

t-PA K1 区的同源模建及 Kringle 区 与 Lysine 相互作用的研究

王国力 刘士辉 徐秀英 黄培堂 黄翠芬

(北京生物工程研究所 北京 100071)

摘 要 利用同源模建方法预测了 t-PA K1 区的三维结构。通过结构叠合确定了 t-PA K1、K2 区 纤溶酶原 K1、K4 区及 UK K 区的 Lysine 结合口袋。静电势计算及疏水性分析表明,在 t-PA K2 区以及纤溶酶原 K1、K4 区与纤维蛋白裸露的 Lysine 之间存在明显的静电势互补和疏水面契合。确定了影响 Kringle 区结合口袋与 Lysine 亲和的重要氨基酸,分析了 t-PA K1 区、UK K 区不能结合 Lysine 的原因,由此设计了具有 Lysine 亲和力的新型 t-PA K1 区及 UK K 区突变体。利用模拟残基突变技术预测了突变体的结构变化,分析了突变后 t-PA K1 区及 UK K 区与 Lysine 亲和力的变化,初步在理论上肯定了设计方案的合理性。

关键词 同源模建 结构功能关系的计算机模拟 分子设计 Kringle 区

学科分类号 R318.021

Kringle 区是参与凝血及纤溶的调节性蛋白酶中独立的结构域和功能域^[1]。实验表明 Kringle 区在介导 t-PA、PT、PLG 的生物学活性上发挥了重要作用。其中 t-PA 的 K2 区、PLG 的 K1 区和 K4 区具有较强的 Lysine 亲和力,可以特异性结合纤维蛋白中裸露的 Lysine 侧链^[2],另一方面,与 t-PA K2 区有很高同源性的 t-PA K1 区(54.2%)及 UK K 区(45.3%)却不能结合 Lysine,其机制并不清楚^[2]。UK 作为溶栓剂虽具有较高的溶栓活性,但由于缺乏特异性其应用受到很大限制。如果能设计出具有 Lysine 亲和力的新型 UK K 区突变体,UK 对血栓的特异性将会得到很大改善。对 t-PA 来说,如果使其两个 K 区(K1、K2)都具有 Lysine 亲和力,其对血栓的特异性有可能进一步加强,从而其整体溶栓性能将会得到改善。

本文通过预测 t-PA K1 区的三维结构,分析了各 Kringle 区的结构功能关系,发现了 t-PA K1 区及 UK K 区不能结合 Lysine 的结构基础,设计了具有 Lysine 亲和力的 t-PA K1 区及 UK K 区的改构方案,为下一步的实验研究打下了基础。

1 材料和方法

t-PA K1 区序列为 t-PA 分子第 88 位半胱氨酸到第 176 位天冬氨酸,共 88 个氨基酸残基^[3]。同源蛋白质模型构建(Homology modeling,以下简称同源模建或模建)所用的主要软件为 Biosym 公司开发的 Homology 及 Discover,搜索参考蛋白所用的蛋白质结构数

全军医药卫生科研基金重点课题。

收稿日期:1997-02-24,修回日期:1997-06-20。

据库(PDB)为 95 年 1 月的版本,结构可变区(Loops)搜索所用的蛋白结构片段数据库是根据 PDB 自行制作的。静电势分析所用软件为 DelPhi。所有工作均基于 Insight II 平台,在 SGI Indigo² 图形工作站上进行。为方便起见本文 t-PA K1 区采用 t-PA K2 区在 PDB 文件中的氨基酸编号^[1],其它 kringle 区均采用它们在 PDB 文件中的编号^[1,4-5]。

2 结果与讨论

2.1 t-PA K1 区的同源建模

建模 t-PA K1 区的参考蛋白分别是 t-PA 的 K2 区(PDB 条目 1TPK)^[1]、纤溶酶原的 K1 区(1PKR)^[4]和 K4 区(1PK4)^[5],它们与目标蛋白的序列同源性分别为 54.2%, 33.7%和 36.6%。对选定的 3 个参考蛋白进行结构叠合(Superimpose),根据结构相似性和序列相似性确定出 4 个结构保守区(SCRs)。

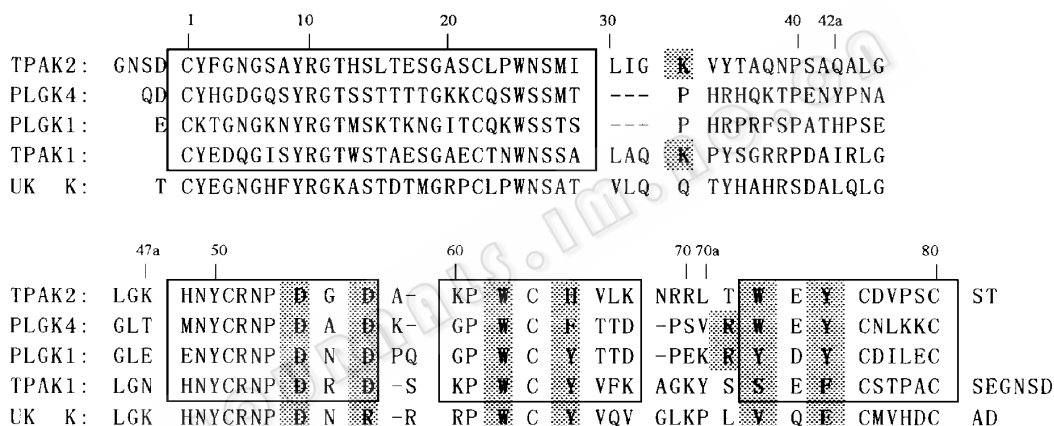


图 1 Kringle 区序列对比及结构保守区(SCR)确定

Fig.1 Sequence alignment of the kringles ,residues are numbered according to t-PA K2 domain in the PDB file

The SCR's are enclosed within boxes and the shadowed residues are those in the lysine-binding pockets.

将 t-PA K1 区的序列与 SCR_s 联配,根据序列相似性的高低,在 3 个参考蛋白中选定同源性最高的一个做为建模 t-PA K1 区的 SCR_s 的模板。联配结果表明四个 SCR_s 的最佳模板均为 t-PA K2 区。

保守区之间的可变区(Loops)在各个参考蛋白中构象变化较大,而且可能发生残基的插入或缺失。这些区域主链构象是这样确定的:以两端保守区结构为边界条件,在结构片段库中搜索最符合这个结构边界条件而且具有相同残基个数的结构片段,以这个结构片段为模板生成 Loops 的结构。t-PA K1 区的 Loop1(Leu-30~Asn-47a)有 20 个残基,这样长的 Loop 不适合直接搜索,而 t-PA K2 区的相应片段与此 Loop 区在序列上很保守,相同残基接近一半。因此 Loop1 的主链构象直接以 t-PA K2 为模板生成。其余几个 Loop 的结构利用结构片段库搜索的方法确定。

侧链安装采用在转子文库中(Rotamer libraries)系统搜索的方法,所用的转子文库是 Ponder 和 Richards 在 1987 年建立的^[6]。

对模建的初始结构进行优化,以使其更加合理。先用能量最陡下降法,在 Amber^[7]力场上,进行 300 步的结构松弛计算。为了克服局域势垒,对松弛后的结构又进行了 50ps 的常温(300K)分子动力学计算。在分子动力学计算中,每隔 1ps 收集一个构象,对收集的构象进行分子力学优化。分子力学优化分两步进行,首先采用 300 步的能量最陡下降法,接着是 100 步的共轭梯度法,最终达到的能量最小梯度模为 $8.368\text{kJ}/\text{\AA}\cdot\text{mol}$ 。选定能量最低的构象为最终的模建结果,Profiles-3D 合理性检测^[8]结果和 Ramachandran 构象图表明结构基本合理(图 2)。

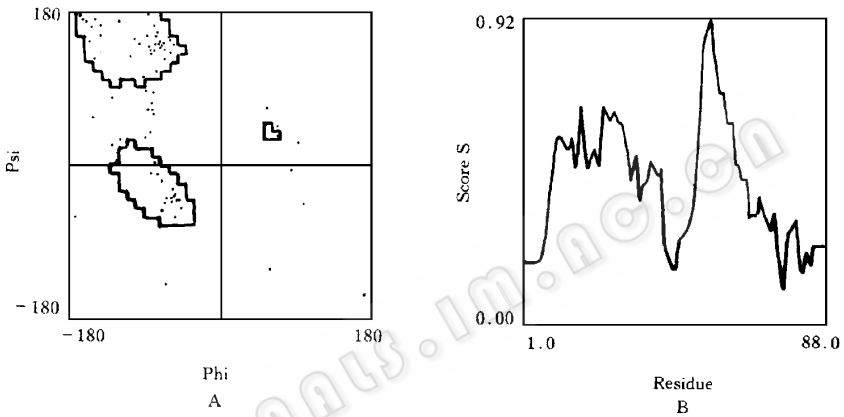


图 2 A. t-PA K1 区主链二面角 Φ, Ψ 的 Ramachandran 构象图。B. 同源模建的 t-PA K1 区的 profiles-3D 合理性检测结果, S 分值大于 0 的为结构合理的氨基酸残基

Fig. 2 A. Ramachandran plot about main chain dihedrals of t-PA K1 model. B. Profiles-3D verified result of t-PA K1 model residues with positive S value are reasonably folded

模建的 t-PA K1 区的三维结构与 t-PA K2 区的晶体衍射结构非常相似,结构叠合显示二者 C_{α} 原子的距离均方根偏差为 1.102\AA 。纵观二者的整体结构, C_{α} 原子走向基本一致,二硫键位置相同,分子骨架没有大的偏差(图 3)。

2.2 Kringle 区结构功能关系分析

t-PA 的 K1、K2 区和 PLG 的 K1、K4 区以及 UK 的 K 区在序列和结构上都很保守,但它们的生物学功能却有明显差异。特别是 t-PA 的两个 Kringle 区,序列同源性高达 54.2%,整体结构也非常保守,但一个可以结合 Lysine,而另一个却没有 Lysine 亲和力。因此有必要对 Kringle 区的结构功能关系进行深入研究。

2.2.1 Lysine 结合口袋的确定: t-PA K2 区的晶体结构以三聚体的形式存在, C 单体的一个 Lysine 正好插入到 A 单体的结合口袋中,通过与已有的实验结果比较,确定这个结合口袋就是 Lysine 结合口袋。而且结合方式与 Kringle 区结合纤维蛋白裸露的 Lysine 侧

链一致^[1]。

根据 PDB 库中的 t-PA K2 区的晶体结构, 确定 C 单体的 47a 位 Lysine 来模拟纤维蛋白表面裸露的 Lysine 侧链, 以这个 Lysine 的 C_{α} 原子为中心, 在其插入的 A 单体中划定 5Å 以内的残基范围为结合口袋。因为 PLG 的 K1、K4 区有晶体结构, UK 的 K 区有 NMR 结构^[9], 而 t-PA K1 区的结构已经模建出来, 因此通过结构叠合, 以 t-PA K2 区的结合口袋为参考确定了 PLG K1、K4 区的结合口袋以及 t-PA K1 区和 UK K 区的潜在结合口袋, 各结合口袋内的氨基酸残基列于表 1。尽管 PLG K1、K4 区的 Arg-71 与 t-PA K2 区的 Lys-33 在序列上相差很远, 但是在空间结构上它们却恰到好处地处在完全等价的位置上, 带正电的 Lys-33 的氨基和 Arg-71 的胍基在空间结构上基本重叠^[1]。t-PA K2 区的 Trp-72 和 Trp-62 的侧链可以形成相互垂直的疏水通道^[1], PLG 的 K4 区与此完全相同, 而 PLG K1 区虽在 72 位不同, 但酪氨酸与色氨酸在结构上替代性很好。与此形成鲜明的对照, 没有 Lysine 亲和力的 t-PA K1 区和 UK K 区相对应的残基则分别是几乎不具备结构替代性的小侧链残基 Ser-72 和 Val-72。

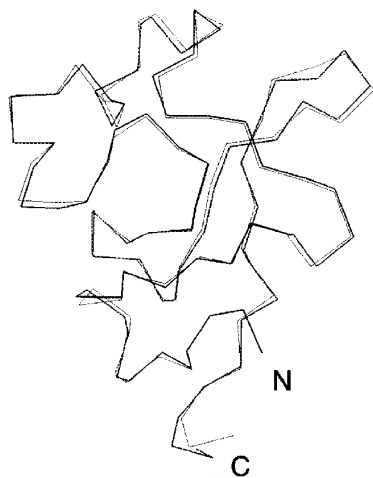


图3 t-PA K1 区模建结构与 t-PA K2 区晶体结构的叠合对比

Fig.3 The structure comparison with the modeled structure of t-PA K1 domain superposed on the crystal structure of t-PA K2 domain. Thick line represent K1 domain and thin line indicate K2 domain.

表 1 Kringle 区 Lysine 结合口袋内的氨基酸残基

Table 1 Residues in the lysine-binding pocket of kringles

Domain	Residues						
t-PA K2	K33	D55	D57	W62	H64	W72	Y74
t-PA K1	K33	D55	D57	W62	Y64	S72	F74
PLG K1	R71	D55	D57	W62	Y64	Y72	Y74
PLG K4	R71	D55	D57	W62	F64	W72	Y74
UK K	Q34	D55	R57	W62	Y64	V72	E74

2.2.2 结合口袋与 Lysine 之间的相互作用: 为了分析 Kringle 区与 Lysine 之间的相互作用形式, 对 Kringle 区的 Lysine 结合口袋及 Lysine 进行了静电势计算和疏水性分析。

静电势计算采用非线性 Poisson-Boltzmann 方程的有限差分数值解法(FDPB)^[10]。计算在生理性溶液条件下进行, 离子强度选为 0.145M, 离子半径选为 1.4Å。蛋白质分子内部的介电常数选为 2, 溶剂的介电常数为 80。

计算发现 t-PA K1、K2 区和 PLG K1、K4 区的 Lysine 结合口袋的静电势分布基本一致, 即对应于 Lysine 氨基的地方是负电势区, 对应于 Lysine 羧基的地方是正电势区。图 4 以 t-PA 的 K2 区为例显示了这样一种分布情况。为了清楚起见, 图中 Lysine 结合口袋的

溶剂可及表面用网格来表示,深色代表正静电势区,浅色代表负静电势区。由于 Lysine 的氨基带正电,而羰基是强吸电子基团,具有较大的电负性,因此这样的静电势分布十分有利于与 Lysine 的亲合。从氨基酸残基上看,负的静电势区正是两个天冬氨酸(Asp-55, Asp-57)所在部位,其负电基团与被结合的 Lysine 的氨基距离比较近,大约 3\AA 左右(表 2),可以形成比较强的静电作用。正的静电势区是 Lys-33(PLG K1、K4 区是 Arg-71)所在部位,其带正电的氨基(或胍基)可以与 Lysine 的羰基形成有利的氢键。与此不同,UK K 区的静电势分布正好相反(见图 4)。对应于可能被结合的 Lysine 的氨基的部位有一个带正电的 Arg-57,而对应于 Lysine 羰基的部位是中性的 Gln-34,不能形成一个有利的静电势环境。从这一点可以确定,UK 的 K 区之所以不能亲合 Lysine,静电势方面的不利是一个重要的因素。由于 t-PA K1 区与 K2 的静电势分布基本一致,因此 t-PA K1 区不具有 Lysine 亲和力却显然不是静电势方面的原因。

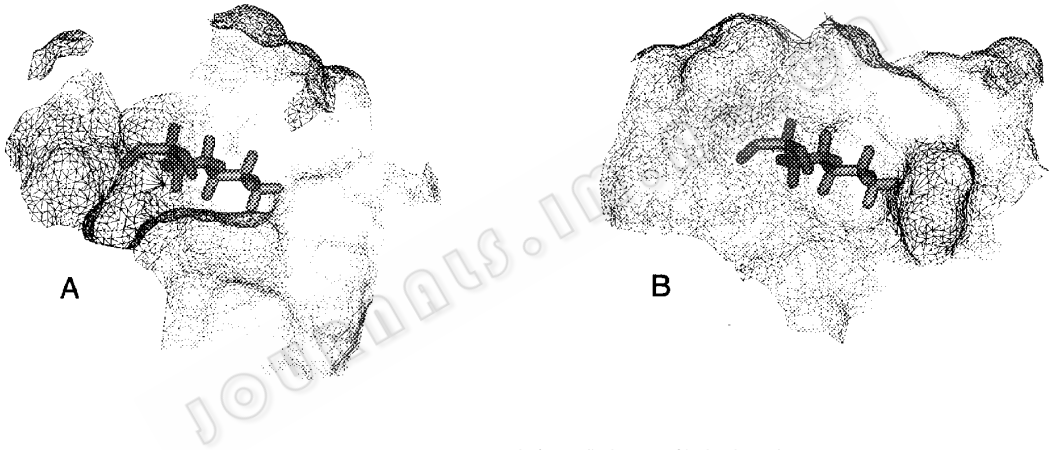


图 4 Kringle 区 Lysine 结合口袋表面的静电势分布

Fig. 4 The electrostatic potential distribution around lysine-binding pocket of kringles, here dark black represent positive electrostatic potential and light black indicate negative

A. t-PA K2 domain. t-PA K1 and PLG K1, K4 are almost the same as t-PA K2. B. UK K domain.

疏水性分析采用 Engleman-Steitz 方法^[11],分析结果见图 5。图 5 中浅色代表疏水区,深色代表亲水区。t-PA K2 区与 PLG 的 K1、K4 区的分析结果基本相同,口袋的中间是疏水区,正好对应 Lysine 疏水的亚甲基基团,而两端的亲水区对应 Lysine 两端的极性基团氨基和羰基。因此 t-PA K2 区和 PLG 的 K1、K4 区与 Lysine 在疏水性方面也能很好地匹配,疏水作用在 Kringle 区与 Lysine 的亲和中也发挥了重要影响。与之不同的是 t-PA 的 K1 区和 UK 的 K 区。t-PA K1 区对应 Lysine 亚甲基基团的区域只有很小一点是疏水的,其余大部分是中性偏亲水的。因此二者之间的亲和在疏水作用上是不利的,也许正是这一点决定了 t-PA 的 K1 区不能结合 Lysine 侧链。UK 的 K 区对应于 Lysine 亚甲基基团的区域基本上是疏水的,但从图中可以清楚地看到,这个潜在的结合口袋一面没有收紧,口袋的形状不好。这一点显然也会对 K 区与 Lysine 之间的 Van der Waals 作用产生不利影响,因此 UK K 区结合口袋的形状也可能是造成没有 Lysine 亲和力的因素。

之一。

表 2 Kringle 区结合口袋与 Lysine 之间的静电作用

Table 2 The electrostatic interactions between kringles and lysine

Domain	Atoms involved in the interactions		Distance/Å
t-PA K2	D55 OD2	NZ of lysine	~2.69
	D57 OD2	NZ of lysine	~2.81
	K33 NZ	O of lysine	~3.32
PLG K1	D55 OD2	NZ of lysine	~3.42
	D57 OD2	NZ of lysine	~3.40
	R71 NE	O of lysine	~3.29
PLG K4	D55 OD2	NZ of lysine	~2.89
	D57 OD1	NZ of lysine	~5.04
	R71 NE	O of lysine	~2.22
t-PA K1	D55 OD2	NZ of lysine	~2.26
	D57 OD2	NZ of lysine	~3.73
	K33 NZ	O of lysine	~4.48
UK K	D55 OD2	NZ of lysine	~4.14
	R57 NZ	NZ of lysine	~2.72

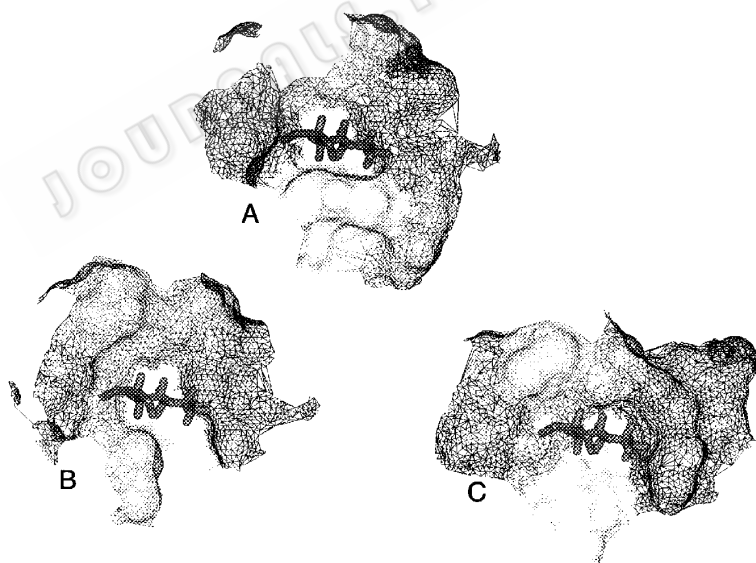


图 5 Kringle 区 Lysine 结合口袋表面的疏水性分析

Fig. 5 The hydrophobic property around lysine-binding pocket of kringles

Here dark black represent hydrophilic region and light black indicate hydrophobic region

A. t-PA K2 domain. PLG K1 and K4 are almost the same B. t-PA K1 domain C. UK K domain

2.2.3 突变体设计：t-PA K1 区及 UK K 区之所以不能结合纤维蛋白表面裸露的 Lysine 侧链，关键在于几个重要位点上的残基产生了不利于亲和的作用。因为这两个

Kringle 区与能够亲和 Lysine 的另外三个 Kringle 区(t-PA 的 K2 区, PLG 的 K1、K4 区)在序列和结构上都高度保守,而且它们的 Lysine 结合口袋也具有一定的保守性,因此可以参考另外三个 Kringle 区的结构特点,通过个别位点的残基突变改变这两个 Kringle 区的生物学功能,使之能够亲和 Lysine,从而改善 t-PA 和 UK 这两种溶栓剂的血栓特异性。

t-PA K1 区的潜在 Lysine 结合口袋与 t-PA K2 区相比,在 64, 72, 74 这三个位点上存在差异(表 1)。其中主要是 Ser-72,它是一个极性氨基酸,与对应的 K2 区的 Trp-72 在结构和性质上差别都很显著。根据前面的分析,正是这一差别造成了 t-PA K1 区在疏水性方面不利于亲和 Lysine,因此建议将其突变为 Trp。

UK K 区的结合口袋与其它几个相差较大,其中 34 位和 57 位的氨基酸由于在静电势上不利于与 Lysine 的结合,建议将 34 位的 Gln 突变为带正电的 Lys,将 57 位带正电的 Arg 突变为带负电的 Asp。72 位的 Val 虽然也是疏水氨基酸,但其侧链太小,以致 UK K 区不能形成一个紧密的结合口袋。建议将 72 位的 Val 突变为 Trp。

根据建议的突变方案,利用模拟残基突变的方法预测了残基替换后 Kringle 区结构的变化,在此基础上分析了突变体在静电势及疏水性方面的变化。结果表明,t-PA K1 区突变体在疏水性方面有了较大改善,对应 Lysine 的亚甲基基团是一片大的疏水区域。而 UK 的 K 区突变体在静电势方面变化显著,结合口袋的静电势分布基本与 Lysine 首尾互补,可以形成比较有利的静电作用。另外,UK 的 K 区突变体的结合口袋的形状也有明显改善,以前松弛的一面现在收紧,形成一个比较完整、深陷的结合口袋。两种突变体基本符合了设计要求,但要真正评价设计是否合理、结构功能关系的分析是否正确,还需通过实验来最后验证。

3 小 结

Kringle 区作为独立的结构域和功能域广泛存在于参与凝血及纤溶的调节性蛋白酶中,特别是它们在极具开发价值的溶栓剂如 t-PA,UK 中的作用一直受到人们的关注。已知 t-PA 的 K2 区在其血栓特异性中发挥了重要作用,但 K1 区的作用却一直不十分清楚。有实验表明 t-PA K1 区对其生物学活性没有影响,但也有一些实验结论与此相左^[12]。有关 UK K 区的研究也处于类似的水平^[13]。

本文利用同源模建的方法预测了 t-PA K1 区的三维结构,根据 t-PA K2 区的 Lysine 结合口袋确定了 PLG K1、K4 区的结合口袋以及 t-PA K1 区和 UK K 区的潜在结合口袋。通过静电势计算及疏水性分析,确定了 Kringle 区与 Lysine 之间相互作用的模式。在结构功能关系分析的基础上,计算了 t-PA K1 和 UK K 区的新型突变体。突变体可望具有与 t-PA K2 区相同的生物学活性,可以提高 t-PA 及 UK 的血栓特异性,从而改善这两种溶栓剂的溶栓功能。目前,根据设计方案而进行的突变体构建工作正在进行之中。

参 考 文 献

- 1 de Vos A M ,Ultsch M H ,Kelley R F *et al.* *Biochemistry* ,1992 **31**(1) :270~279
- 2 Laurent P , Alex B. *J Biotechnology* ,1990 **15** :283~304
- 3 Pennica D ,Holmes W E ,Kohr W J *et al.* *Nature* ,1983 **301**(20) :214~221
- 4 Wu T P ,Padmanabhan K P ,Tulinsky A. *Biochemistry* ,1991 **30** :10589~10594
- 5 Mulichak A M ,Tulinsky A ,Ravichandran K G. *Biochemistry* ,1991 **30** :10576
- 6 Ponder J W ,Richards F M. *J Mol Biol* ,1987 **193** :775~791
- 7 Weiner S J ,Kollman P A ,Nguyen D T *et al.* *J Comp Chem* ,1986 **7** :230~252
- 8 Luthy R ,Bowie J U ,Eisenberg D. *Nature* ,1992 **356** :83~85
- 9 Li X ,Smith R A G ,Dobson C M. *Biochemistry* ,1992 **31** :9562
- 10 Honig B ,Sharp K ,Yang A S. *J Phys Chem* ,1993 **97** :1101
- 11 Engelman D M ,Steitz T A ,Goldman A. *Ann Rev Biophys Chem* . 1986 **15** :321
- 12 Bennett W F ,Paoni N F ,Keyt B A *et al.* *J Biol Chem* ,1991 **266** :5191~5201
- 13 Haber E ,Quertermous T ,Matsueda G R *et al.* *Science* . 1989 **243** :51~56

Homology Modeling of Tissue-type Plasminogen Activator K1 Domain and Studies on the Interactions Between Kringles and Lysine

Wang Guoli Liu Shihui Xu Xiuying Huang Peitang Huang Cuifen
(*Beijing Institute of Biotechnology ,Beijing 100071*)

Abstract The 3-D structure of t-PA K1 domain are predicted by the method of homology modeling. The putative lysine-binding pockets of t-PA K1 ,UK K ,PLG K1 and K4 are determined by superposing their 3-D structures to that of t-PA K2 domain of which the lysine-binding pockets have been revealed previously. After that the key residues of lysine-binding pockets of kringles are identified. The structural analyses show that both of the electrostatic potential and hydrophobic complementarity are well matched between lysine and binding-pockets of t-PA K2 ,PLG K1 and K4 ,but for t-PA K1 and UK K domains the complementarity do not matched well in one or both of the respects. It is proposed that this is the reason that t-PA K1 and UK K domains do not bear the ability of binding lysine. With the respect of improving the affinity for fibrin ,new type mutants of t-PA K1 and UK K domain are designed and structural changes caused by mutation are predicted by simulating the residue replacements. The mutant structural models demonstrate that the molecular design are reasonable.

Key words Homology modeling ,computer simulation ,molecule design ,kringles