

## 树鼩载脂蛋白 CI 蛋白质结构分析

吕新跃 王克勤<sup>①</sup> 陈保生<sup>①</sup>

(武警医学院病理学教研室, 天津 300162)

### Analysis of the Protein Structure of Tree Shrew Apolipoprotein CI

LU Xin-Yue, WANG Ke-Qin and CHEN Bao-Sheng  
(Department of Pathology, Medical College of Chinese People's Armed Police Forces, Tianjin 300162, China)

#### ABSTRACT

**Aim** To obtain the protein structure deduced from tree shrew (TS) apolipoprotein CI (apo CI) cDNA sequence.

**Methods** Using Chou's method modified and some related computer softs, not only TS apo CI cDNA sequence cloned in a constructed cDNA library of TS liver tissue was analyzed and compared with those of other species, but also the deduced amino acid (AA) sequence, its secondary structure and hydrophobicity of the protein were determined and predicted.

**Results** TS apo CI cDNA sequence consists of 380 nucleotides encoding an 88 AA apo CI precursor (a 26 AA signal peptide and a 62 AA mature protein). The function domains of the protein were predicted to locate in two regions of 13~26 AA and 32~57 residues.

The secondary structure may contain several kinds of conformations. There is an obvious hydrophobic region at the C-terminal of the amino acid sequence.

**Conclusion** The present results shows TS apo CI is more hydrophobic and may have a stronger ability of binding lipids than human apo CI.

**KEY WORDS** Tree shrew; Apolipoprotein CI; Gene; Protein structure

**摘要** 为了获得树鼩载脂蛋白 CI 的蛋白质结构信息,利用相应的计算机软件及改良的 Chou 等方法,对从构建的树鼩肝组织 cDNA 文库中克隆出的载脂蛋白 CI cDNA 序列进行了分析和比较,并对编码的蛋白质氨基酸序列及其二级结构、亲疏水性进行了分析和预测。结果表明,克隆出的树鼩载脂蛋白 CI cDNA 序列(新基因已被 GenBank 接受)由 380 个核苷酸构成,编码翻译 88 个氨基酸的载脂蛋白 CI 前体(含 26 个氨基酸构成的信号肽和 62 个氨基酸组成的成熟蛋白),并推测出该蛋白的功能域、蛋白质的二级结构及 C 末端有一明显的疏水区。结果提示:树鼩载脂蛋白 CI 的疏水性及结合脂质的能力强于人的载脂蛋白 CI。

**关键词** 树鼩; 载脂蛋白 CI; 基因; 蛋白质结构

树鼩是良好的不易感动脉粥样硬化(atherosclerosis, As)动物,其不易感性可能与它血浆中高水平高密度脂蛋白(high density lipoprotein, HDL)及 HDL 中各种载脂蛋白(apolipoprotein)的组成和代谢特点密切相关<sup>[1,2]</sup>。载脂蛋白 CI 是一种分子量最小、碱性较强的载脂蛋白,亦存在于 HDL 中,其功能还不十分明确。有研究报道,载脂蛋白 CI 在体外能激活卵磷脂胆固醇酰基转移酶<sup>[3]</sup>、结合运输磷脂、抑制载脂蛋白 E 与低密度脂蛋白受体相关蛋白及低密度脂蛋白受体的结合。由此推测载脂蛋白 CI 也在富含甘油三酯的脂蛋白代谢方面起调控作用<sup>[4]</sup>。而不同种属间载脂蛋白 CI 的分子结构及其比较研究较少,在 HDL 及其他脂蛋白代谢过程中的作用也不明了。最近,我们首次报道了利用分子生物学技术,克隆出完整的树鼩载脂蛋白 CI cDNA 结构<sup>[5]</sup>。本研究主要利用相应的计算机软件及改良的 Chou 等<sup>[6]</sup>方法,对从构建的树鼩肝组织 cDNA 文库中克隆出

的载脂蛋白 CI cDNA 序列进行分析和比较,并对编码的蛋白质氨基酸序列及其二级结构、亲疏水性、功能域等进行分析和预测,从中获得树鼩载脂蛋白 CI 蛋白质结构的重要信息。

## 1 材料和方法

### 1.1 材料

树鼩(滇西种)从解放军 302 医院实验动物中心购进。cDNA 合成试剂盒(time saver cDNA synthesis Kit)为 Pharmacia 公司产品;寡聚脱氧胸苷酸纤维素(oligo-dT cellulose)购自德国 Boehringer Mannheim 公司;包装蛋白、噬菌体载体( $\lambda$ gt11)及限制性内切酶和连接酶均购自 Promega 公司;其余试剂为国产分析纯。DNASIS、PROSIS 软件包为 HITACHI 公司产品;737 型 DNA 自动测序仪为美国 PE 公司产品。

### 1.2 方法

**1.2.1 树鼩载脂蛋白 CI cDNA 的克隆及测序** 按一步法快速提取树鼩肝组织总 RNA,再经 oligo(dT)-纤维素柱层析,分离出 mRNA。以 3  $\mu$ g mRNA 为模板,oligo(dT)<sub>12-18</sub> 为引物,参照改进的方法<sup>[7]</sup>,构建 cDNA 的文库。再用制备的兔抗树鼩载脂蛋白 CI 多抗血清(效价为  $1 \times 10^5$ )稀释至 1:500 作为一抗,对构建的表达型 cDNA 库进行三轮筛选。将获得的两个阳性克隆再亚克隆,然后将其连接重组于 pGEM 3Zf(+)质粒,转化 DH5 $\alpha$  大肠杆菌,提取重组质粒 DNA,用双脱氧末端终止法自动荧光测序。

**1.2.2 cDNA 核酸序列的比较和分析**<sup>[8]</sup> 对已测出全长的 cDNA 阳性克隆序列,通过 blast 方式,向 GenBank 数据库查询比较所测序列的同源性,初步确认 cDNA 序列的性质。序列比较发现,未发表的新序列通过电子表格等形式向 GenBank 申报登记新序列,Gen-

Bank 审核证实后将给予接受号(accession number)。

**1.2.3 编码氨基酸序列的分析和比较** 采用 Clone 和 DNASIS 软件,分析已测出 cDNA 序列中的酶切位点、开放阅读框架及推译出的氨基酸序列。

**1.2.4 蛋白质二级结构及亲疏水性预测** 按改良的 Chou 等<sup>[6]</sup>方法,用 PROSIS 蛋白序列分析软件对由 cDNA 序列编码的树鼩载脂蛋白 CI 的亲疏水性及二级结构进行预测,并对载脂蛋白 CI 的氨基酸构成进行分析及同源性比较。

## 2 结果

### 2.1 树鼩载脂蛋白 CI cDNA 序列及其分析

测序发现两个克隆中的大插入子(T-6 克隆)全长 380 个碱基对(base pair,bp),它含 21 bp 的 5' 端非翻译区、264 bp 组成的开放阅读框架及 95 bp 组成 3' 非翻译区(图 1, Figure 1)。

序列比较显示,树鼩载脂蛋白 CI cDNA 仅与人及其它种属动物的载脂蛋白 CI cDNA 和载脂蛋白 CI 基因组 DNA 的同源性较高,它与人的载脂蛋白 CI cDNA 的同源性为 72.7%,从而证实克隆的这一基因为树鼩载脂蛋白 CI cDNA 序列,GenBank 已接收这一新基因(accession number:U86690)。

### 2.2 树鼩载脂蛋白 CI 蛋白序列分析及比较

由 cDNA 序列推译出的树鼩载脂蛋白 CI 的前体蛋白(88 个氨基酸组成)经氨基酸序列分析及同源性比较,确定该前体蛋白包括一个 26 个氨基酸构成的信号肽和一个由 62 个氨基酸组成的成熟肽(图 1, Figure 1)。

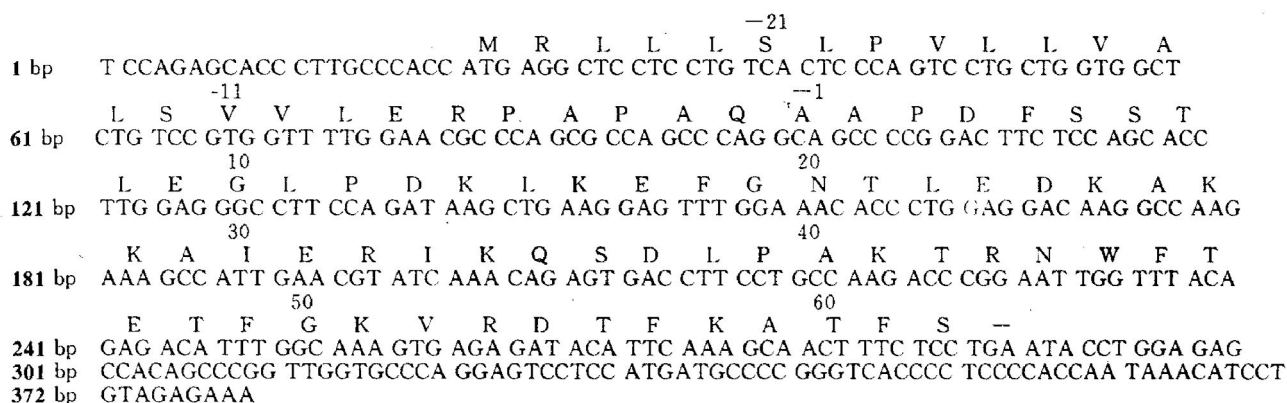


Figure 1. Tree shrew apo CI cDNA sequence and its deduced amino acid sequence.

树鼩载脂蛋白 CI 在蛋白水平上与人、狒狒、大鼠、小鼠及狗的同源性分别为 66%、67%、65%、64.5%和 73%。通过计算各种氨基酸的组成,获得其分子量为7 051 kDa。

2.3 树鼩载脂蛋白 CI 二级结构及亲疏水性

成熟的树鼩载脂蛋白 CI 可形成  $\alpha$ -螺旋、 $\beta$ -折叠、 $\beta$ 转角及无规则卷曲等多种二级结构形式(图 2, Figure 2)。形成  $\alpha$ -螺旋的主要区间在第 22~35 位氨基酸和第 45~51 位氨基酸区间,这两个  $\alpha$ -螺旋区域分别与文献报道的人载脂蛋白 CI 激活卵磷脂-胆固醇乙酰转移酶 (lecithin-cholesterol acetyl transferase, LCAT)和结合脂质的部位有部分重叠。

树鼩载脂蛋白 CI 和人载脂蛋白 CI 的亲疏水性区域分布大部分近似,但树鼩载脂蛋白 CI 的 C 末端呈现出明显的疏水性,而人的该蛋白却无此疏水区域,反映出二者在亲疏水性氨基酸构成上的差异(图 3, Figure 3)。

成熟的树鼩载脂蛋白 CI 带正负电荷的极性氨基酸 Lys、Arg、Asp、Glu 的含量占分子构成中总氨基酸的 35.48%(人为 40.35%);但疏水性氨基酸 Leu、Ala、Phe、Ile 等所占的比例为 40.32%,高于人的载脂蛋白 CI(33.33%)。且序列中不含 His、Cys、Met 和 Tyr,为一单链多肽蛋白。

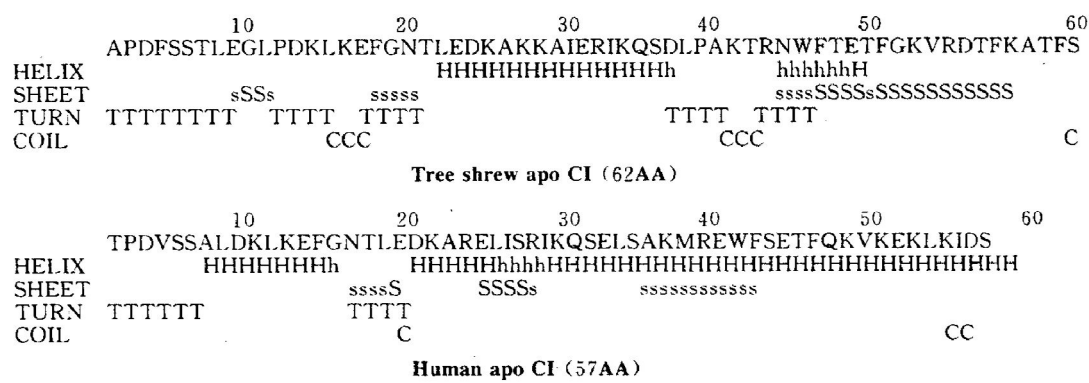


Figure 2. Secondary structure prediction of tree shrew and human apo CI.

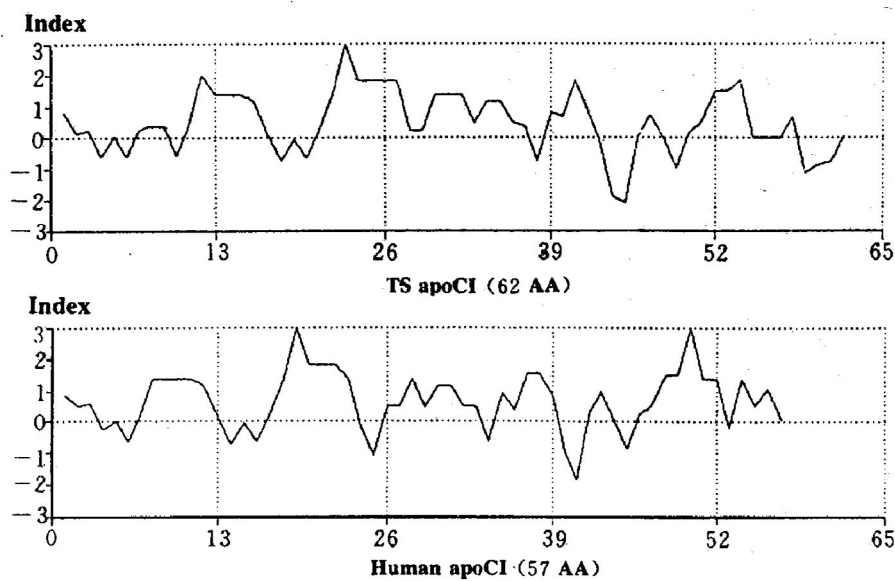


Figure 3. Analysis of the tree shrew apo CI hydrophobicity with PROSIS. Hydrophilic region: above the horizontal line. Hydrophobic region: under the horizontal line. The y-axis shows the hydrophilicity value. The x-axis represents the sequence of tree shrew apo CI amino acid.

### 3 讨论

迄今国际上仅报道了人、狒狒、大鼠、小鼠和狗载脂蛋白 CI cDNA 序列或基因组结构,并由 cDNA 序列的编码区推导出上述种属载脂蛋白 CI 的氨基酸序列。树鼯一直被认为是低等灵长目动物,与上述发表的有关序列比较<sup>[9~13]</sup>,树鼯载脂蛋白 CI 前体蛋白中的信号肽部分与人、狒狒、大鼠、小鼠和狗的氨基酸数目一致,但比人和狒狒载脂蛋白 CI 的成熟肽部分(57 个氨基酸)多 5 个氨基酸,而与狗、大鼠和小鼠的成熟肽部分氨基酸数目相同,均为 62 个氨基酸组成,反映了种属进化上的差异。同时,从蛋白质的进化角度考虑,我们对树鼯的生物学归属提出质疑,它是否属于低等灵长目动物?从我们获得的树鼯载脂蛋白 AI 氨基酸序列的证据<sup>[5]</sup>也支持这一疑问。

通过识别种属间高度保守的区域可以预测蛋白质的功能域<sup>[14]</sup>。我们比较了树鼯、狒狒、大鼠和小鼠载脂蛋白 CI 蛋白质的一级结构,发现第 13~27 位氨基酸和第 32~57 位氨基酸为两个保守的区间,特别是第 13~26 位氨基酸区间,所有种属在该区间组成完全一致,这一极其保守的共有序列为 DKLKEFGN-TLLEDKA,提示这一区域可能与载脂蛋白 CI 的某种功能活性有关。有研究认为此区域是载脂蛋白 CI 激活 LCAT 的功能域,另一相对保守区域(第 32~57 位氨基酸)被认为是载脂蛋白 CI 结合磷脂的主要部位<sup>[3,15]</sup>。因此推测树鼯载脂蛋白 CI 序列中的这些保守区域也具有人及其他哺乳类动物载脂蛋白 CI 类似的功能,即第 13~27 位氨基酸区间为激活 LCAT 的功能域,C 端保守区可能与结合脂质有关。蛋白质二级结构预测这两个区域可形成  $\alpha$ -螺旋,从另一个侧面也提示这两处可能为树鼯载脂蛋白 CI 的功能活性区所在。

成熟的树鼯载脂蛋白 CI 带正负电荷的极性氨基酸和疏水性氨基酸的含量较高,可形成载脂蛋白特有的两性  $\alpha$ -螺旋。疏水性氨基酸的含量高于人,特别是在其 C 末端第 50~61 位氨基酸区间含有 50% 的疏水性氨基酸,而人载

脂蛋白 CI 的此区域内疏水性氨基酸仅占 30%。亲疏水性分析显示:树鼯载脂蛋白 CI 的 C-端也有一显著的疏水区,而人的该蛋白则无这一疏水性分布,鉴于 C 末端为结合脂质的活性区,故提示树鼯载脂蛋白 CI 结合和转运脂质的能力要强于人的载脂蛋白 CI。这为继续研究不易感 As 动物树鼯载脂蛋白 CI 的功能提供了重要线索。

综上所述,我们在国内外首次成功地克隆出不易感 As 动物树鼯载脂蛋白 CI cDNA 序列,由此推译出它的蛋白质一级结构,并对其功能区、二级结构及亲疏水性等进行了分析。上述结果为深入研究树鼯载脂蛋白 CI 的分子结构奠定了基础,对其结构和功能的深入研究将有助于阐明树鼯载脂蛋白 CI 在脂质代谢中的作用。

### 参考文献

- 1 Wang KQ, Chen BS, Xie YH, et al. Are tree shrews and Beijing ducks good models for the study of the anti-atherogenicity HDL? In: Woodford FP, Davignon J, Sniderman A, eds. *Atherosclerosis X*, Amsterdam: Elsevier Press, 1995. 126~131.
- 2 She MP, Xia RY, Ran BF, et al. HDL and prevention of experimental atherosclerosis with special reference to tree shrews. *Ann NY Acad Sci*, 1990, **598**: 339~351.
- 3 Soutar AK, Sigler GF, Smith LC. Lecithin: cholesterol acyltransferase activation and lipid binding biosynthetic fragments of apolipoprotein CI. *Scand J Clin Lab Invest*, 1978, **38** (Suppl): 53~58.
- 4 Clavey V, Lestavel-Delattre S, Copin C, et al. Modulation of lipoprotein B binding to the LDL receptor by exogenous lipids and apolipoproteins CI, CII, CIII and E. *Arterioscler Thromb Vasc Biol*, 1995, **15**(5): 963~971.
- 5 Lu XY, Chen BS, Wang KQ. Study on cDNA sequences of tree shrew apolipoprotein AI and CI. *Atherosclerosis*, 1997, **134**(1~2): 19.
- 6 Chou PY, Fasman GD. Empirical prediction of protein conformation. *Ann Rev Biochem*, 1978, **47**: 251~276.
- 7 吕新跃, 陈保生, 王克勤, 等. 不易感动脉粥样硬化动物北京鸭肝 cDNA 文库的快速构建. *中国动脉硬化杂志*, 1997, **5** (4): 304~308.
- 8 Altschul SF, Gish W, Miller W, et al. Basic local alignment search tool. *J Mol Biol*, 1990, **215**: 403~

- 410.
  - 9 Knott T J, Priestley LM, Urdea M, et al. Characterisation of mRNAs encoding the precursor for human apolipoprotein CI. *Nucleic Acid Res*, 1984, **12**: 3 909~915
  - 10 Pastorcic M, Birnbaum S, Hixon JE. Baboon apolipoprotein CI cDNA and gene structure and evolution. *Genomics*, 1992, **13**(2): 368~374.
  - 11 Luo CC, Li WH, Chan L. Structure and expression of dog apolipoprotein AI, E and CI mRNAs: implications for the evolution and functional constraints of apolipoprotein structure. *J Lipid Res*, 1989, **30**(11): 1 735~746.
  - 12 Shen PY, Howlett GJ. Nucleotide sequence of cDNA for rat apolipoprotein CI. *Nucleic Acid Res*, 1989, **17** (15): 6 405.
  - 13 Hoffer MJV, van Eck MM, Havekes LM, et al. The mouse apolipoprotein CI gene: Structure and expression. *Genomics*, 1993, **18**: 4 257~262.
  - 14 Weinberg RB. Identification of domains in the plasma apolipoproteins by analysis of inter-species sequence variability. *J Lipid Res*, 1994, **35**: 2 212~222.
  - 15 Sparrow JT, Gotto AM. Phospholipid binding studies with synthetic apolipoprotein fragments. *Ann NY Acad Sci*, 1980, **348**: 187~211.
- (1998-04-11 收到, 1998-07-24 修回. 编辑: 文乐珊)

(1998-04-11 收到, 1998-07-24 修回。编辑: 文乐珊)

# 征 订 启 示

1999 年报刊杂志的征订工作又开始了。在过去的一年里,感谢广大作者和读者朋友对《中国动脉硬化杂志》不断的关怀和支持,正是这份关怀和支持,才使得这份你我共同拥有的杂志一直保持着良好的发展势头,能不断丰富自己的内涵,提高办刊质量。为了给广大读者带来更多更新的科研及临床信息,与国际接轨,本刊决定自明年(1999 年)起改为符合国际标准的 A<sub>4</sub>(即大 16)开本,以扩大内容量。并对所设栏目有所调整,以实验研究为主,同时加强临床方面的报道,共开设了专家述评、论著、研究简报、临床经验、短篇与病例报道、临床病例讨论、方法技术、研究成果综述、文献综述、国外动态、学术争鸣、简讯和消息等栏目,将以崭新的面貌来感谢广大朋友的支持。

《中国动脉硬化杂志》为季刊,改版后每期定价 8.50 元,全年 34.00 元。全国各地邮局均可订阅,邮发代号 42-165,并由中国国际图书贸易总公司(北京 399 信箱)对国外总发行,代号 4723Q。本刊热忱欢迎大家到各地邮局订阅或直接与本编辑部联系订购,地址:湖南省衡阳医学院内中国动脉硬化杂志编辑部,邮编 421001,电话:(0734)8281289。

另本编辑部现有少量 1993~1997 年各期杂志,需者可直接与编辑部联系,平均每期定价 6.50 元(含邮资)。

编辑部