad. Sci. USA 79, 7679-7683.

122. Rossman M. G., Liljas, A., Branden, C.-I., and Banaszak, L. J. (1975) in The Enzymes (P. D. Boyer, ed.), 3rd ed., Vol. 11, pp. 61-102, Academic Press, New York.

123. Warren, J. C., and Mueller, J. R. (1977) Methods. Enzym. 46, 447-460.

124. Shoham, M., and Steitz, T. A. (1982) Biochim. Biophys. Acta. 705, 380-384.

125. Maley, F., and Lardy, H. A. (1955) J. Biol. Chem. 214, 765-773.

126. Cheng, S.-Y., Wilchek, M., Cahnmann, H. J., and Robbins, J. (1977) J. Biol. Chem. 252, 6076-6081.

127. McDonald, R. C., Steitz, T. A., and Engleman, D. M. (1979) Biochem. 18, 338-342.

128. Craik, C. S., Rutter, W. J., and Fletterick, R. (1983) Science 220, 1125-1129.

129. Kaneko, M., Kurokawa, M., and Ishibashi, S. (1985) Arch. Biochem. Biophys. 237, 135-141.