

免疫信息学预测在治疗蛋白研究中的应用与发展

张林涵, 陈建华*

中国药科大学生命科学技术学院, 江苏 南京

收稿日期: 2023年4月20日; 录用日期: 2023年5月19日; 发布日期: 2023年5月26日

摘要

治疗性蛋白存在的免疫原性问题, 能够潜在地引起抗药物抗体的产生或细胞介导的免疫反应, 因此免疫原性评价是治疗蛋白开发过程中至关重要的一步。由于免疫系统的复杂性, 直接的实验方法是昂贵且耗时的, 并且实验结果需要反复认证。目前国内外存在多种免疫信息学工具, 可针对治疗蛋白的B细胞表位以及T细胞表位进行预测。利用多种免疫信息学工具进行治疗蛋白免疫原性研究, 能够预测出治疗蛋白的高免疫原性位点并对其进行改造, 降低免疫原性。免疫信息学的应用避免了传统免疫学验证的缺点, 为蛋白药物免疫原性的探究提供了新思路。

关键词

治疗蛋白, 免疫原性, 免疫信息学

Application and Development of Immunoinformatics Prediction in Therapeutic Protein Research

Linhan Zhang, Jianhua Chen*

College of Life Science and Technology, China Pharmaceutical University, Nanjing Jiangsu

Received: Apr. 20th, 2023; accepted: May 19th, 2023; published: May 26th, 2023

Abstract

Therapeutic proteins present immunogenicity issues that can potentially cause the production of

*通讯作者。

anti-drug antibodies or cell-mediated immune responses, making immunogenicity evaluation a critical step in the development of therapeutic proteins. Due to the complexity of the immune system, direct experimental methods are expensive and time-consuming, and the results need to be repeatedly certified. Various immunoinformatics tools exist both domestically and internationally to predict B-cell epitopes as well as T-cell epitopes for therapeutic proteins. Immunogenicity studies of therapeutic proteins using multiple immunoinformatics tools can predict highly immunogenic sites of therapeutic proteins and modify them to reduce immunogenicity. The application of immunoinformatics avoids the drawbacks of traditional immunological validation and provides new ideas for the exploration of protein drug immunogenicity.

Keywords

Therapeutic Protein, Immunogenicity, Immunoinformatics

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 免疫信息学对治疗蛋白的免疫原性预测

治疗蛋白是利用活细胞或病毒开发的生物来源药物, 其应用于临床治疗的主要障碍之一便是免疫原性问题。免疫原性是治疗蛋白开发的关键, 传统的免疫原性测量方法将整个抗原进行鉴定, 然后评估其片段与免疫细胞或抗体的反应性。尽管高通量方法在定性和功能性分析方面取得了进展, 但这些方法仍然是费力、耗时和昂贵的。随着免疫学的发展, 对于免疫与相关疾病、抗原抗体的相互作用等研究也不断加深, 利用研究过程中积累的大量免疫相关信息数据搭建的数据库[1] [2] [3], 能够对于多种类的蛋白进行免疫原性预测与分析。

当外来物质暴露在免疫系统中时, 它会引起特定的免疫反应, 这种能力被称为免疫原性, 免疫原性由抗原上存在的抗原表位决定, 抗原表位根据与不同免疫细胞的相互作用, 分为 B 细胞表位以及 T 细胞表位, 免疫信息学数据库针对蛋白的免疫原性预测, 主要也是指针对 B、T 细胞抗原表位的预测。

B 细胞表位能与 B 细胞表面的 B 细胞表位受体(B Cell Receptor, BCR)结合, 通过 BCR 的信号转导, 激活 B 细胞内的多种通路, 最终导致 B 细胞的增殖、活化, 合成和分泌免疫球蛋白[4]。B 细胞表位可被进一步分为线性表位和构象表位。B 细胞线性表位是由抗原中一段连续的氨基酸残基组成, 位于抗原的表面且结构简单。因此线性表位的预测更多需要考虑氨基酸性质, 如氨基酸极性、表面可及性、亲水性等性质。常用的线性 B 细胞表位预测工具如 Bcepred、ABCpred 等。Bcepred 对氨基酸残基的柔韧性、亲水性、表面可及性和极性等方面进行分析, 利用获得的 1029 个非冗余 B 细胞表位和从随机获得的相同数量的非表位对数据库进行训练, 预测准确度为 58.70% [5]; ABCpred 使用标准前馈(FNN)和递归神经网络(RNN)预测抗原序列中的 B 细胞表位, 该数据集由从数据库中获得的 700 个非冗余 B 细胞表位和从 Swiss-Prot 数据库中随机获得的相同数量的非表位组成。最终网络的总体预测准确率为 65.93% [6]。B 细胞构象表位是抗原的三维结构表位, 由序列中可能不相邻, 但是蛋白的折叠可以使它们在空间上接近的氨基酸残基组成。由于 B 细胞构象表位具有空间结构, 预测过程需要建立三维模型。常用的 B 细胞构象表位数据库如: DiscoTope、ElliPro。DiscoTope 利用抗体-抗原蛋白复合物的 X 射线晶体学确定构象表位的分析方法, 对氨基酸残基种类、空间结构和氨基酸残基特性等信息进行汇总分析, 并创建分数评估

体系[7]。ElliPro 利用 Thornton 等人提出的一种在蛋白质球状表面突出的蛋白质区域中识别连续表位的方法, 在抗体-蛋白质结合位点计算椭圆外氨基酸残基的突出指数进行预测[8]。

T 细胞表位指能与 T 细胞表面的主要组织相容性复合体(Major Histocompatibility Complex, MHC)分子结合, 被 T 细胞受体(T Cell Receptor, TCR)识别并产生免疫应答的表位。根据抗原表位的来源, 可被进一步分为 MHC I 类分子结合表位和 MHC II 类分子结合表位。T 细胞上 MHC I 类分子针对内源性抗原表位识别, 常用预测工具如: RankPep。RankPep 将输入的肽段与设定的 MHC 分子的结合潜力, 与一组已知且与设定的 MHC 分子的能够结合的相似肽段进行比较评估, 从已知与特定 MHC 结合的肽集合中导出的比对图谱。将这些谱应用于肽段与 MHC I 结合的预测, 并开发新算法对输入蛋白中的所有可能的肽进行排序[9]。MHC II 类分子针对外源性抗原识别, 常用预测工具如: PREDIVAC。PREDIVAC 是利用高亲和力和结合数据开发出的预测工具, 实现了 T 细胞表位定位并且预测结果实现了 MHC II DR 等位基因 95% 的覆盖[10]。

2. 免疫信息学预测的应用

目前随着免疫信息学的发展, 市场上已经出现了海量不同种类不同算法的预测工具, 其预测逻辑也在不断完善。免疫信息学的应用一方面能够通过预测, 改造具有高免疫原性的治疗蛋白, 降低免疫原性提高药物效率; 另一方面在于通过预测特定蛋白的高免疫原性肽段, 制备多肽疫苗, 注射到机体后刺激机体产生抗体, 预防和治疗特定疾病。由于免疫信息学可以有效地将研究成果转化应用, 在免疫学研究中发挥了越来越重要的作用, 也使得越来越多的科学家将注意力集中到了这种成本低、耗时少的新方法上。

尿酸是一种由嘌呤代谢产生的小分子物质, 在生理 pH 值和温度下, 人体血清尿酸浓度高于(408 $\mu\text{mol/L}$ [6.8mg/dL])时, 称为高尿酸血症。在此浓度下尿酸盐易形成晶体, 在全身各处沉积, 引发痛风。痛风进程主要分为: 高尿酸血症、无症状的尿酸盐晶体沉积, 晶体沉积无症状痛风, 晶体沉积伴急性痛风片, 以及痛风石、慢性痛风关节炎和放射学侵蚀为特征的晚期痛风[11]。由于人类在进化过程中尿酸酶基因的沉默, 导致尿酸在体内难降解难排泄。尿酸酶将尿酸转化为易溶解以及易被肾脏排的尿囊素, 是具有潜力的降尿酸药物[12]。首次利用尿酸酶进行的降尿酸治疗, 采用的是黄曲霉来源尿酸酶, 治疗和预防人肿瘤溶解综合征引发的高尿酸血症[13] [14], 为了改善异源蛋白带来的免疫原性问题, 2010 年 Savient 公司研发了新的尿酸酶产品(Krystexxa), 嵌合尿酸酶与聚乙二醇(PEG)相连接, 进一步降低免疫原性[15]。但是在静脉滴注过程中, 发现 PEG 的引入虽然降低了尿酸酶抗体的产生, 但是多达 40% 的患者体内产生了高滴度抗 PEG 抗体, 导致了后续治疗的失败[16]。因此该产品只被 FDA 批准用于传统治疗无效的痛风病人及难治性成人慢性痛风。对于尿酸酶的免疫原性, 还需要更进一步的研究。目前在尿酸酶领域, 免疫原性已经成为阻碍尿酸酶药物研究的关键。免疫信息学预测针对尿酸酶的应用也在逐渐兴起, 如 Nelapati 等人利用 IEDB 已整合的数据库, 对 *Arthrobacterglobiformis* 以及 *Bacillus fastidiosus* 来源的尿酸酶, 进行了亲水性、表面可及性、延展性 3 个维度的氨基酸特性预测以及利用 Discotope 进行 B 细胞的构象表位分析, 同时也评估了 T 细胞表位。并通过多序列比对和基序分析, 避开进化过程中的保守残基。最终确定了 159、169、264、203 位氨基酸进行突变, 突变后的尿酸酶在上述数据库中均表现为免疫原性降低。最后通过分子对接、自由能计算和分子动力学模拟评估了诱变对 Uricase 催化活性和结构稳定性的影响[17]。Tripathi 等人也利用了 IEDB 数据库对 *Aspergillus flavus*、*Bacillus subtilis* 以及猪来源尿酸酶进行了 T、B 细胞表位研究, 同时进行了致敏性预测, 最终构建了完整的尿酸酶表位图谱[18]。这些研究对于未来尿酸酶药物的发展奠定了重要基础, 解释了尿酸酶免疫原性可能的来源。

L-天冬酰胺水解酶(EC 3.5.1.1)能够催化 L-天冬酰胺脱酰胺, 反应产生 L-天冬氨酸和氨。恶性白血病细胞通过剥夺细胞外来源的 L-天冬酰胺实现快速增殖, 与正常细胞相反, 恶性白血病细胞只能缓慢地合

成 L-天冬酰胺, 其生存依赖于外源供应。L-天冬酰胺酶作为药物, 可以分解血液中的天冬酰胺, 其抗肿瘤活性是通过耗尽血液循环池中的 L-天冬酰胺, 使得肿瘤细胞缺乏这种特定的氨基酸而导致细胞凋亡来实现的。患者血清中 L-天冬酰胺的快速消耗确保了通过细胞凋亡实现最佳的白血病杀伤, 因此 L-天冬酰胺水解酶常被用作治疗急性儿童淋巴细胞白血病和淋巴瘤的治疗剂[19]。目前临床作为药物使用的 L-天冬酰胺均来源于 *E. coli*、*Erwiniachrysanthemi* 菌属以及其 PEG 形式, 但是由于异源蛋白的使用, 患者很大概率出现了过敏反应, 甚至出现过敏性休克[20], 这对治疗有着巨大的影响。Ramya L N 等人通过预测, 鉴定出了将 L-天冬酰胺酶的 B 细胞和 T 细胞表位, 进行了抗原表位点突变并通过分子动力学模型进行了验证, 突变后 L-天冬酰胺酶的免疫原性评分降低 50% [21]。实验结果将为进一步的酶工程和临床研究铺平道路, 以获得更有效的 L-天冬酰胺酶, 提高患者生存率。

除了通过免疫信息学预测降低治疗蛋白免疫原性, 利用预测出的高免疫原性肽段进行疫苗研发, 也是免疫信息学的热门应用方向。肺炎克雷伯菌在人类胃肠道定植, 易在免疫功能低下的儿童和成人中引起严重感染, 是全球住院感染的主要原因之一, 也是在儿科病房的新生儿和婴儿中脑膜炎和败血症的主要元凶, 同时也是社区获得性肺炎的病原体之一。随着抗生素的广泛使用, 肺炎克雷伯氏菌的耐药性问题也愈演愈烈[22]。为了避免耐药性的产生, Zargarani F N 等人利用 IEDB 已整合的免疫信息学数据库预测出了肺炎克雷伯菌菌毛蛋白上潜在的 B 细胞表位, 同时也利用 IEDB、SYFPEITHI、Pro-PredI 工具进行了 T 细胞表位预测, 对获得的表位进行了分子动力学研究, 预测和分析得到的表位均具有较高的亲和力, 具有适合作为疫苗开发的潜力[23]。

3. 免疫信息学的潜在问题

尽管免疫信息学预测的应用逐渐兴起, 数据库准确性仍然受到数据库的构架、训练数据库的参考数据、抗原抗体结合过程的影响, 目前依旧存在着各种问题。在 B 细胞表位预测中, 由于 B 细胞线性表位简单的结构以及只与氨基酸特性有关的特点, 使得 B 细胞线性表位的预测较为简单, 此类数据库的建立数量大, 常用数据库的预测准确性较高; 于此相比较, 由于 B 细胞构象表位涉及空间构象, 其预测过程需要对蛋白进行建模, 对数据库的训练要求也更高, 因此市场上存在 B 细胞构象表位数据库的数量少, 不同训练逻辑下的数据库预测准确性差距大, 总体上均低于 B 细胞线性表位。抗原与 T 细胞表位的结合由于涉及到抗原呈递细胞处理再呈递给 T 细胞的过程, 同时 T 细胞 MHC 分子结合槽与表位结合的研究尚不充分, 导致 T 细胞表位预测准确性要整体低于 B 细胞表位预测。其中 MHC I 类分子由于其结合槽结构较为封闭保守且存在锚定残基, 对于构建预测数据库有益, 因此预测难度低于 MHC II 分子表位; MHC II 分子结合槽较为开放, 并且表位识别过程与结合槽外侧链基团也有紧密联系, 导致 MHC II 分子与抗原表位结合相关数据不足或质量低, 抗原表位识别较为困难, 因此预测难度大大高于 MHC I 类分子结合表位[24]。

免疫信息学的参与, 为治疗蛋白的免疫原性研究提供了强大的助力, 但是免疫信息学本身发展并不完善, 在应用过程中少数数据库的预测结果存在着很大不确定性。因此使用多种在不同参考数据、不同算法逻辑下的不同数据库的预测信息进行综合评估, 才能保证预测结果的准确性。同时需将预测结果在治疗蛋白的三维结构中复现, 选择满足成为抗原表位的基本原则的肽段, 从分子动力学、热力学等角度深层次揭示可能使得肽段成为表位的条件, 并进一步优化和氨基酸改造才能有效减小治疗蛋白的免疫原性。

参考文献

- [1] Bahrami, A.A., Payandeh, Z., Khalili, S., et al. (2019) Immunoinformatics: *In Silico* Approaches and Computational

- Design of a Multi-Epitope, Immunogenic Protein. *International Reviews of Immunology*, **38**, 307-322. <https://doi.org/10.1080/08830185.2019.1657426>
- [2] Doneva, N., Doytchinova, I. and Dimitrov, I. (2021) Predicting Immunogenicity Risk in Biopharmaceuticals. *Symmetry-Basel*, **13**, Article No. 388. <https://doi.org/10.3390/sym13030388>
- [3] Faraji, F., Karjoo, Z., Moghaddam, M.V., *et al.* (2018) Challenges Related to the Immunogenicity of Parenteral Recombinant Proteins: Underlying Mechanisms and New Approaches to Overcome It. *International Reviews of Immunology*, **37**, 301-315. <https://doi.org/10.1080/08830185.2018.1471139>
- [4] Treanor, B. (2012) B-Cell Receptor: From Resting State to Activate. *Immunology*, **136**, 21-27. <https://doi.org/10.1111/j.1365-2567.2012.03564.x>
- [5] Saha, S. and Raghava, G.P.S. (2004) BcePred: Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-Chemical Properties. In: Nicosia, G., Cutello, V., Bentley, P.J., *et al.*, Eds., *Artificial Immune Systems, Proceedings*, Springer, Berlin, 197-204. https://doi.org/10.1007/978-3-540-30220-9_16
- [6] Saha, S. and Raghava, G.P.S. (2006) Prediction of Continuous B-Cell Epitopes in an Antigen Using Recurrent Neural Network. *Proteins-Structure Function and Bioinformatics*, **65**, 40-48. <https://doi.org/10.1002/prot.21078>
- [7] Andersen, P.H., Nielsen, M. and Lund, O. (2006) Prediction of Residues in Discontinuous B-Cell Epitopes Using Protein 3D Structures. *Protein Science*, **15**, 2558-2567. <https://doi.org/10.1110/ps.062405906>
- [8] Ponomarenko, J., Bui, H.H., Li, W., *et al.* (2008) ElliPro: A New Structure-Based Tool for the Prediction of Antibody Epitopes. *BMC Bioinformatics*, **9**, Article No. 514. <https://doi.org/10.1186/1471-2105-9-514>
- [9] Reche, P.A., Glutting, J.P. and Reinherz, E.L. (2002) Prediction of MHC Class I Binding Peptides Using Profile Motifs. *Human Immunology*, **63**, 701-709. [https://doi.org/10.1016/S0198-8859\(02\)00432-9](https://doi.org/10.1016/S0198-8859(02)00432-9)
- [10] Oyarzún, P., Ellis, J.J., Bodén, M., *et al.* (2013) PREDIVAC: CD4+ T-Cell Epitope Prediction for Vaccine Design That Covers 95% of HLA Class II DR Protein Diversity. *BMC Bioinformatics*, **14**, Article No. 52. <https://doi.org/10.1186/1471-2105-14-52>
- [11] Perez-Ruiz, F., Calabozo, M., Erauskin, G.G., *et al.* (2002) Renal Underexcretion of Uric Acid Is Present in Patients with Apparent High Urinary Uric Acid Output. *Arthritis and Rheumatism*, **47**, 610-613. <https://doi.org/10.1002/art.10792>
- [12] Maiuolo, J., Oppedisano, F., Gratteri, S., *et al.* (2016) Regulation of Uric Acid Metabolism and Excretion. *International Journal of Cardiology*, **213**, 8-14. <https://doi.org/10.1016/j.ijcard.2015.08.109>
- [13] Cortes, J., Moore, J.O., Maziarz, R.T., *et al.* (2010) Control of Plasma Uric Acid in Adults at Risk for Tumor Lysis Syndrome: Efficacy and Safety of Rasburicase Alone and Rasburicase Followed by Allopurinol Compared with Allopurinol Alone—Results of a Multicenter Phase III Study. *Journal of Clinical Oncology*, **28**, 4207-4213. <https://doi.org/10.1200/JCO.2009.26.8896>
- [14] Navolanic, P.M., Pui, C.H., Larson, R.A., *et al.* (2003) Elitek (TM)-Rasburicase: An Effective Means to Prevent and Treat Hyperuricemia Associated with Tumor Lysis Syndrome, a Meeting Report, Dallas, Texas, January 2002. *Leukemia*, **17**, 499-514. <https://doi.org/10.1038/sj.leu.2402847>
- [15] Shannon, J.A. and Cole, S.W. (2012) Pegloticase: A Novel Agent for Treatment-Refractory Gout. *Annals of Pharmacotherapy*, **46**, 368-76. <https://doi.org/10.1345/aph.1Q593>
- [16] Lipsky, P.E., Calabrese, L.H., Kavanaugh, A., *et al.* (2014) Pegloticase Immunogenicity: The Relationship between Efficacy and Antibody Development in Patients Treated for Refractory Chronic Gout. *Arthritis Research & Therapy*, **16**, R60. <https://doi.org/10.1186/ar4497>
- [17] Nelapati, A.K., Das, B.K., Ettiyappan, J.B.P., *et al.* (2020) *In-Silico* Epitope Identification and Design of Uricase Mutein with Reduced Immunogenicity. *Process Biochemistry*, **92**, 288-302. <https://doi.org/10.1016/j.procbio.2020.01.022>
- [18] Tripathi, S., Parmar, J. and Kumar, A. (2020) Structure-Based Immunogenicity Prediction of Uricase from Fungal (*Aspergillus flavus*), Bacterial (*Bacillus subtilis*) and Mammalian Sources Using Immunoinformatic Approach. *Protein Journal*, **39**, 133-144. <https://doi.org/10.1007/s10930-020-09886-0>
- [19] Labrou, N.E., Papageorgiou, A.C. and Avramis, V.I. (2010) Structure-Function Relationships and Clinical Applications of L-Asparaginases. *Current Medicinal Chemistry*, **17**, 2183-2195. <https://doi.org/10.2174/092986710791299920>
- [20] Yen, H.J., Chang, W.H., Liu, H.C., *et al.* (2016) Outcomes Following Discontinuation of *E. coli* L-Asparaginase upon Severe Allergic Reactions in Children with Acute Lymphoblastic Leukemia. *Pediatric Blood & Cancer*, **63**, 665-670. <https://doi.org/10.1002/pbc.25869>
- [21] Ramya, L.N. and Pulicherla, K.K. (2015) Studies on Deimmunization of Antileukaemic L-Asparaginase to Have Reduced Clinical Immunogenicity—An *in Silico* Approach. *Pathology and Oncology Research*, **21**, 909-920. <https://doi.org/10.1007/s12253-015-9912-0>

-
- [22] Mohd Asri, N.A., Ahmad, S., Mohamud, R., *et al.* (2021) Global Prevalence of Nosocomial Multidrug-Resistant *Klebsiella pneumoniae*: A Systematic Review and Meta-Analysis. *Antibiotics (Basel, Switzerland)*, **10**, Article No. 1508. <https://doi.org/10.3390/antibiotics10121508>
- [23] Rostamian, M., Farasat, A., Lorestani, R.C., *et al.* (2022) Immunoinformatics and Molecular Dynamics Studies to Predict T-Cell-Specific Epitopes of Four *Klebsiella pneumoniae* Fimbriae Antigens. *Journal of Biomolecular Structure & Dynamics*, **40**, 166-176. <https://doi.org/10.1080/07391102.2020.1810126>
- [24] Neeffjes, J., Jongsma, M.L.M., Paul, P., *et al.* (2011) Towards a Systems Understanding of MHC Class I and MHC Class II Antigen Presentation. *Nature Reviews Immunology*, **11**, 823-836. <https://doi.org/10.1038/nri3084>