

doi:10.3969/j.issn.1672-5565.2015.03.06

血清白蛋白序列中的相似氨基酸突变特点分析

陆会宁^{1*}, 齐燕姣², 冯树华², 李世玺², 赵娅敏², 金能智³

(1. 西北民族大学生命科学与工程学院, 兰州 730124;

2. 西北民族大学化工学院, 兰州 730124;

3. 甘肃省计算中心, 兰州 730000)

摘要:血清白蛋白是必不可少的生命物质,在生命运动和发展延续中起着重要的作用。它不仅能维持正常的血浆渗透压,最重要的是能够储存和运输众多的内源性和外源性物质。本文利用生物信息学的方法分析了几种不同物种的血清白蛋白的结构信息和疏水性特点,研究表明人、牛、猴、兔、狼、猫的血清白蛋白序列均属于亲水性蛋白质,在100 bp以内的疏水性值差别比较明显。通过对血清白蛋白进行多序列比对分析,发现兔血清白蛋白的氨基酸突变的数目是最多的。在这几种血清白蛋白序列中,氨基酸突变更容易发生在结构相似、极性相似和能量比较接近的氨基酸之间,如D和L、E和D。对于人血清白蛋白来说,从疏水性的丙氨酸A到酸性的谷氨酸E的突变比较多,使得人血清白蛋白在进化过程中的亲水性增强,是个很好的储存和运输小分子的载体。这些基于生物信息学方面的血清白蛋白的突变及其进化关系的研究,为进一步研究药物与血清白蛋白的相互作用在其他物种中表现和特点提供了良好的基础。

关键词:血清白蛋白;序列比对;突变;疏水性

中图分类号:Q-03 **文献标志码:**A **文章编号:**1672-5565(2015)03-179-07

Mutational characteristics analysis of the similar amino acid of serum albumin sequences

LU Huining^{1*}, QI Yanjiao², FENG Shuhua², LI Shixi², ZHAO Yamin², JIN Nengzhi³

(1. Department of Life Sciences and Biological Engineering, Northwest University for Nationalities, Lanzhou 730124, China;

2. Department of Chemical Engineering, Northwest University for Nationalities, Lanzhou 730124, China;

3. Gansu Province Computing Center, Lanzhou 730000, China)

Abstract: Serum albumin is essential to life and plays an important role in the life movement and development. It can not only maintain normal plasma osmolality, but also can store and transport many endogenous and exogenous substances. In this paper, the structure information and hydrophobic characteristics of serum albumins from several different species were analyzed by using bioinformatics. Results indicated that serum albumin sequences from different species were all hydrophilic, such as Homo sapiens, Bos Taurus, Rabbit, Felis catus, Macaca mulatta and Canis lupushuman. Hydrophobic value was relatively obvious within 100 bp of these serum albumin sequences. By using multiple sequence alignment analysis, we found that the rabbit had the most number of amino acid mutations. Among these serum albumin sequences, the amino acid mutations were more likely to occur between amino acid with similar structure, similar hydrophilic and energy, such as D and L, E and D. For the human serum albumin, mutations from hydrophobic alanine A to glutamic E were relatively much more than others, which may be related with its good hydrophily and carrier of preserving or transporting small molecules. The analysis of mutation and evolutionary relationships based on bioinformatics could provide good foundation for further studying the interaction between drugs and serum albumin sequences in other species.

Keywords: Serum albumin; Sequence alignment; Mutation; Hydrophobic

收稿日期:2015-06-10;修回日期:2015-07-14.

基金项目:国家民委科研项目(14XBZ021);中央高校基本科研业务费专项资金项目(NO.31920130035);中央高校基本科研业务费专项资金项目(zyz2011064)。

*通信作者:陆会宁,男,讲师,研究方向:分子生物学与生物信息学;E-mail:lzz.316@163.com.

血清白蛋白(Serum albumin, SA)是人和动物体血液循环体系中含最丰富的蛋白质,约占血清总蛋白的60%^[1-2]。血清白蛋白除了能维持正常的血浆渗透压外,最基本的作用是能够储存和转运众多的内源性和外源性物质,如金属离子、脂肪酸、氨基酸、胆红素、酶、激素和各种各样的代谢物,以及诸如抗凝血剂、安神药、常用麻醉剂和抗癌抗肿瘤药物^[3-8]等等。在生物体中,血清白蛋白是必不可少的生命物质,在生命运动和发展延续中起着重要的作用。有关血清白蛋白的构象、组成、结构及与小分子的相互作用的机理,是目前生命科学、化学和临床医学共同关注的课题。

由于牛血清白蛋白和人血清白蛋白具有较高的相似性,且来源广泛、廉价等优良条件,常常代替人血清白蛋白被用于研究各种小分子、大分子与血清白蛋白的相互作用^[9-12],期望能在分子水平上揭示相关生命活动过程的奥秘。随着科技的发展,越来越多的实验方法和技术被用来研究血清白蛋白在生命体中重要作用和意义。特别是随着生物技术的发展,对其他哺乳动物的研究也逐渐热门起来^[13-16]。但是对这些物种的蛋白质序列特点和差异等还没有深刻的认识。因此,本论文通过生物信息学方法分析人类(Homo sapiens)、牛(Bos taurus)、兔(Rabbit)、猫(Felis catus)、猕猴(Macaca mulatta)和狼(Canis lupus)的血清白蛋白序列之间的差异,了解序列之间的氨基酸发生突变的特点,从而认识血清白蛋白序列的进化特点,为进一步深入研究其他物种的生理特点和药物与血清白蛋白的相互作用提供了一定的基础。

1 材料和方法

1.1 蛋白质序列

从NCBI网站(<http://www.ncbi.nlm.nih.gov/pubmed>)获得载蛋白质序列,其中来自人类Homo

sapiens、牛Bos taurus、兔Rabbit、猫Felis catus、猕猴Macaca mulatta和狼Canis lupus的血清白蛋白序列ID分别是1MBO、AAN17824.1、NP_001075813.1、NP_001009961.1、NP_001182578.1和BAC10663.1。长度分别为585 aa、607 aa、584 aa、608 aa、608 aa和608 aa。

1.2 蛋白质序列分析

多序列的比对是采用默认的参数设置在ClustalW软件中进行的。利用Lasergene软件分析蛋白质的氨基酸序列长度、组成以及等电点;通过MegAlign软件中Lipman-Pearson方法进行序列相似性分析;TMHMM Server (<http://www.cbs.dtu.dk/services/TMHMM/>)和TMPred (http://www.ch.embnet.org/software/TMPRED_form.html)预测跨膜区。用ExPASy的ProtScale程序进行蛋白质的疏水性分析。基于clustal软件的多序列比对和Jone-Taylor-Thornton(JTT)模型,采取缺失序列删除的方式,用MEGA软件的Neighbor-Joining(NJ)方法构建进化树。

1.3 分子结构优化

在Pubchem上下载氨基酸分子的结构,然后通过Chem3D软件的MM2法做构型优化,精确度为0.001。通过Gaussian软件的密度泛函理论B3LYP的方法和6-31G(D)的基组做进一步的几何优化。

2 结果与分析

2.1 不同物种血清白蛋白序列的相似性

从表1血清白蛋白序列的相似性分析可以看出,人类与猕猴的血清白蛋白序列相似性最高,达到了93.3%。其次就是猫和狼的,是87.3%。人类与牛的序列相似性仅仅是74.5%。由于牛的血清白蛋白序列容易获得,且与人类的有较高的相似性,因此常常在很多研究中用来替代人血清白蛋白^[17-20]。其中,相似性最低的是牛和兔的血清白蛋白序列,只有71.5%。

表1 血清白蛋白序列的相似性

Table 1 Sequence similarity of serum albumin

	BSA	CSA	FSA	HSA	MSA	RSA	%
BSA	100	76.4	77.2	74.5	74.7	71.5	
CSA		100	87.3	80.8	81.4	74.6	
FSA			100	81.8	80.5	75.2	
HSA				100	93.3	74.3	
MSA					100	74.7	
RSA						100	

注:BSA-牛血清白蛋白;CSA-狼血清白蛋白;FSA-猫血清白蛋白;HAS-人血清白蛋白;MSA-猕猴血清白蛋白;RSA-兔血清白蛋白。

Notes:BSA-Bovine Serum Albumin;CSA-Canis lupus familiaris Serum Albumin;FSA-Felis catus Serum Albumin;HAS-Human Serum Albumin;MSA-Macaca mulatta Serum Albumin;RSA-Rabbit Serum Albumin.

2.2 蛋白质序列的疏水性预测

疏水性预测对于蛋白质结构和性质的研究具有重要的作用^[21-22]。通过对这几种血清白蛋白序列进行疏水性预测(见图1),发现除了FSA和CSA外,其他蛋白质序列在100 bp以内的疏水性值跨度比较大。其中RSA具有最大值2.811和最小值-3.489。对于亲缘关系比较近的HAS和MSA来说,

也许正是由于较多的极性氨基酸到疏水性氨基酸之间的E→A突变,使得MSA的疏水性值较HSA的要高。BSA的疏水性要比HSA的较高些,尤其是在100 bp以内的差别比较明显。总体来讲,这几种血清白蛋白都属于亲水性蛋白,没有跨膜区。因此,血清白蛋白作为药物运输载体在生理活动中起着非常重要的作用^[23-27]。

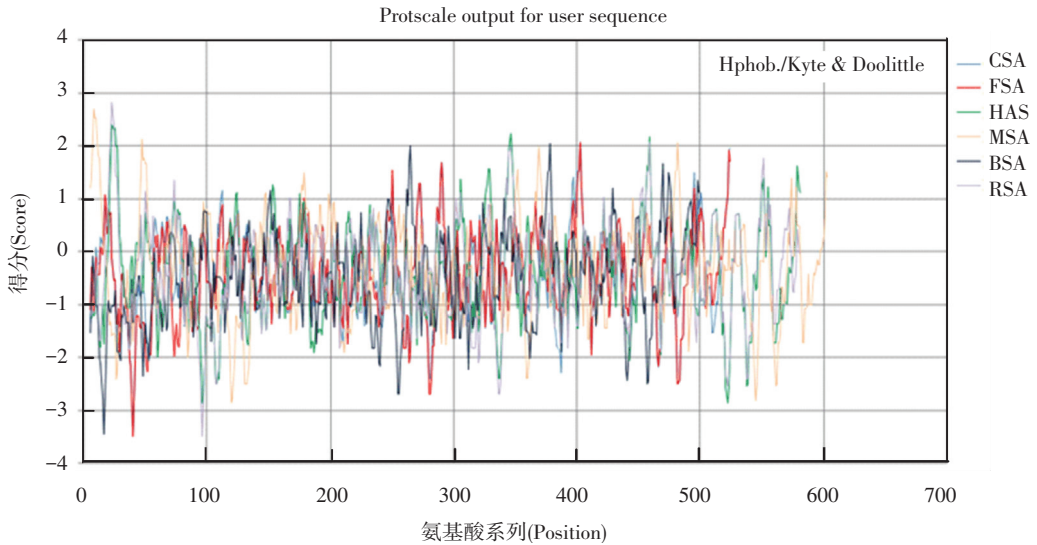


图1 不同物种血清白蛋白序列的疏水性预测

Fig.1 Hydrophobic prediction of serum albumin sequences for different species

2.3 蛋白质序列之间的突变规律

为了了解氨基酸序列的结构特点,本文详细分析了这些不同物种的血清白蛋白序列之间的异同点。基于氨基酸多序列比对^[28-29],我们定义是在同一个位点中,如果有三个以上的氨基酸是相同的D,则认为其公共的氨基酸是D(序列比对具体数据未列)。这样,通过统计分析发现,总共的突变数目是293个,类型有99种,如:E→G(QADHKV)、A→G

(EPTLIVSYK)、D→E(NAGST)、T→D(EAKSG)、F→L(Y)、S→T(APVMKN)、M→V、Q→K(ASETH)、N→S(KDHR)、L→F(GTIVKWMQ)、P→G(QR-KATHS)、V→K(ATLIGSP)、H→Q(KP)、G→K(NDE)、K→H(RNEQTH)、I→V(LM)、Y→F、R→Q(E)。其中RSA的突变位点最多,占27.30%。而CSA、FSA、HAS和MSA的突变数目比较接近,FSA的最少,仅占11.94%(见图2)。

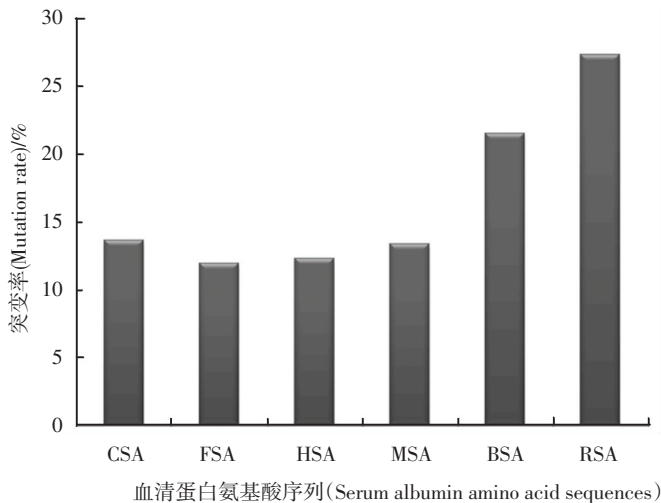


图2 不同物种的血清白蛋白氨基酸序列的突变率

Fig.2 Mutation rate of serum albumin amino acid sequence for different species

在图3中列举了CSA、FSA、HAS、MSA、BSA和RSA氨基酸序列不同类型突变的概率。对于CSA来说,这些突变中,概率最高的是从L→V和从A→G突变;对于FSA来说,概率较高的是D→E和L→V;对于HSA来说,概率较高的是A→E;对于MSA来说,概率较高的是E→A和F→L;对于BSA来说,概率较高的是S→T K→R A→T D→N E→D;RSA是突变最多的一个,其中E→G、E→D、A→G、D→E、F→L、Q→K、K→H、V→L、L→I和S→A是比较多的。在这其中68.2%的氨基酸的突变是发生在极性相似的两

个氨基酸之间。例如,亮氨酸L到缬氨酸V都是疏水性的氨基酸,丝氨酸S和苏氨酸T都属于中性氨基酸;谷氨酸E和天冬氨酸D都属于酸性的氨基酸。对于极性不相同的氨基酸之间的突变,大多数更容易发生从疏水性的氨基酸到极性氨基酸之间突变,例如HSA的丙氨酸A到谷氨酸E的突变。而这正是HSA发生较多的突变,这种序列进化的特点使得人类血清白蛋白的亲水性增强。而与HSA亲缘关系(见图4)较远的CSA和FSA来说,更容易发生的是疏水性氨基酸与疏水性氨基酸之间的突变。

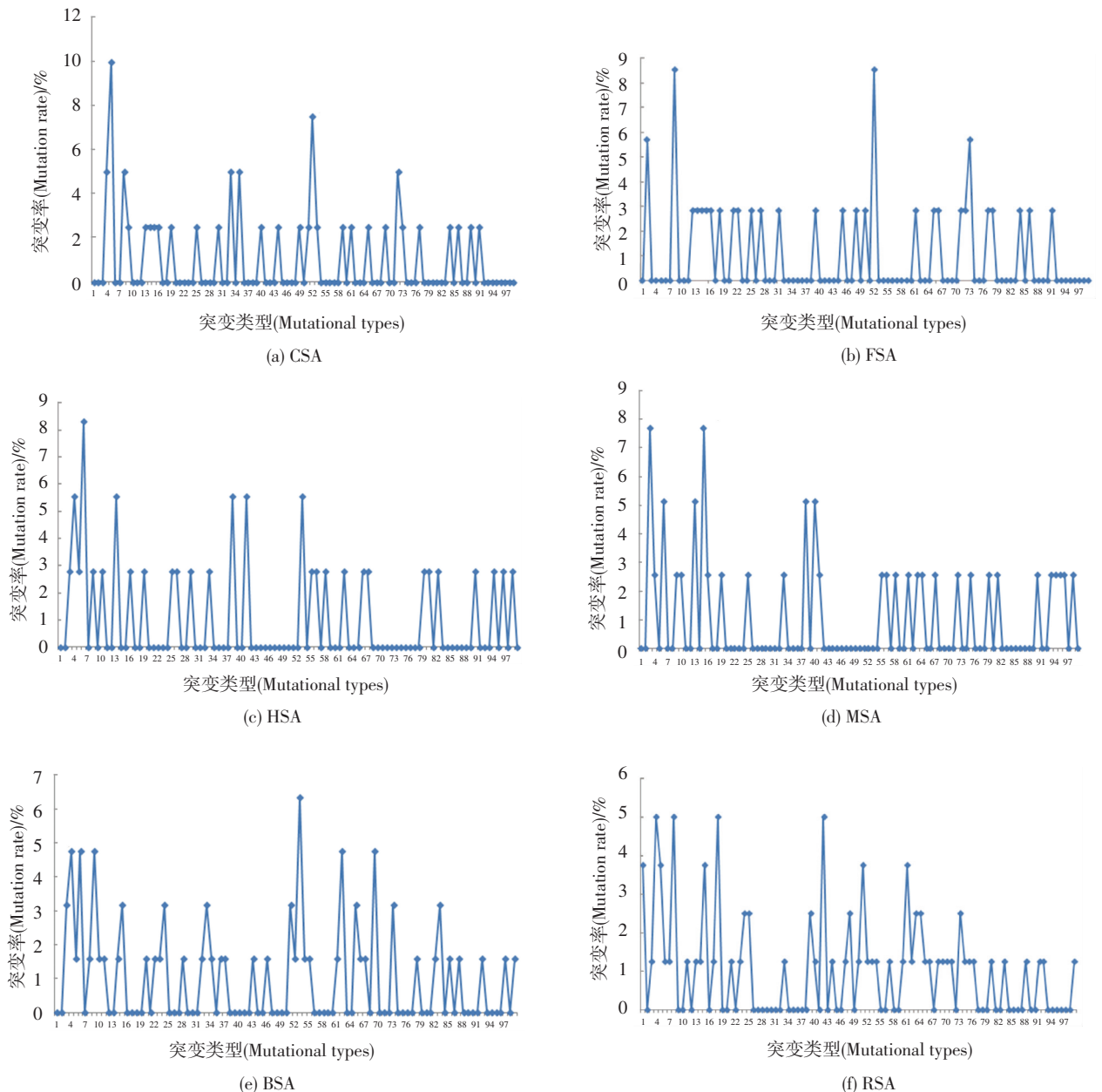


图3 CSA、FSA、HAS、MSA、BSA和RSA氨基酸序列不同类型突变的概率

Fig.3 Different types of mutations rates of CSA、FSA、HAS、MSA、BSA and RSA amino acid sequences

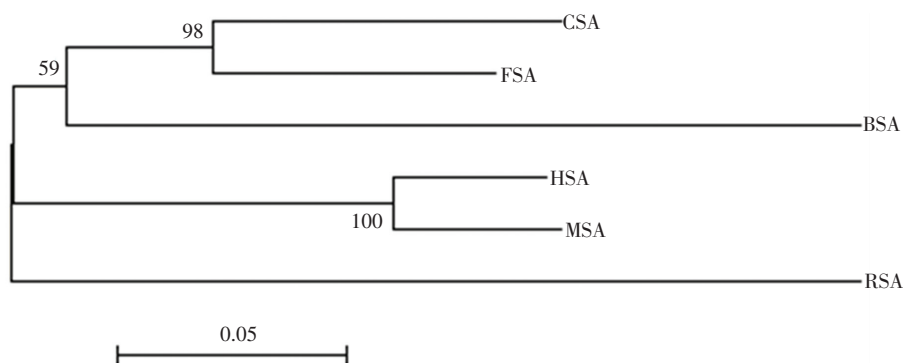


图4 血清白蛋白氨基酸序列的系统发育树

Fig.4 The phylogenetic tree of serum albumin amino acid sequences

从氨基酸的结构特点来看(见图5),突变更容易发生结构相似的两个氨基酸之间。例如,丙氨酸A和谷氨酸E之间、缬氨酸V和亮氨酸L之间、谷氨酸E和天冬氨酸D之间。经过B3LYP/6-31G*条件下对这几种氨基酸的机构进行优化^[30-33],发现丙氨酸A(见图5a)和谷氨酸E(见图5b)之间的稳定构象很相似,差别在于左侧R基团CH₃和CH₂CH₂COOH的不同。正是由于这个R基团的差异,导致这两个分子之间的能量差是 5.981×10^5 kJ/mol(A的是 -8.497×10^5 kJ/mol,E的是 -1.448×10^6 kJ/mol),而且A是疏水性的氨基酸,E是酸性氨基酸。因此,可以

认为这两个氨基酸分子之间的突变是比较难于发生的,主要是在MSA和BSA上。谷氨酸E和天冬氨酸D(见图5c)的稳定构象差别较为明显,但是这两个分子都属于酸性氨基酸,最低能量差仅为 1.032×10^5 kJ/mol。缬氨酸V和亮氨酸L的分子结构差别仅仅是一个CH₂基团,但是稳定构象差别比较明显(见图5d和图5e)。这两种氨基酸都属于疏水性氨基酸,分子的最低能量分别是 -1.056×10^6 kJ/mol和 -1.159×10^6 kJ/mol,差也是 1.032×10^5 kJ/mol。因此可以推断氨基酸E和D之间、V和L之间的突变要比氨基酸A和E之间的突变容易发生。

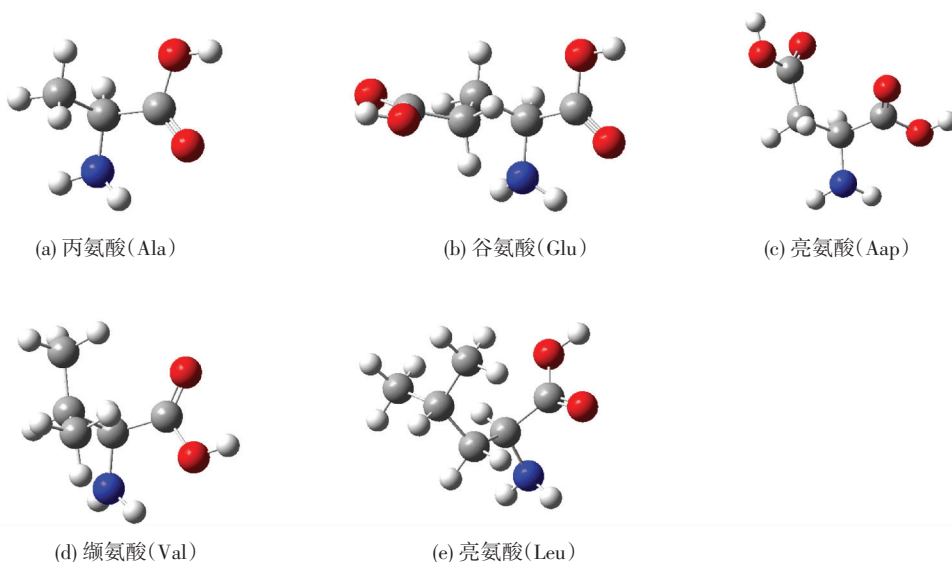


图5 经B3LYP/6-31G*条件下优化的氨基酸分子结构

Fig.5 Optimized molecular structure of amino acids under the condition of the B3LYP/6-31G*

3 结论

用生物信息学的方法讨论血清白蛋白的结构和功能,通过疏水性预测发现人、牛、猴、兔、狼、猫的血

清白蛋白序列均属于亲水性蛋白质。通过对血清白蛋白进行多序列比对分析氨基酸的突变点和插入/缺失,发现Rabbit的血清白蛋白突变是最多的。突变容易发生在结构相似、极性相似和能量比较接近的氨基酸D和L以及氨基酸E和D之间。对于人

类来说,其血清白蛋白氨基酸序列中发生了从疏水性的丙氨酸 A 到酸性的谷氨酸 E 的突变,尽管这两种氨基酸分子的稳定结构比较相似,但是它们之间的能量差距比较大。这些基于生物信息学方面的血清白蛋白的研究,对于进一步研究药物与血清白蛋白的相互作用,以及在其他物种中表现和特点奠定了一定的基础。

参考文献

- [1] YANG F, ZHANG Y, LIANG H. Interactive association of drugs binding to human serum albumin[J]. *Int J Mol Sci.*, 2014, 15(3): 3580–3595.
- [2] YASSEEN Z J, HAMMAD J H, ALTALLA H A. Thermodynamic analysis of thymoquinone binding to human serum albumin [J]. *Spectrochim Acta A Mol Biomol Spectrosc.*, 2014, 124: 677–681.
- [3] ZHANG L, EISENBERG A. Multiple morphologies of “Crew-Cut” aggregates of polystyrene-b-poly (acrylic acid) block copolymers [J]. *Science*, 1995, 268 (5218): 1728–1731.
- [4] WANG Q, SUN Q, MA X, et al. Probing the binding interaction of human serum albumin with three bioactive constituents of *eriobotrya japonica* leaves: Spectroscopic and molecular modeling approaches [J]. *J Photochem Photobiol B.*, 2015, 148:268–276.
- [5] CHEN T, ZHU X, CHEN Q, et al. Interaction between Z-ligustilide from *Radix Angelica sinensis* and human serum albumin[J]. *Food Chem.*, 2015, 186:292–297.
- [6] YU X, JIANG B, LIAO Z, et al. Study on the interaction between Besifloxacin and bovine serum albumin by spectroscopic techniques[J]. *Spectrochim Acta A Mol Biomol Spectrosc.*, 2015, 149:116–121.
- [7] CHEN M, GUO H, LIU Y, et al. Structural changes of human serum albumin induced by cadmium acetate[J]. *J Biochem Mol Toxicol.*, 2014, 28(6):281–287.
- [8] ROY S, DAS T K. Spectroscopic studies of interaction between biologically synthesized silver nanoparticles and bovine serum albumin [J]. *J Nanosci Nanotechnol.*, 2014, 14(7):4899–905.
- [9] SHI J H, CHEN J, WANG J, et al. Binding interaction of sorafenib with bovine serum albumin: Spectroscopic methodologies and molecular docking [J]. *Spectrochim Acta A Mol Biomol Spectrosc.*, 2015, 149:630–637.
- [10] FAN H, JI N, ZHAO M, et al. Interaction of bovine serum albumin with starch nanoparticles prepared by TEMPO-mediated oxidation[J]. *Int J Biol Macromol.*, 2015, 78:333–338.
- [11] PENG X, QI W, HUANG R, et al. Elucidating the influence of gold nanoparticles on the binding of salvianolic Acid B and rosmarinic Acid to bovine serum albumin[J]. *PLoS One.*, 2015, 10(4):e0118274.
- [12] BLAZKOVA I, SMERKOVA K, BLAZKOVA L, et al. Doxorubicin interactions with bovine serum albumin revealed by microdialysis with on-line laser-induced fluorescence detection at subpicogram level [J]. *Electrophoresis.*, 2015, doi: 10.1002/elps.201400564.
- [13] LEGROS C, CHESNEAU D, BOUTIN J A, et al. Melatonin from cerebrospinal fluid but not from blood reaches sheep cerebral tissues under physiological conditions[J]. *J Neuroendocrinol.*, 2014, 26(3):151–163.
- [14] NUNAMAKER E A, LEE D R, LAMMEY M L. Chronic diseases in captive geriatric female Chimpanzees (Pan troglodytes) [J]. *Comp Med.*, 2012, 62(2):131–136.
- [15] NAIDU S, PETERSON M L, SPEAR B T. Alpha-fetoprotein related gene (ARG): a new member of the albumin gene family that is no longer functional in primates[J]. *Gene.*, 2010, 449(1–2):95–102.
- [16] NADEZHDIRIN E V, VINOGRADOVA T V, SVERDLOV E D. A comparative analysis of regulatory regions of the transthyretin gene in the mouse, human, and chimpanzee genomes[J]. *Bioorg Khim.*, 2004, 30(4):383–388.
- [17] CHOW Y H, YAP Y J, ANUAR M S, et al. Interfacial partitioning behaviour of bovine serum albumin in polymer-salt aqueous two-phase system[J]. *J Chromatogr B Analyt Technol Biomed Life Sci.*, 2013, 934:71–78.
- [18] KLAUMÜNZER M, WEICHSEL U, MAČKOVIĆ M, et al. Transmission electron microscopy and time resolved optical spectroscopy study of the electronic and structural interactions of ZnO nanorods with bovine serum albumin [J]. *J Phys Chem B.*, 2013, 117(33):9683–9689.
- [19] 王君, 丁娜, 张朝红, 等. 光谱法研究纳米二氧化硅 (SiO₂) 催化超声波照射对牛血清白蛋白 (BSA) 的损伤 [J]. *光谱学与光谱分析.*, 2009, 29(4): 1069–1073. WANG Jun, DING Na, ZHANG Chaohong, et al. Investigation on damage of bovine serum albumin (BSA) catalyzed by Nano-sized silicon dioxide (SiO₂) under ultrasonic irradiation using spectral methods[J]. *Spectroscopy and Spectral Analysis.*, 2009, 29(4): 1069–1073.
- [20] 周菊峰, 胡钞粟, 蒋建宏, 等. 槲皮素铜配合物与牛血清白蛋白相互作用的荧光光谱分析[J]. *山东化工.*, 2015, 44(4): 55–57. ZHOU Jufeng, HU Chaoli, JIANG Jianghong, et al. Investigation on the interaction of quercetin with bovine serum albumin [J]. *Shandong Chemical Industry.*, 2015, 44(4): 55–57.
- [21] 曾茂茂, 王霄, 陈洁. 蛋白质疏水性测定方法的相关性及适用性[J]. *食品科学.*, 2011, 32(15): 117–120. ZENG Maomao, WANG Xiao, CHEN Jie. Correlation and applicability of different methods for determining protein

- hydrophobicity [J]. *Food Science*, 2011, 32(15): 117-120.
- [22] 彭新俊, 王翼飞. 基于残基序列信息的蛋白质相互作用位点预测[J]. *计算机与应用化学*, 2008, 25(6): 649-653.
PENG Xinjun, WANG Jifei. Predicting protein interaction sites with residue sequence information [J]. *Computers and Applied Chemistry*, 2008, 25(6): 649-653.
- [23] ZHANG Y, LEE P, LIANG S, et al. Structural basis of non-steroidal anti-inflammatory drug diclofenac binding to human serum albumin[J]. *Chem Biol Drug Des*, 2015, doi: 10.1111/cbdd.12583.
- [24] LI Y, WANG Y, WANG A, et al. Spectroscopic study on the interaction of bovine serum albumin with zinc (II) phthalocyanine [J]. *Luminescence*, 2015, doi: 10.1002/bio.2908.
- [25] REHMAN M T, KHAN A U. Understanding the interaction between human serum albumin and anti-bacterial/anti-cancer compounds [J]. *Curr Pharm Des.*, 2015, 21(14):1785-1799.
- [26] ZHANG Q, NI Y, KOKOT S. Competitive interactions of ionic surfactants with salbutamol and bovine serum albumin: a molecular spectroscopy study with implications for salbutamol in food analysis [J]. *J Agric Food Chem.*, 2013, 61(32):7730-7738.
- [27] TANG L, JIA W, ZHANG D. The effects of experimental conditions of fluorescence quenching on the binding parameters of apigenin to bovine serum albumin by response surface methods[J]. *Luminescence*, 2014, 29(4):344-351.
- [28] BAWONO P, VAN DER VELDE A, ABELN S, et al. Quantifying the displacement of mismatches in multiple sequence alignment benchmarks[J]. *PLoS One.*, 2015, 10(5):e0127431.
- [29] 苏进展, 王想兰, 冯世秀, 等. 叶子花重瓣品种 DOD 基因序列比对分析[J]. *草原与草坪*, 2014, 34(2): 7-10.
SU Jinzhan, WANG Xianglan, FENG Shixiu, et al. Comparison of DOD gene sequence in multi-bracted cultivars of *Bougainvillea* [J]. *Grassland and Turf*, 2014, 34(2): 7-10.
- [30] XAVIER T S, KENNY P T, MANIMARAN D, et al. FT-IR and Raman spectroscopic and DFT studies of anti-cancer active molecule N-{(meta-ferrocenyl) Benzoyl} -l-Alanine -Glycine ethyl ester[J]. *Spectrochim Acta A Mol Biomol Spectrosc*, 2015, 145:523-30.
- [31] HOU M, MEI Q, HAN B. Solvent effects on geometrical structures and electronic properties of metal Au, Ag, and Cu nanoparticles of different sizes[J]. *J Colloid Interface Sci.*, 2015, 449:488-93.
- [32] FU H, CAO M, SHE Y, et al. Electronic effects of the substituent on the dioxygen-activating abilities of substituted iron tetraphenylporphyrins: a theoretical study[J]. *J Mol Model.*, 2015, 21(4):92.
- [33] SOLIMAN S M, HAGAR M, IBID F, et al. Experimental and theoretical spectroscopic studies, HOMO-LUMO, NBO analyses and thione-thiol tautomerism of a new hybrid of 1, 3, 4-oxadiazole-thione with quinazolin-4-one [J]. *Spectrochim Acta A Mol Biomol Spectrosc*, 2015, 145:270-279.