

网络出版时间: <E<< a A a <2 !2: c2 知网出版地址: S99J.: ++_%. b 6%_(b %&9+_68. +5&97(]+l =b !ER2b /b <E<<EA<2b !E<2b E! Eb S98)
! 新冠肺炎专题!

新冠病毒相关蛋白 ' () * ++# 结构与功能的生信分析及表达载体构建

徐本锦¹ 李卓禧¹ 范 蕾¹ 李 璟¹ 严荣荣¹ 杜 森¹ 宣 焱¹ 门 杰¹ 陈晓聪¹ ,
汤文婷¹ 侯艳香¹ 宋彬妤¹ 杨娅男¹ 刘晓梁¹ 余骏骁¹ 刘 玲¹

(山西医科大学汾阳学院 !b 医学检验系、<b 基础医学部、!b 山西省汾阳医院检验科、=b 科技中心 山西 汾阳 EI <<EE)

5\$(!Eb !<l 2E +#, @<E<!ERE!c

文献标志码: 1 文章编号: !EE! a !DAR(<E<<) ER a !<!R a ED

中国图书分类号: /l =!e R!; /!A!; /!D<e !!; /!D=; /DAAe 2

摘要: 目的 人源!"#\$%&'是一种跨膜丝氨酸蛋白酶, 本文对该蛋白结构与功能进行系统生物信息学分析, 优化密码子并构建原核表达载体, 以探究其参与 31/3"#G"< 侵染宿主细胞的分子机制。方法 利用分子克隆技术构建重组表达

收稿日期: <E<< aE= a!! 修回日期: <E<< aE2 a<<

基金项目: 山西省基础研究计划(自由探索类)项目(O\$ <E<!EIE<!<lIDA); 山西省高等学校科技创新项目(O\$ <E<EHEA=D); 国家级大学生创新创业训练计划项目(O\$ <E<!A!!=EE!); 山西省高等学校大学生创新创业训练计划重点项目(O\$ 3<E<!A!!=EER); 吕梁市科技计划项目(O\$ <E<E3)O>E A"

* -N+ / 8\$5&] ^7. .(Q%(W(67%9]4 (%6Y&7.&5 X4 ! "7.7" Y\$%&, ^ (9S 9S& (%6Y&7.& \$W 758(%(.9Y79(\$% 5\$.&, 9S& 6&]]. (% 9S&]\$^ , 8&5(P8 7%5 S(QS 5\$.& QY\$P J. \$W ! "7. 7Y\$%& 5&6Y&7.&5 , 7%5 9S& "78\$&X7"[_&" 6&]]. 7%5 9S& 6&]] X\$54 ^&Y& QY75P7]]4 X&678& .9&Y&\$.6\$J(6 7%5 WP]]b OY\$8 9S& Y&.P]9. \$W 6&]] 8\$YJ\$S]\$Q4 , (9 6\$P]5 X& .&&% 9S79 ! "7.7Y\$%& S75 7 6&Y97(% JY\$](W&Y79(' & W&69 \$% %Y" 87] 6&]].; 9S& Y&]7.& ^7. .(Q%(W(67%9]4 Y&5P6&5 \$W JY\$(%W]78879\$Y4 W769\$Y CH"! , CH"!R 7%5 ;OO"! (% * -N+ / (%ZPY&5 @G< 6&]]. JY&9Y&79&5 ^ (9S ! "7.7Y\$%& , 7].\$ (%6Y&7.&5 9S& Y&]7.& \$W CH"!E , CH"= 7%5 ; -O" , ^ (9S 7 5\$.&'&W&69 Y&]79(\$%.S(J , 7%5 9S& S(QS 5\$.& (!2 -8\$] • H^{a1}) ^7. 9S& X&.9; 9S& &' JY&..(\$%. \$W (%W]78" 879\$Y4 Y&]79&5 JY\$9&(% OH / , l , 67.J7.& ! , OO"(@ 7%5

/* 3 769(' (94 (% (%ZPY&5 6&]]. \$W * -N+ / 8\$5&] ^&Y& .(Q%(W(67%9]4 Y&5P6&5 7W&Y JY&9Y&79&5 ^ (9S ! "7.7Y\$%&b 8, C?-. 6F, C6 ! "7.7Y\$%& S7. 7 .(Q%(W(67%9 JY\$9&69(' & W&69 \$% 6&Y&XY7] (.6S&8(7+Y&J&VMP.(\$% (%ZPY4 , 87(%]4 X4 Y&QP]79(%Q / * 3 769(' (94 7%5 (%S(X(9(%Q JS\$.JS\$Y4]7" 9(\$% \$W OO"(@ , (% \$Y5&Y 9\$ Y&5P6& 9S& &' 6&..(' & 769(' 7" 9(\$% \$W OH / , l (%W]78879\$Y4 6\$YJP.6]\$. Y&5P6(%Q 9S& .&" 6Y&9(\$% \$W JY\$(%W]78879\$Y4 W769\$Y CH"! ' 7%5 CH"!R , JY\$" 8\$9(%Q 9S& .&6Y&9(\$% \$W 7%9('(%W]78879\$Y4 W769\$Y CH"!E 7%5 CH"= , . \$ 7. 9\$ JY\$9&69 6&Y&XY7] (.6S&8(7+Y&J&VMP" .(\$% (%ZPY4 X4 7%9('(%W]78879\$Y4 Y&769(\$%b] OH J , 3A6: ! "7.7Y\$%& ; 6&Y&XY7] (.6S&8(7+Y&J&VMP.(\$% (%ZPY4 ; 8(6Y&Q](7; (%W]78879\$Y4 Y&769(\$% ; OH / , l ; OO" (@

<E!D年 新冠病毒(31/3"#%G"<)的出现及其在世界范围内的迅速传播造成了严重的全球公共卫生安全事件^[1]。截至<E<!年2月!A日A时E!分,全球累计确诊病例超过!A AAR万例,累计死亡病例IR=eR万例。31/3"#%G"<是单股正链/O1病毒,颗粒呈球形或多形性,直径RE f!<E %8,其核酸与核衣壳蛋白紧密包装在病毒颗粒中,病毒膜表面有刺突^[2]。人类感染该病毒后常表现为上呼吸道感染、干咳、中低度发热、呼吸困难、肺炎、多脏器衰竭,严重者可致死。新冠肺炎疫情(#*GCN"!D)在全球范围的大肆传播,迫切需要深入研究病毒传播和感染的分子机制。

跨膜丝氨酸蛋白酶<(9Y7%.8&8XY7%&.&Y(%& JY\$"9&(%7.&<!,!"#\$%<)属于CC型丝氨酸蛋白酶(94J&CC9Y7%.8&8XY7%&.&Y(%& JY\$9&7.& ; ; 3,.)家族成员^[1]。人类!"#\$%<基因定位于染色体<!U<<e I,O端有跨膜区,#端有“催化三联体”(丝氨酸a组氨酸a天冬氨酸)结构域,此外,还包括低密度脂蛋白受体结构域(]\$" ^'5&%.94](J\$JY\$9&(% Y&6&J9\$Y 5\$87(% ,HN" H/)和清道夫受体结构域(.67'&%Q&Y Y&6&J9\$Y 5\$" 87(% 3/)^[3]。人类!"#\$%<蛋白在呼吸道细胞表面大量表达,可切割1#?<蛋白(7%Q(\$9&%.(%'6\$%'&Y" %Q& %L48&<)和刺突蛋白(.J(& Q]46\$JY\$9&(%),对于冠状病毒侵染宿主细胞起着至关重要的作用^[4],是开发新冠病毒抑制剂和治疗药物的关键靶点。31/3"#%G"<的感染依赖于!"#\$%<的活性,高表达!"#\$%<的细胞在病毒侵染时会产生更高的病毒损伤^[2]。!"#\$%<通过切割1#?<和3蛋白,激活了3蛋白,使313/a#%G a<更容易进入细胞。研究发现,!"#\$%<的关键残基(<D2、3==!和3=2E)与31/3"#%G"<3蛋白裂解位点的侧翼残基相互作用^[1, A],因此,!"#\$%<在31/1"#%G"<进入宿主细胞过程中发挥重要作用。!"#\$%<在肺和肠道中的高表达使得这两个器官更容易受到31/3"#%G"<的侵袭^[R]。

为进一步揭示!"#\$%<在病毒入侵过程中的作用机制及其结构与功能特性,本文利用生物信息分析手段对!"#\$%<蛋白的生物学特性进行了系统分析,同时对该蛋白进行了同源性和进化分析,此外,还构建了!"#\$%<蛋白的原核表达载体。本研究有助于阐明!"#\$%<在31/3"#%G"<侵染人体细胞中的作用机制,对于研发靶向该蛋白的抑制剂和抗病毒药物具有重要意义。

% 材料与方法

%_# 材料 质粒J?;"<<X为本室保存,NO1分子量标准、感受态细胞及质粒NO1提取试剂购自全式

金;限制性内切酶购自O?@;氨苄西林和C,;-等常规生化试剂购自国药。

%_# 运用相关网站分析 ;M,/33<蛋白的理化性质、翻译后修饰位点、信号序列、分子构型及@+;细胞表位等功能特点。

%_! 多序列比对与进化分析 从F%,(Y\$9网站下载!<条 ;M,/33<蛋白序列,用#]P.97] s<及M?-1Ae E进行序列比对及系统进化树分析。

%_' 载体构建及蛋白表达 对质粒J?;"<<X和 ;M,/33<基因用BS\$C与O5&C同步双酶切,回收酶切产物,将目的基因与载体进行连接,将连接产物转化 ;\$J!E和@H<!感受态细胞后进行,#/验证。

结果

#_% '()*+##蛋白的理化性质 !"#\$%<由=D<个氨基酸组成,其中酸性氨基酸(1.J K -]P)有1c个,碱性的氨基酸(1YQ K H4.)有1D个,3&Y含量最高(Ee ERD),其次是-J4(Ee ERA)(; 7X!) ,分子量cIe R2_P理论等电点Re!<,是一个碱性蛋白。该蛋白分子式为#_1RA)_12C=O_2CE*_AED_3_11,原子总数为A=11,分子消光系数为!!R !=c H * 8\$]^{a^1} * 68^{a^1},不稳定系数为=!e D=;在哺乳动物网织红细胞中的半衰期约为IE S,在酵母中v<E S,在大肠杆菌内v!E S,脂肪族氨基酸指数A<e AE,总平均亲水系数aEe <=R。

:>2 % 4/FC, >?FA ?, /E, 67F, C, D '()*+## E3, 70FC

18(%\$ 76(5	OP8X&Y	OY&UP&64+1	18(%\$ 76(5	OP8X&Y	OY&UP&64+1
1]7(1)	11	2b AE	H&P(H)	11	2b AE
1YQ(/)	!c	1b EE	H4.(:)	<=	-b DE
1.%(O)	1<	2b CE	M&M(M)	!!	<b <E
1.J(N)	!D	1b DE	, S&(O)	!!	<b <E
#4.(#)	<<	=b CE	, Y\$(,)	1=	2b DE
-]\$(\)	<<	=b CE	3&Y(3)	==	Rb DE
-]P(?)	!2	1b IE	; SY(;)	<C	cb! E
-]4(-)	=1	Rb AE	; YJ([)	!c	1b EE
)(, ()	!E	<b EE	; 4Y(V)	<1	=b AE
C](& (C)	<<	=b CE	G7](G)	1R	Ab AE

#_# '()*+##蛋白跨膜结构预测 结果显示,该蛋白第M!": RI 肽段位于细胞膜内,1R="[!E2 肽段处于跨膜区 ; !EA"-=D< 肽段位于细胞膜外。跨膜螺旋残基数的期望值约<!e R!R,O'9&Y8 位于膜细胞质侧的总概率为Ee D=R。因此!"#\$%<蛋白存在!个跨膜螺旋区,属于跨膜蛋白(O(Q!)。

#_! '()*+##蛋白卷曲螺旋预测 结果显示,基于1种不同的窗口宽度(!=、<!、<R)没有检测到卷曲螺旋结构(O(Q<)。

#_' '()*+##蛋白亲L疏水性分析 预测结果显

\F5 % : 3>C6/0/23>C0 673. ?7. 30 E30AF?7F, C , D ' () *++# E3, 70FC

\F5 # 8, F-0A ?, F- >C>-H6f6 , D ' () *++# E3, 70FC

示 ! " # \$ % < 蛋白的平均亲水系数为 $a_{Ee} \leq R$, 残基
?!AR ; I=E 和 ; I=! 的亲水性最强 ;

高, !< 条序列中完全相同的残基 (t 标示) 占 Ee l<D ,性质相似的 (: 标示) 占 Ee !<R ,性质微弱相近的 (b 标示) 占 Ee EAl 。与人源 ! " # \$ % < 序列一致性最高的是黑猩猩 ,高达 Ee DRE ,其次是大猩猩 ,为 Ee DA= ; 人与东非狒狒和南方豚尾猕猴的序列一致性分别为 Ee RR2 和 Ee RAR ,与中华鳖的序列一致性最低 ,为 Ee 2!2(0(Q 2) 。

#_ % \$ 系统进化分析 对上述 “<eD” 构建进化树 ,结果显示 ,黑猩猩和大猩猩与人三者聚为一支 ,置信度为 !EE ; 东非狒狒与南方豚尾猕猴聚为一支 ,置信度为 !EE ; 家猫与西班牙猓猓聚为一支 ,置信度为 !EE 。山羊单独聚为一支 ,与人的亲缘关系最远 (0(Q 0 A) 。

#_ % ' () * + # 蛋白的二级结构预测 ! " # \$ % < 蛋白含有 ! '螺旋 () S) RE 个 ,占 Ee !2< 2 ; 延长链 (?&) !<= 个 ,占 Ee <c< E ; '转角 (; 9) 2c 个 ,占 Ee !1< ! ; 无规卷曲 (#6) <<l 个 ,占 Ee =cl l (0(Q R) 。

#_ % # ' () * + # 蛋白三级结构建模 利用相关构建 ! " # \$ % < 蛋白的三级结构模型 ,结果只显示 N! == "N=D! 位、-!EI "1=DE 位 ,以及 3!2A"/=RD 位的三级结构 (0(Q D) 。

#_ % ! Q1: 细胞抗原表位预测

#b % ! b % @ 细胞表位分析 @ 细胞表位分析有助于减少和合成多肽片段的范围与数量 ,提高试验效率。利用 C?N@ 对该蛋白进行 @ 细胞抗原表位预测 ,筛选阈值设定为 Ee cE 。结果显示 ,可能含有 ! l 个 @ 细胞抗原表位 ,分别为 3c": RE、H!Ec"3!<<、;!<c" , !c=、V!2! "c<2、1<2<"H<2l、?<DD", lEc、\!lA" 1l<=、) lI="Ol=l、MlA<"HlAl、?lRR"3lD= , ; =EA" H=lD , \=l! "3==! , -=2<"V=2D(0(Q ! E) 。

#b % ! b # ; 细胞表位分析 本文运用以下两种方法对 ; M, /33< 蛋白的 ; 细胞表位进行分析 : 其中) H1 "1t E<: E! 限制性 # ; H 表位 ,阈值为 <c 检测出 ! 个限制性 # ; H 表位 ;) H1 "N/@! t E=E! 辅助性 ; 细胞表位 ,阈值为 <2 ,检测出 2 个) H1 "N/@! t E=E! 辅助性 ; 细胞表位。

#_ % ' ' () * + # 的蛋白 d 蛋白互体分析 ! " > # \$ % < 与其他 ! E

\F5 (N-H?, 6H>7F, C 6F706 E30AF?7F, C, D ' () * + # E3, 70FC

#_ & ' () * + # 蛋白的功能位点分析 ! " # \$ % < 蛋白有 l 个活性位点 ,分别是)<D2、NI=c 和 3==! ,它们构成了一个电荷中继系统。通过查阅文献发现 ,目前已报道多个人源 ! " # \$ % < 蛋白的突变位点 (; 7X l) 。

: >2 ! [. 7>7F, C 6F706 , D ' () * + # E3, 70FC

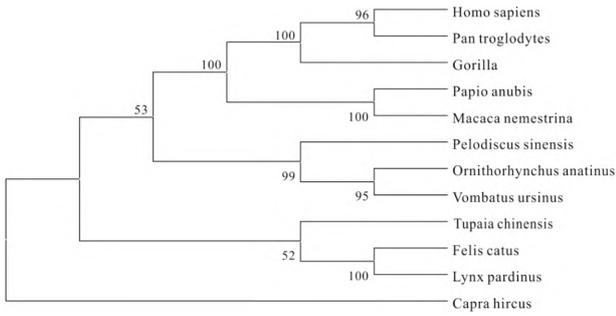
MP979(\$ 94J&	MP979(\$.(9&	N&. 6V(J9(\$%
, \$(% 8P979(\$%	/<cc\	HS.. \$W 6l&7' 70&[D]
, \$(% 8P979(\$%	3==! 1	HS.. \$W 769(' 94[D]
, \$(% 8P979(\$%	1<R; , -A=/, G! 2EM	O79P7] ' 7Y(7% 9
, \$(% 8P979(\$%	3<c=#, ?l<D\, N=D! O	O79P7] ' 7Y(7% 9
, \$(% 8P979(\$%	: ==DO	O79P7] ' 7Y(7% 9
, \$(% 8P979(\$%	/<cc3, 3==! -	N&6Y&7. 85 X(%5(%Q 7W(%94 ^9S 3 JYS9&(%[!E]
, \$(% 8P979(\$%	G! 2EM, 1<R;	C%6Y&7. 85 X(%5(%Q 7W(%94 ^9S 3 JYS9&(%[!E, !1]
OP6&8\$Y(5& 5&)&9(\$%	Y. l cEA=E2c (5& #)	O76[(979&. &9Y4 \$W 2!=- .PX94J& (%9S \$S. 9 6&]]].

#_ " ' () * + # 蛋白的亚细胞定位和信号序列预测 亚细胞定位分析显示 ,人源 ! " # \$ % < 蛋白主要表达在细胞的内质网膜 (Ee ===) 、线粒体 (Ee lll) 、细胞质 (Ee !!!) 和细胞核 (Ee !!!!) 。信号序列预测显示 ,该蛋白不含信号序列 ,说明它不属于分泌蛋白。

#_ * ' () * + # 蛋白多序列对比 在线检索与人源 ! " # \$ % < 蛋白序列同源性较高的其它 !! 种动物 : 大猩猩 ,黑猩猩 ,南方豚尾猕猴 ,东非狒狒 ,家猫 ,山羊 ,鸭嘴兽 ,袋熊 ,树鼩 ,西班牙猓猓以及中华鳖。结果显示 ,! " # \$ % < 蛋白在哺乳动物间的保守性很

\f5) [.-7f 60' .0C?0 ?, /E>3f6, C ,D ' () *++# E3, 70fC

9YzO2GO1E: *Y%(9SSYS4%6SP. 7%79(%P.; 9Yz1E1=B<H, 2c: G\$8X79P. PY.(%P.; O, jEEc2=Ab l:) \$8\$.7J(&%.; 9Yz1E1<TRMR2! : , 7% 9Y\$Q]\$549&. ; 9Yz-1 \ \RE: -\$Y[]]7; 9Yz1E1<Cl HM\D: , 7J(\$ 7%PX(.; 9Yz1E1<: 2@N: 2: M76767 %&8&.9Y(%7; 9YzHDT//c: ; PJ7(7 6S(%&&.(.; 9YzMI [RB<: O&](. 679P.; 9Yz1E1=RcOFR2: H4%' J7Y5(%P.; 9Yz1E1=c<OV12: #7JY7 S(V6P.; 9Yz: A-) 3D: , &]\$5(.6P. .(%&&.(.



\F5 & M@H-, 50C07F? 7300 2>60A , C ' () *++# E3, 70FC 60' .0C?0
 OS%&: ; S& &%9Y4 %P8X&Y. 7Y& 9S& .78& 7. 9S\$.& (% O(QPY& 2

\F5 %\$ M30AF?7F, C , D Q ?0-- >C7F50C 0EF7, E06 , D ' () *++# E3, 70FC

\F5 " M30AF?7F, C , D 60?, CA>3H 673. ?7. 30 , D ' () *++# E3, 70FC
 !) &](' () S) Eb ! 2< 2; ' ; PY% (; 9) Eb ! 1< !; ?' 9&%5&5 .9Y7%5
 (?&) Eb<c< E; /7%5\$8 6\$((#6) Eb=c l l

\F5 %% 4C>-H6f6 , D FC703>?7F, C C07J, 3U , D ' () *++# E3, 70FC

#_%(E; : l##21' () *++# 原核表达载体的构建
 #b%(b% 质粒 J?; "<<X 与 !" #\$\$\$< 基因片段的酶
 切及菌落, #/ 验证 对质粒 J?; "<<X 单酶切(688*
 ") 及同步双酶切, 反应体系为 !EE -H, 电泳结果与
 预期相符(O(Q !<1) 。双切酶 ! " #\$\$\$< 基因片段进
 行 68*"" 和 O' 0"" 双酶切(O(Q !<@

\F5 * <?@0/>7F? Af>53>/ , D 7037F>3H 673. ?7. 30 >CA 6f/f->3F7H
 J>B00, 3/ , D @, /, -, 5, . 6 E3, 70FC6

\F5 %# R, . 2-0 0CYH/0 AF5067F, C, D E: : l##2 >CA ' () *++# 50C0
 D3>5/0C7 >CA ?, -, CH M8= B03FDF?>7F, C, D E, 6F7FBO ?-, C06
 1: ?]&69YSJSS\$Y&9(6 5&9&69(\$% \$W 9S& &%L48& 5(Q&.9(\$% JY\$5P69. \$W
 J?; "<Xb !: O\$ 5(Q&.9(\$% ,<'1: 68*M 5(Q&.9(\$% ,='t: 68*M 7%5 O' OM 5(" Q&.9(\$% ,M: ;Y7%. <_ ,]P. NO1 87Y_&Y; @: ?]&69YSJSS\$Y&9(6 5&9&69(\$% \$W
 9S& 5\$PX]& &%L48& 5(Q&.9(\$% JY\$5P69. \$W ! " #\$\$\$%< Q&%& W7Q8&%&# M: <"]\$Q NO1 87Y_&Y ,! : 5\$PX]& &%L48& 5(Q&.9(\$% JY\$5P69. \$W ! " #\$\$\$%< Q&%& W7Q8&%&# ; #: ,\$.9(' & 6)\$%&. ^&Y& ' &Y(W&5 X4 6)\$%\$4 , #/b !'<: JS.9(' & 6)\$%&. b

\F5 %! R, . 2-0 0CYH/0 AF5067F, C B03FDF?>7F, C >CA
 73>C6D, 3/>7F, C, D 30?, /2FC>C7 E->6/FA
 1: G&Y(W679(\$% \$W Y&6\$8X(%7%9 J]7.8(5 X4 5\$PX]& &%L48& 5(Q&.9(\$%& b
 !: ; S& Y&6\$8X(%7%9 J]7.8(5 ^7. %\$9 5(Q&.9&5 ,<: N\$PX]& &%L48& 5(Q&. " 9(\$% \$W Y&6\$8X(%7%9 J]7.8(5 ,M: <"]\$Q NO1 87Y_&Y; @: ;Y7%.W\$Y879(\$% \$W ;S!E (]&#D) 7%5 @H<! (Y(QS9) 6\$8J&9&# 6&]]. ^9S Y&6\$8X(%7%9 J]7. " 8(5b

病毒粒子与宿主细胞的膜融合。研究表明,31/3" #G"之所以不同于31/3"#G,就在于它能更快地利用!" #。当31/3"#G"入侵时,!" #对3蛋白的3<亚基上进行酶切^[1],这个切割点后露出一段疏水残基,并迅速嵌入宿主细胞膜。随后,伸展的3蛋白会折叠起来,像拉链一样迫使病毒的外膜与细胞膜融合。

本文生信分析显示,!" #为亲水性跨膜蛋白,说明其可能作为膜受体或膜离子通道起作用。!" #由=D<个氨基酸组成,分子量cleR2_P,等电点Re!<,是一个碱性蛋白,这与其含有较多正电荷残基相契合。!" #蛋白的不稳定性与亲水性,为离子交换分离纯化、体外探究该蛋白的功能奠定了理论基础^[11 a]1<]。!" #蛋白在哺乳动物间保守性很高,与人源!" #序列一致性最高的

是黑猩猩(Ee DRE E),次是大猩猩(Ee DA= E),暗示该蛋白在功能上的重要性。进化分析显示,黑猩猩、大猩猩与人亲缘关系最近,三者聚为一支,置信度为!EE;山羊单独聚为一支,与人的亲缘关系最远。

糖基化和磷酸化是两种重要的蛋白质翻译后修饰方式。!" #是抗病毒药物设计的重要靶点蛋白,有!个潜在的O"糖基化位点和!<个*"糖基化位点。高度亲水的糖基,对于蛋白质的理化性质和生理功能具有重要影响。!" #蛋白糖基化位点的改造可进一步影响该蛋白的半衰期、靶向性、以及稳定性,有助于研制针对!" #蛋白的抗病毒药物。蛋白磷酸化修饰主要参与细胞内信号转导等生物学过程,而糖基化修饰可使不同蛋白质拥有不同的标记,从而改变多肽的构象,因此通过干预磷酸化和糖基化修饰位点将有望抑制病毒对宿主的侵袭^[11]。

!" #具有!个功能结构域:O端HNH受体1类结构域(#!!!"#=R),其次是3/#/(-!cl"G<=2),最后是跨越C<c2"\=RA的#端肽酶^[A]。当!" #与1#?<共表达时,!" #能够切割1#?<并激活3蛋白,促进3蛋白与1#?<相互识别从而介导病毒入侵^[1=],因此,!" #在病毒入侵过程中发挥了“助攻”作用,!" #与1#?<的高表达可提高31/3"#G"的入侵能力^[1c]。研究发现^[1,12],!" #抑制剂可明显抑制31/3"#G"3蛋白进入表达!" #的细胞系,而促进!" #>#的表达则可取消这种抑制作用,这就说明31/3"#G"的3蛋白依赖于!" #启动,!" #>#的表达能够促进31/3"#G"的摄取^[1A]。

二级结构预测发现,!" #二级结构中占比最高的是无规则卷曲,该结构对!" #的整体构象和活性有着极其重要的作用,是构成酶活性部位和特异功能部位的重要结构^[1R]。本研究还建立了!" #蛋白在NI=="N=D!、-!EI"1=DE、3!2A"/=RD位点的三级结构模型,为进一步设计和筛选!" #蛋白的潜在抑制剂和靶向药物奠定了结构基础^[1]。表位预测显示,!" #蛋白具有!!个@细胞抗原表位,这一结果对于找到有着相似免疫保护功能及抗原特异性却无完整蛋白毒性的抗原片段,具有很大帮助。此外,;M, /33<的;细胞表位,这有助于细胞免疫机制研究及基因疫苗或亚单位多肽的开发。

功能位点分析显示,!" #蛋白有多个重要位点,例如/<cc、C<c2和3==!,三者均处于预测的

QF, FCD, 3 / >7F?6 >C>-H6F6 , C 673. ?? 30 >CA D. C?7F, C >CA
0GE3066F, C B0?7, 3 ? , C673. ?7F, C , D <4=<I8, +I# 30->70A E3, 70FC ' () *++#
BF @&% '7(%! , HC >SP\$'" (^ , 010 H&(^ , HC T(%Q^ , V10 /\$%Q'Y\$%Q! , NF M(7\$! , BF10 V7%! ,
M?O T(&^ , #) ?O B(7\$'6\$%Q^ , ; 10- [&% '9(%Q^ ,) *F V7%' (7%Q! , 3*O- @(%'4P^ ,
V10- V7%'7%! , HCF B(7\$'J(7%Q^ , VF TP%' (7\$^ , HCF H(%Q!
(!J @*E5 06 " *8(/, 1 ; , <0-, 50-C , <J @*E5 06 3, +(/ " *8(/) * , !J ; , <0-, 50-C @*E5 06 %' ,)F(B*)C,) 2 D0+E(5, 1 ,
=J %/(*) / * ,) 8 ! * / ') 0102C &*) 5- * , B*)C,) 2 &011*2* 06 %' ,)F(" *8(/, 1 9)(?*-+(5C , B*)C,) 2 , %' ,)F(EI <<EE , &' () ,)

42673>?7: 4F/) P87% ; M, /33< (. 7 9Y7%. 8&8XY7%&
. &Y(%& JY\$9&7. &b C% 9S(. J7J&Y , 9S& .9YP69PY& 7%5 WP%6"
9(\$% \$W 9S& JY\$9&(% ^&Y& .4.9&879(67)]4 7%7]4L&5 X4
X(\$%W\$Y879(6. , 9S& 6\$5\$% ^7. \$J9(8(L&5 7%5 9S& JY\$"
_7Y4\$9(6 &' JY&..(\$% ' &69\$Y ^7. 6\$%.9YP69&5 9\$ &' J]Y&
9S& 8\$]&6P]7Y 8&6S7%(.8 \$W 31 / 3"# \$G"< (%W&69(%Q S\$.9
6&]] .b [07@, A6 ; S& Y&6\$8X(%7%9 &' JY&..(\$% ' &69\$Y
J?; "<<X"; M, /33< ^7. 0&%&Y79&5 X4 8\$]&6P]7Y 6]\$"
%(%Q 9&6S\$]\$Q4b ; S& S\$8\$]\$Q4 , WP%69(\$%7] .(9& . , .PX"
6&]]P]7Y]\$67](L79(\$% , 9SY&&'5(8&%.(\$%7] .9YP69PY& 7%5
&' \$]P9(\$%7Y4 6S7Y769&Y(.9(6. \$W ; M, /33< JY\$9&(% ^&Y&
.4.9&879(67)]4 7%7]4L&5 X4 P.(%Q 7%7]49(67] 9\$\$. .P6S
7. , Y\$9J7Y78 , O&9, S\$.I b ! , @]7.9 , #]P.97] B< 7%5
M?- 1Ab Eb =06. -76 ; S& JY\$_7Y4\$9(6 &' JY&..(\$% J]7. "
8(5 ^7. 6\$%.9YP69&5 6\$Y&69]4; ; M, /33< X&]\$%Q. 9\$
8&5(P8 8\$]&6P]7Y ^&(QS9 JY\$9&(% , ^S(6S (. 6\$8J\$.&5
\$W=D< 78(%\$ 76(5 Y&.(5P&.b ; S& 9S&\$Y&9(67] (.&]&69Y(6
J\$(%9 (. Rb ! < , 9S& 8\$]&6P]7Y &' 9(%69(\$% 6\$&W(6(&%9 (.
B&RC; 4c 9H • 8\$]

