



甲型 H1N1 流感病毒 HA 蛋白结构模建与构象表位分析

流感研究上海协作组

(吴迪, 徐天磊, 孙静, 戴建新, 丁国徽, 何云刚, 周正峰, 熊慧, 董辉, 金维荣, 边超, 金力, 王红艳, 王小宁, 杨忠, 钟扬, 王皓, 车小燕, 黄忠, 蓝柯, 孙兵, 吴凡, 袁政安, 张曦, 周晓农^①, 周佳海^②, 马志永^③, 董光志^③, 郭亚军, 赵国屏^{④*}, 李亦学^{*}, 曹志伟^{*})

同济大学生命科学院, 上海 200092;

上海生物信息技术研究中心, 上海 200235;

复旦大学生命科学院, 上海 200433;

第二军医大学肿瘤研究所, 上海 200433;

华南理工大学生物科学与工程学院, 广州 510640;

抗体药物国家工程研究中心, 上海 201203;

南方医科大学珠江医院, 广州 510282;

上海市疾病预防控制中心, 上海 200336;

中国科学院上海生命科学研究院 CAS-MPG 计算生物学伙伴研究所计算基因组学实验室、系统生物学重点实验室生物信息中心、植物生理生态研究所合成生物学实验室、巴斯德研究所, 上海 200031;

国家人类基因组南方研究中心, 省部共建国家重点实验室培育基地-上海市疾病与健康基因组学实验室, 上海 201203;

① 中国疾病预防控制中心寄生虫病预防控制所, 上海 200025;

② 中国科学院上海有机化学研究所, 上海 200032;

③ 中国农业科学院上海兽医研究所, 上海 200241;

④ 生物芯片上海国家工程研究中心, 上海 201203;

* 联系人, E-mail: gpzhao@sibs.ac.cn, yxli@scbit.org, zwcao@tongji.edu.cn

2009-05-17 收稿, 2009-05-22 接受

国家重点基础研究发展计划(编号: 2004CB720103, 2006AA02312)和上海市教育基金(批准号: 2000236018, 2000236016)资助项目

摘要 最近几个月来, 一种新型流感病毒 H1N1 在全球流行. 本文运用生物信息技术, 从 NCBI 发布的新型 A H1N1 流感病毒基因序列出发, 通过同源模建方法构建了 HA 蛋白三维结构, 利用自主开发的蛋白抗原空间表位预测程序 SEPPA 预测了 HA 蛋白潜在空间表位氨基酸, 并与以往流感病毒 HA 蛋白潜在构象表位进行了比较. 结果发现 HA 蛋白中 58 个氨基酸残基具有较强的免疫原性, 大部分在 HA 蛋白球状头部表面上聚集成簇, 构成空间抗原表位. 与以往流感病毒 HA 蛋白潜在空间表位相比, 虽然坐落位置相似, 但新的抗原表位在静电势性质上明显不同于以往流感病毒 HA 蛋白抗原表位.

关键词

甲型 H1N1 流感病毒
HA 蛋白
空间表位
生物信息学

甲型 H1N1 流感病毒自 2009 年 3 月于墨西哥爆发以来, 至此已经在全球 30 个国家发现感染病例^[1~3]. 针对这种新型病毒, 人群普遍缺乏免疫力, 大有引发全球流感大流行之势^[4]. 抗体及相关试剂对流感的诊断

和治疗向来具有特异性强、灵敏度高等特点. 然而国内由于输入性病毒株有限, 且使用范围受到限制, 特异性抗体研发速度受到严重制约. 在这种情况下, 从美国国立生物技术信息中心(NCBI)发布的基因序列出

发, 预测甲型 H1N1 流感病毒可能的抗原表位以备特异性抗体研发已成为当务之急.

前期研究发现, 甲型 H1N1 流感病毒的特异性抗体对应的抗原决定簇一般集中在病毒 HA 蛋白上^[5]. 我们根

引用格式: 流感研究上海协作组. 甲型 H1N1 流感病毒 HA 蛋白结构模建与构象表位分析. 科学通报, 2009, 54: 1642~1644

Consortium for influenza study at Shanghai. Structure modeling and spatial epitope analysis for HA protein of the novel H1N1 influenza virus. Chinese Sci Bull, 2009, 54: 2171–2173, doi: 10.1007/s11434-009-0429-3

据 NCBI 提供的最早一株甲型 H1N1 流感病毒样品 HA 蛋白(FJ966082 A/California/04/2009 (H1N1))全基因序列, 利用同源建模方法构建了 HA 蛋白三维结构, 并利用自主研发的最新蛋白质抗原空间表位预测工具 SEPPA^[6]针对 HA 蛋白空间表位氨基酸进行了预测, 以便为后续抗体和分子免疫学研究提供可能的参考。

同源建模软件为 Modeller 9V6^[7]。因为蛋白结构数据库 PDB 中 HA 蛋白结构被分为两段独立结构。因此, 根据所选模板分别构建了 2 个独立片段, 并最后生成完整 HA 蛋白结构。同源建模模板包括 Swine H9

流感病毒 HA 蛋白(PDB ID: 1JSD), Avian H5 流感病毒 HA 蛋白(PDB ID: 1JSM), 1930 Swine H1 流感病毒 HA 蛋白(PDB ID: 1RV0), 1934 H1 流感病毒 HA 蛋白(PDB ID: 1RVX), H3N2 流感病毒 HA 蛋白(PDB ID: 2VIU), 以及 H5 流感病毒 HA 蛋白(PDB ID: 2IBX)。选择为模板的 6 个 HA 蛋白, 是 PDB 数据库中蛋白质序列经过 BLAST 程序比对, 按照序列相似性在 95%以上聚类之后, 由不同组中选出的。甲型 H1N1 流感病毒 HA 蛋白 HA1 片段与模板蛋白的同源性为 40%~84%, 对应 E 值为: 0.0, HA2 片段与模板蛋白的同源性为

63%~92%, 对应 E 值为: 0.0。甲型 H1N1 流感病毒 HA 蛋白与模板蛋白的序列比对结果如图 1(a)所示。为避免使用单一的模建结构对于后续分析带来偏差, 我们选取了 10 个可能的同源模建结构(网络版附录 1)。模建生成结构采用 CHARMM, InsightII 软件进行结构优化。优化之后的结构通过 PROCHECK 进行检测^[8,9], 详细的评价结果包含在补充材料中(网络版附录 2)。

模建的 HA 蛋白通过 SEPPA 程序对构象表位进行预测, 采用默认参数 1.8。对于 HA 蛋白潜在空间表位预测, 要求在 10 个模建结构中都能被

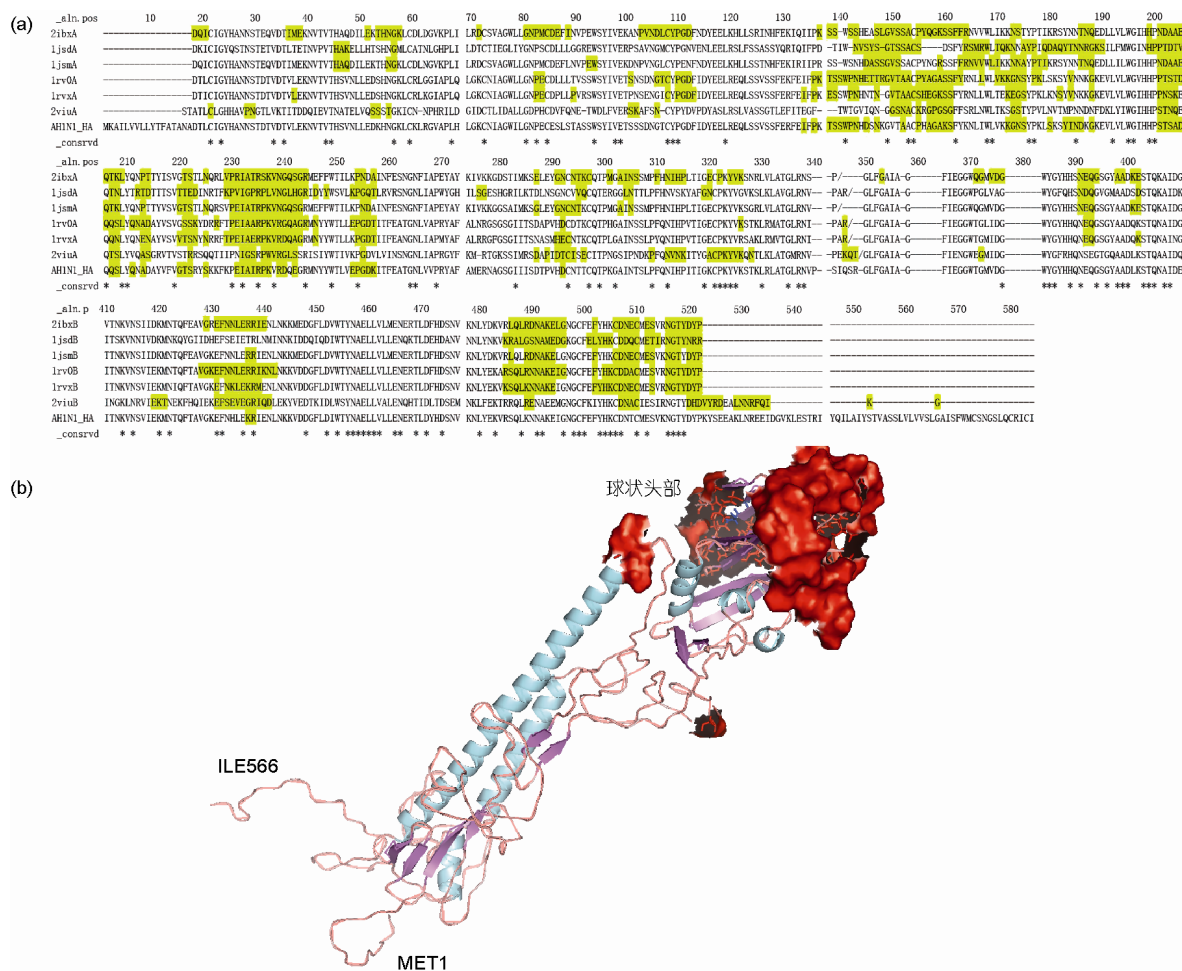


图 1 甲型 H1N1 流感病毒 HA 蛋白模建结构

HA 蛋白模建结构头部(globular head)为暴露在外部分。(a) 甲型 H1N1 流感病毒 HA 蛋白与模板蛋白序列比对结果。HA 蛋白及模板蛋白潜在构象表位用黄色突出显示, 模板蛋白保守位点氨基酸残基以“*”标出。(b) HA 蛋白模建结构, 蛋白质结构使用丝带(solib ribbon)形式显示, 构象表位使用表面(surface)形式显示

预测到的才被认定为潜在的空间表位氨基酸。抗原表位预测结果初步显示,空间表位氨基酸主要分布在 HA 蛋白球状头部和细长尾部顶端。因为 HA 蛋白球状头部通常被认为凸出在病毒表面而细长尾部一般位于病毒内部^[5],所以我们只选取了处在头部的潜在表位氨基酸。这样,抗原表位氨基酸有 58 个,在图 1(a)中使用黄色突出显示。构象表位氨基酸主要位于 HA 蛋白的球状头部表面,聚集成簇,构成空间表位,如图 1(b)所示。

为了比较甲型 H1N1 流感病毒与以往流感病毒的异同点,HA 蛋白

潜在构象表位与以往流感病毒潜在构象表位进行了比较。由于同源建模时依据数据库中所有 HA 蛋白结构的序列聚类选取模板,所选模板蛋白具有一定的代表性,因此将模板蛋白的可能空间表位同样用 SEPPA 在相同参数下进行了预测(构象表位见网络版附录 3),并与建模 HA 蛋白潜在构象表位进行了对比。如图 2 所示,所有 HA 蛋白的空间表位氨基酸在 HA 蛋白结构上所坐落的位置相似,但是化学特征已经有了明显变化。通过计算构象表位残基的 Gasteiger charge 参数,并根据

静电势不同,我们绘制了构象表位 surface 形状。已往流感病毒 HA 蛋白表位中往往带有正电势(蓝色),或者是正电势(蓝色)与负电势(红色)氨基酸残基共存于表位表面,而在甲型 H1N1 流感病毒 HA 蛋白潜在表位整体呈现弱正电势,未见有强负电势氨基酸残基(如 Glu 和 Asp)处在表面。由于抗原表位氨基酸的电性差异,可能直接影响到抗体分子的识别和结合。这有可能就是目前甲型 H1N1 流感病毒人体缺乏免疫力的原因。对于该预测的进一步验证则需要后续分析及实验结果的支持。

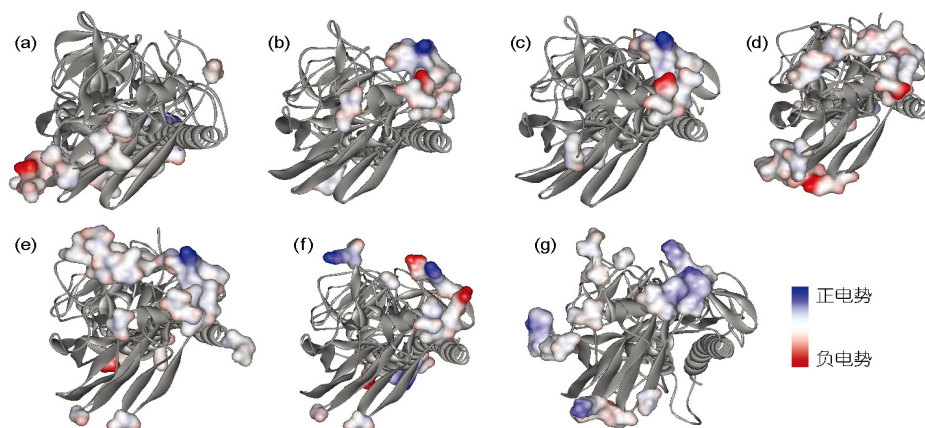


图2 甲型 H1N1 流感病毒 HA 蛋白及已往流感病毒 HA 蛋白潜在构象表位

(a) H3N2 流感病毒 HA 蛋白(PDB ID: 2VIU); (b) H5 流感病毒 HA 蛋白(PDB ID: 2IBX); (c) Avian H5 流感病毒 HA 蛋白(PDB ID: 1JSM); (d) Swine H9 流感病毒 HA 蛋白(PDB ID: 1JSD); (e) 1930 Swine H1 流感病毒 HA 蛋白(PDB ID: 1RV0); (f) 1934 H1 流感病毒 HA 蛋白(PDB ID: 1RVX); (g) 甲型 H1N1 流感病毒 HA 蛋白。以上均为头部(globular head)俯视图,整个蛋白质结构使用丝带(solid ribbon)形式表示,潜在构象表位使用表面(surface)形式表示。潜在表位表面(surface)颜色根据静电势着色,红色为负电势,蓝色为正电势,白色为中性

参考文献

- 1 WHO Influenza A (H1N1)-update 29. http://www.who.int/csr/don/2009_05_15/en/index.html, 15 May 2009
- 2 Jonathan R K. Swine influenza. *J Clin Pathol*, 2009, doi:10.1136/jcp.2009.067710
- 3 Fraser C, Donnelly C A, Cauchemez S, et al. Pandemic potential of a strain of influenza A (H1N1): Early findings. *Science*, 2009, doi: 10.1126/science.1176062
- 4 Irvine R M, Brown I H. Novel H1N1 influenza in people: Global spread from an animal source. *Vet Rec*, 2009, 164: 577—578
- 5 Ekiert D C, Bhabha G, Elsliger M A, et al. Antibody recognition of a highly conserved influenza virus epitope. *Science*, 2009, 324: 246—251, doi: 10.1126/science.1171491
- 6 Sun J, Wu D, Xu T L, et al. SEPPA: a computational server for spatial epitope prediction of protein antigens. *Nucleic Acids Res*, 2009, doi: 10.1093/nar.gkp417
- 7 Sali A, Potterton L, Yuan F, et al. Evaluation of comparative protein modeling by MODELLER. *Proteins: Struct Funct Bioinformatics*, 1995, 23: 318—326, doi: 10.1002/prot.340230306
- 8 Laskowski R A, MacArthur M W, Moss D S, et al. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst*, 1993, 26: 283—291, doi: 10.1107/S0021889892009944
- 9 Morris A L, MacArthur M W, Hutchinson E G, et al. Stereochemical quality of protein structure coordinates. *Proteins*, 1992, 12: 345—364, doi: 10.1002/prot.340120407