

蛋白激酶的五配位磷化学结构分析*

李 武 麻 远 赵玉芬**

(生命有机磷化学及化学生物学教育部重点实验室 清华大学化学系, 北京 100084)

摘要 基于实验合成的五配位磷氨基酸晶体的平方锥和三角锥的结构, 以及氨基酸侧链参与形成五配位磷化学结构实验的基础上, 设计了软件来搜索国际蛋白数据库中可能形成五配位磷结构的含磷蛋白. 利用自主编写的软件研究发现, 国际蛋白质数据库(PDB 库)收录的 398 个含磷的蛋白激酶中, 有 382 个蛋白激酶(96%)在活化过程中可能经历了五配位磷化学过渡态或中间体. 例如 1H8E 蛋白的 Lys16 的 N_{α} 原子与 GNP2 的 P 原子形成了可能的五配位化学结构, 将该结构与实验得到的氨基酸小分子晶体的实际五配位磷结构进行叠加, 结构符合很好, 相关结构的偏差 RMSD 值为 0.71 Å 文中的软件下载地址为: <ftp://166.111.28.228>, 软件版权登记号 2002SR2943.

关键词 五配位磷化学结构 蛋白激酶

早在 17 世纪人们就在有机体中发现了磷元素, 最近半个世纪的研究发现磷元素在生命系统的诸多方面有至关重要的作用, 如蛋白质的可逆磷酸化问题现在仍是生命化学的重要研究课题^[1], 然而蛋白质磷酸化过程的确切作用机理尚未研究清楚^[2]. 目前生物体系中发生的对磷上亲核取代反应的途径仍有争议, 一种认为形成了五配位磷的中间体, 另一种则认为经过了三配位的偏磷酸负离子中间体, 酶与底物作用的五配位磷和偏磷酸晶体结构在最近均有报道^[3,4]. 我们在大量实验基础上倾向于认为五配位磷化学结构是分析这些过程的关键, 本文报道了用一种自主编写的软件对国际蛋白质数据库中收录的含磷激酶进行分析, 寻找其可能形成五配位磷化学结

构的结果.

1 蛋白质五配位磷化学结构的证据

在以前的研究工作中, 我们发现 N-磷酸化- α -氨基酸可以通过形成五配位磷化学结构中间体被活化, 发生自组装成寡肽、磷上酯交换、磷酸基的 N-O 迁移等反应^[5,6].

为了捕获五配位磷过渡态, 可利用硅醚衍生化的氨基酸对其进行细致的分析. 五配位磷酰氨基酸的生成可通过 ^{31}P NMR 谱跟踪堆积图进行观察^[7].

Zhao 等研究并总结了氨基酸侧链参与形成五配位磷化合物的动力学数据^[8]. 发现侧链含有亲核性官能团的氨基酸, 如组氨酸(侧链为咪唑基), 天冬氨酸

2003-12-25 收稿, 2004-04-20 收修改稿

* 国家自然科学基金(批准号: 20132020)、科技部、教育部和清华大学资助项目

** E-mail: yfzhao@mail.tsinghua.edu.cn

(侧链为 $-\text{CO}_2\text{H}$), 苏氨酸(侧链为 $-\text{OH}$), 丝氨酸(侧链为 $-\text{OH}$)相对较快。

进一步详细研究表明, 这是因为含有亲核性侧链的氨基酸五配位磷化合物可以形成第二种五配位磷化合物, 如图 1 中硅醚化的丝氨酸五配位磷化合物可以分别形成磷酸羧酸混合酸酐形式的五配位磷化合物 4 的结构(羧基参与形成五配位磷)和螺磷杂噁唑

烷形式的五配位磷化学结构 5(羟基参与形成五配位磷)。实验发现, 温度对五配位磷化合物形成的影响非常显著, 在 -18°C 以下, 只形成 4, 高于该温度, 侧链羟基与羧基竞争形成两种五配位化合物 4 和 5, 40 时, 5 的含量占五配位磷化合物总量的 20%^[9]。5 的形成有两种可能途径, 一种途径由 4 经过六配位磷过渡态, 另一途径是由 3 直接形成 5。

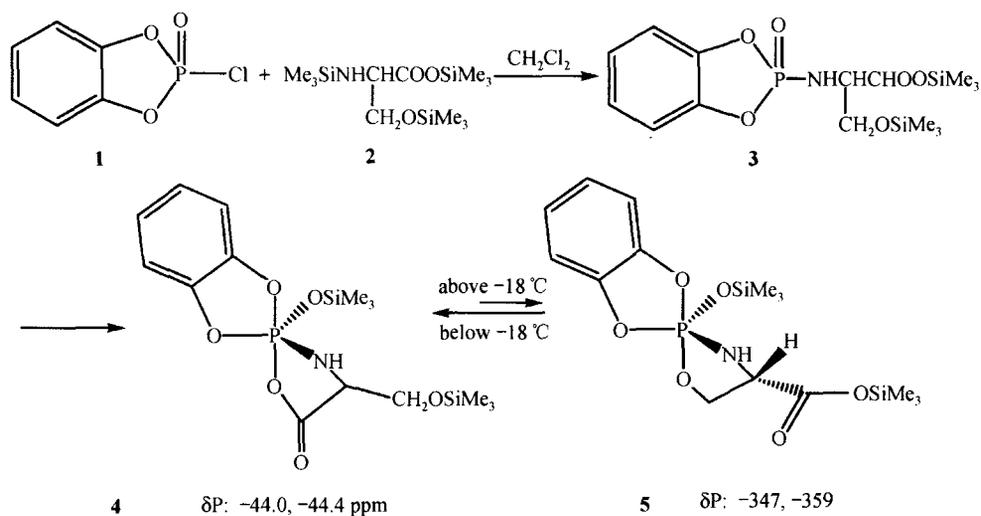


图 1 O,O'-亚苯基磷酰氯与 N,O,O'-三甲硅基丝氨酸的反应方程式

Li 等在实验中还观察到了六配位磷化合物的证据^[10]。在对硝基-O,O'-亚苯基磷酰氯(NO-PPC)与三甲硅基丝氨酸(N,O,O'-Tis-Ser)、苏氨酸和半胱氨酸反应过程中, 在磷谱中可以观察到五配位及六配位磷化合物的信号, 氨基酸侧链的羟基, 羧基, 氨基, 巯基等均可以参与形成五配位及六配位磷化合物结构, 这也是本文软件设计出发点之一。

在蛋白质化学中, 蛋白质激酶本身被 ATP 磷酸化后, 能够使其他底物蛋白磷酸化, 它在细胞的信号传导, 基因调控等生物过程中扮演着重要角色^[11-19]。在上述实验基础上, 我们推测五配位磷化学结构可能是蛋白激酶被磷酸化时的过渡态或中间体, 基本模型如图 2 所示。蛋白激酶中活性位点氨基酸残基的侧链亲核进攻 ATP 中的 *g*-磷原子, 形成活化的含有五配位磷化学结构的蛋白激酶, 它可能是一种亚稳态的中间体, 或者是一种势能较高的过渡态, ATP 中

的 *g*-磷酸基被转移到蛋白激酶上, 激酶被磷酸化, 同时失去一分子 ADP。同样磷酸化的激酶在使底物蛋白质磷酸化时, 也是由底物蛋白质中活性位点氨基酸残基的侧链亲核进攻蛋白质激酶中的磷酸基, 经五配位磷过渡态或中间体完成磷酸基的转移(图 3)。上述两步中(图 2 和 3)均可能经历了五配位磷过渡态。

2 蛋白质激酶的五配位磷化学结构搜索软件的设计

目前公认的五配位磷化学结构, 有三角双锥(TBP)和四方锥(SP)两种构型。三角双锥有三个 e 键和两个在主轴的 a 键, a 键与 e 键二面角成 90° , 三个 e 键互成 120° , a 键比 e 键长。四方锥中, 有一个 a 键, 四个 b 键。a 键与 b 键二面角成 105° , 两个 b 键成 88° 夹角。由于键角较小, SP 中各价键电子对之间相互排斥较大, 使得 SP 构型较 TBP 构型稳定性稍差^[20]。

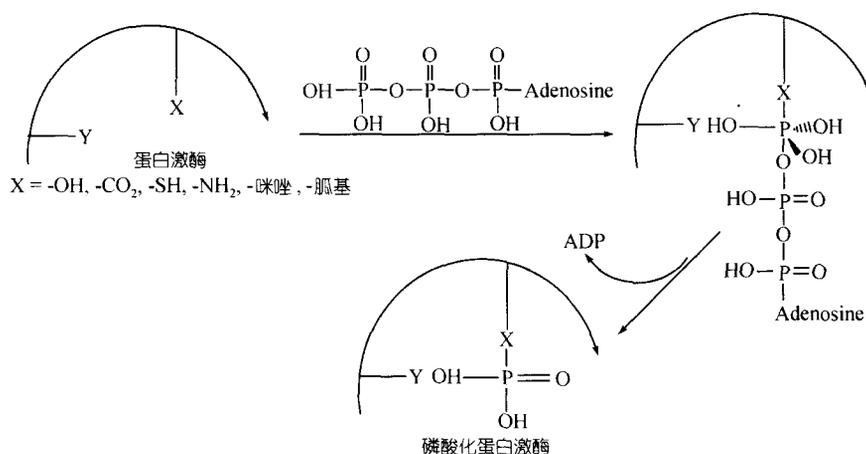


图 2 蛋白激酶被磷酸化的示意图

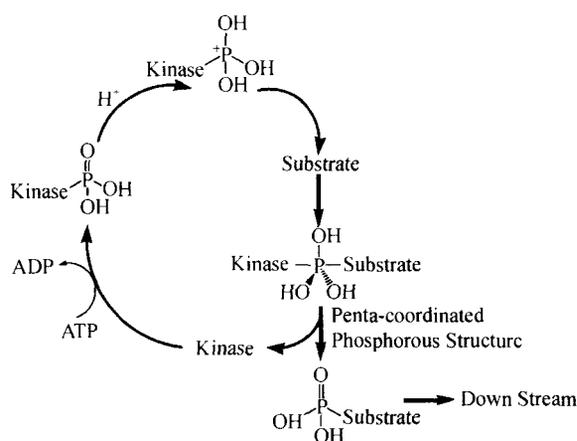


图 3 磷酸化的蛋白激酶使底物蛋白质磷酸化的机理示意图

我们实验室分别培养得到了氨基酸五配位磷化学结构的平方锥和三角双锥晶体结构^[21,22]。图 4 中, 双氨基酸五配位平方锥晶体结构(图 4(a))接近理想五配位磷化合物的平方锥结构(图 4(b)), 氨基酸五配位磷化合物晶体的三角双锥结构(图 4(c))接近理想五配位磷化合物的三角双锥结构(图 4(d))。这两个五配位磷化合物晶体是说明氨基酸侧链参与形成五配位磷化学结构以及蛋白活化过程模型的有效证据。

目前有文献报道的五配位磷晶体结构中, P-O 键长一般在 1.57~2.82 Å 之间, 多数在 1.57~1.76 Å 之间^[23]。

虽然 TBP 与 SP 晶体结构的键长在 1.57~2.82 Å

之间, 键角在 88°~120° 之间^[23], 但考虑到晶体数据是静态相对稳定结构, 实际分子在运动中键长与键角与晶体中会有所变化, 有一定范围的偏差, 应灵活考虑。因此基于我们五配位模型化合物与晶体, 我们设计了软件(含磷蛋白及核酸数据库知识发掘系统)搜索可能形成的五配位磷化学结构, 软件搜索设计的方法可以细分如下:

方法 1(Method 1): 磷原子与第 5 个可能成键的原子(氨基酸残基侧链中的 N, S, O 原子, 例如赖氨酸的氨基 N 原子, 组氨酸咪唑的 N 原子, 丝氨酸、苏氨酸的羟基 O 原子, 半胱氨酸的巯基 S 原子)的坐标距离在 1.8~3.2 Å 之间, 键角 XPY(X 代表 N, S, O 原子, P 代表磷原子, Y 代表那些和磷原子已经成键的原子)范围在 85°~125° 之间;

方法 2(Method 2): 磷原子与第 5 个可能成键的原子(氨基酸残基侧链中的 N, S, O 原子)的坐标距离在 1.8~4.2 Å 之间, 键角 XPY 范围在 75°~125° 之间。

目前软件由 2 个主要程序构成, 可以运行在 Microsoft windows® 98, 2000, NT, XP 操作系统及普通微机下 (CPU 2GHz 以上, 内存 512M 以上)。

3 对蛋白质晶体库搜索的结果与讨论

利用上述软件, 我们分析了国际蛋白质数据库^[24,25](PDB)中现有的能够形成五配位磷化学结构的 382 个含磷蛋白激酶(包括磷酸化的蛋白激酶和蛋白激酶及其磷酸化抑制剂共结晶的数据)。以表 1 为例说明软件搜索的结果: 表 1 选取的 12 个蛋白代表了

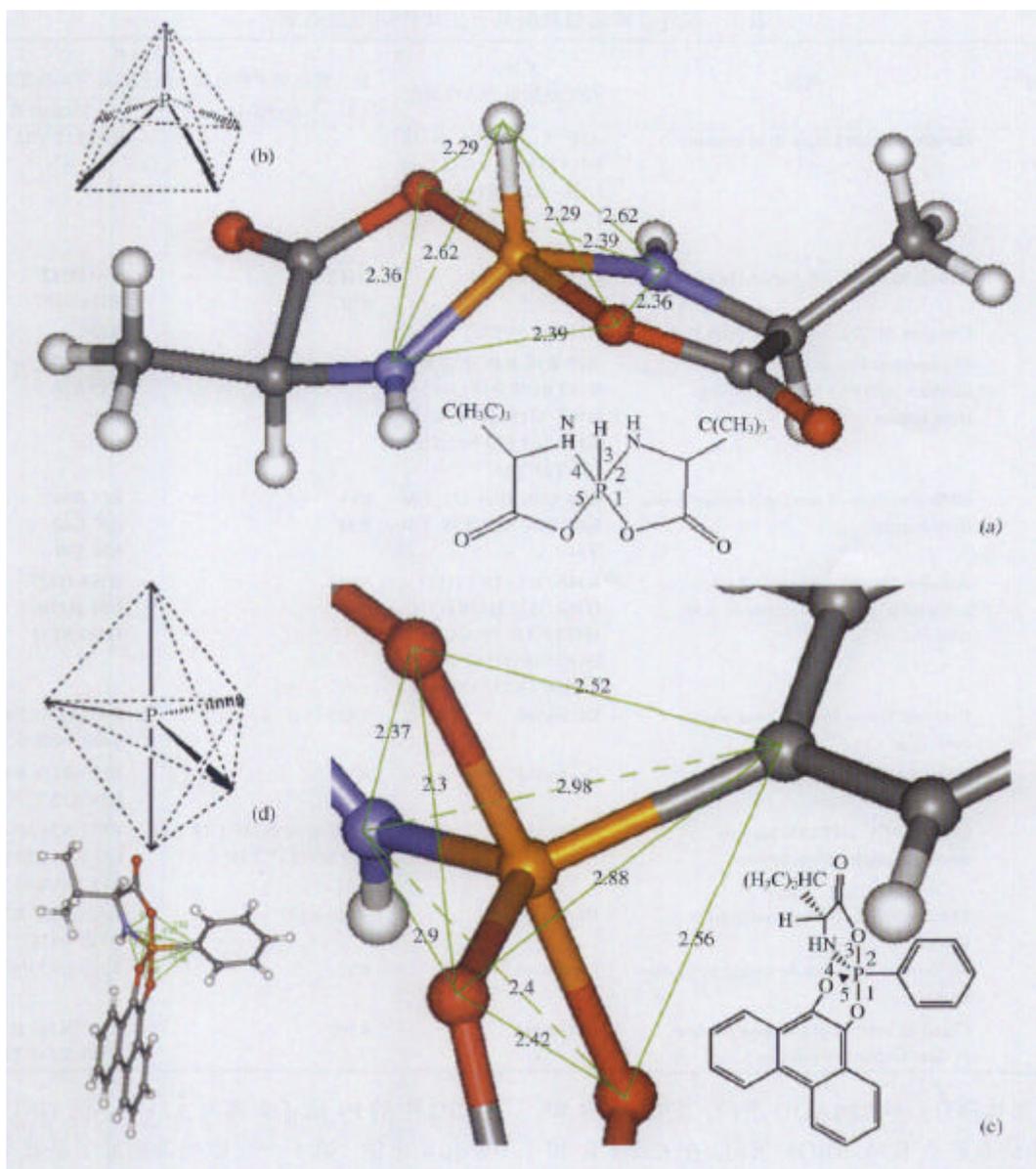


图 4 氨基酸五配位磷化合物晶体的平方锥(a)、理想平方锥(b)、氨基酸五配位磷化合物体的三角双锥(c)理想三角双锥(d)的示意图

(a) 中 P 原子的 5 个键长如下(Å): P-O(1)为 1.74; P-N(2)为 1.627; P-H(3)为 1.455; P-N(4)为 1.627; P-O(5)为 1.74. (c) 中 P 原子的 5 个键长如下(Å): P-O₂(1)为 1.735; P-C(2)为 1.794; P-O₃(3)为 1.749; P-N(4)为 1.627; P-O(5)为 1.641. 图 4.5 的绘图程序为 WebLab ViewerLite 4.0

不同来源(如细菌, 酵母, 老鼠, 小鸡, 人类等), 不同生物活性(如 HIV-1 抑制剂等)的蛋白体系, 由于全部 382 个含磷蛋白激酶的详细结果篇幅较多, 在此从略.

在表 1 中, Class A 列是当前 PDB 数据文件中研究者已报道的有重要生物功能的活性位点. Class A

列中的许多位点是由实验方法得到的. 而对于尚无报道的位点, 在本表 Class A 列中被标明 Uncharted. Class B 列是通过本软件搜索到的可能形成五配位磷化学结构的氨基酸残基.

表 1 中 Class A 列与 Class B 列中有部分共同的

表 1 部分含磷蛋白激酶具有生化作用的位点表

PDB 登记号 ²⁵	简称	Class A 有文献报道的活性部位	Class B 软件搜索到可能形成五配位磷化学结构的氨基酸残基	
			Method I	Method II
1HE8	Phosphoinositide kinase from human	A11 A18 A146 N116 D119 D30 G13 G15 G60 L120 K16 K117 K647 F28 P34 S145 S17 T35 Y32		K16 S17 Y32
1AV5	Protein Kinase C inhibitor from human	H51 H112 H114	H112 S107	N99 H112 H114 S107
1BO5	Complex of Glycerol kinase from E. Coli	R236 G234 T235		R236
1E9A	Thymidylate kinase complex with inhibitor AZTP5A in HIV prodrug from human	A17 R16 R45 R76 R97 R143 R192 D15 Q157 G102 G18 I184 K182 K19 F105 F72 F42 S183 S20 T21 Y151		R16 R97 K19 S20 T21
1AOU	NMR Structure of protein-tyrosine kinase from human	R34 G95 H60 I73 L96 K62 S44 S36 T38 T74 Y61	R14 R34	R14 R34 E37 K62 S36 T38
1EW8	Alkaline Phosphatase (E.C.3.1.3.1) complex with phosphonoacetic acid from E.Coli	R166 D153 D51 D327 D369 E322 H370 H331 H412 K328 MO3452 PAE556 PO4557 T155 ZN450 ZN451 ZO3453	S102	R166 D327 D51 H370 H412 S102
1A3W	Pyruvate kinase from saccharomyces cerevisiae	Uncharted	S402	R459 D266 S402 S404 T403 T407 W452
1AKE	Adenylate kinase complex with the inhibitor from E coli	Uncharted	R167	R123 R156 R167 R36 R88 K13 T15
1CF4	CDC42/ACK GTPASE-binding domain complex from human	Uncharted	R537 N26 N545 C18 E31 S22 S30 T17 T35 Y32	R537 N26 N545 C18 E31 K27 S22 S30 T17 T35 Y32
1CRK	Mitochondrial creatine kinase from chicken	Uncharted	R125 R127	R125 R127 R287 R315 R336
1F6T	Nucleoside diphosphate kinase from slime mold	Uncharted	R92	R109 R92 H59 T98
1JLU	Catalytic subunit of Camp-dependent protein kinase from mouse	Uncharted	R165	R165 N340 D166 H87 K168 K189 T195

位点(表中黑体部分)。例如 1AOU 蛋白, 在 Class B 和 Class A 中均有位点 R34。1BO5 蛋白, 在 Class B 和 Class A 列中均有位点 R236。另外表中的一些蛋白, 例如 1CF4、1F6T、1JLU, 在 PDB 库数据文件中没有有关的活性位点信息, 而通过该软件, 可以发现可能的活性位点, 这些信息无疑对今后的实验有重要指导意义。

图 5 所示为人源 Phosphoinositide kinase (PDB 登记号 1HE8) 的晶体与其抑制剂 GNP2 共结晶数据的分析示意图: GNP2 中磷原子附近氨基酸残基搜索的范围被限制如下: 原子坐标距离在 1.8 ~ 4.2 Å 之间, 键角范围在 75°~125° 之间。从软件搜索的结果可以看出, 1HE8 蛋白的氨基酸残基 Lys16 的 N_Z 原子与

GNP2 中的 P_a 原子距离为 3.58 Å, 而 GNP2 的 P_a 原子有 4 个键, 第 5 个可能成键的原子是氨基酸 Lys16 中的 N_Z 原子, 可以形成的构型与三角双锥非常接近。图 5 中的绿线表示三角双锥的边, 数值为边长, 单位为 Å。图 5 中的虚线框为缩略图。

图 5 中相关键长、键角与五配位磷化学结构中的理想三角双锥(TBP)的构型非常接近。将此结构与晶体五配位磷三角双锥结构(图 4(c))进行比较, 将 6 个可能(已经)成键原子(其中第 6 个原子为磷原子)进行结构相关部分叠加, 结果如图 6。图 6 中细线表示的结构是图 5 的结构, 粗线和球表示的结构是图 4(c)的结构。均方根偏差(RMSD, Root Mean Square Deviation)

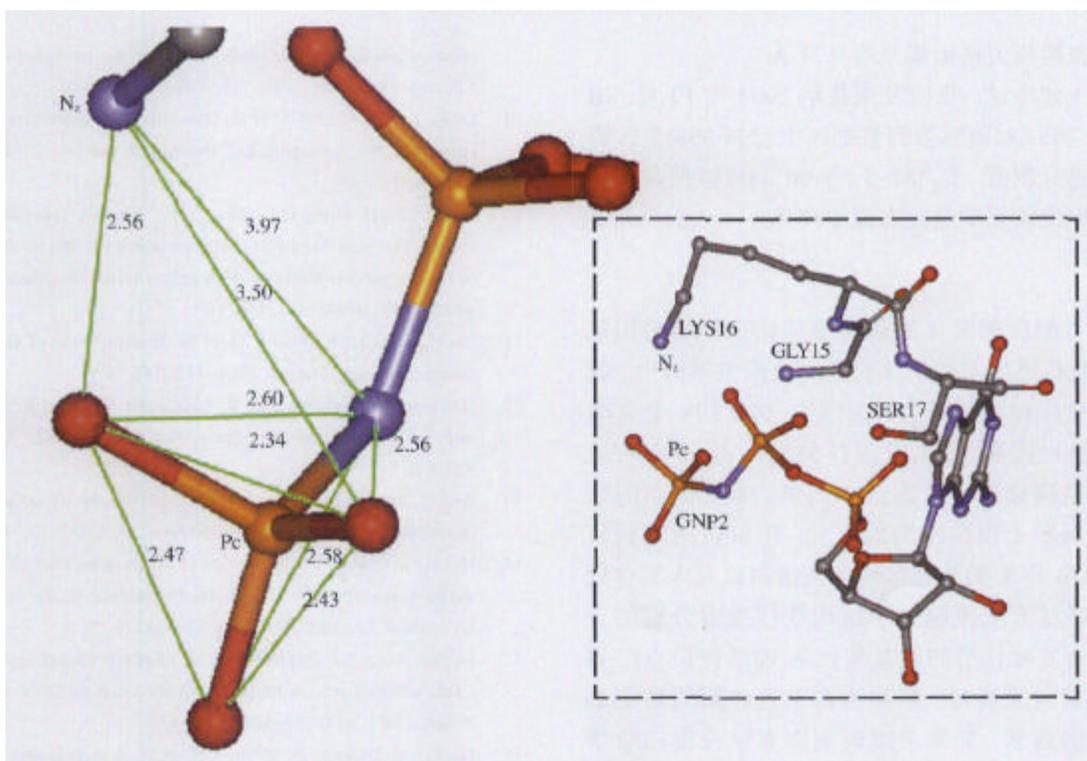


图 5 用软件搜索到的 1HE8 蛋白体系中 Lys16 可能参与形成的五配位磷化学结构(图中距离数值单位为 \AA)

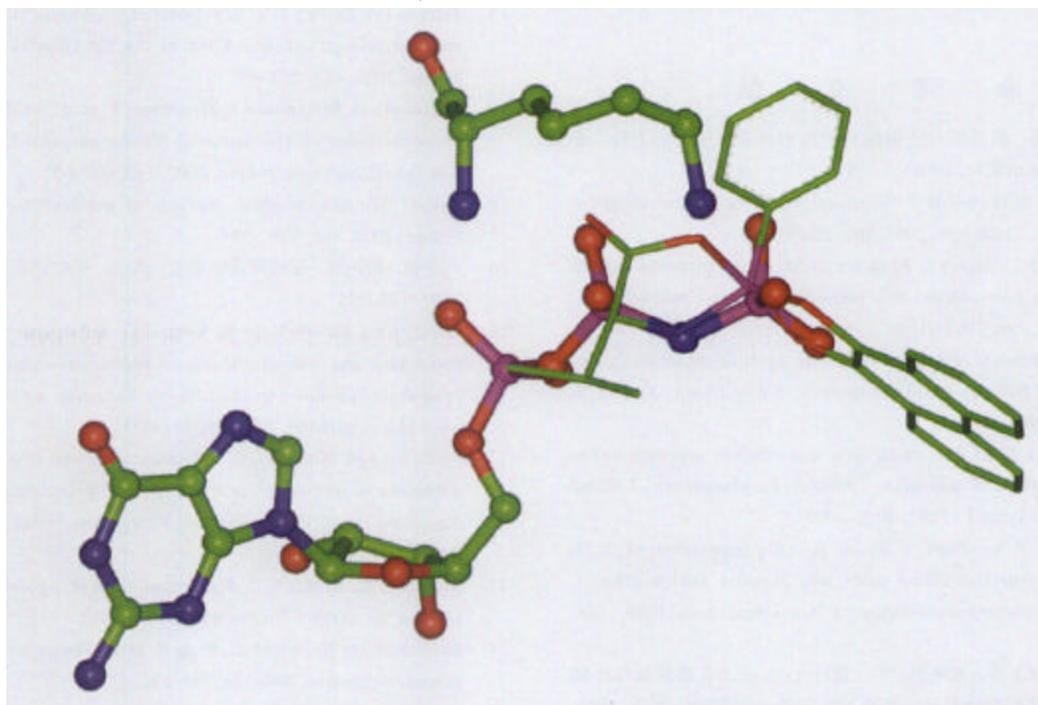


图 6 五配位磷晶体的三角双锥结构(图 4(c))与潜在的激酶五配位磷结构(图 5)的叠加示意图
两者相关结构的 RMSD(Root Mean Square Deviation)值为 0.71 \AA 图 6 中的绘图程序为 Insight II 2000.01®

计算结果表明相关结构偏差为 0.71 Å

通过上述方法, 我们发现截至 2003 年 12 月, 用 Method II 方法对国际蛋白数据库中已知 398 个含磷蛋白激酶进行搜索, 有 382 个(占 96%)可以搜索到可能形成五配位磷化学过渡态或中间体.

4 结论

基于五配位磷酰化氨基酸晶体化合物的结构以及五、六配位磷氨基酸衍生物的实验研究基础上, 我们设计了软件对国际蛋白数据库中已知 398 个含磷蛋白激酶进行搜索, 结果发现有 382 个(占 96%)可能形成五配位磷化学过渡态或中间体. 本文用到的软件不仅仅局限于蛋白激酶的研究, 还可以应用到研究现有 PDB 库中的其他含磷天然蛋白以及人工合成的蛋白. 通过五配位磷化学结构软件搜索分析的方法来发现有重要作用的氨基酸残基(即活性位点), 省时省力而且效果良好, 能够为将来设计酶的抑制剂提供良好的线索. 如果考虑时间因素以及蛋白组学中对一组相关含磷蛋白质同时进行比较研究的需要, 五配位磷化学结构软件搜索分析的方法具有强大的优势.

参 考 文 献

- 敖世洲, 等. 蛋白质可逆磷酸化对细胞活动的调节. 上海: 上海科学技术出版社, 1994
- Johnson L N, Lewis R J. Structural basis for control by phosphorylation. *Chem Rev*, 2001, 101: 2209~2242
- Sushmita D L, Zhang G F, Debra D M, et al. The pentavalent phosphorus intermediate of a phosphoryl transfer reaction. *Science*, 2003, 299: 2067~2071
- Choe J Y, Iancu C V, Fromm H J, et al. Metaphosphate in the active site of fructose-1,6-bisphosphatase. *J Biol Chem*, 2003, 278: 16015~16020
- Fu H, Tu G Z, Li Z L, et al. New and efficient approach to the synthesis of penta-coordinate spirobicyclic phosphorus. *J Chem Soc, Perkin Trans*, 1997, 2021~2022
- Fu H, Li Z L, Zhao Y F, et al. Oligomerization of N,O-bis(trimethylsilyl)- α -amino acids into peptides mediated by *o*-phenylene phosphorochloridate. *J Am Chem Soc*, 1999, 121: 291~295
- 王倩, 赵玉芬, 安逢龙, 等. 通过 O,O-亚苯基磷酰氯与 α -氨基酸三甲衍生物的反应观察五配位磷. *中国科学, B 辑*, 1995, 25: 58~63
- Zhao Y F, Yan Q J, Wang Q. Phosphoryl transfer reaction regulated by amino acid side chains: A model for phosphoproteins. *Int J Peptide Protein Res*, 1996, 47: 276~281
- Lu K, Tu G Z, Guo X F, et al. Structure and isomerization of O,O-phenylene Penta-coordinated Phosphoryl Serine. *J Mol Struct*, 2002, 610: 65~72
- Li Z L, Fu H, Gong H G, Zhao Y F. ^{31}P NMR Spectral Evidence for the Hexacoordinated phosphorus intermediates in the reaction of oxyphosphorochloridate with amino acids. *Phosphorous Sulfur and Silicon*, 1999, 147: 215~216
- Gu H, Saito K, Klamann L D, et al. Essential role of Gab2 in the allergic response. *Nature*, 2001, 412: 186~190
- Himanen J, Rajashankar K R, Lackmann M, et al. Crystal structure of an Eph receptor-eprin complex. *Nature*, 2001, 414: 933~937
- Asai T, Tena G, Plotnikova J, et al. MAP kinase signalling cascade in arabidopsis innate immunity. *Nature*, 2002, 415: 977~983
- Haj H G, Verweir P J, Squire A, et al. Imaging sites of receptor dephosphorylation by PTP1B on the surface of the endoplasmic reticulum. *Science*, 2002, 295: 1708~1711
- Nash P, Tang X J, Stephen O, et al. Multisite phosphorylation of A CDK inhibitor sets A threshold for the onset of DNA replication. *Nature*, 2001, 414: 514~521
- Gachet Y, Tournier S, Millar J B, et al. A map kinase-dependent actin checkpoint ensures proper spindle orientation in fission yeast. *Nature*, 2001, 412: 352~355
- Tercero J A, Diffley F X. Regulation of DNA replication fork progression through damaged DNA by the Mec1/Rad53 checkpoint. *Nature*, 2001, 412: 553~557
- Masumoto H, Muramatsu S, Kamimura Y, et al. S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. *Nature*, 2002, 415: 651~655
- Zaslloff M. Antimicrobial peptides of multicellular organisms. *Nature*, 2002, 415: 390~394
- 刘纶祖, 刘钊杰. 有机磷化学导论. 武汉: 华中师范大学出版社, 1991. 182~215
- Liu Z, Yu L, Chen Y, et al. Interesting differences between the protonated and sodium adducts of pentacoordinated bisaminoacylspiroporphoranes in electrospray ionization mass spectrometry. *J Mass Spectrom*, 2002, 38: 231~233
- Fu H, Xu J H, Wang R J, et al. Synthesis, crystal structure and diastereomeric transfer of pentacoordinated phosphoranes containing valine or iso-leucine residue. *Phosphorous Sulfur and Silicon*, 2003, 178: 1963~1971
- Swamy K C, Kumar N S. Pentacoordinated phosphorus in action. *Current Sci*, 2003, 85: 1256~1258
- Berman H M, Westbrook J, Feng Z, et al. The protein data bank. *Nucleic Acids Res*, 2000, 28: 235~242
- Authors of Scientific Community Worldwide, as specified in PDB files' author sections, which available on [HTTP://WWW.RCSB.ORG/PDB](http://www.rcsb.org/pdb)