

ИМПБ РАН 50 лет. Лаборатория кристаллографии макромолекул

Лунин В.Ю.

ИМПБ РАН – филиал ИПМ им. М.В. Келдыша

lunin@impb.ru

Ключевые слова: структурная биология, рентгеновская дифракция, фазовая проблема, уточнение атомных моделей, максимизация правдоподобия.

Основная деятельность лаборатории связана с определением пространственной структуры биологических макромолекул и их комплексов. Возникновение в 1976 году рентгеновской тематики в НИВЦ АН СССР (с 1992 г. ИМПБ РАН) было связано с началом освоения в СССР методов определения атомной структуры биологических макромолекул методом рентгеноструктурного анализа (РСА). Целью первого этапа работы были освоение методики и адаптация к доступным в то время компьютерам сложных программных комплексов, поддерживающих различные этапы расшифровки пространственной структуры белков. Эта работа проводилась в тесном сотрудничестве с группой сотрудников Института Белка АН СССР, возглавляемой Ю.Н. Чиргадзе, и Отделом вычислительной техники НИВЦ и завершилась созданием в Пушчино вычислительной среды, позволяющей осуществлять основные этапы этой работы. В частности, впервые в СССР были введены в эксплуатацию наиболее сложные из используемых в этой области программ – программы уточнения атомной структуры макромолекул и проведено уточнение структур белков и вирусов. В дальнейшем сформировались три основные направления деятельности лаборатории: разработка математических методов кристаллографии макромолекул; создание программных средств для определения структуры методами РСА; участие в проектах по определению структуры биологических макромолекул разных классов. За время существования лаборатории ее сотрудниками опубликовано более 200 работ, из них более 100 статей в научных журналах. Работа лаборатории поддерживалась грантами Российского фонда фундаментальных исследований, Лондонского королевского общества, Американской кристаллографической ассоциации, Министерства образования Франции, программами межкакадемического Российско-Британского и Российско-Французского научного сотрудничества, Национальными центрами научных исследований Франции (CNRS) и Германии (DFG). Общее представление о тематике работы лаборатории в разные годы дают заголовки основных публикаций сотрудников лаборатории [1–18] и приведенные ниже примеры задач, решаемых в лаборатории.

Алгоритм быстрого дифференцирования и его использование в задачах кристаллографии макромолекул. Начало этой тематики в лаборатории датируется началом 1983 г, когда на ежегодной научной конференции НИВЦ АН СССР сотрудник ЦЭМИ АН СССР К.М. Ким сообщил в своем выступлении обескураживающий факт: все компоненты градиента функции любого числа переменных могут быть вычислены за то же время, что и одно значение этой функции. Чтобы оценить значение этого факта для кристаллографии белка, следует учесть, что при работе с макромолекулами часто приходится иметь дело со сложными функциями, зависящими от десятков (и сотен) тысяч переменных. Парадоксально то, что потребность вычислить все десятки тысяч компонент градиента требует, при правильной организации вычислений, того же самого времени, что и расчет одной из его компонент! Адаптация этой идеи к задачам РСА позволило принципиально пересмотреть принципы построения программ минимизации, используемых на различных этапах расшифровки структуры. Интересно отметить, что, при всей неожиданности его существования, этот алгоритм базируется на чрезвычайно простом правиле дифференцирования сложной функции.

Статистическое моделирование в оптимизационных задачах биологической кристаллографии. Заключительным этапом рентгеноструктурного исследования является уточнение предварительной атомной модели исследуемого объекта посредством минимизации расхождения экспериментально измеренных величин с их теоретически рассчитанными по модели значениями. Такой подход является вполне разумным, если предварительная модель объекта полна, т.е. содержит все или почти все атомы, реально присутствующие в кристаллической ячейке. Однако если существенная доля атомов не включена в модель (что является обычным на начальных стадиях работы), то рассчитанные по такой частичной модели теоретические значения могут быть далеки от истинных значений, даже если координаты атомов частичной модели определены без ошибок. В таком случае попытка подгонки значений величин, рассчитываемых по частичной модели, к их экспериментальным аналогам будет вводить атомы частичной модели с правильных позиций. Одним из возможных путей преодоления этой проблемы является вероятностное моделирование источников несовершенства модели. При таком подходе может, например, ставиться задача нахождения такой предварительной модели, которая имеет наивысший шанс быть улучшенной до правильной полной

модели путем случайного добавления в модель недостающих атомов. С формальной точки зрения этот подход есть не что иное, как определение параметров вероятностных распределений путем максимизации их статистического правдоподобия. В обсуждаемом случае, этими параметрами являются координаты атомов предварительной модели. Предложенная схема может быть использована и в других случаях наличия "неустраняемых" ошибок в модели (т.е. ошибок, не устраняемых варьированием параметров модели). Она занимает промежуточное положение между полным игнорированием факта наличия такого рода ошибок и существенным изменением модели путем расширения набора параметров, описывающих модель.

Разработка методики ab-initio решения фазовой проблемы кристаллографии белка, стартуя с низкого разрешения. Рентгеновский дифракционный эксперимент дает значения модулей комплексных коэффициентов Фурье (структурных факторов) функции, описывающей распределение электронов в исследуемом объекте. Восстановление утерянных в эксперименте значений фаз представляет собой центральную. «фазовую» проблему РСА. В основе разработанной методики лежит итерационная процедура Монте-Карловского типа, сопряженная с многоуровневой фильтрацией и специальными процедурами кластерного анализа и усреднения. Первым этапом этой процедуры является случайная генерация большого числа потенциально возможных наборов фаз (вариантов). Сгенерированные варианты образуют стартовую, случайную "популяцию". Дальнейшая задача состоит в "обогащении" стартовой популяции, то есть в повышении процентного содержания в ней "хороших" вариантов. Для этого отобранные варианты фильтруются на основе «слабых» критериев, каждый из которых не позволяет найти правильное решение однозначно. Такое обогащение приводит к тому, что варианты начинают группироваться в окрестности центров, отвечающих решению фазовой проблемы. Были предложены и изучены различные «слабые» критерии отбора: гистограммы электронной плотности, статистическое правдоподобие, связность областей высокой электронной плотности и т.д.

Список литературы

1. Chirgadze Yu.N., Oreshin V.D., Sergeev Yu.V., Nikonov S.V., Lunin, V.Yu. Structure of gamma-crystalline IIIb from calf lens at 5A resolution. *FEBS Letters*. 1980. V. 118. P. 296–298. doi: [10.1016/0014-5793\(80\)80242-0](https://doi.org/10.1016/0014-5793(80)80242-0)
2. Lunin V.Yu., Urzhumtsev A.G. Improvement of Protein Phases by Coarse Model Modification. *Acta Crystallographica A*. 1984. V. 40. P. 269–277. doi: [10.1107/S0108767384000544](https://doi.org/10.1107/S0108767384000544)
3. Lunin V.Yu., Urzhumtsev, A.G. Program construction for macromolecule atomic model refinement based on the fast Fourier transform and fast differentiation algorithms. *Acta Crystallographica A*. 1985. V. 41. P. 327–333. doi: [10.1107/S010876738500071X](https://doi.org/10.1107/S010876738500071X)
4. Urzhumtsev A.G., Lunin V.Yu., Vernoslova E.A. FROG - high-speed restraint-constraint refinement program for macromolecular structure. *J. Applied Crystallography*. 1989. V. 22. P. 500–506. doi: [10.1107/S0021889889004905](https://doi.org/10.1107/S0021889889004905)
5. Lunin V.Yu. Electron-Density Histograms and the Phase Problem. *Acta Crystallographica D*. 1993. V. 49. P. 90–99. doi: [10.1107/S0907444992009247](https://doi.org/10.1107/S0907444992009247)
6. Lunin V.Yu., Skovoroda T.P. R-free Likelihood-Based Estimates of Errors for Phases Calculated from Atomic Models. *Acta Crystallographica A*. 1995. V. 51. P. 880–887. doi: [10.1107/S010876739500688X](https://doi.org/10.1107/S010876739500688X)
7. Shlyapnikov S.V., Lunin V.V., Perbandt M., Polyakov K.M., Lunin V.Yu., Levnikov V.M., Betzel Ch., Mikhailov A.M. Atomic structure of the *Serratia marcescens* endonuclease at 1.1 Å resolution and the enzyme reaction mechanism. *Acta Crystallographica D*. 2000. V. 56. P. 567–572. doi: [10.1107/S090744490000322X](https://doi.org/10.1107/S090744490000322X)
8. Lunin V.Y., Lunina N.L., Petrova T.E., Skovoroda T.P., Urzhumtsev A.G., Podjarny A.D. Low-resolution ab initio phasing: problems and advances. *Acta Crystallographica D*. 2000. V. 56. P. 1223–1232. doi: [10.1107/S0907444900010088](https://doi.org/10.1107/S0907444900010088)
9. Lunin V.Y., Lunina N.L., Ritter S., Frey I., Berg A., Diderichs K., Podjarny A.D., Urzhumtsev A., Baumstark M.W. Low-resolution data analysis for low-density lipoprotein particle. *Acta Crystallographica D*. 2001. V. 57. P. 108–121. doi: [10.1107/S0907444900014608](https://doi.org/10.1107/S0907444900014608)
10. Lunin V.Y., Afonine P.V., Urzhumtsev A.G. Likelihood-based refinement. I. Irremovable model errors. *Acta Crystallographica A*. 2002. V. 58. P. 270–282. doi: [10.1107/S0108767302001046](https://doi.org/10.1107/S0108767302001046)
11. Afonine P.V., Lunin V.Y., Muzet N., Urzhumtsev A. On the possibility of the observation of valence electron density for individual bonds in proteins in conventional difference maps. *Acta Crystallographica D*. 2004. V. 60. P. 260–274. doi: [10.1107/S0907444903026209](https://doi.org/10.1107/S0907444903026209)
12. Petrova T., Ginell S., Mitschler A., Hazemann I., Schneider T., Cousido A., Lunin V.Y., Joachimiak A., Podjarny A. Ultrahigh-resolution study of protein atomic displacement parameters at cryotemperatures obtained with a helium cryostat. *Acta Crystallographica D*. 2006. V. 62. P. 1535–1544. doi: [10.1107/S0907444906041035](https://doi.org/10.1107/S0907444906041035)
13. Petrova T., Lunin V.Y., Ginell S., Hazemann I., Lazarski K., Mitschler A., Podjarny A., Joacjimiak A. X-Ray-Radiation-Induced Cooperative Atomic Movements in Protein. *J. Molecular Biology*. 2009. V. 387. P. 1092–1105. doi: [10.1016/j.jmb.2009.02.030](https://doi.org/10.1016/j.jmb.2009.02.030)

14. Lunin V.Y., Grum-Grzhimailo A.N., Gryzlova E.V., Sinitsyn D.O., Petrova T.E., Lunina N.L., Balabaev N.K., Tereshkina K.B., Stepanov A.S., Krupyanskii Y.F. Efficient calculation of diffracted intensities in the case of non-stationary scattering by biological macromolecules under XFEL pulse. *Acta Crystallographica D*. 2015. V. 71. P. 293–303. doi: [10.1107/S1399004714025450](https://doi.org/10.1107/S1399004714025450)
15. Lunin V.Y., Lunina N.L., Petrova T.E., Baumstark M.W., Urzhumtsev A.G. Mask-based approach to phasing of single-particle diffraction data. *Acta Crystallographica D*. 2016. V. 72. P. 147–157. doi: [10.1107/S2059798315022652](https://doi.org/10.1107/S2059798315022652)
16. Pflüger T., Hernández C.F., Lewe P., Frank F., Mertens H., Svergun D., Baumstark M.W., Lunin V.Y., Jetten M.S.M., Andrade S.L.A. Signaling Ammonium across Membranes through an Ammonium Sensor Histidine Kinase. *Nature Communications*. 2018. V. 9. Article No. 164. doi: [10.1038/s41467-017-02637-3](https://doi.org/10.1038/s41467-017-02637-3)
17. Urzhumtsev A.G., Lunin V.Y. Introduction to crystallographic refinement of macromolecular atomic models. *Crystallography Reviews*. 2019. V. 25. P. 164–262. doi: [10.1080/0889311X.2019.1631817](https://doi.org/10.1080/0889311X.2019.1631817)
18. Urzhumtsev A., Lunin, V.Y. Analytic modeling of inhomogeneous-resolution maps in cryo electron microscopy and crystallography. *IUCrJ*. 2022. V. 9. doi: [10.1107/S2052252522008260](https://doi.org/10.1107/S2052252522008260)