CLONING AND EXPRESSION OF ISOCITRATE LYASE FROM HUMAN ROUND WORM *STRONGYLOIDES STERCORALIS*

SIDDIQUI A.A.*,**, STANLEY C.S.* & BERK S.L.*,**

Summary :

A full length cDNA (1463 bp) encoding isocitrate lyase (EC 4.1.3.1) of *Strongyloides stercoralis* is described. The nucleotide sequence of this insert identified a cDNA coding for the isocitrate lyase. The conceptually translated amino acid sequence of the open reading frame for *S. stercoralis* isocitrate lyase encodes a 450 amino acid residue protein with an apparent molecular weight of 50 kDa and a predicted *pI* of 6.39. The sequence is 69 % A/T, reflecting a characteristic A/T codon bias of *S. stercoralis*. The amino acid sequence of *S. stercoralis* isocitrate lyase is compared with bifunctional glyoxylate cycle protein of *Caenorhabditis elegans* and isocitrate lyases from *Chlamydomonas reinhardtii* and *Myxococcus xanthus*. The full length cDNA of *S. stercoralis* was expressed in pRSET vector and bacteriophage T7 promoter based expression system. *S. stercoralis* lyase recombinant protein, purified via immobilized metal affinity chromatography, showed a molecular mass of 50 kDa on polyacrylamide gels. The role of isocitrate lyase in the glyoxylate cycle and energy metabolism of *S. stercoralis* is also discussed.

KEY WORDS : *Strongyloides*, parasite, isocitrate lyase, glyoxylate, enzyme.

Résumé : CLONAGE ET EXPRESSION DE L'ISOCITRATE LYASE DE *STRONGYLOIDES STERCORALIS*

Un ADN recombinant complet codant pour l'isocitrate lyase (EC 4.1.3.1.) de Strongyloides stercoralis *a été déterminé. La lecture de cette séquence donne une protéine de 450 acides aminés avec un poids moléculaire apparent de 50 kDa et un* pI *calculé de 6.39. Cette séquence possède un rapport A/T de 69 % caractéristique du rapport A/T de* S. stercoralis. *La séquence d'acides amminés de cette isocitrate lyase est comparée avec celle de la protéine du cycle fonctionnel du glyoxal de* Coenorhabditis elegans *et des isocitrates lyases de* Chlamydomonas reinhardtii *et* Myxococcus xanthus. *La totalité de cet ADN recombinant a été exprimée à l'aide d'un vecteur pRSET et du système d'expression du bactériophage T7. La protéine recombinante de* S. stercoralis *a été purifiée par chromatographie d'affinité sur métal et a présenté une masse moléculaire de 50 kDa en gel de polyacrylamide. L'importance de l'isocitrate lyase dans le cycle du glyoxal et dans le métabolisme de* S. stercoralis *est discutée.*

MOTS CLÉS : Strongyloides, *isocitrate lyase, glyoxal, enzyme.*

INTRODUCTION

Isocitrate lyase (EC 4.1.3.1) is an enzyme that catalyzes the conversion of isocitrate to succinate and glyoxylate. This is the first step in the glyoxylate bypass, an alternative to the tricarboxylic acid cycle in bacteria, fungi, higher plants and nematodes (Beeching, 1988; Atomi *et al.*, 1990; Liu *et al.*, 1995). Among nematodes, only *Caenorhabditis elegans* (Wadsworth & Riddle, 1989), hookworms (Singh *et al.*, 1992) and *Ascaris suum* (Patel & McFadden, 1978) have been shown to contain an active glyoxylate cycle. Apart from *C. elegans* (Liu *et al.*, 1995), no information is available on nucleotide and amino acid sequences of cDNAs of isocitrate lyase from parasitic nematodes. Herein we report the isolation and expression of a cDNA encoding isocitrate lyase from the human parasite, *Strongyloides stercoralis.*

MATERIALS AND METHODS

During the screening of a *S. stercoralis* cDNA library with anti-isocitrate dehydrogenase antibodies (Siddiqui *et al.*, 2000a) a clone was serendipitously picked which contained a cDNA coding for isocitrate lyase of *S. stercoralis* (SsICL). Briefly, a cDNA library was constructed as follows: double-stranded, size-selected (> 0.4 kb), high molecular weight cDNA was prepared using oligo-dT primed *S. stercoralis* mRNA (L_3 larvae) which was isolated using magnetic beads (Dynal, Lake Success, NY). The cDNA ends were modified by ligation of linkers encoding *EcoR*I and *Xho*I sites and then cloned into compatible sites present in the multiple cloning region of the Uni-ZAP XR vector (Stratagene, La Jolla, CA). *S. stercoralis* cDNA expression library was screened with the antibodies as described earlier (Karcz *et al.*, 1991). Pri-

* Department of Internal Medicine, James H. Quillen College of Medicine, East Tennessee State University, Johnson City, TN.
** James H. Quillen Veterans Affair Medical Center, Mountain Home, TN.
Correspondence: Afzal A. Siddiqui. Department of Internal Medicine, James H. Quillen College of Medicine, Box 70622, East Tennessee State University, Johnson City, TN 37614.
Tel: (423) 439-6376 – Fax: (423) 439-6387 – Email: siddiqui@etsu.edu

mary screening with anti-*S. stercoralis* isocitrate dehydrogenase (Siddiqui *et al.*, 2000a) yielded several positive clones, one such clone (14B-7) containing a ~ 1.5 kb insert was selected because of its very high immunoreactivity. This clone was plaque-purified and subcloned into pBluescript. The 1.5 kB insert was sequenced using T3, lyase 1 (5'-GCA AGT TGC TGC TGA TGG TA-3') and lyase 2 (5'-TTG CAT ATG CTC CAT ATT GTG A-3') primers at the Molecular Genetics Facility, Athens, GA. The sequence was then compared with the GenBank database by BLAST analysis (Altschul *et al.*, 1997). The 1.5 kB insert contained the full-length cDNA for SsICL (Fig. 1). The SsICL sequence has been deposited in GenBank (Accession # AF187048). Multiple sequence alignment of SsICDH with other sequences was performed via MULTALIN software (Corpet, 1988).

For expression work, a *Bam*HI and *Kpn*I restriction fragment from the clone 14B-7 containing *S. stercoralis* isocitrate lyase was excised and subcloned into the same restriction sites of the pRSET A vector containing N-terminal polyhistidine and anti-Xpress epitope tags (Invitrogen, Carlsbad, CA). The recombinant *S. stercoralis* isocitrate lyase-pRSET plasmid construct was transformed into BL21(DE3)pLysS according to the manufacturer's instructions. The expression of isocitrate lyase was induced with IPTG (1mM, final concentraion) at 37° C. The recombinant protein was purified using Xpress System ProBond resin columns (Invitrogen, Carlsbad, CA), under denaturing conditions. The purified protein was separated on 10% SDS- polyacrylamide gels and blotted onto nitrocellulose. Recombinant protein was visualized using 1:50,000 dilution of anti-Xpress primary antibody (Invitrogen, Carlsbad, CA) and a 1:6,000 dilution of alkaline phosphatase conjugated goat anti-mouse IgG.

RESULTS AND DISCUSSION

An open reading frame for SsICL potentially encodes a 450 amino acid residue protein with an apparent molecular weight of 50 kDa and an isoelectric point of 6.39. SsICL sequence contains a typical poly (A) addition site (AATAAA) 13 bases upstream of poly (A) tail which is composed of 19 bases. Further, SsICL sequence has a characteristic A/T bias (Moore *et al.*, 1996; Siddiqui *et al.*, 2000a, b) and is 69 % A/T. The full length sequence of SsICL exhibits 69 % identity in amino acid composition with bifunctional glyoxylate cycle protein of *C. elegans* (U23159). SsICL showed 64 % and 56 % identities in amino acid sequence with isocitrate lyases of *Myxococcus xanthus* (AF013216) and *Chlamydomonas reinhardtii* (U18765), respectively (Fig. 2).

Protein Families Data Base of Alignments and Hidden Markov Models (Pfam HMM) search revealed, SsICL amino acid sequence contains two conserved isocitrate lyase domains (amino acid residues 19-248 and 254-431) (Fig. 1). The ICL signature pattern, K-[KR]-C-G-H- [LMQ] (Beeching, 1988; Atomi *et al.*, 1990), was detected in SsICL sequence (amino acid residues 188-193; KKCGHM) with C as a putative active site residue (Fig. 1). A cysteine, a histidine and a glutamate or aspartate have been found to be important for enzyme activity. Only one cysteine residue is conserved among the sequences of the fungal, plant, bacterial and *S. stercoralis* enzymes; it is located in the middle of a conserved hexapeptide that can be used as a signature pattern for ICL (Beeching, 1988; Atomi *et al.*, 1990). Prosite pattern search revealed that SsICL amino acid sequence has several phosphorylation sites which are as follows: amino acid residues 15-17, 50-52, 124-126, 186-188 and 232-234 (Protein Kinase C phosphorylation sites); residues 98-101, 113-116, 169-172, 228-231, 238-241, 266-269, 330-333, 360-

```
ggcatgaggttaaaagctagtaatttttaccatattgtaaaaaattctcctaaaggaaga
    M  R  L  K  A  S  N  F  Y  H  I  V  K  N  S  P  K  G  R
tggacaggaattaaacgtaattatgaaattaaagatgtcctaaaattacgtggttcttta
 W  T  G  I  K  R  N  Y  E  I  K  D  V  L  K  L  R  G  S  L
aatattgaatatacacttgctaataaaacatcaaataaactttggcatctcattcataca
 N  I  E  Y  T  L  A  N  K  T  S  N  K  L  W  H  L  I  H  T
gaaccatatgtagctgctttaggtgcacaaacaggtaatcaagctgttcaaatgattaaa
 E  P  Y  V  A  A  L  G  A  Q  T  G  N  Q  A  V  Q  M  I  K
gctggattaaaagcaatttatctttctggatggcaagttgctgctgatggtaatactgct
 A  G  L  K  A  I  Y  L  S  G  W  Q  V  A  A  D  G  N  T  A
ggtgatatgtatccagatcaatcattatatccttcaaattctggacctgaattagcaaga
 G  D  M  Y  P  D  Q  S  L  Y  P  S  N  S  G  P  E  L  A  R
cgtattaataaatcattaagaagagctgatcaaattgaatgtgcagagtctgatgatatg
 R  I  N  K  S  L  R  R  A  D  Q  I  E  C  A  E  S  D  D  M
caaccatatagagattattatgcccctattgtagctgattgtgaagctggttttggaggt
 Q  P  Y  R  D  Y  Y  A  P  I  V  A  D  C  E  A  G  F  G  G
agtttaaattgttttgaaattacaaaatcatacattgaatctggtgttgctggtgtacat
 S  L  N  C  F  E  I  T  K  S  Y  I  E  S  G  V  A  G  V  H
tttgaagatcaattaggatcagaaaaaaaaatgtggtcatatgggaggaaaagttttaatt
 F  E  D  Q  L  G  S  E  K  K  C  G  H  M  G  G  K  V  L  I
ccaatatctgaacatattcgtcatttaaatgctgcacgtttagctgctgatgtatgtgat
 P  I  S  E  H  I  R  H  L  N  A  A  R  L  A  A  D  V  C  D
acaccaacaattattgtagctagaactgatgctgaaagtgctagactcttaacaagtgat
 T  P  T  I  I  V  A  R  T  D  A  E  S  A  R  L  L  T  S  D
atagatgaaagagatcatccatttattgatagaaaagctggtagaacatcagaaggtttt
 I  D  E  R  D  H  P  F  I  D  R  K  A  G  R  T  S  E  G  F
tatagacttaaagattctacatcaatggaagcatgtattcaaagaggtattgcatatgct
 Y  R  L  K  D  S  T  S  M  E  A  C  I  Q  R  G  I  A  Y  A
ccatattgtgatatgatatggatggaaacatcatatccatctttatcccaagctaaagaa
 P  Y  C  D  M  I  W  M  E  T  S  Y  P  S  L  S  Q  A  K  E
tttgctgaaggtgtaaaaagagaatttccagataaattatttgcatataactgttcacca
 F  A  E  G  V  K  R  E  F  P  D  K  L  F  A  Y  N  C  S  P
tcatttaattggggtaaacatttaaaaaagagtgatatggaaaaatatcaaagagaactt
 S  F  N  W  G  K  H  L  K  K  S  D  M  E  K  Y  Q  R  E  L
ggagctatgggatttaaatatcaatttattactcttgctggttatcatacaaatagtttt
 G  A  M  G  F  K  Y  Q  F  I  T  L  A  G  Y  H  T  N  S  F
tcaatatttgatttagctaaaaattatcgtgaacgtggtatggctgcttatgctgaactt
 S  I  F  D  L  A  K  N  Y  R  E  R  G  M  A  A  Y  A  E  L
caaaaagctgaatttgatgcagaaaaatcaggatatactgctgtaaaacatcaacgtgaa
 Q  K  A  E  F  D  A  E  K  S  G  Y  T  A  V  K  H  Q  R  E
gttggaactggatattttgatgttcttggaaatgcttgtgctggtggttctagttctaca
 V  G  T  G  Y  F  D  V  L  G  N  A  C  A  G  G  S  S  S  T
actgcacttacaggaagtacagaagaggcacaatttaaaacaacggccgttgaagatgat
 T  A  L  T  G  S  T  E  E  A  Q  F  K  T  T  A  V  E  D  D
gaagttatgacggtgagtaaaaatgctactgtctaattttagtttatttttattttcgtg
 E  V  M  T  V  S  K  N  A  T  V
tgtatttatttaaataatttaacaaaaatttttaactatataaaaataaaaatataatt
attgaaaaaaaaaaaaaaaaaaa
```

Fig. 1. – Nucleotide and deduced amino acid sequence of the cDNA encoding a *Strongyloides stercoralis* isocitrate lyase. Two conserved isocitrate lyase domains (amino acid residues 19- 248 and 254-431) are underlined. The ICL signature pattern KKCGHM (amino acid residues 188-193) is shown in bold.

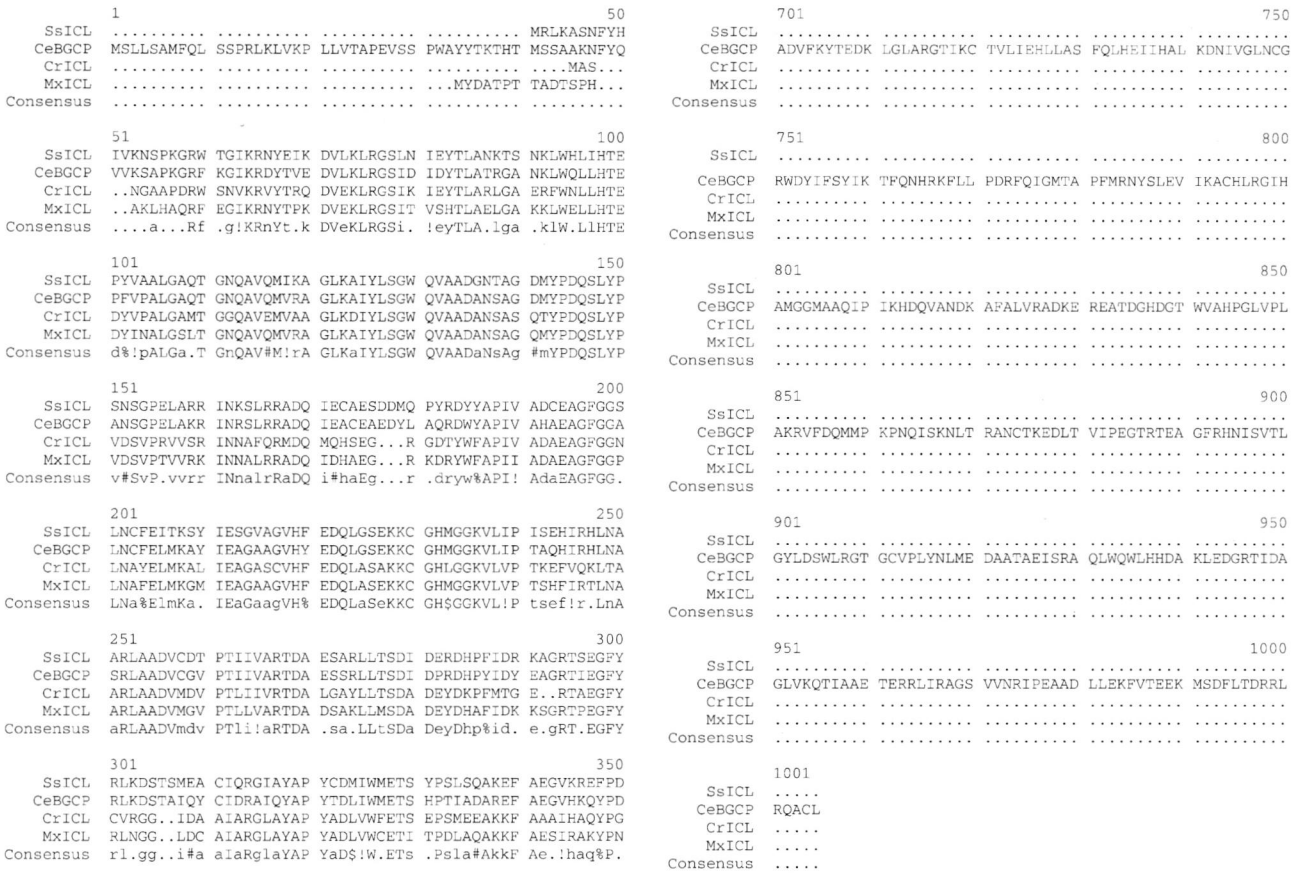

Fig. 2. – Alignment of the *Strongyloides stercoralis* isocitrate lyase with isocitrate lyases from other species. The alignment was performed with Multalin software. SsICL = *Strongyloides stercoralis* isocitrate lyase (AF187048); CeBGCP = *Caenorhabditis elegans* bifunctional glyoxylate cycle protein (U23159); CrICL = *Chlamydomonas reinhardtii* isocitrate lyase (U18765); MxICL = *Myxococcus xanthus* isocitrate lyases (AF013216). Consensus levels: high = 90 % low = 50 %.Consensus symbols: ! is anyone of IV, $ is anyone of LM, % is anyone of FY, # is anyone of NDQEBZ.

363, 425- 428 and 434-437 (Casein Kinase II phosphorylation sites); 254-260 and 328-335 (Tyrosine Kinase phosphorylation sites). More importantly, SsICL also has an eukaryotic putative RNA- binding region RNP-1 signature [RK]-G-{EDRKHPCG}-[AGSCI]-[FY]-[LIVA]-x-[FYLM] (Kamada & Miwa, 1992) which is RGIAYAPY and is located from amino acid residues 274-281 (Fig. 1). Transfac search (Quandt *et al.*, 1995) of the SsICL nucleotide sequence showed the presence of eight nucleotides (1324-1313, 368-379, 503-514, 317-328, 421-410, 956-967, 240- 229 and 1060-1071) consensus sequences (NWWWATCATNNN) of *C. elegans* skn-1 motif, a maternal gene product (Blackwell *et al.*, 1994). Furthermore, the prediction for SsICL according to the neural networks method for cytoplasmic/nuclear discrimination was found to be "cytoplasmic" with 55.5% reliability (Reinhardt & Hubbard, 1998).

SsICL recombinant protein was found to have an apparent molecular weight of 50 kDa which is consistent with isocitrate lyase from several other systems including *Mycobacterium avium* and *Mycobacterium tuberculosis* (Honer Zu Bentrup *et al.*, 1999).

In *C. elegans*, isocitrate lyase and malate synthase convert acetyl CoA derived from the degradation of yolk fatty acids into succinate from which carbohydrate is synthesized (Khan & McFadden, 1980). The enzyme activities increase greatly during embyrogenesis and fall considerably during L1 development (Wadsworth & Riddle, 1989), suggesting that these activities may be related to the metabolic requirement of embryos and early larvae and are regulated in a developmentally specific manner (Liu *et al.*, 1995). Both of these enzymes are contained within a single polypeptide, known as bifunctional glyoxylate cycle protein (Liu *et al.*, 1995). Conversely, SsICL is coded by a separate cDNA, as is the case with all of the microbial and plant ICLs (Vanni *et al.*, 1990). Cloning and expression of SsICL is a first step in elucidating the metabolic requirements and/or shifts in pathways during *S. stercoralis* development. This may provide additional insights into mechanisms of control of the two different *S. stercoralis* life cycles.

REFERENCES

Altschul F., Madden T.L., Schäffer A.A., Zhang J., Zheng Z., Miller W. & Lipman D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research,* 1997, *25,* 3389-3402.

Atomi H., Ueda M., Hishida T., Teranishi Y. & Tanaka A. Peroxisomal isocitrate lyase of the n-alkane-assimilating yeast *Candida tropicalis*: gene analysis and characterization. *Journal of Biochemistry (Tokyo),* 1990, *107,* 262-266.

Beeching J.R. High sequence conservation between isocitrate lyase from *Escherichia coli* and *Ricinus communis*. *Protein Sequence and Data Analysis,* 1989, *2,* 463-466.

Blackwell T.K., Bowerman B., Priess J.R. & Weintraub H. Formation of a monomeric DNA binding domain by Skn-1 by ZIP and homedomain elements. *Science,* 1994, *266,* 621- 628.

Corpet F. Multiple sequence alignment with hierarchial clustering. *Nucleic Acids Research,* 1988, *16,* 10881-10890.

Honer Zu Bentrup K., Miczak A., Swenson D.L. & Russell D.G. Characterization of activity and expression of isocitrate lyase in *Mycobacterium avium* and *Mycobacterium tuberculosis*. *Journal of Bacteriology,* 1999, *181,* 7161-7167.

Kamada S. & Miwa T. A protein binding to CArG box motifs and to single-stranded DNA functions as a transcriptional repressor. *Gene,* 1992, *119,* 229-236.

Karcz S.R., Podesta R.B., Siddiqui A.A., Dekaban G.A., Strejan G.H. &. Clarke M.W. Molecular cloning and sequence analysis of a calcium-activated neutral protease (calpain) from *Schistosoma mansoni*. *Molecular and Biochemical Parasitology,* 1991, *49,* 333-336.

Khan F.R. & McFadden B.A. Embyrogenesis and glyoxylate cycle. *FEBS Letter,* 1980, *115,* 312-314.

Liu F., Thatcher J.D. BarraL J.M. & Epstein H.F. Bifunctional glyoxylate protein in *Caenorhabditis elegans*: A developmentally regulated protein of intestine and muscle. *Developmental Biology,* 1995, *169,* 399-414.

Moore T.A, Ramachandran S., Gam A.A. Neva F.A., Lu W., Sanders L., Williams S.A. & Nutman T.B. Identification of novel sequences and codon usage in *Strongyloides stercoralis*. *Molecular and Biochemical Parasitology,* 1996, *79,* 243-248.

Patel T.R. & McFadden B.A. *Caenorhabditis elegans* and *Ascaris suum*: fragmentation of isocitrate lyase in crude extracts. *Experimental Parasitology,* 1978, *44,* 72-81.

Quandt K.F., Kayas H., Wingender E. & Werner T. New fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acid Research,* 1995, *23,* 4875-4884.

Reinhardt A. & Hubbard T. Using neural networks for prediction of the subcellular location of proteins. *Nucleic Acids Research,* 1998, *26,* 2230-36.

Siddiqui A.A., Stanley C.S. & Berk S.L. A cDNA encoding the highly immunodominant antigen of *Strongyloides stercoralis*: γ-subunit of isocitrate dehydrogenase (NAD^+). *Parasitology Research,* 2000a, *86,* 279-283.

Siddiqui A.A., Stanley C.S., Skelly P.J. & Berk S.L. A cDNA encoding a nuclear hormone receptor of the steroid/thyroid hormone receptors superfamily from the human parasitic nematode, *Strongyloides stercoralis*. *Parasitology Research,* 2000b, *86,* 24-29.

Singh S.P., Katiyar J.C. & Srivastava V.M. Enzymes of the tricarboxylic acid cycle in *Ancylostoma ceylanicum* and *Nippostrongylus brasiliensis*. *Journal of Parasitology,* 1992, *78,* 24-29.

Vanni P., Giachetti E., Pinzauti G. & McFadden B.A. Comparative structure, function and regulation of isocitrate lyase, an important assimilatory enzyme. *Comparative Biochemistry and Physiology,* 1990, *B95,* 431-458.

Wadsworth W.H. & Riddle D.L. Developmental regulation of energy metabolism in *Caenorhabditis elegans*. *Developmental Biology,* 1989, *132,* 167-173.

Reçu le 25 avril 2000
Accepté le 22 juin 2000