Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 2563–2567, April 1989
Biochemistry

Complete amino acid sequence of rat brain hexokinase, deduced from the cloned cDNA, and proposed structure of a mammalian hexokinase

(yeast hexokinase/homology/evolution/protein structure)

DAVID A. SCHWAB AND JOHN E. WILSON

Biochemistry Department and the Neuroscience Program, Michigan State University, East Lansing, MI 48824

Communicated by Irwin A. Rose, December 22, 1988

ABSTRACT **The complete amino acid sequence for the type I isozyme of hexokinase from rat brain has been deduced from the nucleotide sequence of cloned cDNA. The nucleotide sequence of 91 bases in the 5′ untranslated region as well as that of the entire 3′ untranslated region preceding the poly(A) sequence have also been determined. The N- and C-terminal halves of brain hexokinase show extensive sequence similarity to each other and to yeast hexokinase. These results provide direct support for the proposal that the mammalian hexokinases of ≈100 kDa have evolved by a process of duplication and fusion of a gene encoding an ancestral hexokinase similar to the yeast enzyme of ≈50 kDa. Taking this similarity in sequence to indicate basic similarity in structure between the N- and C-terminal regions of brain hexokinase and the yeast enzyme, a proposed structure for the mammalian hexokinase has been developed by fusing two molecules of yeast hexokinase, whose structure has previously been determined by x-ray crystallographic studies. Various features of the model are shown to be consistent with experimental observations bearing on the structure of the brain enzyme.**

Hexokinase (ATP:D-hexose 6-phosphotransferase, EC 2.7.1.1) catalyzes the phosphorylation of glucose using MgATP as phosphoryl donor. In mammalian tissues, there are three isozymes of "low K_m" hexokinase, commonly referred to as types I, II, and III (reviewed in ref. 1). These isozymes have several properties in common, including a K_m for glucose in the submillimolar range and a marked sensitivity to inhibition by the product glucose 6-phosphate; the latter is generally accepted as being a major factor in regulation of hexokinase activity, although additional factors may be involved (1). Isozymes I–III are also similar in that they consist of a single polypeptide chain of ≈100 kDa. These properties contrast with those of a fourth type of hexokinase found in mammalian liver and pancreas, called type IV or, more commonly, "glucokinase." The latter enzyme is similar to the hexokinase found in yeast, consisting of a single polypeptide chain of ≈50 kDa and being insensitive to inhibition by physiologically relevant levels of glucose 6-phosphate. Such comparisons led several investigators (2–7) to suggest that the mammalian isozymes I–III might have evolved by a process involving duplication and fusion of a gene encoding an ancestral hexokinase similar to present-day glucokinase and yeast hexokinases.

According to this scenario, one of the catalytic sites in the fusion protein evolved to take on regulatory function while the other retained its catalytic role. The location of catalytic function in the C-terminal half (8, 9) and the regulatory site for glucose 6-phosphate in the N-terminal half (10) of the type I isozyme from rat brain is consistent with this proposal. Furthermore, there is now direct evidence for similarity in the amino acid sequence of mammalian and yeast hexokinases (11–13). Two groups (14, 15) have cloned the genes for yeast isozymes A and B, which differ only slightly in their deduced amino acid sequence. More recently, we have cloned and sequenced a cDNA encoding the entire C-terminal half of the rat brain enzymes (13), in which catalytic function resides (8, 9). Comparison of the deduced amino acid sequence with that of yeast hexokinase demonstrated extensive similarities in sequence, including retention of several residues that, based on x-ray crystallographic studies of the Steitz laboratory (16), are thought to be important in binding of the substrate glucose. Also conserved are several residues proposed to be involved in binding of MgATP (13).

Here we describe the isolation and sequencing of cDNA clones providing the entire sequence of rat brain hexokinase.* Thus, it has become possible to directly test another prediction based on the proposed evolutionary pathway—namely, that extensive sequence similarity will exist between the N- and C-terminal regions of a "low K_m" mammalian hexokinase. This is shown to be the case. Perhaps more importantly, the similarity in sequence can reasonably be presumed to indicate an overall conservation of secondary and tertiary structure (17–20). We propose a specific structure for the mammalian enzyme based on that of yeast hexokinase (16, 21–23). Certain existing experimental observations are shown to be consistent with the proposed structure, which should prove useful in planning and analysis of future experiments aimed at elucidation of structure–function relationships in mammalian hexokinases.

MATERIALS AND METHODS

Total RNA was isolated from adult rat brains and poly(A) mRNA selected on oligo(dT)-cellulose as described (24). A rat brain cDNA library was constructed in λgt10 following the procedure of DeWitt and Smith (25). Northern blot analysis and plaque hybridizations were performed as described by Maniatis *et al.* (26) using as probe the previously described (13) clone, designated HKI 12.4-4 and corresponding to the sequence in the C-terminal half of the molecule. DNA was isolated from positive recombinant phage (27) and cloned into M13mp18 for sequencing. Nonrandom deletions were generated as described by Henikoff (28, 29) and were sequenced by the dideoxynucleotide chain-termination method (30).

Rat brain hexokinase was subjected to limited digestion with trypsin as described by Polakis and Wilson (31). Under these conditions, in which native structure of the enzyme is preserved, cleavage occurs primarily at two sites, previously designated T_1 and T_2; this gives rise to fragments of 10, 50, and 40 kDa, which correspond to extreme N-terminal, more

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*The sequence reported in this paper is being deposited in the EMBL/GenBank data base (accession no. J04526).

central, and C-terminal regions, respectively. Partial cleavage intermediates of 90 kDa (50-plus 40-kDa fragments) and 60 kDa (10-plus 50-kDa fragments) are also seen (31). Alternatively, hexokinase was partially denatured in 0.6 M guanidine hydrochloride followed by limited digestion with trypsin as described by White and Wilson (10); under the latter conditions, cleavage occurs at a new site, designated T_3, which is located near the middle of the molecule and gives rise to fragments of 52 and 48 kDa, corresponding to N- and C-terminal halves, respectively. Tryptic cleavage products were separated by SDS/PAGE (31), fragments of interest were isolated by electroelution, and sequences were determined by automated Edman degradation at either the Protein Structure Laboratory (University of California, Davis) or the Macromolecular Structure Facility (Michigan State University).

Computer modeling and secondary structure analysis of yeast hexokinase were done at the computational chemistry facility of Upjohn (Kalamazoo, MI), using the software package MOSAIC and based on the "closed" structure of the yeast enzyme (16, 22, 23) using x-ray crystallographic coordinates from the Brookhaven Protein Data Bank.

RESULTS AND DISCUSSION

Two cDNA clones are of interest in the present context. One, designated HKI 1.4-7, contains a 3.7-kilobase (kb) insert that was found to start 91 bases upstream from the translation initiation codon and to extend 32 bases past the 3′ end of the previously isolated and sequenced clone, HKI 12.4-4 (13). HKI 1.4-7 therefore includes all of the coding sequence plus untranslated sequence at both 5′ and 3′ ends. It was sequenced in its entirety, and in both directions, with the sequencing strategy depicted in Fig. 1. A second clone, designated HKI 1.1 and containing a 2-kb insert, was found to extend an additional 13 bases beyond the 3′ end of HKI 1.4-7 followed by a sequence of 27 adenine residues. The latter clone was sequenced from its 3′ end, providing the 40-base sequence not included in HKI 1.4-7 followed by >500 bases, whose sequence agreed perfectly with that determined from bidirectional sequencing of HKI 1.4-7. The nucleotide sequence obtained from these clones accounts for 3.7 kb of the 4.3-kb mRNA detected by Northern blotting (data not shown).

The nucleotide sequence determined from HKI 1.4-7 and HKI 1.1 is shown in Fig. 2, along with the deduced amino acid sequence. Several segments of the deduced sequence match that determined by direct sequencing of the protein or fragments thereof. Starting with the initiating methionine, a

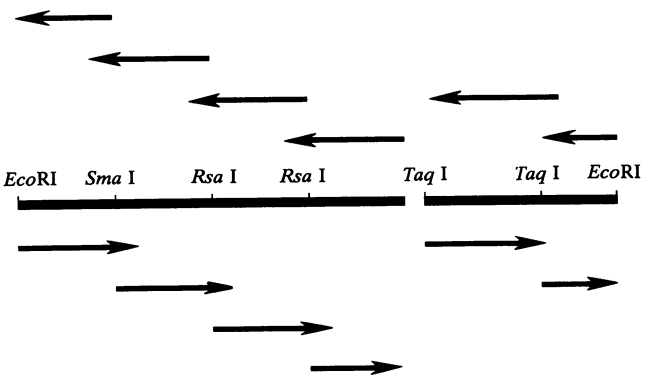

FIG. 1. Sequencing strategy with relevant restriction sites for clone HKI 1.4-7. The 3.7-kb insert was cloned in the *Eco*RI site of λgt10. Arrows indicate direction and extent of sequencing of subcloned fragments, including nonrandom deletions generated according to Henikoff (28, 29).

9-residue segment is in good agreement with the N-terminal sequence determined by Polakis and Wilson (33); the only discrepancy is the deduced Ala-Ala-Gln, which could not be unambiguously identified (for technical reasons) by Polakis and Wilson and which had been reported as (Ala, Gln)-Ala. Another 20-residue segment, corresponding to positions 102–121 in the deduced sequence, agrees completely with the sequence determined directly from the 50-kDa and 90-kDa (which includes the 50-kDa fragment as its N-terminal segment) fragments produced by tryptic digestion under the conditions of Polakis and Wilson (31). Residues at positions 463–471 in the deduced sequence are identical to a 9-residue sequence determined directly from the 48-kDa fragment produced by tryptic cleavage at T_3 under the partially denaturing conditions of White and Wilson (10). In addition, as has been noted (13), the amino acid sequence deduced from the nucleotide sequence of HKI 12.4-4, which is included within the sequence of HKI 1.4-7, also shows complete agreement with that obtained by direct sequencing of other tryptic fragments derived from the C-terminal half of rat brain hexokinase and with the C-terminal sequence of the enzyme. Finally, immediately adjacent to the N terminus of each of these tryptic fragments is an arginine or lysine residue, as is to be expected considering the specificity of this protease. In short, the agreement between deduced and directly determined amino acid sequence has been demonstrated at several segments ranging from N terminus to C terminus of the enzyme. These results confirm the authenticity of these clones and thus that the deduced sequence is that of the type I isozyme of rat hexokinase.

Alignment of the deduced amino acid sequences for the N- and C-terminal halves of brain hexokinase and that of the yeast hexokinase isozymes A and B is presented in Fig. 3. The alignment was carried out by first matching regions in which similarities were unequivocal and then aligning intervening regions while attempting to minimize gaps. Although the resulting alignment is necessarily somewhat arbitrary, we believe it to be conservative and reasonable. It surely is abundantly clear that there is extensive sequence similarity between the N- and C-terminal halves of the brain enzyme and between these and the yeast hexokinase isozymes. In the alignment shown in Fig. 3, 49% of the amino acid residues are identical and 17% are conservative substitutions when the N- and C-terminal sequences are compared. Such remarkable internal sequence similarity, in conjunction with the obvious similarity to yeast hexokinase, clearly supports the proposal (2–7) that this mammalian hexokinase represents the product of duplication and fusion of a gene encoding an ancestral hexokinase of ≈50 kDa.

Secondary structural features of yeast hexokinase were identified by the algorithm intrinsic to the MOSAIC program; these are shown beneath the corresponding amino acid sequences in Fig. 3. Several of the insertions or deletions that have occurred in the course of evolution of yeast and mammalian hexokinases from their common ancestor correspond to surface regions that are devoid of, or near the ends of, identified secondary structure. Frequently, such changes can be accommodated with minimal disruption of overall structure (17, 34). However, we note certain insertions/deletions that seem particularly likely to result in substantial differences between the structures of the yeast and mammalian enzymes: the absence in the mammalian enzyme of residues corresponding to residues 254–260 and 439–445 of yeast hexokinase and the presence of a 6-residue segment in both N- and C-terminal halves (residues 404–409 and 852–857, respectively) of the mammalian enzyme, which has no counterpart in yeast hexokinase. All of these segments are located in or near the hinge region linking the large and small lobes of yeast hexokinase (16), and the relative disposition of

```
                                     CGCCGATCTGCCGCTGGAGGACCACTGCTCACCAGGGCTACTGAGGAGCCACTGGCCCCACACCTGCTTTTCCGCATCCCCCACCGTCAGC   91
ATG ATC GCC GCG CAA CTA CTG GCC TAT TAC TTC ACC GAG CTG AAG GAT GAC CAA GTC AAA AAG ATT GAC AAG TAT CTG TAC GCC ATG CGG  181
Met Ile Ala Ala Gln Leu Leu Ala Tyr Tyr Phe Thr Glu Leu Lys Asp Asp Gln Val Lys Lys Ile Asp Lys Tyr Leu Tyr Ala Met Arg   30
CTC TCT GAT GAG ATT CTG ATA GAT ATC CTG ACA CGA TTC AAG AAA GAG ATG AAG AAT GGC CTC TCC CGG GAT TAT AAT CCA ACA GCC TCC  271
Leu Ser Asp Glu Ile Leu Ile Asp Ile Leu Thr Arg Phe Lys Lys Glu Met Lys Asn Gly Leu Ser Arg Asp Tyr Asn Pro Thr Ala Ser   60
GTC AAG ATG CTG CCC ACC TTG CTC CGG TCC ATT CCG GAC GGC TCA GAA AAG GGG GAT TTC ATT GCC CTG GAT CTC GGC GGG TCT TCC TTT  361
Val Lys Met Leu Pro Thr Leu Leu Arg Ser Ile Pro Asp Gly Ser Glu Lys Gly Asp Phe Ile Ala Leu Asp Leu Gly Gly Ser Ser Phe   90
CGA ATC CTG CGG GTG CAG GTG AAC CAC GAG AAG AAC CAG AAC GTC AGC ATG GAG TCT GAG ATC TAC GAC ACC CCA GAG AAC ATC GTG CAT  451
Arg Ile Leu Arg Val Gln Val Asn His Glu Lys Asn Gln Asn Val Ser Met Glu Ser Glu Ile Tyr Asp Thr Pro Glu Asn Ile Val His  120
GGC AGT GGA ACC CAG CTT TTC GAT CAT GTC GCT GAC TGC CTG GGA GAC TTC ATG GAG AAA AAG AAG ATC AAG GAC AAG AAG TTA CCC GTG  541
Gly Ser Gly Thr Gln Leu Phe Asp His Val Ala Asp Cys Leu Gly Asp Phe Met Glu Lys Lys Lys Ile Lys Asp Lys Lys Leu Pro Val  150
GGA TTC ACA TTT TCC TTC CCC TGC CGA CAA TCC AAG ATA GAT GAG GCT GTA CTG ATC ACG TGG ACA AAG CGG TTC AAA GCC AGT GGC GTG  631
Gly Phe Thr Phe Ser Phe Pro Cys Arg Gln Ser Lys Ile Asp Glu Ala Val Leu Ile Thr Trp Thr Lys Arg Phe Lys Ala Ser Gly Val  180
GAA GGA GCG GAT GTG GTC AAG TTG CTG AAT AAA GCC ATT AAG AAG CGA GGG GAC TAT GAT GCT AAC ATT GTC GCC GTG GTG AAT GAC ACA  721
Glu Gly Ala Asp Val Val Lys Leu Leu Asn Lys Ala Ile Lys Lys Arg Gly Asp Tyr Asp Ala Asn Ile Val Ala Val Val Asn Asp Thr  210
GTA GGG ACC ATG ATG ACC TGC GGT TAT GAT GAC CAA CAG TGT GAA GTC GGC CTG ATC ATT GGC ACA GGC ACC AAT GCT TGC TAC ATG GAG  811
Val Gly Thr Met Met Thr Cys Gly Tyr Asp Asp Gln Gln Cys Glu Val Gly Leu Ile Ile Gly Thr Gly Thr Asn Ala Cys Tyr Met Glu  240
GAA CTG CGA CAC ATC GAC CTG GTG GAA GGC GAC GAG GGG AGG ATG TGT ATT AAC ACG GAA TGG GGA GCC TTT GGG GAT GAT GGG TCC CTG  901
Glu Leu Arg His Ile Asp Leu Val Glu Gly Asp Glu Gly Arg Met Cys Ile Asn Thr Glu Trp Gly Ala Phe Gly Asp Asp Gly Ser Leu  270
GAA GAC ATC CGA ACC GAG TTT GAC AGA GAG TTA GAC CGT GGA TCT CTC AAC CCT GGG AAG CAG CTG TTC GAG AAG ATG GTG AGC GGC ATG  991
Glu Asp Ile Arg Thr Glu Phe Asp Arg Glu Leu Asp Arg Gly Ser Leu Asn Pro Gly Lys Gln Leu Phe Glu Lys Met Val Ser Gly Met  300
TAC ATG GGG GAG CTG GTC CGG CTA ATC CTG GTG AAG ATG GCC AAG GAA GGC CTC TTA TTC GAA GGG CGC ATC ACT CCA GAG CTG CTC ACG 1081
Tyr Met Gly Glu Leu Val Arg Leu Ile Leu Val Lys Met Ala Lys Glu Gly Leu Leu Phe Glu Gly Arg Ile Thr Pro Glu Leu Leu Thr  330
AGG GGA AAG TTC AAC ACT AGT GAC GTG TCC GCC ATT GAA AAG GAT AAG GAA GGC ATT CAA AAT GCC AAG GAA ATC TTA ACC CGC TTG GGA 1171
Arg Gly Lys Phe Asn Thr Ser Asp Val Ser Ala Ile Glu Lys Asp Lys Glu Gly Ile Gln Asn Ala Lys Glu Ile Leu Thr Arg Leu Gly  360
GTG GAG CCG TCT GAT GTT GAC TGT GTG TCG GTC CAG CAC ATC TGC ACG ATC GTC TCC TTC CGA TCA GCC AAC CTG GTG GCC GCC ACG CTC 1261
Val Glu Pro Ser Asp Val Asp Cys Val Ser Val Gln His Ile Cys Thr Ile Val Ser Phe Arg Ser Ala Asn Leu Val Ala Ala Thr Leu  390
GGT GCC ATC TTG AAC CGC CTG CGG GAC AAC AAG GGC ACA CCC AGC CTG CGG ACC ACG GTT GGC GTG GAC GGT TCT CTC TAC AAG ATG CAC 1351
Gly Ala Ile Leu Asn Arg Leu Arg Asp Asn Lys Gly Thr Pro Ser Leu Arg Thr Thr Val Gly Val Asp Gly Ser Leu Tyr Lys Met His  420
CCA CAG TAC TCC CGG CGG TTC CAC AAG ACC CTG AGG CGG GTG GTG CCT GAC TCC GAC GTC CGT TTC CTC CTC TCA GAG AGT GGC ACG GGC 1441
Pro Gln Tyr Ser Arg Arg Phe His Lys Thr Leu Arg Arg Val Val Pro Asp Ser Asp Val Arg Phe Leu Leu Ser Glu Ser Gly Thr Gly  450
AAG GGG GCC GCC ATG GTG ACG GCA GTA GCC TAC CGC CTG GCT GAG CAG CAC /1490 - 3597/ TTTAGTGAGCCATTGTTGTACGTCTAGTAAACTTTGT 3634
Lys Gly Ala Ala Met Val Thr Ala Val Ala Tyr Arg Leu Ala Glu Gln His

ACTGATTCAAAAAAAAAAAAAAAAAAAAAAAAAAA  3669
```

FIG. 2. Nucleotide sequence for clone HKI 1.4-7, with additional 3′ sequence obtained from clone HKI 1.1. Not shown is the sequence for nucleotides 1490–3597, which correspond to nucleotides at positions 4–2111 in the previously reported (13) sequence of clone HKI 12.4-4 and encode the C-terminal half of the molecule. (The first three nucleotides in the previously reported sequence were derived from the *Eco*RI linker but had been mistakenly indicated as part of the coding sequence for hexokinase, an error corrected in the present report.) The deduced amino acid sequence for the N-terminal half of the molecule is indicated below the nucleotide sequence; regions corresponding to sequence determined directly from the protein are underlined. The presumed polyadenylylation signal (32) in the 3′ region is also underlined.

the latter would surely be significantly affected by the noted differences.

Residues corresponding to the catalytically important (16, 21–23, 35) Ser-158, Asn-210, Asp-211, Glu-269, and Glu-302 of yeast hexokinase, as well as residues proposed to be important in binding of MgATP (13), are all found at equivalent positions in the sequence of both N- and C-terminal halves of the brain enzyme, despite evidence suggesting that catalytic activity is restricted to the C-terminal half (8–10). Such extensive conservation of these residues in a region of the molecule no longer serving catalytic function seems surprising and suggests that they may play additional roles. An alternative but more complicated interpretation might be that the C-terminal half represents the site of the initial catalytic event, and thus the primary site of labeling when reactive substrate analogs are added individually (8, 9), but that resulting conformational changes induce a catalytically competent structure in the N-terminal half—i.e., although only one-half would be active at a given time, the N- and C-terminal halves may alternate as functional catalytic sites, a reciprocating mechanism that is, in some respects, analogous to the "half of the sites reactivity" seen with some oligomeric enzymes (36).

The alignment in Fig. 3 suggests that the C-terminal half of brain hexokinase lacks the sequence corresponding to the N-terminal segment and first helix of the yeast enzyme. We have generated a model of the mammalian enzyme (Fig. 4), based on the x-ray structure of the yeast enzyme (16, 21–23), by fusing the C terminus of one complete yeast hexokinase molecule with a second, truncated molecule lacking this region and then rotating the now-joined molecules to eliminate steric conflict while retaining sufficient proximity to permit noncovalent interactions between them. There was surprisingly little latitude in permissible alignments, and thus the model presented is far from arbitrary despite the somewhat subjective manner in which it was generated. It should be noted here that the first 20 residues of the yeast sequence could not be located by Steitz and coworkers (16, 21–23) due to localized disorder in this region. Hence, not shown in Fig. 4 but preceding the N-terminal helix, is a flexible peptide segment.

Several features of this model are consistent with existing experimental results. For example, brain hexokinase binds to the outer mitochondrial membrane via a hydrophobic N-terminal sequence (33), which inserts into the membrane (37). It is apparent that the flexible N-terminal peptide segment, mentioned above, would be well suited to serve in tethering the enzyme to the mitochondrial membrane, and the protrusion of this segment from the enzyme's surface is consistent with its notable susceptibility to proteolytic attack (33). The absence of hydrophobic character in this segment of the yeast enzyme accounts for the inability of yeast hexokinase to bind to mitochondria (38); there is precedent for evolution of functionally important, hydrophobic N-terminal segments from nonhydrophobic precursor sequences (39, 40).

If the two halves of the enzyme exist in similar conformations, it might be expected that the tryptic cleavage sites, T_1 and T_2 (31), are located in analogous positions in the N- and C-terminal halves. This is the case. Direct sequencing of the resulting fragments indicates that T_1 and T_2 correspond to cleavage at Lys-101 and Arg-551, respectively, which are found at virtually identical positions in the aligned N- and C-terminal sequences (Fig. 3). Based on assumed structural similarity to the yeast enzyme, these residues are in, or near, β-turns at the surface of the molecule, and hence their marked sensitivity to tryptic attack is consistent with their location in the proposed structure (Fig. 4).

The N- and C-terminal regions of brain hexokinase interact strongly by noncovalent forces, so much so that even after cleavage by trypsin, the resulting fragments remain tightly associated with retention of properties characteristic of the intact enzyme (31). However, in 0.6 M guanidine hydrochloride, these interactions appear to be greatly weakened and a new tryptic cleavage site, designated T_3 and located approximately in the middle of the molecule, is revealed (10). N-terminal sequencing of the resulting C-terminal fragment

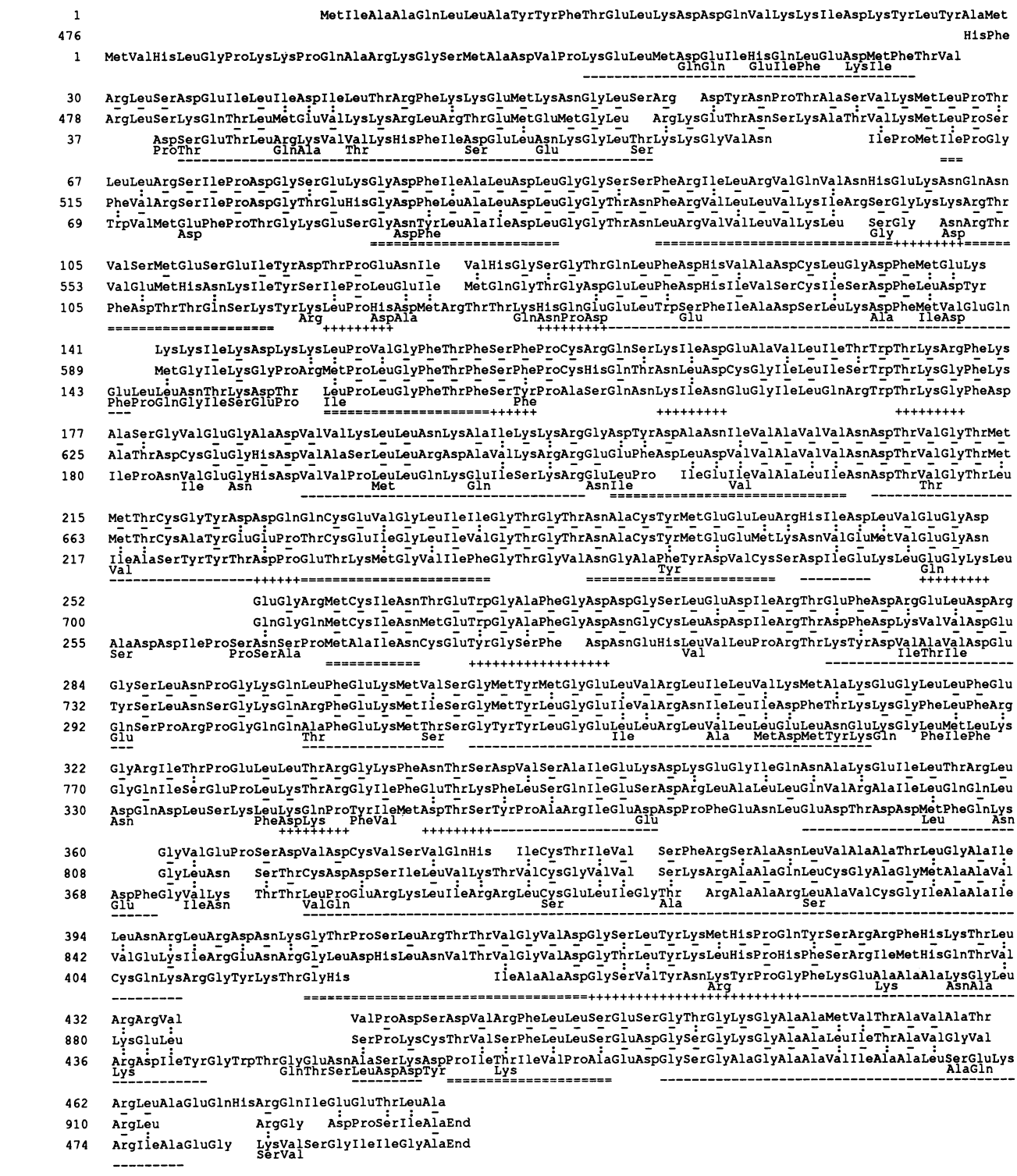

FIG. 3. Comparison of deduced amino acid sequences for the N- and C-terminal halves of rat brain hexokinase and yeast isozymes A and B. The top line provides the sequence for the N-terminal half of brain hexokinase (Fig. 2), and aligned beneath it is the corresponding sequence as deduced for the C-terminal half (13). The third line provides the deduced sequence for yeast isozyme A; the sequence of isozyme B is identical to this except for changes at the positions shown on the fourth line (15). Secondary structural features of the yeast enzyme are represented beneath the yeast sequence. α-Helix is represented by a single dashed line, β-strands are shown by a double dashed line, and β-turns are indicated by ++++.

reveals that cleavage occurs at Arg-462, only a few residues from the point at which the N- and C-terminal halves are fused. Based on the proposed structure (Fig. 4), this site is located between the two strongly interacting halves and would be expected to be inaccessible to trypsin. When those interactions are diminished by addition of denaturant, increased susceptibility to tryptic attack can be expected. Hence, manifestation of T_3 under these conditions is consistent with its location in the proposed structure.

It is apparent that the structure shown in Fig. 4 cannot be considered to represent that of brain hexokinase in detail. For example, we have not attempted to introduce structural changes that might result from the insertions/deletions discussed above. In addition, more refined coordinates of the

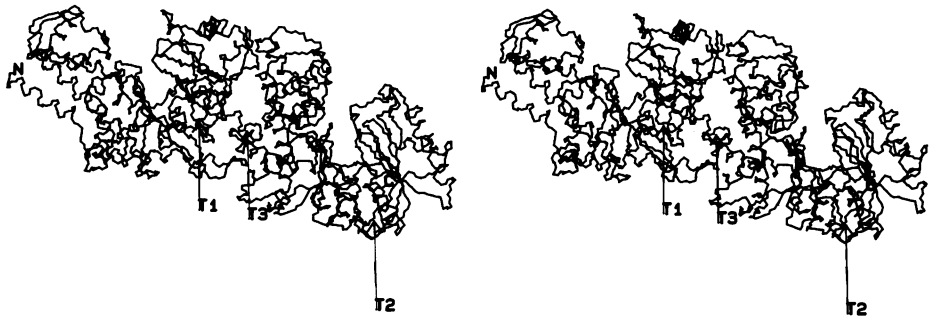

FIG. 4. Proposed structure for rat brain hexokinase. This stereo representation of the polypeptide backbone was based on the x-ray crystallographic structure of yeast hexokinase. The locations of tryptic cleavage sites T_1, T_2, and T_3 are indicated by vertical arrows.

yeast structure (16) are not yet available through the Brookhaven data base, although these are not likely to result in any major changes from the structure proposed earlier by Steitz and colleagues (21–23), upon which the present model for the mammalian hexokinase is based. Despite its limitations, we believe that the proposed structure is a reasonable one, developed on a rational basis and consistent with extant information, and that it will prove useful in the design and interpretation of future experiments aimed at elucidation of structure–function relationships in brain hexokinase and in other mammalian hexokinases.

Note. After this manuscript was submitted, Nishi *et al.* (41) reported the complete deduced amino acid sequence for the type I isozyme of human hexokinase. There is 92% identity between the sequences of the rat and human enzymes. Still more recently, Andreone *et al.* (42) reported the deduced sequence for the type IV isozyme ("glucokinase") from rat. Rat glucokinase shows 51% identity with the sequence of the C-terminal half of the rat type I isozyme (using a slightly different alignment, Andreone *et al.* calculated 53% identity) and 47% identity with the sequence of the N-terminal half.

We are pleased to acknowledge the expert assistance of Dr. David L. DeWitt in construction of the λgt10 cDNA library. We are grateful to the Upjohn Co., Kalamazoo, MI, for allowing us to use their computer graphics facility, and to Dr. Gerry Maggiora and his colleagues in the Computational Chemistry Center at Upjohn for their advice and assistance. This work was made possible by the financial support provided by Grant NS 09910 from the National Institutes of Health.

1. Wilson, J. E. (1984) in *Regulation of Carbohydrate Metabolism*, ed. Beitner, R. (CRC, Boca Raton, FL), pp. 45–85.
2. Colowick, S. P. (1973) in *The Enzymes*, ed. Boyer, P. D. (Academic, New York), 3rd Ed., Vol. 9, pp. 1–48.
3. Easterby, J. S. & O'Brien, M. J. (1973) *Eur. J. Biochem.* **38**, 201–211.
4. Rose, I. A., Warms, J. V. B. & Kosow, D. P. (1974) *Arch. Biochem. Biophys.* **164**, 729–735.
5. Holroyde, M. J. & Trayer, I. P. (1976) *FEBS Lett.* **62**, 215–219.
6. Ureta, T. (1982) *Comp. Biochem. Physiol. B* **71**, 549–555.
7. Gregoriou, M., Trayer, I. P. & Cornish-Bowden, A. (1983) *Eur. J. Biochem.* **134**, 283–288.
8. Nemat-Gorgani, M. & Wilson, J. E. (1986) *Arch. Biochem. Biophys.* **251**, 97–103.
9. Schirch, D. M. & Wilson, J. E. (1987) *Arch. Biochem. Biophys.* **254**, 385–396.
10. White, T. K. & Wilson, J. E. (1987) *Arch. Biochem. Biophys.* **259**, 402–411.
11. Marcus, F. & Ureta, T. (1986) *Biochem. Biophys. Res. Commun.* **139**, 714–719.
12. Schirch, D. M. & Wilson, J. E. (1987) *Arch. Biochem. Biophys.* **257**, 1–12.
13. Schwab, D. A. & Wilson, J. E. (1988) *J. Biol. Chem.* **263**, 3220–3224.
14. Kopetzki, E., Entian, K.-D. & Mecke, D. (1985) *Gene* **39**, 95–102.
15. Stachelek, C., Stachelek, J., Swan, J., Botstein, D. & Konigsberg, W. (1986) *Nucleic Acids Res.* **14**, 945–963.
16. Harrison, R. (1985) Ph.D. Thesis (Yale Univ., New Haven, CT).
17. Creighton, T. E. (1983) *Proteins* (Freeman Publications, New York), pp. 252–262.
18. Keim, P., Heinrikson, R. L. & Fitch, W. M. (1981) *J. Mol. Biol.* **151**, 179–197.
19. Rossman, M. G., Liljas, A., Branden, C. I. & Banaszak, L. J. (1975) in *The Enzymes*, ed. Boyer, P. D. (Academic, New York), 3rd Ed., Vol. 11, pp. 61–102.
20. Weber, I. T., Takio, K., Titani, K. & Steitz, T. A. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7679–7683.
21. Anderson, C. M., Stenkamp, R. E. & Steitz, T. A. (1978) *J. Mol. Biol.* **123**, 15–33.
22. Anderson, C. M., Stenkamp, R. E., McDonald, R. C. & Steitz, T. A. (1978) *J. Mol. Biol.* **123**, 207–219.
23. Bennett, W. S., Jr., & Steitz, T. A. (1980) *J. Mol. Biol.* **140**, 211–230.
24. Pittler, S. J., Kozak, L. P. & Wilson, J. E. (1985) *Biochim. Biophys. Acta* **843**, 186–192.
25. DeWitt, D. L. & Smith, W. L. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 1412–1416.
26. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
27. Davis, R. W., Botstein, D. & Roth, J. R. (1980) *A Manual for Genetic Engineering: Advanced Bacterial Genetics* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 80–82.
28. Henikoff, S. (1984) *Gene* **28**, 351–359.
29. Henikoff, S. (1987) *Promega Notes* (Promega, Madison, WI), No. 8, pp. 1–3.
30. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
31. Polakis, P. G. & Wilson, J. E. (1984) *Arch. Biochem. Biophys.* **234**, 341–352.
32. Wickens, M. & Stephenson, P. (1984) *Science* **226**, 1045–1051.
33. Polakis, P. G. & Wilson, J. E. (1985) *Arch. Biochem. Biophys.* **236**, 328–337.
34. Craik, C. S., Rutter, W. J. & Fletterick, R. (1983) *Science* **220**, 1125–1129.
35. Shohan, M. & Steitz, T. A. (1982) *Biochim. Biophys. Acta* **705**, 380–384.
36. Herzfeld, J., Ichiye, T. & Jung, D. (1981) *Biochemistry* **20**, 4936–4941.
37. Xie, G. & Wilson, J. E. (1988) *Arch. Biochem. Biophys.* **267**, 803–810.
38. Kovacs, L., Nelson, B. D. & Ernster, L. (1986) *Biochem. Biophys. Res. Commun.* **134**, 285–291.
39. Vassarotti, A., Stroud, R. & Douglas, M. (1987) *EMBO J.* **6**, 705–711.
40. Bibus, C. R., Lemire, B. D., Suda, K. & Schatz, G. (1988) *J. Biol. Chem.* **263**, 13097–13102.
41. Nishi, S., Seino, S. & Bell, G. I. (1988) *Biochem. Biophys. Res. Commun.* **157**, 937–943.
42. Andreone, T. L., Printz, R. L., Pilkis, S. J., Magnuson, M. A. & Granner, D. K. (1989) *J. Biol. Chem.* **264**, 363–369.