

Рациональный дизайн терминальной дезоксинуклеотидилтрансферазы человека

Укладов Е.О.^{1, 2*}, Тюгашев Т.Е.¹, Кузнецов Н.А.^{1, 2}

¹ Институт химической биологии и фундаментальной медицины СО РАН, Новосибирск, Россия

² Факультет естественных наук, Новосибирский государственный университет, Новосибирск, Россия

* e.ukladov@g.nsu.ru

Ключевые слова: полимеразы; терминальная дезоксинуклеотидилтрансфераза; молекулярное моделирование; молекулярная динамика; водородные связи

Мотивация и цель: Терминальная дезоксинуклеотидилтрансфераза (TdT) вместе с полимеразами β , λ , μ и X принадлежит к X -семейству ДНК-полимераз. TdT является уникальной полимеразой для позвоночных и участвует в присоединении случайных нуклеотидов на концы ДНК в процессе V(D)J-рекомбинации [1]. Благодаря уникальной способности к 5'-3'-безматричному синтезу за счет структуры петли 1, TdT нашла применение в молекулярной биологии и биотехнологии [2]. Однако существенным ограничением применения данного фермента является выраженная селективность TdT человека к присоединяемым нуклеотидам, падающая в ряду: dGTP > dCTP $\approx$ dATP > dTTP [3]. Поэтому целью исследования стало *in silico* моделирование мутантных форм TdT, обладающих измененной селективностью по отношению к dNTP, а также увеличенной каталитической эффективностью.

Методы и алгоритмы: С помощью программы Modeller были получены пре- и посткаталитические комплексы TdT с гексамерным ДНК-праймером на основании кристаллографических структур (PDB ID: 4I27 и 4I29 соответственно). Для параметризации dNTP в прекаталитическом комплексе использовался сервер R.E.D. [4]. Для ионов Mg^{2+} в активном центре (АЦ) была взята dummy-model с разнесенными зарядами. Молекулярная динамика проводилась в ПО GROMACS с силовыми полями 14SB для белка и OL15 для ДНК с моделью воды TIP3P и добавлением ионов Na и Cl в концентрации 0.1 М. Моделирование проводили при температуре 300 К. Результаты анализировали на языке Python с использованием пакетов pandas, matplotlib.pyplot, seaborn и scipy.stats.

Результаты: Исследование включает две основные части: 1) изучение роли ряда аминокислотных остатков в петле 1 и АЦ фермента; 2) предсказание замен, потенциально изменяющих селективность и каталитическую эффективность TdT. На основании филогенетического анализа, кристаллографических структур и моделирования фермента дикого типа для проведения функционального анализа отобраны остатки Asp345, Asp395, Leu397, Phe400, Glu456, Asp472, входящие в состав АЦ или петли 1. Для фермента дикого типа и мутантных форм, содержащих замены данных аминокислотных остатков, были получены модельные структуры двойных белок-ДНК и тройных белок-ДНК-dNTP комплексов.

Полученные данные выявили, что замена D345E, меняющая один из аминокислотных остатков, координирующих каталитические ионы Mg^{2+} в АЦ, в одной из траекторий приводит к стабильному увеличению расстояния между Ра-

фосфата присоединяемого нуклеотида и 3'-концевым кислородом ДНК-праймера. В других траекториях наблюдается выворачивание бокового радикала His342 из АЦ. Известно, что замена его позиционного гомолога H429A в полимеразе *μ* человека приводит к снижению независимой от матрицы активности [5], а мутантная форма TdT, содержащая замену D345E, имеет на два порядка меньшую по сравнению с диким типом активность [6]. В связи с этим можно заключить, что TdT обладает стабильным активным центром и даже незначительное изменение его геометрии может приводить к уменьшению ферментативной активности.

На следующем этапе было проведено аланиновое сканирование трех аминокислотных остатков Leu397, Phe400 и Asp472. Согласно кристаллографическим структурам, Leu397 «расклинивает» азотистые основания (АО) последнего и предпоследнего нуклеотидов ДНК, нарушая их стэкинг-взаимодействие и стабилизируя стэкинг-взаимодействие АО последнего и встраиваемого нуклеотидов. При замене L397A описанный эффект исчезает: АО dNTP, последнего и предпоследнего нуклеотидов располагаются параллельно, при этом несколько меняется положение dNTP в АЦ, что, видимо, является каталитически некомпетентным состоянием. Остаток Phe400 участвует в стэкинг-взаимодействии с Phe384 и Trp449 и опосредованно стабилизирует положение АО dNTP. При замене F400A нарушается цепочка стэкинг-взаимодействий, а АО пиримидиновых нуклеотидов (dCTP, dTTP) в 1/3 траекторий принимают перпендикулярное положение по отношению к АО 3'-концевого нуклеотида, что, по-видимому, приводит к каталитически некомпетентному состоянию.

При моделировании TdT дикого типа было обнаружено множество ионных мостиков и водородных связей между петлей 1 и остовом белка. Наиболее стабильными оказались Asp395-Arg457 (образует в среднем 1.5 водородной связи), Asp398-Trp449 (1.3), Gln401-Leu436 (1.5), Ser391-Asp472 (1.6). Предполагалось, что при замене D472A исчезнет один из контактов и положение петли 1 станет менее стабильным. Однако при исчезновении водородной связи Ser391-Asp472 в составе мутантной формы D472A другие связи стали стабильнее: Asp395-Arg457 (2), Asp398-Trp449 (1.7), Gln401-Leu436 (1.9). Таким образом, общая подвижность петли 1 не изменилась при внесении данной замены.

При моделировании комплексов фермента дикого типа было обнаружено, что остатки Asp395 и Glu456 образуют стабильные водородные связи с dGTP (наиболее эффективно встраиваемым нуклеотидом) и практически не образуют с другими dNTP. Было выдвинуто предположение, что данные остатки могут играть ключевую роль в механизме субстратной селективности фермента, и, таким образом, для выравнивания селективности необходимо увеличить количество водородных связей с другими нуклеотидами. Боковые цепи аминокислотных остатков Asp и Glu выступают акцепторами водородной связи, потому были предложены и проанализированы их замены на Asn и Gln, способные выступать и акцепторами, с целью увеличения количества водородных связей между АО dNTP и белком. Среди ортологов TdT в соответствующем Asp395 положении у четверти от общего числа последовательностей представлен Glu. Дополнительно, по аналогии с TdT *Cottoperca gobio*, была выбрана форма D395K+E456Q. Итого были получены модели фермент-субстратных комплексов, содержащих замены: D395E, D395N, D395Q, E456Q, E456N, D395K+E456Q, D395N+E456N, D395Q+E456Q, D395Q+E456N.

Анализ количества водородных связей между белком и АО (рис. 1) показал, что среди одиночных мутантных форм наибольшее количество водородных связей по сравнению с ферментом дикого типа образовывали D395N и E456N, среди двойных – D395N+E456N. Двойная аспарагиновая мутантная форма единственная показала статистически увеличение количества водородных связей относительно дикого типа с тремя нуклеотидами: dATP, dTTP и dCTP.

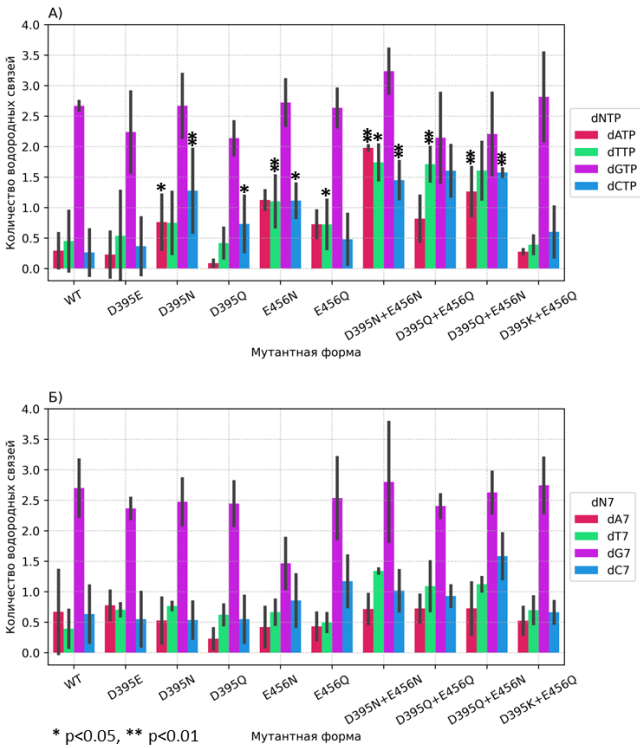


Рис. 1. Среднее количество водородных связей между белком и АО встраиваемого/3'-концевого нуклеотида для пре- (А) и пост-каталитических (Б) комплексов соответственно

Выводы: Выяснена роль аминокислотных остатков Asp345, Asp395, Leu397, Phe400, Glu456, Asp472 в связывании и стабилизации положения dNTP. Предложены аминокислотные замены D395N, E456N и D395N+E456N, потенциально повышающие субстратную селективность к dATP, dCTP и dTTP.

Финансирование: Исследование поддержано грантом РФФИ № 21-64-00017.

Rational design of human terminal deoxynucleotidyl transferase

Ukladov E.O.^{1, 2*}, Tyugashev T.E.¹, Kuznetsov N.A.^{1, 2}

¹ Institute of Chemical Biology and Fundamental Medicine, SB RAS, Novosibirsk, Russia

² Department of Natural Sciences, Novosibirsk State University, Novosibirsk, Russia

* e.ukladov@g.nsu.ru

Key words: polymerases; terminal deoxynucleotidyltransferase; molecular modelling; molecular dynamics; hydrogen bonds

Motivation and Aim: Terminal deoxynucleotidyl transferase (TdT), along with polymerases β , λ , μ , and X, belongs to the X-family of DNA polymerases. TdT serves as a unique polymerase for vertebrates and participates in the addition of random nucleotides at DNA termini during V(D)J recombination [1]. Due to its distinctive ability for 5'-3' template-independent synthesis facilitated by loop 1 structure, TdT has found applications in molecular biology and biotechnology [2]. However, a significant limitation in the application of this enzyme is the pronounced selectivity of human TdT towards the added nucleotides, ranking as follows: dGTP > dCTP $\approx$ dATP > dTTP [3]. Therefore, the research objective encompasses *in silico* modeling of mutant forms of TdT, possessing altered selectivity towards dNTPs as well as increased catalytic efficiency.

Methods and Algorithms: Pre- and post-catalytic complexes of TdT with a hexameric DNA primer were obtained using the Modeller program based on crystallographic structures (PDB ID: 4I27 and 4I29 respectively). The parameterization of dNTP in the pre-catalytic complex utilized the R.E.D. server [4]. For Mg^{2+} ions in the active site (AS), a dummy model with distributed charges was employed. Molecular dynamics simulations were conducted using the GROMACS software package with 14SB force fields for the protein and OL15 for DNA, employing the TIP3P water model, and adding Na^+ and Cl^- ions at a concentration of 0.1 M. The simulations were performed at a temperature of 300 K. Results analysis was carried out using the Python language with the utilization of the pandas, matplotlib.pyplot, seaborn, and scipy.stats packages.

Results: This study comprised two main parts: (1) investigating the role of several amino acid residues within loop 1 and the active site (AS) of the enzyme, and (2) predicting substitutions potentially altering the selectivity and catalytic efficiency of TdT.

Based on phylogenetic analysis, crystallographic structures, and wild-type enzyme modeling for functional analysis, residues Asp345, Asp395, Leu397, Phe400, Glu456, and Asp472, which are part of the AS or loop 1, were selected. Model structures of double protein-DNA and triple protein-DNA-dNTP complexes were obtained for both the wild-type enzyme and mutant forms containing substitutions of these amino acid residues.

The obtained data revealed that the substitution D345E, altering one of the amino acid residues coordinating catalytic Mg^{2+} ions in the AS, led to a stable increase in the distance between the $P\alpha$ phosphate of the incoming nucleotide and the 3'-end oxygen of the DNA primer in one trajectory. In other trajectories, a flipping of the sidechain of His342 from the AS was observed. It is known that the substitution of its positional homolog H429A in human polymerase μ leads to a decrease in template-independent activity [5], and the mutant form of TdT containing the D345E substitution exhibits activity two orders of magnitude lower compared to the wild type [6]. Hence, it can be concluded that TdT possesses a stable active site, and even minor changes in its geometry may lead to decreased enzymatic activity.

In the next stage, alanine scanning of three amino acid residues, Leu397, Phe400, and Asp472, was conducted. According to crystallographic structures, Leu397 “wedges” the nitrogenous bases (NB) of the last and penultimate nucleotides of DNA, disrupting their stacking interaction and stabilizing the stacking interaction between the NB of the last nucleotide and the embedded nucleotide. Upon substitution with L397A, this effect disappears: the NB of dNTP, the last, and penultimate nucleotides align parallelly, with a slight alteration in the position of dNTP in the AS, which appears to be a catalytically incompetent state.

The residue Phe400 participates in stacking interactions with Phe384 and Trp449, indirectly stabilizing the position of dNTP NB. Upon substitution with F400A, the stacking interaction chain is disrupted, and in one-third of the trajectories, the NB of pyrimidine nucleotides (dCTP, dTTP) adopt a perpendicular position with respect to the NB of the 3'-end nucleotide, presumably leading to a catalytically incompetent state.

During the modeling of the wild-type TdT enzyme, numerous ionic bridges and hydrogen bonds between loop 1 and the protein backbone were identified. The most stable ones were found to be Asp395-Arg457 (averaging 1.5 hydrogen bonds), Asp398-Trp449 (1.3), Gln401-Leu436 (1.5), and Ser391-Asp472 (1.6). It was hypothesized that upon substitution D472A, one of the contacts would disappear, rendering the position of loop 1 less stable. However, when the hydrogen bond Ser391-Asp472 vanished in the mutant form D472A, other bonds became more stable: Asp395-Arg457 (2), Asp398-Trp449 (1.7), Gln401-Leu436 (1.9). Thus, the overall mobility of loop 1 remained unchanged following this substitution.

During the modeling of wild-type enzyme complexes, it was discovered that residues Asp395 and Glu456 formed stable hydrogen bonds with dGTP (the most efficiently incorporated nucleotide) and practically did not form bonds with other dNTPs. It was postulated that these residues might play a crucial role in the enzyme's substrate selectivity mechanism, and thus, to equalize selectivity, it was necessary to increase the number of hydrogen bonds with other nucleotides. The side chains of Asp and Glu amino acid residues act as hydrogen bond acceptors, hence substitutions with Asn and Gln, capable of acting as acceptors, were proposed and analyzed to increase the number of hydrogen bonds between dNTP NBs and the protein. Among the TdT orthologs, Glu was present at the corresponding position to Asp395 in a quarter of the total sequences. Additionally, following the analogy with TdT *Cottoperca gobio*, the D395K+E456Q form was selected. As a result, models of enzyme-substrate complexes containing substitutions D395E, D395N, D395Q, E456Q, E456N, D395K+E456Q, D395N+E456N, D395Q+E456Q, D395Q+E456N were obtained.

The analysis of the number of hydrogen bonds between the protein and NBs (Fig. 1) revealed that among the single mutant forms, D395N and E456N formed the highest number of hydrogen bonds compared to the wild-type enzyme. Among the double mutants, D395N+E456N showed the highest increase in the number of hydrogen bonds relative to the wild type with three nucleotides: dATP, dTTP, and dCTP.

Conclusion: The role of the amino acid residues Asp345, Asp395, Leu397, Phe400, Glu456, and Asp472 in binding and stabilizing the position of dNTP has been elucidated. Amino acid substitutions D395N, E456N, and D395N+E456N have been proposed, potentially enhancing substrate selectivity towards dATP, dCTP, and dTTP.

Funding: The study is supported by the grant from the Russian Science Foundation, No. 21-64-00017.

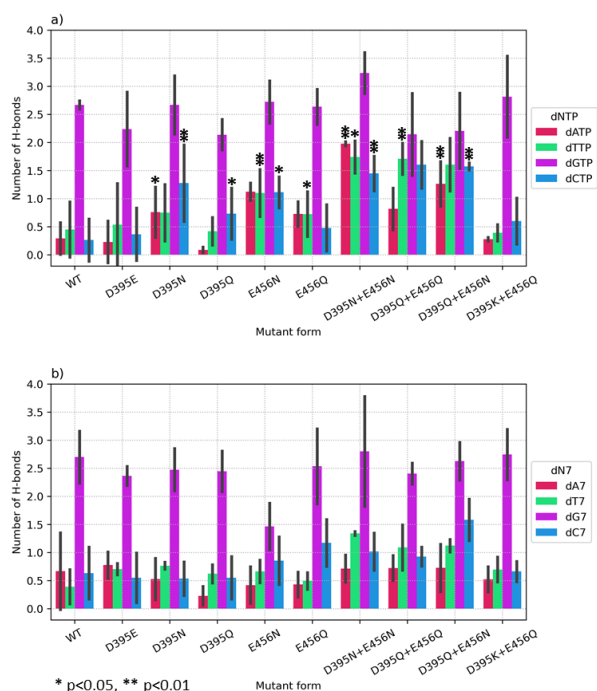


Fig. 1. Mean number of hydrogen bonds between the protein and the NB of the incoming/3'-end nucleotide for pre- (ternary) (A) and post-catalytic (binary) complexes (B) respectively

Список литературы/References

1. Schatz D.G. Recombination centres and the orchestration of V(D)J recombination. *Nat Rev Immunol.* 2011;11(4):251-263
2. Ashley J., Potts I.G., Olorunniji F.J. Applications of Terminal Deoxynucleotidyl Transferase Enzyme in Biotechnology. *ChemBioChem.* 2022;2022:e202200510
3. Kuznetsova A.A., Tyugashev T.E., Alekseeva I.V. et al. Insight into the mechanism of DNA synthesis by human terminal deoxynucleotidyltransferase. *Life Sci Alliance.* 2022;5(12):e202201428
4. Vanquelef E., Simon S., Marquant G. et al. R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res.* 2011;39:W511-W517. doi 10.1093/nar/gkr28
5. Moon A.F., Garcia-Diaz M., Batra V.K. et al. The X family portrait: structural insights into biological functions of X family polymerases. *DNA Repair.* 2007;6(12):1709-1725. doi 10.1016/j.dnarep.2007.05.009
6. Yang B., Gathy K.N., Coleman M.S. Mutational analysis of residues in the nucleotide binding domain of human terminal deoxynucleotidyl transferase. *J Biol Chem.* 1994;269(16):11859-11868