

赵惠民：蛋白质科学版 ChatGPT 来了？

● 本报记者 张晴丹

凝视着手中的试管，又望向桌上那些依旧在运转的大大小小的实验仪器，赵惠民的心像被人用力攥着。他明白，已经到了不得不放弃的时候。读博第三年，学校资格考试的巨大压力如海啸般袭来，这项奋战了两年但依然没走通的课题，被他亲手按下了停止键。

事情虽已过去 30 年，但对美国伊利诺伊大学香槟分校教授赵惠民而言，那时的痛苦仍然记忆犹新。

赵惠民并未因一次羁绊而止步。“科学探索是无止境的，而且有很多的不可预测性，这就是做科研的魅力所在。”23 年间，他带领团队取得累累硕果。

近日，赵惠民团队再次在《科学》发表论文。这次，他们发明了一种新的人工智能 (AI) 工具，可以根据酶的氨基酸序列预测酶的功能。就像 ChatGPT 使用文字数据创建预测文本一样，他们正在利用蛋白质的语言预测它们的功能。

30 多年后又回到原点

20 世纪 80 年代流传着一句顺口溜：“穷清华，富北大，不要命的上科大。”赵惠民就是“不要命的”那一个。1987 年，他以优异成绩从浙江嘉兴一中考入中国科学技术大学（以下简称中国科大）生物系。他对科研的兴趣从这里萌生。

“那时候，中国科大的训练非常严格，学风在高校中名列前茅，五年制教学课程任务重，科研氛围很浓。”赵惠民告诉记者。前 3 年他就掌握了数学、物理、化学、生物等学科知识。

后来，赵惠民成了施蕴渝（1997 年当选中国科学院院士）的学生。彼时，施蕴渝刚从国外进修回来，率先在国内开展蛋白质结构与功能的计算生物学研究。受老师指引，赵惠民开启了分子动力学方面的计算模拟研究。

为了开阔眼界、提升自己，本科毕业后，23 岁的赵惠民在 1992 年到美国加州理工学院读博，第一次接触到了酶工程。从那以后，他便与酶结下了不解之缘。

酶有多重要呢？它是一种生物催化剂，在生命体系中扮演着关键角色，

控制和调节各种代谢途径，保障它们高效工作。

研究酶的人都希望把它的功能改造得更好，得到更高效、更稳定、更优质的催化剂。赵惠民独辟蹊径，想预测酶的功能。

8 年前，他的实验室自主建立了一套全集成自动化系统。6 年前，他开始涉猎 AI 领域，想把 AI 技术与合成生物学、自动化结合起来，以更快地改造酶工程或者代谢工程。在中国科大积累的计算模拟知识重新派上了用场。

“没想到兜兜转转 30 多年后，又回到了原点。”赵惠民感慨说，“可见，很多东西你学的时候觉得没多大用，但说不定将来哪个时候就能发挥价值。”

赵惠民团队在 2014 年搭建了第一台原型机，在 2020 年完成了强化版。一切正式起步。他带领团队用了 3 年时间，研发出被称为 CLEAN 的 AI 工具。论文在线后，通过网络界面，研究人员只需在搜索框中输入序列就可以看到结果，如同使用搜索引擎一样。

“我们不是第一个使用 AI 工具预测酶功能的，但我们是第一个使用这种被称为对比学习的新深度学习算法预测酶功能的。我们发现，这种算法比其他 AI 工具更准确、更可靠、更灵敏。”赵惠民表示。

无论是生物学界、医学界还是工业界，这种工具都将帮助研究人员快速识别合成化学物质和材料所需的最合适的酶。

投顶刊比较顺的一次

2022 年 10 月 8 日，一切准备就绪，赵惠民将论文投给了《科学》。有着丰富投稿经历的他，预感这次不会遭遇审稿人“刁难”，他对这项成果有信心。

果不其然，文章很快通过编辑这一关，进入外审环节。不到两个月，赵惠民就收到了审稿人的意见。

审稿人都给出了积极正面的评价，认为这项研究成果中的 AI 算法非常新颖。最关键的是，很多有关 AI 的文章几乎都是用数据组测试 AI 工具有多好，而赵惠民团队的论文中不仅有计算机的论证，还提供了强有力的

实验数据支持。

计算和体外实验验证显示，该工具不仅可以预测以前未表征的酶的功能，还可以纠正被领先软件错误标记的酶，并正确识别具有两种或多种功能的酶。

当然，研究并非尽善尽美，审稿人还是提出了一些小意见。因为预测并非百分百准确，审稿人希望作者进行更仔细的研究，以回答预测的准确率到底是多少。

赵惠民带领团队紧锣密鼓地进行了两个月补充实验，加入了很多新数据。改进后的文章在 2023 年 1 月底再次送审，不到一个月文章就被接收，从投稿到发表仅耗时 5 个月。

“这是我投顶刊比较顺的一次。”赵惠民说。

在赵惠民看来，论文在《科学》上发表只是迈出第一步。他们计划扩大 CLEAN 背后的 AI，进一步开发机器学习算法，以表征其他蛋白质，如结合蛋白。

“有很多未被表征的结合蛋白，如受体和转录因子，我们也想预测它们的功能。”赵惠民有一个宏大的愿望，他想预测所有蛋白质的功能，这样就可以对细胞中所有蛋白质了如指掌，以便更好地研究或设计整个细胞用于生物技术或生物医学。

为了让这款工具越发强大，该团队还在访问页面专门设置了一个键，用户可以点进去输入反馈意见。“希望大家多提意见，我们会根据反馈再增加一些新的功能。”

在美国做科研一点也不容易

闲暇时，赵惠民常跟学生很走心地分享过往经历，想让他们明白做科研本就不会一帆风顺。

30 年前，他就曾因为考核压力，不得不放弃一项两年都未走通的研究。他不吝与学生分享科研路上曾有过的至暗经历，因为科研本就是一条“光荣



2018 年，在瑞典斯德哥尔摩诺贝尔颁奖典礼后，赵惠民（右）向导师、当期诺奖得主 Arnold 表示祝贺。受访者供图

的荆棘路”，接受失败是必修课。

“常听朋友说在国内做科研很‘卷’，有些人会以为在美国做科研很轻松，从我的亲身经历来说，一点也不容易。”赵惠民表示。

他读博的时候，就有好几名同年进实验室的同学因为顶不住压力，没拿到学位就心灰意冷地离开了。好在他跨过黑暗，选择再度出发，成为了美国加州理工学院教授 Frances H. Arnold 最有成就的弟子之一。

2018 年，Arnold 因为“酶定向进化”上的突出成就获得诺贝尔化学奖。在 Arnold 所获诺贝尔化学奖的科学背景材料中，赵惠民在 5 篇关键论文里都有署名，其中有 4 篇是第一作者。

导师对赵惠民的影响除了科研上的飞跃外，还有做人做事的改变。Arnold 对学生总是直来直去，毫不留情地当面指出缺点。“我们中国人说话比较委婉，当面批评多少有点让人下不了台。所以刚去的时候，我不太适应，她的严厉，有时候连美国学生也吃不消。”

但相处久了，他发现，导师的很多反馈就像酶一样，促使他迅速成长和进步。后来他带学生时，也习惯就事论事、不绕弯子，并且鼓励学生勇于表达自己的意见和想法，最好跟他争个面红耳赤。

现在，他的很多学生已经成长为一些高校、科研院所的中坚力量。其中，33 人成为了教授，在国内做教授的就 15 人。如今，他的团队有近 50 名成员，来自世界多个国家，中国人占了大多数。

相关论文信息：<https://www.science.org/doi/10.1126/science.adf2465>