

doi:10.3969/j.issn.1672-5565.2013.02.10

构建和评估猪氨基酰化酶 I 的三维结构

刘志刚, 邓 贝, 杨 波, 胡 征

(湖北工业大学生物工程学院, 湖北 武汉 430068)

摘要:运用同源模建的方法构建了 pACY1 三维结构模型,并在能量最小化后对模型进行分子动力学模拟和结构合理性评估。同源模建生成了 50 个原始模型,经过 PROCHECK 评测后,筛选出模型 A、B 进行能量最小化,并得到模型 A1、B1。分子动力学模拟结果表明模型 B1 二聚体结构较稳定。PROCHECK、ProSa 以及 WHATIF 检测结果验证了模型 B1 属于合理性结构。得到的猪氨基酰化酶 I (pACY1) 的三维结构,为研究其结构与功能关系打下基础。

关键词: 氨基酰化酶 I; 同源模建; 能量最小化; 分子动力学模拟

中图分类号: Q518.2 **文献标识码:** A **文章编号:** 1672-5565(2013)-02-130-06

Construction and evaluation of a three-dimensional structure of porcine aminoacylase I

LIU Zhi-gang, DENG Bei, YANG Bo, HU Zheng

(College of Biotechnology, Hubei University of Technology, Wuhan 430068, China)

Abstract: A three-dimensional structure of porcine Aminoacylase I (pACY1) was constructed by homology modeling. The models after energy minimization were under the procedures of molecular dynamics (MD) simulation and rational evaluation. 50 models were built by homology modelling. Among them the best two models (model A, B) were chosen by PROCHECK and were further refined by energy minimization. The refined models were signed as model A1 and model B1 from model A and model B respectively. Model B1 has the better stability than model A1 during MD simulation. The evaluations illustrated model B1 was a reasonable structure by PROCHECK, ProSa and WHATIF. We get a reasonable 3D structure of pACY1, which could be further applied to study the relation between structure and function of pACY1.

Key words: Aminoacylase I; Homology Modeling; Energy Minimization; Molecular Dynamics Simulation

氨基酰化酶 I (Aminoacylase I, EC 3.5.1.14; ACY1) 特异性催化水解 N-酰基-L-氨基酸,生成 L-氨基酸。其生理功能被认为是挽救代谢后氨基酸以及抑制癌细胞生长的作用^[1-4]。已发现一些遗传性疾病是由人 ACY1 基因突变导致酶活性降低或丧失引起的^[5-7]。ACY1 属于肽酶 M20 家族,是一个含 Zn²⁺ 的同源二聚体蛋白。猪氨基酰化酶 I (Porcine aminoacylase I, pACY1) 被用于氨基酸的手性拆分,它的单个亚基由 406 个氨基酸和单个 Zn²⁺ 组成。Zn²⁺ 是 pACY1 水解活性所必须的,很多研究者认为底物直接结合于 Zn²⁺ 并进行水解作用。但

是,一些研究者认为锌离子仅发挥维持结构稳定性的作用^[8-9]。这就需要 pACY1 的结构信息才能明确锌离子的功能及 ACY1 的催化机制。

1997 年,Zhao 等报道得到了猪氨基酰化酶 I 的晶体^[10],但并没有解析的结构。2003 年,Lindner 报道了人氨基酰化酶 I (hACY1) T347G 突变体不完整的晶体结构。二聚体结构域在形成晶体之前缺失了,而仅得到锌结合结构域的结构^[11]。此结构中 Zn²⁺ 结合一个甘氨酸分子。然而定点突变的结果显示,二聚体结构域上的一些氨基酸残基对酶的活性及底物结合非常重要^[11]。另外,通过比较猪和人氨

收稿日期:2012-08-22;修回日期:2012-09-07.

基金项目:湖北省教育厅基金项目(Q20081411)和湖北工业大学高层次人才基金项目(20062016)的资助。

作者简介:刘志刚, E-mail: Liuzg77@yahoo.com.cn.

基酰化酶 I 的序列差异进行的定点突变研究表明底物不结合于 Zn^{2+} 上^[12]。

到目前为止,由于缺乏完整的结构信息,人们对 ACY1 的底物结合部位及催化机理还不了解。为了解决这个问题,我们对 pACY1 的结构进行了模拟和优化。通过合理性评测,筛选得到了合理的 pACY1 二聚体蛋白三维结构。这为研究 pACY1 的结构与功能关系打下了基础。

1 方法

1.1 序列与结构比对

通过 BLASTp 选择了与 pACY1 同家族且同源性最高的两个已知结构的蛋白为模板,它们是 hACY1 的 T347G 突变体和 *Neisseria meningitidis* mc58 succinyldiaminopimelate desuccinylase, 其结构在 PDB 中的编号分别为 1Q7L 和 1VGY。运用 CLUSTALW 对 pACY1 的氨基酸序列与这两个蛋白的氨基酸序列进行比对^[13]。同时,利用 PHYRE 对 pACY1 的二级结构进行预测并与模板进行结构比对^[14]。两种比对结果通过最小间隔和插入进行人为调节。

1.2 同源建模

根据同源比对的结果,利用 MODELLER9v7 (<http://www.salilab.org/modeller/>) 程序,构建 pACY1 分子的三维空间结构模型。对二级结构不匹配和无模板区段采用相关二级结构限制,对 Zn^{2+} 进行距离限制,对单体进行对称限制。将得到的 50 个原始模型与二级结构预测以及多重序列比对的结果进行对照,选取合理的模型通过 PROCHECK3.5 分析,得到最合理的起始模型^[15]。

1.3 结构优化

将起始模型加全氢和水分子,然后用 AMBER7 软件中的 SANDER 模块,进行能量最小化的优化计算。采用 PARM99 力场和锌原子四面体配位力场,介电常数为 1.0, 12Å 的 cutoff,先运行 1 000 步的最陡梯度下乘法的模拟,紧接着是 2 500 步的共厄梯度下乘法的模拟^[16-17]。在模拟的过程中采用 40 kJ/mol 每 Å² 的力场对所有的原子进行约束。将优化后的模型进行合理性分析。

1.4 分子动力学模拟

用 AMBER7 软件中的 GROMACS 模块,进行分子动力学模拟(molecular dynamics simulation),检测优化后模型的稳定性。在常温常压的条件下,运行非约束性分子动力学模拟。采用步时 2.0 fs,模拟

温度 300 K,时长 200 ps,8.0 Å 的 non-bonded cutoff 的参数设置。通过分析单体或二聚体模型分子动力学模拟前后重原子构象的 RMSD (root-mean-square deviation) 来评价模型的稳定性。

1.5 合理性评估

上面优化后的模型首先进行 PROCHECK 分析,选择蛋白质主链和侧链构象都在合理区域的模型。另外,运用 ProSa2003 对模型进行位能分析,选择残基能量都分布于负值的模型^[18]。最后,利用 WHA-TIF6.0 分析模型的残基包装特性^[19]。

2 结果与讨论

2.1 序列及结构比对

运用 NCBI 的 BLASTp 工具分析 pACY1 全氨基酸序列的同源性,分析的结果表明,pACY1 和肽酶 M20 家族中的其他水解酶具有很高的同源性。挑选已知结构且同源性较高的蛋白作为待选模板。用 PHYRE 程序预测 pACY1 全序列 406 个氨基酸序列的二级结构,预测的结果表明 pACY1 由 10 个主要的 α -螺旋和 15 个 β -折叠组成。这和肽酶 M20 家族中大多数酶的主体结构相符合。

将待选模板的二级结构与 PHYRE 预测 pACY1 的二级结构进行对比,选择了 hACY1 的 T347G 突变体锌结合结构域(PDB ID: 1Q7L)和 *Neisseria meningitidis* mc58 succinyldiaminopimelate desuccinylase (nmDAPE) (PDB ID: 1VGY) 作为模板。hACY1 的 T347G 突变体的二聚体结构域在形成晶体之前降解了,所解析的结构为锌结合结构域。在 Swiss-PdbViewer 中,1Q7L 能很好地重叠在 nmDAPE 催化结构域上,除了与二聚体结构域连接的小部分氨基酸残基外,其他二级结构元件基本重叠在一起。计算 C α 原子的 RMSD 值,可得知在 263 个进行重叠的氨基酸残基中有 201 个残基的 C α 原子 RMSD 值为 1.43 Å。

利用 ClustalW 程序对模板序列与 pACY1 的氨基酸序列进行比对。做序列比对时同时需要参照 PHYRE 二级结构重叠的结果,尽量不让间隔出现在螺旋和折叠中,并最大限度减少间隔的数量。经过手动调节后的最后序列比对结果如图 1 所示。在 1Q7L 中的与二聚体结构域连接的部分氨基酸残基,在二聚体结构域降解后发生了构象的变化,因此这部分残基不能作为模板,并在比对中删除并以斜体字表示。

```

pACY1 ASKGRGEHPSVTLFRQYLRI RTVQFPDYGAAVAFLERARQLGLGQKVEV VPGHVVT 60
1Q7L ---EEHPSVTLFRQYLRI RTVQFPDYGAAVAFFETARQLGLGQKVEV APGYVVT
1VGY ---TETQSELELAKELI SRPSVTP---DDRDCQKLM AERLHKIGFAAEBMHFGN---T
      * * : * : : * * * . : * : : * : : : : : *
pACY1 VLTWPGTNPTLSSILLN SHTDVVPV-FKEHW SHDPFEGFKDADGY IYGRGAQDMKCVSIQ- 119
1Q7L VLTWPGTNPTLSSILLN SHTDVVPV-FKEHW SHDPFEAFKDSEGY IYARGAQDMKCVSIQ-
1VGY KNIWLRRTGKAPVVC FAGHTDVVPTGPEKWDSPFPFAE-RDGRLYGRGAADMK-TSIAC
      * . . . : : .*****. *:* .*** : : * :*.*** ** **
pACY1 YLEAVRRLKVEGHFRPTIHM TFVPDEEVGGHQGMELFVKRPEFQALRAGFALDEGLASPT 180
1Q7L YLEAVRRLKVEGHRFPRTIHM TFVPDEEVGGHQGMELFVQRPEFHALRAGFALDEGLANPT
1VGY FVTACERFVAKHPNHQGSIAL LITSDDEGDALDGTTKVVDVLKARDELIDYCI V-GEPTAV
      : * .*: : . . * : : .*** . . * : .*. : : : : : * . . .
pACY1 DA--FTVFY SERSPWLRVTS TGKPGHGSRF IEDTAAEKLHKV INSI LAFREKQKRLQSN 239
1Q7L DA--FTVFY SERSPWLRV-----
1VGY DKLGDMIKNGRRGSLSGNL TVKKGQCHI-----AYPHLA INPVHTFAPALLELTQEVWD
      *
pACY1 QLKFGAV-TSVNLTMLEGG-VAYNVVPATMSACFDFRVA FVDLKA FEEQLQSWCQAAGEG 298
1Q7Lab -----
1VGYa EGNEYFPPTSFQISNINGGTGATNV IPGELNVKFNFRFSTESTEAGLKQRVHA ILDKH--G

pACY1 VTFEFVQKWMETQVTSTDDSPFWAA FSGVFKDMKLALELEI--CPASTDARY IRAAGVPA 358
1Q7L -----MPWAAFSRVCKDMNLTLEPEI--MPAAGDNRY IRAVGVPA
1VGY VQYDLQWSCS-GQPFLTQAG-KLTDVARAAI AETC-GIEAELSTTGGTSDGRF IKAMAQEL
      . : : : * * : . * * : * .
pACY1 LGFSPMNHTPVLLHDHDERLHEAVFLRGVDI YTQLLSALASVPALPSES 406
1Q7L LGFSPMNRTPVLLHDHDERLHEAVFLRGVDI YTRLLPALASVPALPSDS
1VGY IELGFSNAT--IHQINENVRLNDIPKLSAVYEGILVRL-----
      : . * * * * * : * : * : : : * : * *

```

图 1 CLUSTALW 对 pACY1, T347G mutant of hACY1 (1Q7L), mnDAPE (1VGY) 进行的多重序列排列

Fig. 1 Sequence alignment of pACY1, T347G mutant of hACY1 (1Q7L), mnDAPE (1VGY) produced by CLUSTALW

1Q7L represents the chains A, B and C, D. 1VGY represents the chains A, B. The deleted sequences in 1Q7L are indicated in italics

2.2 同源模拟和优化

pACY1 的单聚体分子中含有一个锌离子^[8,20], 而在已知的 hACY1 的结构中每个亚基都含有两个锌离子^[11], 这说明 pACY1 的锌离子可能有不同于模板蛋白的配位形式。因此, 我们并没有参照模板中锌离子的坐标进行模拟。根据保守性分析, Biagini 和 Puigserver 提出 pACY1 中的锌离子配位残基为 ASP112, His79, 和 Glu147^[21]。我们用 MODELLER 的距离限制的方法模拟锌离子坐标。

hACY1 突变体 T347G 解析的 hACY1 三维结构是锌结合结构域, 由于二聚体结构域在结晶前就缺失, 使得二聚体的锌结合结构域坐标互相靠近。在以 1Q7L 为模板对 pACY1 二聚体结构进行模拟时, 由于模板坐标不满足 MODELLER 所需要的空间限制, 不能在两个催化结构域中间插入一个二聚体结构。我们把 1Q7L 的 AB 链和 CD 链分开作为两个模板, 最终成功地模拟了 pACY1 二聚体蛋白三维结构。

利用 MODELLER 程序, 生成 50 个 pACY1 的原始模型。用 PROCHECK 程序初步检测了这 50 个模型的立体化学特性。如表 1 所示, 选出其中最好的两个模型 A 和 B 并进行能量最小化。优化后的模

型对应为 A1 和 B1。

表 1 PROCHECK 分析模型的立体化学特性

Table 1 PROCHECK analyzes the qualities of the models

Model	Ramachandran plot quality (%)			Goodness factor			
	Core	Allowed	General	Disallowed	Dihe drals	Covalent	Overall
A	90.2	7.0	2.4	0.4	-0.07	-0.41	-0.19
B	91.7	7.6	0.4	0.2	-0.06	-0.47	-0.21
A1 ^a	89.2	9.2	1.5	0.1	-0.18	0.04	-0.10
B1 ^a	90.5	1.3	0.2	0	-0.15	0.04	-0.08

Notes: a 为 Model A1 和 B1 最小化模型 A 和 B 分别。

2.3 分子动力学模拟

对模型 A1 和 B1 的单体及其二聚体结构进行了无约束分子动力学模拟。分析模拟前后结构构象之间的 RMSD 以评价模型的稳定性。结果如图 2 所示, 模型 A1 和 B1 的单体构象在模拟前后存在少许的 RMSD, 而他们的二聚体构象在模拟前后能完全重合, 说明了二聚体结构较单体更加稳定。另外, 模型 B1 的二聚体构象与单体构象之间的 RMSD 的明显小于模型 A1, 这说明模型 B1 较模型 A1 具有更

好的稳定性。

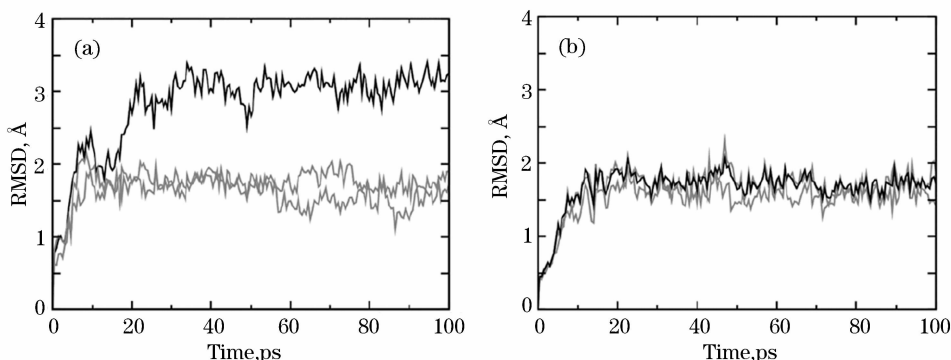


图 2 pACY1 二聚体结构的分子动力学模拟

Fig. 2 Evolution of the dimeric structure of pACY1 during MD simulations

Notes: RMSDs from the starting conformations for the Models A1 and B1 are shown on the plots (a) and (b), respectively. Values for the monomers and the whole dimer are indicated by gray and black lines, respectively

2.4 结构合理性评估

利用 PROCHECK 程序评估了模型的立体化学特性。如表 1 所示, B1 的立体化学特性明显优于其它模型, 并且 B1 的各项参数指标都符合同源模拟结果的要求。Ramachandran plot 结果也表明 (图 3), 模型 B1 90.5% 的残基在核心允许区内, 9.3%

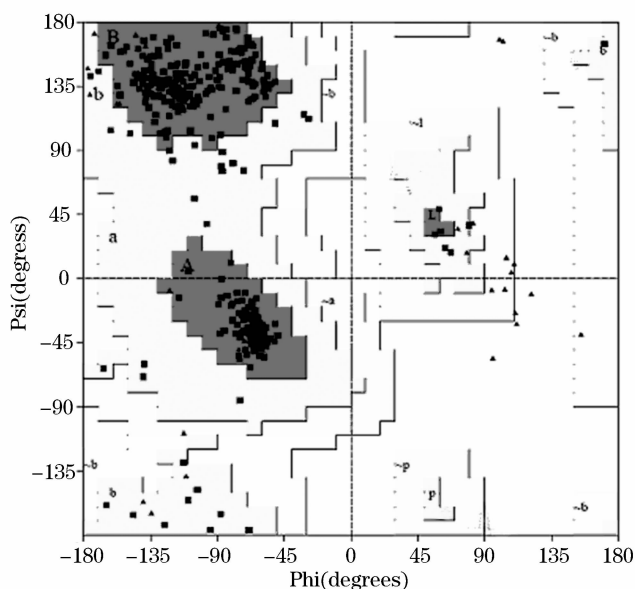


图 3 模型 B1 的 Ramachandran plot 图

Fig. 3 Ramachandran plot the Ψ/Φ distribution of Model B1

Notes: produced by PROCHECK. 90.5% residues are in most favored regions;
8.3% residues are in additional allowed regions;
1.2% residues are in general regions

的残基在最大允许区内, 无残基分布于不允许的区域内。这说明模型 B1 中的每一个残基都处于合理的立体化学构象中, 符合了立体化学 φ, ψ 二面角分布的要求。利用 ProSa 评估模型 A1 和 B1 的残基的能量分布, 如图 4 所示, 模型 A1 有少部分残基的配对能 (pair energy) 大于 0, 而模型 B1 的所有残基的

能量处于负值。这些表明模型 A1 可能存在某些结构错误区域而模型 B1 中的残基的构象都处于能量合理区域。WHATIF 进一步检查了模型 B1 残基的包装质量, 即氨基酸残基之间的相容性。如图 5 所示, 模型 B1 的各个残基的能量都在 WHATIF 规定的合理性范围内 ($Z\text{-core} \geq -5.0$)。

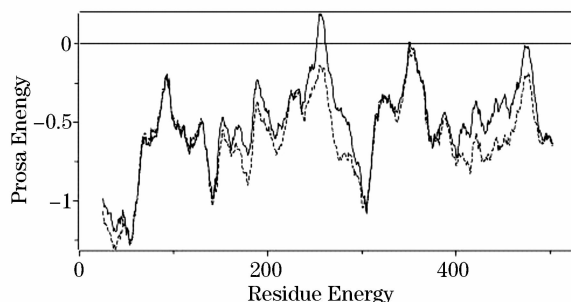


图 4 ProSa 检测模型 A1 和 B1 的能量分布

Fig. 4 ProSa energy profiles calculated for Model A1, B1

Notes: The value of energy should be under 0 for a reliable model. Bold line: Pair energy graphs of Model A1; Dotted line: Pair energy graphs of Model B1

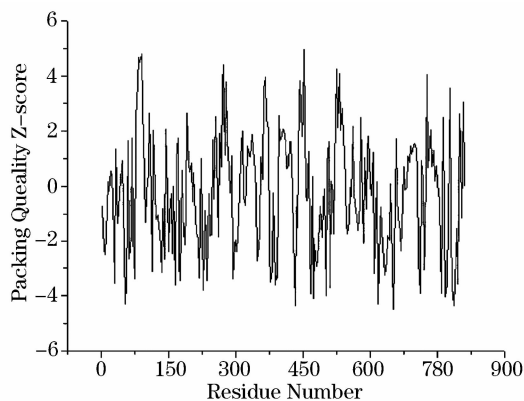


图 5 WHATIF 检测模型 B1 的包装特性

Fig. 5 WHATIF quality control values calculated for Model B1

Notes: The scores should be above -5 for a reliable model

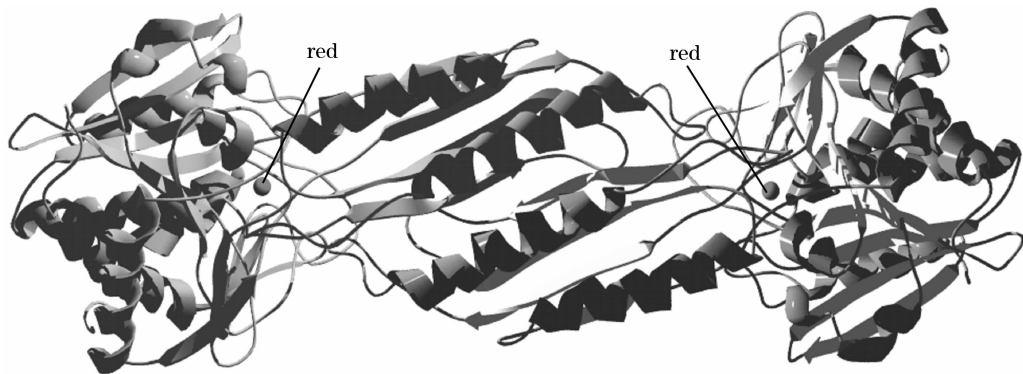


图6 pACY1 二聚体结构模型(模型 B1)

Fig. 6 Ribbon representation of the final pACY1 model

Notes: The backbone of Model B1 is colored by secondary structure, The zinc ion is represented as the red sphere.

3 结论

通过对同源模建的模型进行优化和多项评估,我们筛选出模型 B1 作为对 pACY1 三维结构同源模建的最终结构。如图 6 所示,pACY1 单体由球状的锌离子结合结构域和二聚体结构域组成,每个单体由 10 个 α -螺旋和 15 个 β -折叠组成的。这与我们预测 pACY1 的二级结构相符。其中,锌离子结合结构域的 1~188 和 311~406 氨基酸残基形成了 α -螺旋包夹 β -片层的形状;二聚体结构域的 189~310 氨基酸残基则形成两个 α -螺旋和一个 β -片层。同源模建得到的 pACY1 完整二聚体结构为进一步研究 pACY1 的结构与功能关系打下了基础。另外,由于 hACY1 和 pACY1 在氨基酸水平同源性达到 89%,这个结构的提出有利于探知造成病症的 hacy1 基因突变。

致谢:感谢美国加州大学 Prof. James W. Caldwell 提供的 AMBER7。

参考文献 (References)

- [1] Mitta, M., Kato, I., Tsunasawa, S. The nucleotide sequence of human aminoacylase-1 [J]. *Biochimica et Biophysica Acta*, 1993, 1174: 201-203.
- [2] Tsunasawa, S. Does uncontrolled survival of acetylated peptides lead to cell proliferation? Deletion of N-terminal deacetylating system for protein/peptide in small lung carcinoma cells [J]. *J. Lab. Clin. Med.*, 1992, 120: 505-506.
- [3] Scaloni, A., Jones, W. M., Pospischil, M., Sassa, S., Schneewind, O., Popowicz, AM., Bossa, F., Graziano, SL., Manning, JM. Deficiency of acylpeptide hydrolase in small-cell lung carcinoma cell lines [J]. *J. Lab. Clin. Med.*, 1992, 120: 546-552.
- [4] Maceyka, M., Nava, V. E., Milstien, S., Spiegel, S. Aminoacylase 1 is a sphingosine kinase 1-interacting protein [J]. *FEBS Letters*, 2004, 568: 30-34.
- [5] Sassa, J. O., Olbrich, H., Mohr, V. O., Hart, C., Woldseth, B., Krywawych, S., Bjurulf, B., Lakhani, PK., Buchdahl, RM., Omran, H. Neurological findings in aminoacylase 1 deficiency [J]. *Neurology*, 2007, 68: 2151-2153.
- [6] Tyłki-Szymanska, A., Gradowska, W., Sommer, A., Heer, A., Walter, M., Reinhard, C., Omran, H., Sassa, JO., Juracka, A. Aminoacylase 1 deficiency associated with autistic behavior [J]. *J. Inher. Metab. Dis.*, 2010, 33:1-4.
- [7] Sommer, A., Christensen, E., Schwenger, S., Seul, R., Haas, D., Olbrich, H., Omran, H., Sassa, JO. The molecular basis of aminoacylase 1 deficiency [J]. *Biochimica et Biophysica Acta*, 2011, 1812: 685-690.
- [8] Heese, D., Berger, S., Röhm K. H. Nuclear magnetic relaxation studies of the role of the metal ion in M_n^{2+} -substituted aminoacylase 1 [J]. *Eur. J. Biochem.*, 1990, 188: 175-180.
- [9] Tang, Z. Y., Yu, J. Y., Zhou, Q., He, B., Wang, ZF., Zhou, HM. Secondary structure of holo-and apo-aminoacylase from prediction, circular dichroism, and ft-raman spectroscopy [J]. *J. Biochem.*, 1995, 118: 706-709.
- [10] Zhao, HY., Meng, WY., Lin, ZJ., Zhou HM. Crystallization and preliminary X-ray studies of porcine kidney aminoacylase 1 [J]. *Chinese Science Bulletin*, 1997, 42: 1750-1752.
- [11] Lindner, H., Lunin, V. V., Alary, A., Hecker, R., Cygler, M., Ménard, R. Essential roles of zinc ligation and enzyme dimerization for catalysis in the aminoacylase-1/M20 family [J]. *J. Biol. Chem.*, 2003, 278: 44496-44504.
- [12] Lindner, H., Alary, A., Wilke, M., Sulea, T. Probing the Acyl-Binding Pocket of Aminoacylase-1 [J]. *Biochemistry*, 2008, 47: 4266-4275.
- [13] Larkin, MA., Blackshields, G., Brown, NP., Chenna, R., McGettigan, PA., McWilliam, H., Valentin, F., Wallace, IM., Wilm, A., Lopez, R., Thompson, JD., Gibson, TJ., Higgins, DG. Clustal W and Clustal X version 2.0 [J]. *Bioinformatics*, 2007, 23(21): 2947-2948.
- [14] Kelley, LA., Sternberg, MJE. Protein structure prediction on the web: a case study using the Phyre server [J]. *Nature Protocols*, 2009, 4: 363-371.
- [15] Laskowski, RA., MacArthur, MW., Moss, DS., Thornton, J.

- JM. PROCHECK: a program to check the stereochemical quality of protein structures [J]. *J. App. Cryst.*, 1993, 26: 283-291.
- [16] Wang, J., Cieplak, P., Kollman, PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? [J]. *J. Comput. Chem.*, 2000, 21: 1049-1074.
- [17] Pang, Y. P., Xu, K., Yazal, JE., Prendergas, FG. Successful molecular dynamics simulation of the zinc-bound farnesyltransferase using the cationic dummy atom approach [J]. *Protein Science*, 2000, 9: 1857-1865.
- [18] Wiederstein, M., Sippl, MJ. ProSA-web; interactive web service for the recognition of errors in three-dimensional structures of proteins [J]. *Nucleic Acids Research*, 2007, 35: W407-W410.
- [19] Vriend, G. A molecular modeling and drug design program [J]. *J. Mol. Graph.*, 1990, 8: 52-56.
- [20] Pittelkow, S., Lindner, H., Röhm, KH. Human and porcine aminoacylase 1 overproduced in a baculovirus expression vector system; evidence for structural and functional identity with enzymes isolated from kidney [J]. *Protein Expr. Purif.*, 1998, 12: 269-276.
- [21] Biagini, A., Puigserver, A. Sequence analysis of the aminoacylase-1 family. A new proposed signature for metalloexopeptidases [J]. *Comp. Biochem. Physiol. Part B*, 2001, 128: 469-481.