

DOI: <https://doi.org/10.60797/BIO.2025.8.3>**СОЗДАНИЕ ПРОГРАММНОГО СРЕДСТВА ДЛЯ МОДЕЛИРОВАНИЯ ПОСТТРАНСЛЯЦИОННЫХ МОДИФИКАЦИЙ БЕЛКОВ**

Научная статья

Мискевич О.Е.^{1,*}, Хрусталёв В.В.²¹ORCID : 0009-0003-6465-8039;²ORCID : 0000-0003-1373-3329;^{1,2} Белорусский государственный университет, Минск, Беларусь

* Корреспондирующий автор (olga[at]arusnavi.ru)

Аннотация

Статья посвящена представлению результатов разработки и испытаний программного средства EasyAcetoPhos для добавления посттрансляционных модификаций в трёхмерные структуры белков, а также принципов его работы и примеры применения на реальных данных. Цель исследования — создать оригинальное программное средство для моделирования посттрансляционных модификаций белков на основе геометрических и молекулярно-механических методов. Актуальность обусловлена отсутствием легковесных некоммерческих решений, способных автоматически в интерактивном режиме достраивать фосфатные и ацетильные группы непосредственно в трёхмерных координатах PDB; устойчивость таких модификаций затем проверяется отдельно методом молекулярной динамики.

Реализация включает в себя серверную часть на Python/FastAPI для загрузки и обработки PDB, геометрические правила построения функциональных групп с контролем длин связей, углов и межатомных расстояний, а также веб-интерфейс на базе Mol* с интерактивным выбором аминокислотного остатка из последовательности или напрямую на 3D-модели. На данном этапе реализованы два класса модификаций: фосфорилирование серина, треонина и тирозина с автоматическим формированием тетраэдрической окрестности фосфора и переименованием остатков в стандартизированные формы, а также ацетилирование лизина с конверсией в ALY; во всех случаях выполняется проверка геометрии и блокировка операции при недопустимости достройки модификации. Корректность достраиваемых модификаций оценивалась методами молекулярной динамики в GROMACS на выборке из десяти белков для фосфорилирования и пяти белков для ацетилирования. Анализ продемонстрировал стабильность локальной геометрии модифицированных участков и отсутствие систематических стерических конфликтов.

Полученные результаты показывают корректность геометрического построения, подтверждают применимость предложенного подхода для быстрой подготовки структур к моделированию и анализу, а архитектура программного средства обеспечивает высокую скорость достройки модификаций, а также возможность расширения на новые типы модификаций. Показан воспроизводимый рабочий процесс от интерактивного выбора остатка до получения геометрически согласованной модели, пригодной для дальнейшего вычислительного анализа. В перспективе планируется расширение набора поддерживаемых посттрансляционных модификаций.

Ключевые слова: посттрансляционные модификации, фосфорилирование, ацетилирование, молекулярная динамика, моделирование белков.

DEVELOPMENT OF A SOFTWARE TOOL FOR MODELLING POST-TRANSLATIONAL MODIFICATIONS OF PROTEINS

Research article

Miskevich O.E.^{1,*}, Khrustalev V.V.²¹ORCID : 0009-0003-6465-8039;²ORCID : 0000-0003-1373-3329;^{1,2} Belarusian State University, Minsk, Belarus

* Corresponding author (olga[at]arusnavi.ru)

Abstract

The article presents the results of the development and testing of the EasyAcetoPhos software tool for adding post-translational modifications to three-dimensional protein structures, as well as its operating principles and examples of its application to real data. The aim of the study is to create an original software tool for modelling post-translational modifications of proteins based on geometric and molecular-mechanical methods. The relevance of this research is due to the lack of lightweight non-profit solutions capable of automatically adding phosphate and acetyl groups directly in PDB three-dimensional coordinates in interactive mode; the stability of such modifications is then verified separately using molecular dynamics.

The implementation includes a server component in Python/FastAPI for loading and processing PDB files, geometric rules for constructing functional groups with control of bond lengths, angles, and interatomic distances, as well as a Mol*-based web interface with interactive selection of amino acid residues from a sequence or directly on a 3D model. At this stage, two classes of modifications have been performed: phosphorylation of serine, threonine, and tyrosine with automatic formation of the tetrahedral vicinity of phosphorus and renaming of residues to standardised forms, as well as acetylation of lysine with conversion to ALY; In all cases, geometry is checked and the operation is blocked if the modification cannot be completed. The correctness of the completed modifications was evaluated using molecular dynamics methods in GROMACS on a sample of

ten proteins for phosphorylation and five proteins for acetylation. The analysis demonstrated the stability of the local geometry of the modified sites and the absence of systematic steric conflicts.

The obtained results demonstrate the correctness of the geometric construction, confirm the applicability of the suggested approach for rapid preparation of structures for modelling and analysis, and the architecture of the software tool ensures high speed of modification completion, as well as the possibility of extension to new types of modifications. A reproducible workflow is demonstrated, from interactive selection of the remainder to obtaining a geometrically consistent model suitable for further computational analysis. In the future, an expansion of the set of supported post-translation modifications is planned.

Keywords: post-translational modifications, phosphorylation, acetylation, molecular dynamics, protein modelling.

Введение

Посттрансляционные модификации (РТМ) являются фундаментальным механизмом регуляции белков: они изменяют заряд и полярность боковых цепей, влияют на локальную и глобальную конформацию, устойчивость, взаимодействия и внутриклеточную локализацию, тем самым формируя функциональное разнообразие протеома [1]. В сигнальных путях важную роль играет фосфорилирование остатков серина, треонина и тирозина: добавление или снятие фосфатной группы выступает регулятором, который быстро меняет активность белков и маршруты передачи сигнала [2]. Ацетилирование ε-аминогрупп лизина влияет не только на хроматин и транскрипцию, но и на широкий спектр негистоновых белков, изменяя их устойчивость, комплексообразование и локализацию; при этом процессы ацетилирования и деацетилирования динамичны и обратимы [3]. Масштаб явления подтверждён современными данными: на сегодняшний день описаны сотни типов модификаций, их перечень продолжает расти [1]; РТМ затрагивают почти все стороны клеточной биологии. Для вычислительного моделирования это означает необходимость задавать модификации с корректной локальной трёхмерной геометрией: стереохимические ошибки и тесные межатомные контакты на старте искажают энергетику и динамику системы и приводят к артефактам в молекулярной динамике [4], поэтому качество исходной структуры критично для надёжных расчётов молекулярной динамики [5]. В настоящее время в открытом доступе практически отсутствуют легковесные некоммерческие программные средства, которые автоматически добавляют геометрически корректные РТМ в координаты PDB и позволяют интерактивно выбирать целевой остаток с явной проверкой применимости операции. В данной работе описан инструмент, реализующий эти возможности для двух востребованных модификаций — фосфорилирования остатков серина, треонина и тирозина и ацетилирования лизина; изложены используемые геометрические правила, архитектура программной реализации и результаты проверки на реальных структурах.

Доступные бесплатные решения по работе со структурой белков в первую очередь ориентированы на визуализацию и ручное редактирование. Сходные описываемому программному средству задачи частично решает PyTMs — плагин к настольному приложению PyMOL, предназначенный для удобного введения распространённых посттрансляционных модификаций прямо в окне визуализатора; авторы подчёркивают, что это простой исследовательский инструмент с удобным интерфейсом внутри среды просмотра [6]. Веб-сервис Vienna-PTM решает также схожую, но специализированную на молекулярной динамике задачу: он автоматизирует внесение широкого набора модификаций и сразу выполняет минимизацию энергии, выдавая параметры силовых полей и готовые входные файлы для запуска молекулярной динамики [7]. По сути, это инструмент подготовки конвейера молекулярной динамики, а не интерактивной геометрической сборки в текущей структуре с контекстными проверками вносимых модификаций. Инфраструктура CHARMM-GUI, включая модуль PDB Manipulator, встроена в большой онлайн-конвейер подготовки систем к расчётам и предоставляет множество различных вариантов вносимых модификаций именно как исходный шаг для генерации входных данных для других модулей [8]. Это мощная среда для комплексной подготовки, однако её логика нацелена на сборку полной системы для симуляций, а не на точечную вставку групп с немедленной проверкой геометрии и запретом недопустимых операций. Характерно, что существующие аналоги в значительной степени опираются на библиотечные фрагменты/патчи и дорабатывают геометрию через минимизацию (например, Vienna-PTM использует заранее энерго-минимизированные модифицированные остатки, внедряемые в цельную структуру) [7]. В отличие от вышеуказанных решений, описываемое в данной работе программное средство, реализованное в виде веб-приложения, сосредоточено на другом звене рабочего процесса: интерактивной, воспроизводимой геометрической сборке непосредственно в координатах PDB. Модификации вносятся интерактивно с возможностью выбора целевого остатка как из линейной последовательности, так и непосредственно на просматриваемой трёхмерной структуре исследуемого белка; перед применением модификации выполняются явные проверки её применимости с запретом операции при недопустимости; сама группа строится на основе методов параметрической геометрии по заданным длинам связей и углам с контролем минимальных межатомных расстояний, а результат автоматически стандартизуется по номенклатуре остатков. Такой формат работы снижает операторские ошибки и обеспечивает готовую, геометрически согласованную структуру, которую затем можно без привязки к конкретному программному обеспечению передать в любой конвейер молекулярной динамики.

Цель исследования — создать оригинальное программное средство для моделирования посттрансляционных модификаций белков на основе геометрических и молекулярно-механических методов. Дополнительно цель включает проверку тезиса о том, что при соблюдении заданных геометрических допусков (длины связей, углы, минимальные межатомные расстояния) получаемая структура может рассматриваться как готовая к дальнейшему вычислительному анализу без обязательного запуска молекулярной динамики. Для достижения цели решались следующие задачи: разработать алгоритм геометрической достройки фосфатной группы к остаткам серина, треонина и тирозина; разработать алгоритм геометрической достройки ацетильной группы для остатков лизина; предусмотреть расширение списка внедряемых посттрансляционных модификаций; верифицировать корректность получаемых структур методом молекулярной динамики; сформулировать и зафиксировать критерии геометрической приёмки (целевые значения по длинам и углам, пороги по минимальным межатомным расстояниям) и показать, что при выполнении этих критериев

результат может использоваться как готовая входная структура без обязательной молекулярно-динамической валидации; в качестве дальнейшей прикладной задачи — расширение перечня поддерживаемых посттрансляционных модификаций с сохранением единых геометрических принципов и автоматической проверки применимости.

Описание программного средства

Описываемое программное средство реализовано в виде веб-приложения EasyAcetoPhos и предназначено для точечного добавления фосфатных и ацетильных групп непосредственно в координаты PDB с немедленной проверкой локальной геометрии и понятной обратной связью в браузере. Выбор целевого аминокислотного остатка осуществляется кликом по участку трёхмерной структуры белка, отображаемой в окне просмотра, так и по символу в отображаемой на отдельной панели линейной последовательности; после выбора система оповещает пользователя, допустима ли операция для данного типа остатка и текущего состояния структуры [9]. Выбор вида реализации программного средства как веб-приложения обусловлен доступностью для пользователя без установки, кроссплатформенностью и возможностью интерактивного взаимодействия с программным средством, что снижает порог входа и упрощает пользование инструментом. Результатом работы является PDB-файл с внедрённой модификацией в выбранной позиции; файл пригоден для дальнейшей подготовки и расчётов в любом программном обеспечении для молекулярной динамики.

Сценарий использования

Пользователь загружает PDB-файл в веб-интерфейс, после чего ему становится доступно отображение трёхмерной структуры анализируемого белка. С этого момента целевой аминокислотный остаток можно выбрать либо кликом по интересующему участку трёхмерной структуры, либо через панель линейной аминокислотной последовательности. После выбора целевого остатка запускается автоматическая проверка применимости разных типов модификаций на основе типа остатка, наличия нужного атома присоединения, отсутствия уже внесённой модификации и минимальных межатомных расстояний в локальном окружении. Если условия нарушены, действие блокируется, что исключает ошибочные операции на раннем этапе. При прохождении условий проверки пользователю становится доступен выбор доступной модификации. При выборе применения модификации пользователем, модификация строится параметрически непосредственно в текущих координатах, затем структура обновляется в окне просмотра, пользователю становится доступна выгрузка результирующего PDB-файла. Полученный файл может быть использован в качестве исходных координат для последующей подготовки и расчётов молекулярной динамики в выбранном пользователем программном пакете.

Общие геометрические принципы

Построение новых атомов выполняется на основе параметрической геометрии без внешних шаблонов. Для выбранного аминокислотного остатка определяется «якорный» атом присоединения: OG для серина, OG1 для треонина, OH для тирозина; для ацетилирования — NZ у лизина. Из координат исходного остатка формируется локальный правый ортонормированный базис: ось z задаётся направлением от атома $C\alpha$ выбранного остатка к якорному атому. Направление оси x получается из вектора от $C\alpha$ к $C\beta$ с предварительной ортогонализацией к z и нормируется до единичной длины. Ось y определяется как направление, перпендикулярное одновременно и к x , и к z ; выбор выполняется по «правилу правой руки», что обеспечивает правостороннюю и взаимно ортогональную тройку осей. Если координаты $C\beta$ недоступны или ненадёжны (например, при альтернативных положениях атомов), в качестве второго опорного направления выбирается ближайший тяжёлый атом боковой цепи и проводится ортогонализация к z . Построение в таком локальном базисе делает дальнейшую расстановку новых атомов детерминированной и инвариантной к ориентации всей молекулы в глобальных координатах.

При фосфорилировании точкой присоединения служит гидроксил соответствующего остатка; перед вставкой удаляется связанный с ним водород. Атом фосфора размещается на фиксированном расстоянии от якорного атома кислорода вдоль локальной оси z ; три кислорода фосфатной группы расставляются так, чтобы окружение атома фосфора было близко к тетраэдрическому с углами по 109.5° с равномерным разведением по азимуту. В используемых параметрах длина мостиковой связи O-P к боковой цепи берётся ~ 1.60 – 1.62 Å, длины связей P-O в самой группе — ~ 1.48 – 1.52 Å; такая схема согласуется с классической тетраэдрической координацией P(V) и укорочением терминальных P=O по сравнению с одинарными P-O мостикового типа [10], [11]. После расстановки атомов остаток переименовывается в SEP/TPO/PTP и выполняется контроль минимальных расстояний. При нарушении порогов операция отклоняется.

При ацетилировании точкой присоединения является атом NZ боковой цепи лизина. Перед вставкой удаляется один из его водородов. Формируется амидная группа $-C(=O)-CH_3$, сохраняющая планарность и sp^2 -характер: углы вокруг C-N и C=O близки к 120° , диэдральные параметры выбираются в транс-ориентации относительно карбонила. Для длин связей используются типовые значения: для пептидной (амидной) связи C-N ≈ 1.33 – 1.35 Å и C=O ≈ 1.22 – 1.24 Å; эти параметры согласованы с общеупотребительными ограничениями геометрии, применяемыми при уточнении белковых структур, и с классическим представлением о частично двойственной связи C-N как причине планарности пептидной группы [12], [13]. После построения остаток получает имя ALY и проходит ту же проверку минимальных расстояний.

Контроль применимости модификации и локальной стереохимии выполняется до полноценного внесения изменений. Проверка корректности выполняется последовательно: сначала оценивается применимость операции (тип остатка, наличие якорного атома, отсутствие ранее внесённой модификации), затем — геометрия рассчитанного фрагмента и, наконец, проверка окружения. Геометрические цели формулируются как числовые «точки притяжения» для длин связей и валентных углов; допустимые отклонения заданы узкими допусками, согласованными с практикой рестрейнов при уточнении белковых структур: для длин — порядка ± 0.08 Å, для углов — порядка $\pm 5^\circ$. Если отклонения попадают в расширенные, но ещё допустимые коридоры (например, до ± 0.12 Å по длинам и до $\pm 8^\circ$ по углам),

операция помечается как требующая внимания, иначе — блокируется. Стерические конфликты оцениваются по all-atom-подходу: фиксируются перекрытия между несвязанными атомами с учётами водородов, при превышении порога, принятого в валидации (перекрытие ≥ 0.4 Å), модификация не применяется [12], [14]. Такой порядок позволяет отсеивать случаи с заведомо некорректной локальной геометрией до сохранения результирующего файла, что снижает риск переноса стереохимических ошибок в последующие расчёты молекулярной динамики. Итоговый статус модификации (“PASS”, “Warning” или “Blocked”) определяется по вышеуказанным критериям и отображается пользователю вместе с фактическими длинами, углами и минимальной межатомной дистанцией. Статус “PASS” трактуется как условие достаточности: структура считается готовой к дальнейшим расчётам и анализу без обязательной динамической валидации, статус “Warning” указывает на желательность дополнительной проверки, включая по мере необходимости проверку методом молекулярной динамики, статус “Blocked” исключает применение модификации до устранения причины.

Реализация интерфейса и сервера

Программное средство разделено на серверную часть и интерфейс, что позволяет отделить вычислительную логику от визуализации и обеспечить детерминированность работы при повторных вызовах. Серверная часть реализована в виде REST API с помощью языка программирования Python и фреймворка FastAPI. Разработаны конечные точки, отвечающие соответственно за проверку применимости операции к выбранному остатку, применение модификации с вычислением новых координат, выдачу численных характеристик для интерфейса, экспорт полученной структуры. Выбор FastAPI обусловлен сочетанием высокой скорости и типобезопасной схемы описания маршрутов, что упрощает поддержку и документирование API [14]. Обработка входных PDB-файлов и формирование результата выполняются с опорой на объектную модель пакета Biopython: парсер преобразует файл в иерархию Structure — Model — Chain — Residue — Atom, что гарантирует корректный доступ к атомам, остаткам и нумерации при вставке группы, а запись результата осуществляется теми же средствами обратно в PDB с сохранением сопоставимости с последующей подготовкой к расчётам [15], [16].

Пользовательский интерфейс работает в браузере на основе плагина Mol*, поэтому выбор остатка, к которому будет применена модификация, возможен как по клику на трёхмерной структуре белка, так и через панель линейной последовательности; оба способа синхронизованы и приводят к одному и тому же серверному запросу. После ответа сервера интерфейс сразу подсвечивает выбранный участок, сообщает статус применимости операции и активирует или блокирует соответствующие кнопки. По завершении операции интерфейс автоматически перерисовывает отображение трёхмерной структуры белка с уже модифицированным остатком и предлагает скачать обновлённый PDB-файл. Такой сценарий опирается на стандартные механизмы выбора и уровня детализации в Mol*, делая взаимодействие предсказуемым и воспроизводимым [9], [17]. На рисунках 1 и 2 отображён интерфейс реализованного программного средства.

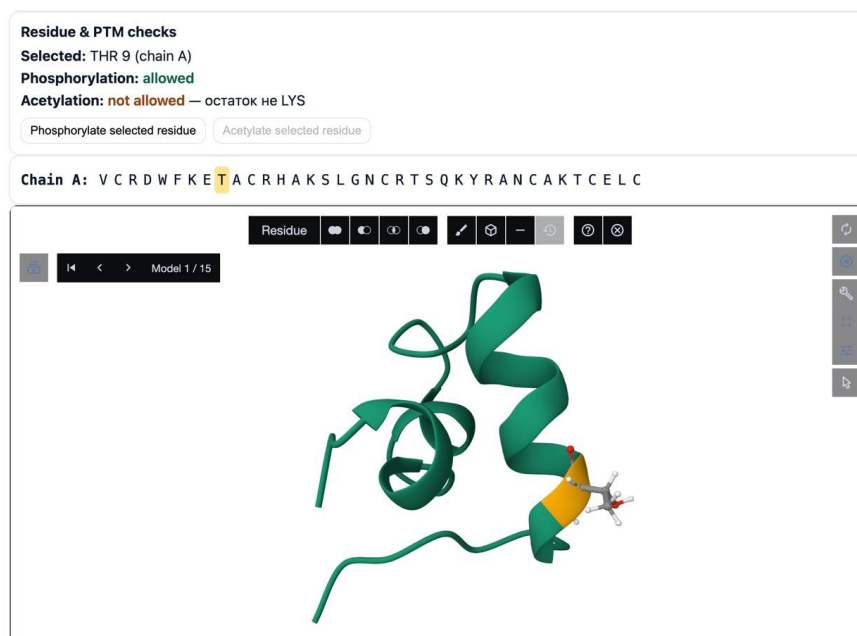


Рисунок 1 - Интерфейс программного средства. Белок без модификации
DOI: <https://doi.org/10.60797/BIO.2025.8.3.1>

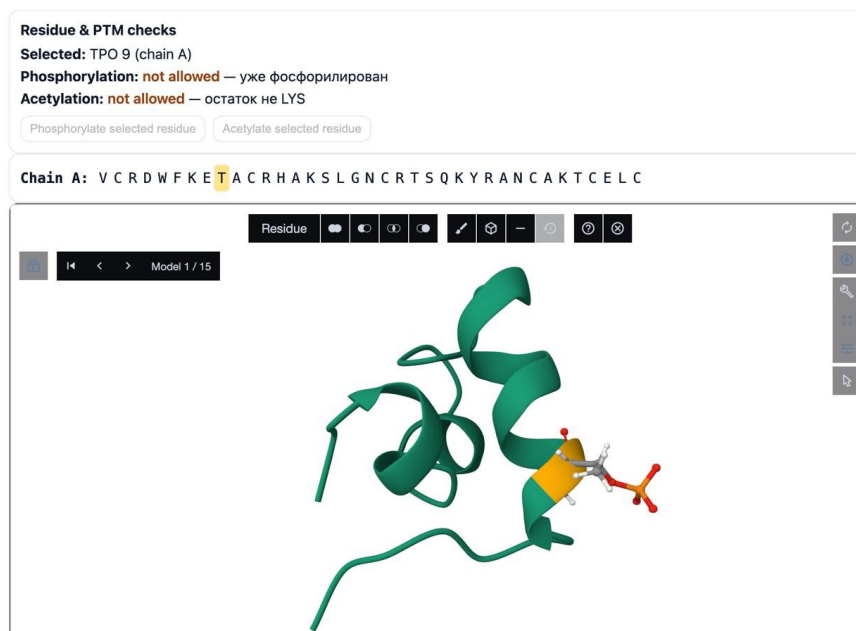


Рисунок 2 - Интерфейс программного средства. Белок с применённым к одному из остатков фосфорилированием
DOI: <https://doi.org/10.60797/BIO.2025.8.3.2>

В совокупности сервер и интерфейс образуют простой и надежный инструмент, в котором параметры модификации рассчитываются на сервере на основе геометрических правил и пороговых значений, а браузерный интерфейс обеспечивает прозрачную обратную связь и контроль корректности на каждом шаге.

Данные и файлы

Входными данными для программного средства являются файлы в формате PDB, содержащие в себе исходные структуры анализируемых белков. При получении на вход от пользователя PDB файла, имя файла нормализуется и проверяется по расширению, после чего сохраняется в рабочем каталоге сервера. Дальнейшая обработка выполняется через стандартную объектную модель пакета Biopython с парсингом файла в соответствующую иерархию, содержащую информацию о структуре, модели, цепях, остатках и атомах. Результат применяемой модификации сохраняется в формате PDB с теми же идентификаторами цепей и исходной нумерацией, изменяется только трёхбуквенный код модифицированного остатка на стандартизованное имя: фосфорилированные остатки серина, треонина и тирозина получают имена SEP, TPO и PTR соответственно, ацетилированный остаток лизина получает имя ALY. Для прозрачного сопоставления с исходными данными используется унифицированная схема имён файлов: к базовому имени исходного файла PDB добавляется код цепи, номер остатка и тэг модификации через двойное подчёркивание, например 1ESX__A15_PTR.pdb. Такая организация облегчает навигацию по результатам.

Внешняя проверка методом молекулярной динамики

Валидация геометрии модифицированных структур выполнялась вне разработанного программного средства, в отдельном расчётном цикле молекулярной динамики в пакете GROMACS с использованием GPU на сервере биологического факультета Белорусского государственного университета. Подготовка систем включала построение топологии в силовом поле CHARMM36 с моделью воды TIP3P, сольватацию и ионизацию до физиологической концентрации, а также стандартные этапы энергетической минимизации, уравнивания при постоянной температуре и давлении и производственного расчёта динамики на 50 наносекунд. Для расчёта дальнедействующих кулоновских взаимодействий применялся метод частично-сеточного Эвальда с гладкой интерполяцией зарядов (smooth PME), являющийся принятым стандартом для биомолекулярных систем в GROMACS [18]. Аппаратно использовалась рекомендованная для видеокарт NVIDIA CUDA-акселерация, обеспечивающая предсказуемое ускорение расчётов без изменения физики модели [19]. Настройки отсечений и форма переключателя ван-дер-Ваальсовского взаимодействия подбирались в соответствии с рекомендациями для сочетания CHARMM36 и TIP3P для обеспечения сопоставимости осуществляемых серий расчётов с опубликованными протоколами [20]. Тестовый набор включал 25 систем: по пять белков с точечным фосфорилированием остатков серина, треонина и тирозина соответственно и десять белков с ацетилированием остатка лизина (табл. 1).

Таблица 1 - Тестовый набор

DOI: <https://doi.org/10.60797/BIO.2025.8.3.3>

Идентификатор системы	Название системы	Название цепи	Номер модифицируемого остатка	Тип модификации
1	1AF8	A	6	TPO

Идентификатор системы	Название системы	Название цепи	Номер модифицируемого остатка	Тип модификации
2	1ERY	A	18	PTR
3	1ACW	A	11	SEP
4	1EQ0	A	4	PTR
5	1EO1	A	7	SEP
6	1EMW	A	32	PTR
7	1EIW	A	10	TPO
8	1EIK	A	41	TPO
9	1EGO	A	33	PTR
10	1ADR	A	36	SEP
11	1E2B	A	16	TPO
12	1E0H	A	38	TPO
13	1DX7	A	21	SEP
14	1DS9	A	121	PTR
15	1DV5	A	20	SEP
16	1A3S	A	14	ALY
17	1A19	A	60	ALY
18	1A6C	A	17	ALY
19	1A5R	A	23	ALY
20	1EGL	A	16	ALY
21	1A03	A	18	ALY
22	1AB3	A	18	ALY
23	1ADR	A	26	ALY
24	1B1A	A	98	ALY
25	1B6F	A	80	ALY

Анализ траекторий был направлен на два аспекта: устойчивость локальной конформации в зоне модификации и сохранность заданной при вставке геометрии самой модифицирующей группы. Сходимость конформации контролировалась по среднеквадратичному отклонению координат (RMSD) относительно опорной структуры, а локальная подвижность — по среднеквадратичным флуктуациям (RMSF) атомов выбранного фрагмента [21], [22]. Сохранность целевых параметров внутри самой модификации отслеживалась по временным рядам длин связей и валентных/диздральных углов, что позволяло убедиться, что значения остаются в диапазонах, заданных геометрическим алгоритмом на этапе вставки [23], [24]. Водородное окружение модификаций описывалось числом и стабильностью водородных связей, рассчитанных по геометрическому критерию, принятому в GROMACS (порог по расстоянию порядка 0.35 нм и по углу порядка 30°), обеспечивая сопоставимость систем по единой шкале [25]. Для интерпретации пространственных конфликтов использовалась доля кадров без тесных межатомных контактов в окрестности модификации; существенные перекрытия между несвязанными атомами (порядка 0.4 Å и более) квалифицировались как неприемлемые и исключались в соответствии с общепринятой атомарной валидацией [5]. Критерии успешности определялись как совместное выполнение трёх условий: стабилизация RMSD модифицированного фрагмента после уравнивания, удержание средних длин и углов внутри допусков, и высокая доля кадров без тесных контактов.

Результаты

В анализ включены 10 систем с ацетилированием лизина и 15 систем с фосфорилированием серина, треонина и тирозина, производственный участок 50 наносекунд с отбрасыванием первых 10 наносекунд как периода уравнивания. Сайты, на которые вносились модификации, отбирались случайно из допустимых позиций, для репрезентативности локальной геометрии соблюдался баланс по типам вторичной структуры для охвата разнообразных локальных окружений. Конкретные параметры отдельных запусков приведены в сопроводительных таблицах 2–4. Целевые термодинамические условия соответствовали стандартной постановке в водной среде TIP3P под управлением термостатирования и баростатирования после предварительной энергетической минимизации и коротких этапов NVT/NPT. По сводным метрикам рассчитанных серий средняя температура оставалась близкой к цели (около 300 К, диапазон от 299.981 до 300.073 К), мгновенные колебания давления соответствуют ожидаемой статистике для малых систем, а плотность системы находилась в узком интервале от 1018 до 1049 кг/м³, что согласуется с выбранной моделью растворителя и подтверждает корректность постановки расчётов (табл. 5).

Таблица 2 - Профиль подвижности и водородного окружения

DOI: <https://doi.org/10.60797/BIO.2025.8.3.4>

Идентификатор системы	RMSD	RMSF	Н-связи с белком, ср. шт./кадр	Н-связи с водой, ср. шт./кадр	SASA (поверхн.), нм ²	Доля кадров без перекрытий, %
1	0.355755	0.0502	2.12047	7.15046	3.36549	100
2	0.439966	0.07311	1.01974	8.49671	4.32449	100
3	0.314584	0.08388	2.00975	8.28468	3.12597	100
4	0.190971	0.09548	2.53723	6.61369	4.31884	100
5	0.185112	0.01767	5.27286	2.34283	3.22822	100
6	0.215961	0.07663	0.813093	7.15192	4.23804	100
7	0.11004	0.05442	2.33783	5.2024	3.35234	100
8	0.432039	0.054591	0.715642	7.42379	3.35729	100
9	0.182194	0.084306	1.05897	8.09295	4.23467	100
10	0.167142	0.06489	4.36232	5.81509	3.12842	100
11	0.125071	0.06774	4.11244	4.15642	3.3639	100
12	0.159583	0.06005	2.75212	7.0035	3.29648	100
13	0.566457	0.08932	2.44978	8.04948	3.13391	100
14	0.515789	0.0743	2.55372	6.88956	4.28749	100
15	0.23091	0.07029	2.1919	6.84958	3.13795	100
16	0.38854	0.09726	1.89805	2.47176	3.97117	100
17	0.67396	0.07659	4.54173	3.65317	4.5002	100
18	0.280092	0.07364	0.413793	3.93553	3.99904	100
19	0.691581	0.0701	0.122493	3.43078	3.58224	100
20	0.171312	0.08287	1.88056	2.11294	4.05843	100
21	1.32952	0.36477	2.50925	5.67166	4.88342	100
22	0.255238	0.01694	0.31472	1.57471	2.65839	100
23	0.434478	0.1019	0.70065	3.46527	4.02019	100
24	0.457804	0.0588	1.37581	2.78811	3.9764	100
25	0.162076	0.0546	1.77911	2.48926	4.06406	100

Таблица 3 - Локальная геометрия фосфата

DOI: <https://doi.org/10.60797/BIO.2025.8.3.5>

Идентификатор системы	Длина связи Р-О, нм	Длина связи Р-О ₁ , нм	Длина связи Р-О ₂ , нм	Длина связи Р-О ₃ , нм	Угол О ₁ -Р-О ₂ , °	Угол О ₁ -Р-О ₃ , °	Угол О ₂ -Р-О ₃ , °	Диэдрал О ₁ -Р-О ₂ -О ₃ , °
1	0.156203	0.147563	0.147433	0.154443	118.653	105.044	104.31	116.508
2	0.160439	0.147637	0.147748	0.155778	119.634	106.549	106.391	120.758
3	0.156558	0.147488	0.147483	0.154787	118.557	104.744	104.677	116.365
4	0.160273	0.147641	0.147647	0.156059	119.565	106.687	106.1	120.868
5	0.156277	0.147769	0.14733	0.154822	118.097	103.849	105.237	115.391
6	0.160403	0.147748	0.147722	0.155746	119.764	106.449	106.449	120.747
7	0.156095	0.14737	0.147299	0.155115	119.227	103.667	104.347	115.119
8	0.156156	0.147452	0.14748	0.15483	118.067	104.011	104.367	114.926
9	0.160202	0.147765	0.147669	0.155752	119.509	106.275	106.416	120.304
10	0.156431	0.147423	0.147309	0.15473	118.669	104.755	104.605	116.385
11	0.155731	0.147466	0.147343	0.154821	118.468	104.519	104.237	115.702
12	0.155795	0.147468	0.147532	0.154818	118.335	104.788	105.043	116.553
13	0.156536	0.147507	0.147378	0.154682	118.429	104.767	104.737	116.365

Идентификатор системы	Длина связи P-O, нм	Длина связи P-O ₁ , нм	Длина связи P-O ₂ , нм	Длина связи P-O ₃ , нм	Угол O ₁ -P-O ₂ , °	Угол O ₁ -P-O ₃ , °	Угол O ₂ -P-O ₃ , °	Диэдрал O ₁ -P-O ₂ -O ₃ , °
14	0.160379	0.147653	0.147672	0.15581	119.782	106.414	106.152	120.46
15	0.156546	0.147471	0.147432	0.15469	118.499	104.875	105.243	116.915

Таблица 4 - Локальная геометрия ацетильной группы

DOI: <https://doi.org/10.60797/BIO.2025.8.3.6>

Идентификатор системы	Длина связи N-C _{карб.} , нм	Длина связи C _{карб.} =O, нм	Длина связи C _{карб.} -C _{метил.} , нм	Угол N-C _{карб.} -O, °	Угол N-C _{карб.} -C _{метил.} , °	Угол O-C _{карб.} -C _{метил.} , °	Диэдрал O-C _{карб.} -N-C _{бок.} , °
1	0.133784	0.122647	0.148359	120.442	119.231	120.128	179.873
2	0.133813	0.122556	0.148245	120.457	119.049	120.304	179.535
3	0.133719	0.122542	0.148366	120.318	119.048	120.443	179.88
4	0.133801	0.122385	0.148359	120.176	119.268	120.352	179.399
5	0.133771	0.122497	0.148281	120.428	119.136	120.247	179.343
6	0.133879	0.122515	0.148342	120.357	119.085	120.366	179.87
7	0.133751	0.122671	0.148572	120.271	119.125	120.401	179.942
8	0.133884	0.122624	0.148198	120.347	119.078	120.38	179.204
9	0.133755	0.122516	0.148173	120.347	118.931	120.527	178.991
10	0.133755	0.122579	0.148484	120.463	119.105	120.224	179.59

Таблица 5 - Термодинамические параметры систем

DOI: <https://doi.org/10.60797/BIO.2025.8.3.7>

Идентификатор системы	Температура, К	Давление, бар	Плотность, кг/м ³
1	300.047	1.438	1034.88
2	299.992	2.921	1049.13
3	300.073	1.994	1037.1
4	299.997	1.517	1046.21
5	300.002	1.483	1034.96
6	299.996	1.489	1036.5
7	299.981	1.729	1045.2
8	300.025	1.765	1030.99
9	299.987	2.106	1042.54
10	300.022	2.028	1035.1
11	300.018	2.35	1037.48
12	299.955	2.139	1045.06
13	300.014	1.355	1018.3
14	299.998	1.216	1032.77
15	300.013	1.752	1048.97
16	299.995	1.497	1039.89
17	299.996	1.422	1036.17
18	300.001	1.149	1032.05
19	300.001	1.209	1022.76
20	299.992	1.976	1033.81
21	299.997	1.451	1039.99
22	299.998	1.77	1031.3
23	299.99	1.879	1033.62
24	300.001	1.785	1042.12

Идентификатор системы	Температура, К	Давление, бар	Плотность, кг/м ³
25	299.996	1.19	1030.01

Для оценки сходимости и подвижности использовались две стандартные метрики: среднеквадратичное отклонение координат (RMSD) и среднеквадратичная флуктуация (RMSF). RMSD показывает, насколько в среднем атомы выбранного фрагмента отклоняются от опорной конфигурации: чем меньше значение RMSD, тем стабильнее форма фрагмента относительно исходной. RMSF характеризует амплитуду колебаний каждого атома вокруг его среднего положения вдоль траектории, низкие значения означают умеренную локальную подвижность. Для оценки выхода на равновесие ориентируются на достижение плато по ряду величин, в том числе энергии и RMSD, именно выход на плато по RMSD служит практическим признаком завершения уравнивания; единых норм для абсолютных значений RMSD/RMSF не существует: интерпретация контекстна, а ключевым критерием служит стабилизация профиля RMSD и умеренные RMSF в рассматриваемом фрагменте [26]. После отбрасывания первых 10 нс уравнивания на оставшемся участке в 40 нс по всем 25 системам средний RMSD составил 0.361 нм (медиана 0.280 нм; диапазон 0.110–1.330 нм), средний RMSF составил 0.0806 нм (медиана 0.0731 нм; диапазон 0.0169–0.3648 нм). Если рассматривать классы модификаций отдельно, то для фосфорилирования остатков серина, треонина и тирозина получены компактные распределения RMSD 0.279 ± 0.149 нм (0.110–0.566 нм) и RMSF 0.068 ± 0.019 нм (0.0177–0.0955 нм), что указывает на устойчивую локальную упаковку после релаксации. Для ацетилирования остатка лизина средние значения выше из-за нескольких подвижных случаев: RMSD 0.484 ± 0.332 нм (0.162–1.330 нм) и RMSF 0.1 ± 0.091 нм (0.0169–0.3648 нм), при этом медианы остаются умеренными (RMSD 0.412 нм; RMSF 0.075 нм), то есть большинство систем стабилизируется, а широкий разброс обусловлен отдельными траекториями с повышенной гибкостью. Во всех рассмотренных системах доля кадров без тесных межатомных контактов в зоне модификации равна 100%, что дополнительно подтверждает отсутствие геометрических проблем после введения модификации.

Геометрия фосфатной и ацетильной групп в динамике. Мостиковая связь между кислородом боковой цепи и фосфором (O-P) в среднем составляет 1.576 ± 0.020 Å (диапазон 1.557–1.604 Å), что близко к целевому диапазону 1.60–1.62 Å, заданному на этапе геометрической вставки. Длины связей внутри самой фосфатной группы имеют значения P-O₁ = 1.476 ± 0.001 Å (1.474–1.478 Å), P-O₂ = 1.475 ± 0.002 Å (1.473–1.477 Å) и P-O₃ = 1.551 ± 0.005 Å (1.544–1.561 Å). Для двух связей P-O значения лежат в пределах целевого диапазона 1.48–1.52 Å, а третья демонстрирует небольшое смещение к верхней границе, оставаясь при этом близкой к ожидаемой для фосфатных фрагментов. Для фосфатных фрагментов типичны тетраэдрические окружения фосфора и различающиеся длины связей: мостиковые P-O порядка 1.60 Å и немостиковые 1.48–1.52 Å [10], [27], [28]. Валентные углы вокруг атома фосфора колеблются в узком диапазоне и в сумме соответствуют тетраэдрической окрестности, средние составляют $118.88^\circ \pm 0.62^\circ$, $105.17^\circ \pm 1.06^\circ$ и $105.22^\circ \pm 0.85^\circ$, при этом диэдральный параметр, характеризующий взаимную ориентацию кислородов, стабилен около $117.6^\circ \pm 2.3^\circ$. Валентные углы вокруг атома фосфора должны быть близки к тетраэдрическому значению 109.5° , наблюдаемая умеренная асимметрия допустима для реальных фосфатных окружений [10], [27]. Наблюдаемая картина указывает, что заданная при вставке локальная геометрия сохраняется после релаксации в молекулярной динамике, а присутствующая легкая асимметрия по одной из связей P-O носит систематический и умеренный характер.

Для амидного фрагмента — связь азота ε-группы лизина с карбонильным углеродом ацетильной группы — средняя длина N-C равна 1.338 ± 0.001 Å (1.337–1.339 Å), что соответствует ожидаемым значениям для частично двойственной амидной связи. Длина C=O карбонила — 1.226 ± 0.001 Å (1.224–1.227 Å) и C(карбонил)–C(метильный) — 1.483 ± 0.001 Å (1.482–1.486 Å) также согласованы с целевыми интервалами, принятыми при геометрической вставке. Планарность амидного фрагмента обусловлена частично двойственным характером связи C-N, типичные длины — около 1.33 Å (C-N) и 1.23 Å (C=O), углы в плоскости должны быть близки к 120° [25], [29], [30]. Наблюдаемые углы в плоскости амидной группы близки к 120° и демонстрируют малый разброс. Диэдральная характеристика самой амидной связки отражает ее планарность и стационарную ориентацию фрагмента. В совокупности эти результаты подтверждают сохранение *sp*²-характера и плоской геометрии ацетильной группы на производственном участке траекторий.

Тесные межатомные контакты оценивались как доля кадров траектории, в которых в окрестности модификации отсутствуют недопустимо малые расстояния между несвязанными атомами. По сводному показателю для всех 25 систем получено одинаковое значение: доля кадров без тесных контактов равна 100% во всех случаях. Иными словами, на всем 40-нс производственном участке всех 25 проанализированных систем ни в одной из траекторий не наблюдались пространственные конфликты вокруг вставленной группы, что согласуется с корректностью изначальной геометрической вставки и ее устойчивостью в динамике.

В целом модификации сохраняют ожидаемое водородное окружение. По всем системам среднее число водородных связей между модифицированной группой и белком составило 2.074 на кадр, а со стороны воды — 5.245. Такая картина согласуется с тем, что фосфатные группы активно гидратируются, тогда как ацетильные группы чаще формируют умеренное число контактов с ближайшими белковыми донорами/акцепторами. В расчётах использован стандартный геометрический критерий водородной связи, рекомендованный в практиках GROMACS [25].

В совокупности результаты по всем траекториям показывают согласованную картину. В зоне модификаций не фиксировались недопустимые сближения несвязанных атомов на всем производственном интервале, то есть геометрическая вставка групп не порождала производственных конфликтов при последующей релаксации. Водородное окружение соответствует химической природе модификаций: фосфатные группы на остатках серина, треонина и тирозина формируют разветвленную сеть контактов с ближайшими донорами и акцепторами белка и

одновременно стабильно гидратируются. Ацетильная группа на лизине удерживает планарную амидную геометрию и дает умеренное число водородных связей, при этом ее взаимодействие с водой и белковым окружением остаются воспроизводимыми от системы к системе. В терминах динамики это выражается в том, что локальная форма модифицированного фрагмента после короткой релаксации стабилизируется: разброс отклонений координат и амплитуд колебаний невелик для фосфатных модификаций и остается умеренным для большинства ацетилированных случаев несмотря на отдельные более подвижные траектории.

Показательные примеры иллюстрируют оба полюса поведения. Для фосфорилирования тирозина (структура 1EQ0, цепь А, остаток 4) наблюдается «спокойная динамика»: форма фрагмента мало меняется, целевые длины и углы фосфатной группы сохраняются, а конфликты отсутствуют. Для ацетилирования лизина (структура 1A03, цепь А, остаток 18) фоновые колебания заметнее, но плоская амидная геометрия и ориентировка ацетильной группы удерживаются, и, как в первом случае, недопустимые сближения не возникает. Эти два сценария — стабильный и более подвижный — демонстрируют главное: построенные на геометрических правилах модификации корректно достраиваются в исходные структуры и сохраняют физически правдоподобную локальную геометрию и окружение в динамике.

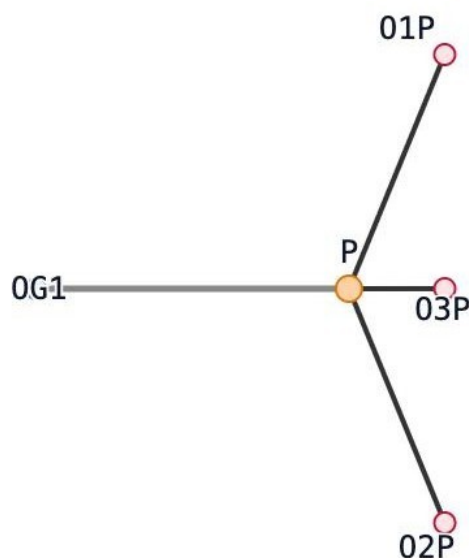
Совокупность наблюдений о сохранении целевых длин и углов, отсутствии недопустимых сближений, стабильности локальной конформации поддерживает вывод о достаточности геометрических критериев приёмки: при статусе “PASS” проверка методом молекулярной динамики может не проводиться и оставляется на случай с повышенной неопределённостью.

Обсуждение

Анализ полученных траекторий показывает, что построенная на основе геометрических правил локальная структура модификации удерживается в динамике без дополнительной ручной подгонки. Длины связей и валентные углы остаются в целевых диапазонах, тесных контактов с несвязанными атомами на производственном участке не наблюдается. Отсутствие атом-атомных перекрытия — тот самый критерий, на который принято опираться при атомарной валидации, — подтверждает, что стартовая геометрия корректна и не создаёт артефактов при релаксации [31]. Для фосфатных групп это означает сохранение тетраэдрической окрестности фосфора и характерного различия мостиковых и немостиковых Р-О, для ацетильной модификации лизина — устойчивую планарность амидной связи и близость углов к 120° , как и ожидается для sp^2 -окружения [7], [32]. Практически это означает, что для рутинных случаев со статусом “PASS” и попаданием всех величин в допуски модифицированная структура в формате PDB может использоваться сразу после применения модификации как вход для любых конвейеров моделирования и анализа, а молекулярная динамика становится опциональным этапом, применяемым по усмотрению исследователя в граничных или спорных ситуациях.

Такая устойчивость локальной геометрии важна с прикладной точки зрения: модифицированная структура в формате PDB получается сразу готовой к расчётам и не привязана к конкретному программному пакету. Её можно напрямую передавать в любой стандартный конвейер молекулярной динамики, не вставляя дополнительных ручных стадий правки координат или локальной подгонки [18]. Это снижает порог входа для пользователей, которым нужен быстрый и воспроизводимый способ добавления модификаций и оперативного запуска дальнейших симуляций.

Сопоставление с ближайшими по задачам инструментами подчёркивает нишу предполагаемого подхода. Плагин PyTMS решает введение распространённых РТМ внутри настольного визуализатора и тем самым удобен для интерактивной ручной работы, но логика его использования остается редакторской — без встроенного сценария немедленной геометрической проверки и без фокуса на детерминированной, воспроизводимой вставке в текущих координатах [6]. Веб-сервис Vienna-PTM, напротив, позиционирован как ступень подготовки к молекулярной динамике: он добавляет очень широкий набор РТМ и сразу проводит минимизацию энергии, генерируя параметры силовых полей и входы для симуляций. Инфраструктура CHARMM-GUI интегрирует химические модификации в большой онлайн-конвейер сборки систем и типично используется как исходный шаг для подготовки к расчётам. Вклад рассматриваемого в данной статье программного средства — в другом звене: точечная, параметрически заданная геометрическая сборка прямо в координатах исходной структуры с немедленной автоматической проверкой применимости и стереохимии, что отражено на рисунке 3, и с нейтральным результатом (стандартный файл PDB), пригодным для любого дальнейшего рабочего процесса.

PO₃ geometry (projection)**Result: PASS****Lengths:**

|OG1–P|: 1.61 Å

|P–O1P|: 1.48 Å

|P–O2P|: 1.48 Å

|P–O3P|: 1.48 Å

Angles at P: [target ≈ 109.5°]

∠OG1–P–O1P: 109.5°

∠OG1–P–O2P: 109.5°

∠OG1–P–O3P: 109.5°

∠O1P–P–O2P: 109.4°

∠O1P–P–O3P: 109.5°

∠O2P–P–O3P: 109.4°

min clash: 0.145 Å

worst pair: O1P ↔ LYS:7:CE (d=2.92 Å)

Рисунок 3 - Автоматическая проверка локальной геометрии
DOI: <https://doi.org/10.60797/BIO.2025.8.3.8>

Ограничения текущей версии разрабатываемого программного средства определяются покрытием модификаций и объемом внешней валидации. Полный набор результатов приведен для фосфорилирования серина, треонина и тирозина, для ацетилирования лизина показана серия на 6 белках, где часть случаев демонстрирует повышенную подвижность окружения при сохранении локальной планарной геометрии. Запуск молекулярной динамики выполняется отдельно от веб-интерфейса, выбор пороговых значений расстояний и углов фиксирован и ориентирован на классические значения, без адаптации под конкретные силовые поля. В текущей реализации не решаются задачи автоматической перепараметризации редких остатков и сложных ковалентных вариантов, как, например, полиацелирование или перекрёстные шивки.

Ближайшие шаги логично продолжить в следующем направлении: расширение класса модификаций на основе тех же параметрических принципов, добавление автоматических проверок стереохимии и локальных контактов для новых типов модификаций, унификация отчётности о статусе применимости для всех модификаций. Параллельно целесообразно добавить более строгую валидацию расстояний, по возможности включить автоматическую калибровку допусков под выбранное силовое поле, экспорт геометрических характеристик для каждой вставленной модификации, чтобы облегчить независимый аудит результатов. Такое развитие программного средства сохранит его центральную

идею — детерминированную геометрическую вставку и прозрачную проверку — при одновременном расширении области применимости и глубины контроля.

Заключение

В работе представлен легковесный веб-сервис EasyAcetoPhos для геометрической вставки посттрансляционных модификаций непосредственно в координаты PDB, обеспечивающий быструю и воспроизводимую подготовку модифицированных структур, с немедленной проверкой локальной геометрии и наглядной обратной связью. Реализовано фосфорилирование остатков серина, треонина и тирозина, а также ацетилирование лизина. Модификации строятся параметрически по целевым длинам связей и углам и автоматически стандартизируются в выходной PDB. Интерфейс позволяет выбирать целевой остаток как в линейной последовательности, так и непосредственно на трёхмерной модели белка, перед применением модификации выполняется проверка допустимости проведения такой операции и окружения, что предотвращает ошибочные действия. Внешняя верификация методом молекулярной динамики на фосфорилированных и ацетилированных системах показала сохранение целевых длин и углов, устойчивость локальной конформации и отсутствие недопустимых тесных межатомных контактов на производственном участке траекторий. При попадании метрик в заданные допуски и получении статуса “PASS” результат работы сервиса следует считать готовой для дальнейших расчётов и анализа структурой, запуск молекулярной динамики не является обязательным и применяется адресно, по исследовательской необходимости. Тем самым инструмент закрывает нишу между визуализаторами, ориентированными на ручное редактирование, и тяжёлыми конвейерами подготовки к расчётам, обеспечивая быстрое и воспроизводимое построение модификаций, совместимое с любым пакетом молекулярной динамики. В дальнейшей работе планируется последовательное расширение поддержки посттрансляционных модификаций и улучшение автоматических критериев приёма.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы на английском языке / References in English

1. Zhong Q. Protein posttranslational modifications in health and diseases: Functions, regulatory mechanisms, and therapeutic implications / Q. Zhong, X. Xiao, Y. Qiu [et al.] // *MedComm*. — 2023. — Vol. 4. — № 3. — e261 p. — DOI: 10.1002/mco2.261.
2. Ardito F. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review) / F. Ardito, M. Giuliani, D. Perrone [et al.] // *International Journal of Molecular Medicine*. — 2017. — Vol. 40. — № 2. — P. 271–280. — DOI: 10.3892/ijmm.2017.3036.
3. Yang H. Concise Review: The Regulatory Mechanism of Lysine Acetylation in Mesenchymal Stem Cell Differentiation / H. Yang, Y. Liu, X. Liu [et al.] // *Stem Cells International*. — 2020. — Vol. 2020. — 11 p. — DOI: 10.1155/2020/7618506.
4. Schreiner E. Stereochemical errors and their implications for molecular dynamics simulations / E. Schreiner, L.G. Trabuco, P.L. Freddolino [et al.] // *BMC Bioinformatics*. — 2011. — Vol. 12. — 9 p. — DOI: 10.1186/1471-2105-12-190.
5. Williams C.J. MolProbity: More and better reference data for improved all-atom structure validation / C.J. Williams, J.J. Headd, N.W. Moriarty [et al.] // *Protein Science*. — 2018. — Vol. 27. — № 1. — P. 293–315. — DOI: 10.1002/pro.3330.
6. Warnecke A. PyTMs: a useful PyMOL plugin for modeling common post-translational modifications / A. Warnecke, T. Sandalova, A. Achour [et al.] // *BMC Bioinformatics*. — 2014. — Vol. 15. — № 1. — 370 p. — DOI: 10.1186/s12859-014-0370-6.
7. Margreitter C. Vienna-PTM web server: a toolkit for MD simulations of protein post-translational modifications / C. Margreitter, D. Petrov, B. Zagrovic // *Nucleic Acids Research*. — 2013. — Vol. 41. — № W1. — P. W422–W424. — DOI: 10.1093/nar/gkt416.
8. Kern N.R. CHARMM-GUI Multicomponent Assembler for modeling and simulation of complex multicomponent systems / N.R. Kern, J. Lee, Y.K. Choi [et al.] // *Nature Communications*. — 2024. — Vol. 15. — № 1. — 5459 p. — DOI: 10.1038/s41467-024-49700-4.
9. Making selections // *Mol* Viewer Documentation*. — URL: <https://molstar.org/viewer-docs/making-selections/> (accessed: 07.09.25).
10. Corbridge D.E.C. The structural chemistry of phosphates / D.E.C. Corbridge // *Bulletin de Minéralogie*. — 1971. — Vol. 94. — № 3. — P. 271–299. — DOI: 10.3406/bulmi.1971.6534.
11. Gamoke B. Nature of PO bonds in phosphates / B. Gamoke, D. Neff, J. Simons // *The Journal of Physical Chemistry A*. — 2009. — Vol. 113. — № 19. — P. 5677–5684. — DOI: 10.1021/jp810014s.
12. Touw W.G. On the complexity of Engh and Huber refinement restraints: the angle τ as example / W.G. Touw, G. Vriend // *Acta Crystallographica Section D: Structural Biology*. — 2010. — Vol. 66. — Part 12. — P. 1341–1350. — DOI: 10.1107/S0907444910040928.

13. Pauling L. The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain / L. Pauling, R.B. Corey, H.R. Branson // *Proceedings of the National Academy of Sciences of the USA*. — 1951. — Vol. 37. — № 4. — P. 205–211. — DOI: 10.1073/pnas.37.4.205.
14. FastAPI framework, high performance, easy to learn, fast to code, ready for production // FastAPI documentation. — URL: <https://fastapi.tiangolo.com/> (accessed: 13.09.25).
15. Bio.PDB.PDBParser module // Biopython documentation. — URL: <https://biopython.org/docs/1.76/api/Bio.PDB.PDBParser.html> (accessed: 15.09.25).
16. Bio.PDB.PDBIO module // Biopython documentation. — URL: <https://biopython.org/docs/1.76/api/Bio.PDB.PDBIO.html> (accessed: 15.09.25).
17. Getting Started // Mol* Viewer Documentation. — URL: <https://molstar.org/viewer-docs/> (accessed: 17.09.25).
18. Long Range Electrostatics // GROMACS documentation. — URL: <https://manual.gromacs.org/current/reference-manual/functions/long-range-electrostatics.html> (accessed: 22.09.25).
19. Installation guide // GROMACS documentation. — URL: <https://manual.gromacs.org/current/install-guide/index.html> (accessed: 23.09.25).
20. Molecular dynamics parameters // GROMACS documentation. — URL: <https://manual.gromacs.org/current/user-guide/mdp-options.html> (accessed: 23.09.25).
21. Gmx rms // GROMACS documentation. — URL: <https://manual.gromacs.org/current/onlinehelp/gmx-rms.html> (accessed: 24.09.25).
22. Gmx rmsf // GROMACS documentation. — URL: <https://manual.gromacs.org/current/onlinehelp/gmx-rmsf.html> (accessed: 24.09.25).
23. Gmx distance // GROMACS documentation. — URL: <https://manual.gromacs.org/current/onlinehelp/gmx-distance.html> (accessed: 24.09.25).
24. Gmx angle // GROMACS documentation. — URL: <https://manual.gromacs.org/current/onlinehelp/gmx-angle.html> (accessed: 24.09.25).
25. Gmx hbond // GROMACS documentation. — URL: <https://manual.gromacs.org/current/onlinehelp/gmx-hbond.html> (accessed: 24.09.25).
26. Ormeño F. Convergence and equilibrium in molecular dynamics simulations / F. Ormeño, I.J. General // *Communications Chemistry*. — 2024. — Vol. 7. — № 1. — 26 p. — DOI: 10.1038/s42004-024-01114-5.
27. Kamerlin S.C. Why nature really chose phosphate / S.C. Kamerlin, P.K. Sharma, R.B. Prasad [et al.] // *Quarterly Reviews of Biophysics*. — 2013. — Vol. 46. — № 1. — 132 p. — DOI: 10.1017/S0033583512000157.
28. Pettersson P. Correlations between the structure and the vibrational spectrum of the phosphate group. Implications for the analysis of an important functional group in phosphoproteins / P. Pettersson, A. Barth // *RSC Advances*. — 2020. — Vol. 10. — № 8. — P. 4715–4724. — DOI: 10.1039/C9RA10366J.
29. Edison A.S. Linus Pauling and the planar peptide bond / A.S. Edison // *Nature Structural And Molecular Biology*. — 2001. — Vol. 8. — P. 201–202. — DOI: 10.1038/84921.
30. Panjikar S. Peptide bonds revisited / S. Panjikar, M.S. Weiss // *IUCrJ*. — 2025. — Vol. 12. — Part 3. — P. 307–321. — DOI: 10.1107/S2052252525002106.
31. Prisant M.G. New tools in MolProbity validation: CaBLAM for CryoEM backbone, UnDowser to rethink "waters," and NGL Viewer to recapture online 3D graphics / M.G. Prisant, C.J. Williams, V.B. Chen [et al.] // *Protein Science*. — 2020. — Vol. 29. — № 1. — P. 315–329. — DOI: 10.1002/pro.3786.
32. Eisenberg D. The discovery of the alpha-helix and beta-sheet, the principal structural features of proteins / D. Eisenberg // *Proceedings of the National Academy of Sciences of the USA*. — 2003. — Vol. 100. — № 20. — P. 11207–11210. — DOI: 10.1073/pnas.2034522100.