

Cationic liposomes represent one of the few examples that can meet these requirements. Currently, there are more than a dozen cationic liposome formulations. These liposomes bind and condense DNA spontaneously to form complexes with high affinity to cell membranes. Endocytosis of the complexes followed by disruption of the endosomal membrane appears to be the major mechanism of gene delivery. The effectiveness

and safety of this DNA delivery method has been established in many studies. Two human gene therapy clinical trials using cationic liposomes have been conducted and more trials will be started in the near future. In the field of gene therapy, cationic liposome is struggling to meet expectations.

Key words cationic liposome, DNA, delivery, express, gene therapy, transfection

哺乳类 C 型凝集素超级家族

陈政良

(第一军医大学免疫学教研室, 广州 510515)

摘要 C 型凝集素超级家族根据其糖识别域 (CRD) 的一级结构分为蛋白聚糖、II 型跨膜受体、胶原凝素、选凝素、自然杀伤细胞受体及多 CRD I 型跨膜受体等 6 个家族。每一家族的成员具有同源的氨基酸序列、相同的总体分子结构及类似的生物学功能。从 CRD 的一级结构至三级结构水平确定了 C 型凝集素选择性糖识别的分子基础。

关键词 C 型凝集素, 糖识别域, 结构, 功能

凝集素是一类能专一识别糖并与之非共价可逆结合的非酶非抗体蛋白质, 动物中一些活性依赖 Ca^{2+} 的凝集素称为 C 型凝集素。自 1906 年报道第一个动物 C 型凝集素牛胶固素 (conglutinin) 以来, 在哺乳动物体内发现了多种多样的 C 型凝集素, 它们构成了一个庞大的超级家族。C 型凝集素分布广泛, 功能复杂, 涉及体内许多重要的生理、病理过程, 因而成为当今令人瞩目的研究热点之一。

C 型凝集素以跨膜蛋白和水溶性蛋白两种形式存在, 都是多结构域分子, 其共同的结构特征是具有一个或多个糖识别域 (carbohydrate recognition domain, CRD)。C 型 CRD 长约 115~130 个残基, 其中有 14 个恒定残基 (包括 4 个 C 残基) 和 18 个保守残基。某些 CRD 含额外 2 个 C 残基, 称作长型 CRD, 而只有 4 个 C 残基者称为短型 CRD。由 C 残基形

成的二硫键对于维系 CRD 的构象及整个分子的活性均十分重要。CRD 的一级结构不同, 决定了 C 型凝集素有不同的糖识别特异性。基于 CRD 一级结构比较确定其在进化树中的位置, Drickamer^[1]于 1993 年将 C 型凝集素超级家族分为蛋白聚糖、II 型跨膜受体、胶原凝素 (collectins)、选凝素 (selectins)、II 型淋巴细胞抗原及巨噬细胞甘露糖受体 (macrophage mannose receptor, MMR) 等 6 组。近年研究已基本确认 II 型淋巴细胞抗原就是自然杀伤 (natural killer, NK) 细胞的识别受体, 并发现树突状细胞表面 205 ku 蛋白 (DEC-205) 和磷脂酶 A₂ 受体 (phospholipase A₂ receptor, PLA₂R) 与 MMR 同属一个多 CRD I 型跨膜受体家族。因此, 本文将第 5、6 组 C 型凝集

素分别更名为 NK 细胞受体 (NK cell receptor, NKR) 和多 CRD I 型跨膜受体. 图 1 总结了 C 型凝集素各家族的成员并示其一级结

构. 下面分别介绍各家族分子的结构与功能, 为方便起见, 叙述未按分类顺序进行.

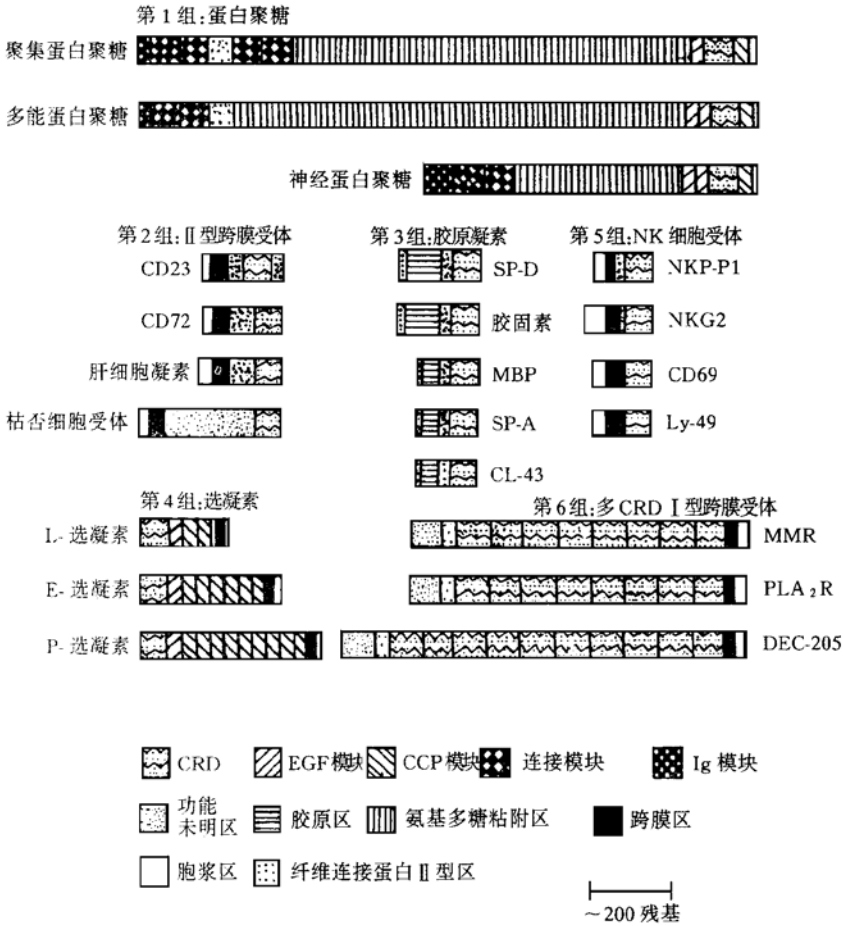


图 1 C 型凝集素各家族成员及其一级结构示意图

1 胶原凝素

胶原凝素为具有胶原样结构的 C 型凝集素. 已确定的该家族成员有五种: 胶固素、甘露聚糖结合蛋白 (mannan-binding protein, MBP)、43 ku 胶原凝素 (CL-43)、肺表面活性物质脱辅基蛋白 A 和 D (lung surfactant proteins A and D, SP-A, SP-D). 前 3 种是血浆蛋白, 由肝细胞合成; 后 2 种存在于肺泡表面, 为 II 型肺泡细胞所分泌. 胶原凝素是由同质 (SP-A 为异质) 三聚体亚基构成的寡聚体, 三聚体中 3 条肽链的近 N 端胶原样顺序互相

缠绕形成三股螺旋结构, C 端的 CRD 则卷曲成为球形头部. 在 MBP 和 SP-A, 6 个亚基平行联接成与补体分子 C1q 类似的花束样结构, 胶固素和 SP-D 分子由 4 个亚基组成十字形结构, CL-43 只见到单体. 胶原凝素的 CRD 可选择性识别多种微生物和病毒感染细胞表面具有末端 Man、ManNAc 及 Fuc 等的糖结构, 并触发其胶原样效应功能区的各种识别和清除机制如激活补体、调理吞噬及中和病毒等而发挥其天然抗感染免疫作用^[2].

大鼠 MBP-A CRD 及其与寡糖配体复合物的晶体结构已经测定, 明确了 MBP 与配体结

合的细节, 并作为模式指导着其他 C 型 CRD 选择性糖识别的分子基础的研究. MBP-A CRD (109~ 221) 大小为 $4.0\text{ nm} \times 2.4\text{ nm} \times 2.5\text{ nm}$, 由两个不同的二级结构区组成: 一部分是 5 段 β 折叠和 2 段 α 螺旋, 另一部分为 4 个环状结构. 该 CRD 可结合 2 个 Ca^{2+} , Ca^{2+} 和糖的结合位点由 3 个环状结构 (L1, 3, 4) 及 1 段 β 折叠 ($\beta 4$) 上的残基构成. 第一个 Ca^{2+} 配位位点由 D161、E165、D188 和 D194 形成, 而 E185、N187、E193、N205 和 D206 则涉及与第二个 Ca^{2+} 的配位结合, 同时构成糖结合位点: E185/N187 和 E193/N205 分别与末端 Man 的 3-OH 和 4-OH 形成氢键. 后 5 个残基在所有结合 Man 或 Glc 的 C 型 CRD 都是保守的, 但 EPN (185~ 187) 中的 185 和 187 位残基的性质则是 CRD 配体结合特异性的主要决定因素. MBP-A 的这两个位置分别是 E 和 N, 而结合 Gal 的 CRD (如肝细胞凝集素) 则分别是 Q 和 D. 定点突变 MBP-A 的 E185 和 N187 为 Q 和 D, 其糖识别特异性就由对 Man 变为对 Gal 了. SP-A 的 187 位是 A (人) 或 R (狗), 这决定了它与糖的结合力较弱且特异性较低. 然而, 许多 C 型凝集素可结合不同的糖类, 故可能还有其他位置的残基影响糖识别特异性^[3,4].

MBP 单个 CRD 与单糖的亲合力很低, 其对天然配体寡糖的高亲和性结合, 依赖于分子中多个 CRD 上适当排列的糖结合位点与配体上多个末端糖基之间的多点相互作用. 结构分析表明, MBP 分子 CRD 与中央螺旋之间及 CRD 之间存在大量的相互作用, 使 CRD 的相对方向固定, 因而糖结合位点的位置及相对距离也不易变. 测定糖结合位点之间的距离, 大鼠 MBP 是 5.3 nm , 人 MBP 为 4.5 nm . 这样大的距离不允许与哺乳类糖蛋白中含 Man 寡糖的多个末端糖基发生接触, 只能与具有相应间隔末端糖基的寡糖相互作用. 此即 MBP 只识别细菌、病毒等病原体表面的外源性糖结构而不与哺乳类细胞表面的内源性糖结合的结构基础. 酵母菌细胞壁上的甘露聚糖及许多细菌细

胞壁上含 GlcNAc 的寡糖均可能以规律重复的间隔提呈单糖, 适合于与 MBP 很大空间的糖结合位点相互作用, 因而能为 MBP 所识别^[4].

2 选凝素

选凝素是介导细胞间相互识别与粘附的一类 I 型跨膜糖蛋白, 该家族包括 L-、E- 和 P- 选凝素 3 个成员. L- 选凝素为 T 淋巴细胞、中性粒细胞及单核细胞所组成性表达, E- 选凝素由炎症性细胞因子诱导内皮细胞表达, 而 P- 选凝素则贮存于血小板 α 颗粒和内皮细胞 Weibel Palade 小体中, 受凝血酶或自由基刺激后转位于细胞膜上. 选凝素是 C 型凝集素超级家族中唯一一个 CRD 位于 N 端的家族. 3 种选凝素的 CRD 均识别唾液酸化、岩藻糖化的 LacNAc, 这类糖结构的典型代表是 sLe^x ($\text{Neu5Ac}\alpha 2 \rightarrow 3\text{Gal}\beta 1 \rightarrow 4 [\text{Fuc}\alpha 1 \rightarrow 3] \text{GlcNAc}\beta 1 \rightarrow \text{R}$) 和 sLe^a ($\text{Neu5Ac}\alpha 2 \rightarrow 3\text{Gal}\beta 1 \rightarrow 3 [\text{Fuc}\alpha 1 \rightarrow 4] \text{GlcNAc}\beta 1 \rightarrow \text{R}$), L- 选凝素还识别唾液酸化、岩藻糖化及硫酸化的含硫酸甘露糖基的聚糖^[5,6]. L- 选凝素为淋巴细胞的归巢受体, 可识别并结合外周淋巴器官小静脉高内皮细胞上的糖结构而介导淋巴细胞归巢; 内皮细胞 E-、P- 选凝素可识别淋巴细胞表面的糖基配体, 在淋巴细胞归巢中也起一定作用^[7]. 3 种选凝素均能通过白细胞与内皮细胞表面相对应分子间的相互识别及结合而在召集白细胞至炎症部位的过程中起关键作用^[8]. 此外, 选凝素与某些病理过程如血栓形成、癌细胞转移等有一定关系.

与 MBP-A CRD 的一级结构比较, 选凝素缺乏第一个 Ca^{2+} 结合位点, 但构成第二个 Ca^{2+} 结合位点的残基则完全一样. 测定 E- 选凝素 CRD 的晶体结构, 发现其总体三维结构与 MBP-A 相似. 其 Ca^{2+} 结合位点中, E80、N82、N105 和 D106 的侧链、D106 的主链及 2 个水分子与 Ca^{2+} 配位结合, 形成一个五边形的配位圈. 此处的 E88 (相当于 MBP-A 中的 E193) 由于位置和方向改变而不能作为 Ca^{2+} 的配体, 该配位位点就由水分子取代了. 突变

研究揭示, E-选凝素 CRD 顶部的 N82、Y94、R97 和 K113 在识别 sLe^x 中起关键作用, 而且, 对于与中性粒细胞粘附至关重要的 Y48、N82、E192、K111 和 K113 在所在选凝素家族成员中都是保守的^[9]. 按照 MBP-寡糖晶体结构的模式分析, sLe^x 可能结合于这些残基, Fuc 以与 Man 结合 MBP 相同的方式结合于类似的位点, 而唾液酸则结合于其邻近上述带正电荷的残基^[8,9]. 然而, 选凝素 CRD 与糖基配体相互作用的细节尚待选凝素 CRD-糖复合物晶体结构分析来予以揭示.

3 II型跨膜受体

本族蛋白为二硫键连接的双肽链 (CD23 为三肽链) 跨膜受体, 肽链的 N 端位于胞浆内, CRD 位于 C 端, 通过长度不等的功能不明序列突出于细胞表面. 该家族成员的典型代表是肝细胞凝集素, 又称为去唾液酸糖蛋白受体. 它们表达于肝实质细胞表面, 对具有末端 GalNAc 的糖结构有选择性高亲和力, 对 Gal 的结合力稍弱. 其功能主要是内吞去末端唾液酸的糖蛋白, 与某些血浆蛋白的转换、代谢有关^[3]. 比较研究发现, 肝细胞凝集素 CRD 与 MBP CRD 的糖结合位点中 Ca²⁺ 周围存在很大差异: MBP 中的 EPN (185~187) 和 H189 在肝细胞凝集素相应位置分别是 QPD (239~241) 和 W243, 而且紧接其后插入了一个富含 G 残基的顺序 YGHGLGG. 这些改变对于形成高亲和力、高选择性的 Gal 结合位点十分重要且已充分, 其中 W243 与糖基之间的相互作用是高亲和力 Gal 结合位点的重要部分^[10]. 巨噬细胞和肝细胞两者的去唾液酸糖蛋白受体一级结构同源性达 77%, 前者对 Gal 和 GalNAc 的亲和力几乎相等, 但后者则优先结合 GalNAc. 最近, 利用这两者的嵌合蛋白鉴定了对这种选择性结合起关键作用的残基 W、R、T 及 H, 它们构成了一个与 GalNAc 及 Ca²⁺ 相互作用的袋状结构^[4].

II 型跨膜受体家族的另一成员是低亲和力 IgE 受体即 CD23, 它可通过其 CRD 与 IgE 的

Fc 区结合. 但 CD23 CRD 相应于 MBP187 位的残基是 T, 导致其与蛋白质的结合力超过与糖的结合力, 故 CD23 与 IgE 的结合并非由 CRD-糖相互作用所介导. 不过, CRD-糖相互作用可能介导 CD23 与另一配体 CD21 的结合. B 细胞表面蛋白 CD72 也是该家族成员, 虽然它能与 T 细胞膜蛋白 CD5 相互作用, 但这不可能是通过 CRD 的糖结合活性实现的, 因为此 CRD 缺乏结合 Ca²⁺ 和糖类所必需的残基^[3].

4 蛋白聚糖

蛋白聚糖的核心蛋白含有 C 型 CRD, 故归属于 C 型凝集素超级家族, 其成员主要有聚集蛋白聚糖 (aggrecan)、多能蛋白聚糖 (versican)、神经蛋白聚糖 (neurocan) 及短小蛋白聚糖 (brevican) 等. 本族蛋白的 CRD 为长型, 能识别 Gal, 已发现聚集蛋白多糖 CRD 可与透明软骨中 II 型胶原上的 Gal 结合. 蛋白聚糖的功能主要是维持细胞外基质的结构完整性, 如透明软骨中的聚集蛋白多糖通过其 N 端连接模块与透明质酸相互作用, 构成巨大的多分子聚集物, 但 CRD 在其中所起的作用尚不清楚^[3].

蛋白聚糖 CRD 与胶原凝集素和选凝素 CRD 的二级结构高度相似. 胶原凝集素和选凝素 CRD 晶体结构中 Ca²⁺ 配位位点的 EPN 在蛋白聚糖里为 QPD, E-选凝素中紧接 EPN 后的 NRQKD 在蛋白聚糖中为 NFFAAG, 后者的疏水性更强^[11]. 与去唾液酸糖蛋白受体对 Gal 高亲和力的 CRD 比较, 蛋白聚糖 CRD 相应于 MBP H189 位不是 W 而是 F, 在 Ca²⁺ 结合位点附近又无富含 G 的插入顺序, 故对 Gal 只显示中等亲和力^[10].

5 自然杀伤细胞受体

NK 细胞的特征是对某些肿瘤细胞和病毒感染细胞具有自然杀伤活性. 这种活性的发挥, 需要效应细胞 (NK 细胞) 与靶细胞直接接触, 必然涉及两者表面分子的特异性相互作用

用. NKR 选择性表达于 NK 细胞表面, 为二硫键连接的双肽链 II 型跨膜蛋白. 已发现人的 NKG2 和 CD69 及啮齿动物的 NKR-P1 和 Ly-49 属于 NKR 家族成员. 这些分子的总体结构与 II 型跨膜受体相同. 序列比较发现, NKR CRD 缺乏构成第二个 Ca^{2+} 结合位点所必需的一些残基, 故曾对它们能否与 Ca^{2+} 和糖类结合表示怀疑, 但近年来的深入研究证实它们都是 Ca^{2+} 依赖的糖结合蛋白. 其突出特征是对 Ca^{2+} 的亲和力特别高, 即便加入 Ca^{2+} 螯合剂, Ca^{2+} 也不容易解离下来, 因此它们对糖类的结合不依赖外源 Ca^{2+} [12~14]. NKR-P1 和 CD69 亲和力最高的单糖配体是 GalNAc 和 GlcNAc [13], 复合寡糖中亲和力最高的配体是衍生于肝素的高度硫酸化的寡糖, 而生理性配体是靶细胞上神经节苷脂和蛋白聚糖中的寡糖顺序. 这些糖类在 NK 敏感肿瘤细胞膜上表达特别丰富, 但 NK 抗性细胞上表达稀少 [12]. Ly-49 的生理性配体是 MHC I 类分子上的糖类 [15, 16]. NKR 根据其功能性质分为两类: 兴奋性 NKR 和抑制性 NKR. 前者以 CD69 和 NKR-P1 为代表, 在识别配体后转导阳性信号入胞, 活化肌醇磷脂系统, 使胞浆 Ca^{2+} 浓度增高, 细胞脱颗粒, 最后导致靶细胞溶解 [17]. 后者如 Ly-49, 在识别靶细胞上自身 MHC I 类分子后转导阴性信号入胞, 抑制 NK 细胞的活化, 使其不能发挥细胞毒作用 [15, 16]. 两类功能不同 NKR 的存在, 赋予 NK 细胞选择性识别靶细胞的能力, 使之能将病变细胞或异体细胞与自身正常细胞区别开来并予以杀伤.

6 多 CRD I 型跨膜受体

MMR、DEC-205 及 PLA_2R 均为含多个 CRD 的单肽链跨膜受体分子, 肽链的 C 端位于胞浆内, 而且它们之间具有明显的结构同源性 [18, 19], 故将其归属为一个共同的家族, 命名为多 CRD I 型跨膜受体. MMR 和 PLA_2R 含 8 个 CRD, DEC-205 则有 10 个 CRD, 这些 CRD 中, 既有短型, 也有长型 [18, 19]. MMR 的 8 个 CRD 中, 只有 CRD4 和 CRD5 含有结合

Ca^{2+} 及 Man 的顺序, 故这 2 个 CRD 形成 MMR 配体结合中心. 然而, 完整 MMR 分子对酵母甘露聚糖的高亲和性结合需要 5 个成串的 CRD (CRD4~8) 参与, 可能那些辅助性 CRD 对于维持糖结合性 CRD 于正确的取向是必要的. MMR 的配体主要是 Man 和 GlcNAc. 含末端 Man 或 GlcNAc 的糖类在哺乳动物体内罕见, 却是细菌、真菌及寄生虫等表面的常见成分. MMR 识别并结合病原体表面的这类糖结构, 介导不依赖调理素的吞噬, 一方面起天然抗感染免疫作用 [3], 另一方面参与对抗原的加工和提呈 [20]. DEC-205 表达于树突状细胞和胸腺上皮细胞. 虽然尚未深入了解其配体结合情况, 但已明确它是一个 Mann -型受体, 树突状细胞和胸腺上皮细胞利用它捕获和内吞带相应糖结构的抗原, 将其运送至含 MHC II 类分子的细胞器中进行加工处理, 因而与抗原提呈有关 [19, 20]. PLA_2R 表达于多种组织细胞, 已知 PLA_2 参与众多的生理活动, 还与某些病理现象及蛇毒毒性有关. PLA_2R 除了结合 PLA_2 外, 还能识别 MMR 的某些配体如 Man 和 GlcNAc [18]. 这一性质的生物学意义未明.

参 考 文 献

- 1 Drickamer K. Calcium-dependent carbohydrate-recognition domains in animal proteins. *Curr Opin Struct Biol*, 1993, **3** (3): 393~400
- 2 陈政良. 胶凝素 (Collectin). 国外医学分子生物学分册, 1996, **18** (1): 24~29
- 3 Day A J. C-type carbohydrate recognition domain (CRD) superfamily. *Biochem Soc Trans*, 1994, **22** (1): 83~88
- 4 Drickamer K. Ca^{2+} -dependent sugar recognition by animal lectins. *Biochem Soc Trans*, 1996, **24** (1): 146~150
- 5 Varki A. Selectin ligands. *Proc Natl Acad Sci USA*, 1994, **91** (16): 7390~7393
- 6 Malhotra R, Taylor N R, Bird M I. Anionic phospholipids bind to L-selectin (but not E-selectin) at a site distinct from the carbohydrate binding site. *Biochem J*, 1996, **314** (1): 297~303
- 7 Butcher E C, Picker L J. Lymphocyte homing and homeostasis. *Science*, 1996, **272** (5259): 60~66
- 8 McEver R P, Moore K L, Cummings R D. Leukocyte trafficking mediated by selectin-carbohydrate interactions. *J Biol Chem*, 1995, **270** (19): 11025~11028
- 9 Graves B J, Crowther R L, Chandran C *et al.* Insight into E-selectin/ligand interaction from the crystal structure and

mutagenesis of the lec/EGF domains. *Nature*, 1994, **367** (6463): 532~ 538

10 Iobst S T, Drickamer K. Binding of sugar ligands to Ca²⁺-dependent animal lectins II. generation of high-affinity galactose binding by site-directed mutagenesis. *J Biol Chem*, 1994, **269** (22): 15512~ 15519

11 Brissett N C, Perkins S J. Molecular modelling analysis of the G-type lectin domain in human aggrecan. *Biochem Soc Trans*, 1996, **24** (1): 99s

12 Bezouska K, Yuen C-T, O'Brien J *et al.* Oligosaccharide ligands for NKR-P1 protein activate NK cells and cytotoxicity. *Nature*, 1994, **372** (6502): 150~ 157

13 Bezouska K, Nepovin A, Horvath O *et al.* CD69 antigen of human lymphocytes is a calcium-dependent carbohydrate-binding protein. *Biochem Biophys Res Commun*, 1995, **208** (1): 68~ 74

14 Brennan J, Takei F, Wong S *et al.* Carbohydrate recognition by a natural killer cell receptor, Ly-49. *J Biol Chem*, 1995, **270** (17): 9691~ 9694

15 Yokoyama W M, Daniels B F, Seaman W E *et al.* A family of murine NK cell receptor specific for target cell MHC class I molecules. *Semin Immunol*, 1995, **7** (2): 89~ 101

16 Yokoyama W M. Hybrid resistance and the Ly-49 family of natural killer cell receptors. *J Exp Med*, 1995, **182** (2): 273~ 277

17 Bezouska K. G-type lectins of natural killer cells: carbohydrate ligands and role in tumor cell lysis. *Biochem Soc Trans*, 1996, **24** (1): 156~ 161

18 Lambeau G, Ancian P, Barhanin J *et al.* Cloning and expression of a membrane receptor for secretory phospholipases A₂. *J Biol Chem*, 1994, **269** (3): 1575~ 1578

19 Jiang W, Swiggard W J, Heufler C *et al.* The receptor DEC-205 expressed by dendritic cells and thymic epithelial

cell is involved in antigen processing. *Nature*, 1995, **375** (6527): 151~ 155

20 Fearon D T, Locksley R M. The instructive role of innate immunity in acquired immunity response. *Science*, 1996, **272** (5259): 50~ 54

Mammalian G-Type Lectin Superfamily. CHEN Zhengliang (*Department of Immunology, The First Military Medical University, Guangzhou 510515, China*).

Abstract Based on primary structure comparisons of their carbohydrate-recognition domains (CRDs), the superfamily of mammalian G-type lectins can be divided into six main families: proteoglycans, type II transmembrane receptors, collectins, selectins, natural killer cell receptors and multi-CRD type I transmembrane receptors, while all members of each family have homologous amino acid sequences, same overall molecular architecture and similar biological functions. The molecular basis for selective carbohydrate recognition by mammalian G-type lectins has been established at the primary to tertiary structure level of CRDs.

Key words G-type lectins, carbohydrate recognition domains, structure, function

一类新内含子的研究状况

刘晓琼 施先宗

(华中师范大学昆虫学研究所, 武汉 430079)

摘要 在真核细胞基因组中发现一类新内含子——AT-AC 内含子, 除结构特征与普遍存在的核 mRNA 前体内含子有明显不同外, 剪接机制上也存在一定差异. 文章介绍了该内含子的分布, 结构特征及剪接机理, 并与主内含子作了相应比较.

关键词 AT-AC 内含子, 主内含子, 剪接体, 剪接位点, snRNP

内含子是不连续基因转录后被剪接去除的 RNA 序列, 或与之对应的 DNA 序列. 在真核细胞中广泛存在的核 mRNA 前体 (pre-mRNA) 内含子属于剪接体介导剪接内含子,

由于其所占比例最大, 分布最广, 因而又被称为主内含子 (major class of introns). 主内含子

收稿日期: 1996-09-09, 修回日期: 1997-04-22