

**НОВИЙ АЛГОРИТМ
ІМІТАЦІЙНОГО ВІДПАЛУ
ДЛЯ ПЕРЕДБАЧЕННЯ
СТРУКТУРИ БІЛКІВ**

Важливою складовою частиною сучасної біології та медицини є дослідження структури білків. Зокрема, це пов'язано з тим, що білки виконують різноманітні функції в організмі, і їхня структура визначає їхню функцію. Тому розуміння структури білків є важливим для розуміння їхньої функції та для розробки нових ліків. Одним з основних методів дослідження структури білків є рентгенівська криystallogрафія. Однак цей метод є досить складним та дорогим. Тому в останні роки значну увагу приділяють розробці нових методів дослідження структури білків. Одним з таких методів є імітаційний відпал. Цей метод ґрунтується на принципах імітації природних процесів, таких як пошук їжі тваринами. Він дозволяє знаходити оптимальну структуру білка, мінімізуючи енергію системи. У статті запропоновано новий алгоритм імітаційного відпалу, який дозволяє більш ефективно знаходити оптимальну структуру білка. Цей алгоритм базується на використанні спеціальних операторів, які дозволяють швидше знаходити оптимальну структуру білка. Для перевірки ефективності запропонованого алгоритму було проведено порівняння його результатів з результатами використання стандартного алгоритму імітаційного відпалу. Результати показали, що запропонований алгоритм дозволяє знаходити оптимальну структуру білка за меншого часу та меншої кількості ітерацій. Крім того, запропонований алгоритм дозволяє знаходити більш різноманітні структури білка, що є важливим для дослідження функції білка. Зокрема, це пов'язано з тим, що білки можуть виконувати різні функції в організмі, і їхня структура визначає їхню функцію. Тому розуміння структури білка є важливим для розуміння їхньої функції та для розробки нових ліків. У статті також наведено результати порівняння запропонованого алгоритму з результатами використання стандартного алгоритму імітаційного відпалу. Результати показали, що запропонований алгоритм дозволяє знаходити оптимальну структуру білка за меншого часу та меншої кількості ітерацій. Крім того, запропонований алгоритм дозволяє знаходити більш різноманітні структури білка, що є важливим для дослідження функції білка. Зокрема, це пов'язано з тим, що білки можуть виконувати різні функції в організмі, і їхня структура визначає їхню функцію. Тому розуміння структури білка є важливим для розуміння їхньої функції та для розробки нових ліків. У статті також наведено результати порівняння запропонованого алгоритму з результатами використання стандартного алгоритму імітаційного відпалу. Результати показали, що запропонований алгоритм дозволяє знаходити оптимальну структуру білка за меншого часу та меншої кількості ітерацій. Крім того, запропонований алгоритм дозволяє знаходити більш різноманітні структури білка, що є важливим для дослідження функції білка.

1994 . . 1960-

CASP,

[1].

(Human Genome Project).

[2],

()

[3].

[4]:

(knowledge-based methods), (modeling from the first principles) . ab initio modeling).

- [5]. Z^3 .
 $\langle X, P, N, V, F \rangle$
 $X = (e_1, e_2, \dots, e_{N-1}), e_i \in \{(1,0,0), (-1,0,0), (0,1,0), (0,-1,0), (0,0,1), (0,0,-1)\};$
 $P: X \rightarrow \{0,1\} -$;
 $x = (x_1, x_2, \dots, x_N), x_i = x_{i-1} + e_{i-1}, i = \overline{2, N}, x_1 = (0,0,0);$
 $P(x) = 0,$ x ;
 $P(x) = 1,$ x ;
 $V = (v_1, v_2, \dots, v_N), v_i \in \{0,1\} -$.

$$F(x) = - \sum_{i,j=1,i>j+1}^N D(x_i, x_j) v_i v_j,$$

$$D(x, y) = \begin{cases} 1, & d(x, y) = 1 \\ 0, & d(x, y) \neq 1 \end{cases}, \quad d(x, y) - Z^3.$$

$$x_{opt} \in X,$$

$$x_{opt} = \arg \min_{x \in X} F(x),$$

$$P(x_{opt}) = 1.$$

$$x, y \in X, \quad D_h(x, y) -$$

$$O_r(x) = \{y : D_h(x, y) \leq r\} -$$

$$r, F(x) -$$

$$O_2(x) \quad r = 2.$$

$$DLS(. 1).$$

[4]).

```

procedure DLS();
  xOptimal = ;
  while O2(xOptimal) do
    x := O2(xOptimal);
    if P(x) = 0 then ;
    if F(x) < F(xOptimal) then
      xOptimal = x;
    endwhile
  return xOptimal;
end

```

. 1.

$\lambda = \text{random}[0,1]$

$T \leq 0.5$.

: 400
 0.0001
 0.9999.

```

procedure SA();
   $x_{Optimal} =$  ;
   $T :=$  < >;
   $x = x_{Optimal}$ ;
  while do
    while do
       $y :=$   $O_2(x)$ ;
      if  $P(y) = 0$  then ;
       $p := \min\{1, e^{-[F(y)-F(x)]/T}\}$ ;
       $tr := \text{random}[0,1]$ ;
      if  $p \geq tr$  then
        if  $F(y) < F(x_{Optimal})$  then
           $x_{Optimal} := y$ ;
           $x := y$ ;
        endif
      endwhile
       $T := T \cdot$  ;
    endwhile
  return  $x_{Optimal}$ ;
end
  
```

.2.

...011010100000001).

$$Fit(x) = \prod_{i=1}^3 (\max\{G_+(x_{1,i}, v_1), G_+(x_{2,i}, v_2), \dots, G_+(x_{N,i}, v_N)\} - \min\{G_-(x_{1,i}, v_1), G_-(x_{2,i}, v_2), \dots, G_-(x_{N,i}, v_N)\} + 1),$$

$$G_+(x, y) = \begin{cases} x, & y = 1, \\ -\infty, & y \neq 1, \end{cases} \quad G_-(x, y) = \begin{cases} x, & y = 1, \\ +\infty, & y \neq 1, \end{cases}$$

$x_{i,j} - j - i - x, x \in X.$

procedure SA();

$xOptimal =$

$T := <$

$x = xOptimal;$

while

while

$y :=$

if $P(y) = 0$ **then**

$$p := \min\left\{1, \frac{e^{-\frac{F(y)-F(x)}{T}}}{Fit(y)/Fit(x)}\right\};$$

$tr := random[0,1];$

if $p < tr$ **then**

if $F(y) < F(xOptimal)$ **then**

$xOptimal := y;$

$x := y;$

endif

endwhile

$T := T;$

endwhile

return $xOptimal;$

end

.3.

48, $F(x_{opt}) -$ [6].
 $x_{opt} \in X,$

101100111101110011001011 01011001100010000000110	-20	-26	-28
111101101111100100110010 000001001000100110011101	-18	-30	-31
010110111111001010010110 101000100110011001010010	-21	-31	-31
010110010111001101100011 111001011010100001001010	-18	-29	-30
001000101111001111011011 100101010010000001101101	-20	-28	-31
111000110101101101101000 000010100100010011111101	-21	-30	-30
010000101110101111011011 000101000111001100110001	-22	-27	-28
011011101111001110000001 011001101000110101011000	-17	-28	-29
010100001010100101111110 011101001011001011100001	-22	-30	-31
011000000110001110100101 100100100110011111110011	-21	-26	-31

(48).

1.

2.

S.A. Chornozhuk

A NEW ALGORITHM FOR SIMULATING THE ANNEALING FOR PREDICTING THE PROTEIN STRUCTURE

The formal statement of the protein structure prediction problem is considered. The classic approaches are investigated as well as the modification of the classic algorithms is firstly presented.

1. Simoni K.N., Hill R.D., Hill R.L. The thermodynamic hypothesis of protein folding: the work of Christian Anfinsen. *Journal of Biological Chemistry*. 2006. 281.14: e11-e11.
2. Gueorguiev V., Kuttel M. Implementation, Validation and Profiling of a Genetic Algorithm for Molecular Conformational Optimization. Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists. ACM, 2016. 16 p.
3. Lipinski-Paes T.A. Multiagent Ab Initio Protein Structure Prediction Tool for Novices and Experts. *International Symposium on Bioinformatics Research and Applications*. Cham: Springer, 2016. P. 163 – 174.
4. “ ” ”, 2016. 142 c.
5. Hulianytskyi L.F., Rudyk V.O. Protein structure prediction problem: formalization using quaternions. *Cybernetics and Systems Analysis*. 2013. 49, N 4. P. 597 – 602.
6. Yue K., Fiebig K.M, Thomas P.D., Chan H.S., Shakhnovich E.I., Dill K.A. A Test of Lattice Protein Folding Algorithms. *Proceedings of the National Academy of Sciences*, 1995. P. 325 – 329.

07.05.2018

Про автора: