1L1715

蛋白質-核酸相互作用の熱力学データベース

○Prabakaran Ponraj¹、安江虹¹、Gromiha M.Michael¹、Selvaraj Samuel¹、上平初穂¹、河野秀俊²、皿井明倫¹（¹理研・筑波研、²Dept. Chem., Univ. Pennsylvania）

Protein-nucleic acid interactions play an important role in the regulation of gene expression. Structural and binding information give insight into the mechanism of protein-nucleic acid recognition. We have developed the Thermodynamic Database for Protein-Nucleic Acid Interactions which contains currently more than 1000 entries to provide the opportunity for using the thermodynamic data of protein-nucleic acid binding. It includes information about protein, nucleic acid, the complex, experimental condition, binding data and literature information. Also, the details of secondary structure and accessible surface area of mutant residues are given for complexed and free proteins. All other related archives such as PDB, NDB, PMD, EC, PIR, ProTherm [1] and NCBI PUBMED are hypertext-linked. The molecular structures, base-amino acid interactions etc., can be visualized and analyzed through 3DinSight [2] developed in our laboratory. The WWW interface allows users to search data based on different conditions with several display and sorting options. The database will be freely accessible at <http://www.rtc.riken.go.jp/jouhou/recognition.html>.

References

1. M.M. Gromiha et al. ProTherm: Thermodynamic Database for Proteins and Mutants. *Nucleic Acids Res.* 1999, 27, 286-288; *Nucleic Acids Res.* 2000, 28, 283-285.
2. J. An et al. 3DinSight: An integrated database and search tool for the structure, function and properties of biomolecules. *Bioinformatics* 1998, 14, 188-195.

P.Prabakaran, J.An, M.M.Gromiha., S.Selvaraj, H.Uedaira, H.Kono and A.Sarai : Thermodynamic database for protein-nucleic acid interactions

1L1745

転写因子のターゲット予測のための統合ツール

○Wang Yifei、皿井明倫（理研・筑波研）

Recognition is a fundamental process for information transduction in biological system not only at the molecular level but also at the level of a network of assembled molecules. In particular, the recognition of specific DNA sites by protein is critical for the transcription, replication, and regulation of genetic information. Various sequence-based prediction have been developed (1). These methods often yield a huge amount of binding sites depending on the quality of sequence information used, and thus make it difficult to verify by experiments. It is necessary to develop some new approaches to improve the precision for prediction. We are developing a strategy for predicting the potential binding sites and target genes of transcription factors. We created computer programs based on Hidden Markov Model and neural network, and integrated them into a system (2) for searching binding sites and target genes of transcription factors. Using this system, the DNA sequence can be screened by three prediction methods independently. These results were used to screen the potential binding sites and target genes. We used this integrated search system to screen the actual target genes, and found that the prediction precision was improved significantly. We believe that the predicted results will provide a useful guide for further experiments.

1. K. Quandt et al., *Nucleic Acids Res.*, 23,4878-4884 (1995).
2. Y. Wang and A. Sarai, *Genome Informatics* 1999, 251-252 (1999).

Y. Wang, A. Sarai : An Integrated Search Tool for Binding Sites and Target Genes of Transcription Factors

1L1730

構造情報に基づく蛋白質・DNA 認識特異性の定量化とターゲット予測への応用

○皿井明倫¹、Selvaraj Samuel¹、Prabakaran Ponraj¹、河野秀俊²（¹理研・筑波研、²ペンシルバニア大）

転写因子などの DNA 結合蛋白質による特異的なターゲット配列の認識は、遺伝子発現の時間的空間的な制御において重要な役割を果たしている。これまでに多くの蛋白質と DNA の複合体の構造が明らかにされている。しかしながら、蛋白質による DNA 配列の認識機構についてはまだよくわかっていないし、認識配列を精度よく予測することもできていない。このことは構造情報がまだ十分に利用されていないことを意味している。そこで我々は、蛋白質と DNA の複合体の構造データを解析することにより、配列認識の特異性特異性を定量化し、構造との関係を明らかにしたり、ゲノムレベルでのターゲット配列の予測をしようと試みている。この方法ではまず、蛋白質・DNA 複合体の構造データベース中の塩基とアミノ酸の相互作用を統計解析することにより、それらの間の経験的相互作用ポテンシャルを導出する (1,2)。このポテンシャルを用いて実際の蛋白質・DNA 複合体に対して全体の相互作用ポテンシャルに相当するものを計算する。次に、複合体中の DNA 配列を多くのランダムな配列に置き換えて相互作用ポテンシャルを計算しそれと比較することにより、複合体中の特異 DNA 配列に対する相互作用の特異性スコア (Z-score) を計算する。今回、最新の構造データを用いて相互作用ポテンシャルの更新を行った。このポテンシャルを用いて、構造と特異性の相関について解析した。特に、DNA の変形の効果、複数の転写因子による協同的結合の効果について解析を行った。一方、この方法を用いると、転写因子のターゲット配列を予測することができる。発表では、これらの結果について報告する。

1. H. Kono and A. Sarai, *Proteins* 35, 114-131 (1999).
2. Sarai and H. Kono in *Handbook of Computational Biology*, Marcel Dekker Inc. New York (2000) in press.

A. Sarai, S. Selvaraj, P. Prabakaran, H. Kono : Quantification of Specificity in Protein-DNA Recognition Based on Structural Information and Its Application to Target Prediction

1M0900

海底熱水環境下におけるアラニンとグリシンの重合

○今井栄一、本多元、羽鳥晋由、松野孝一郎（長岡技術科学大学・生物）

海底熱水噴出孔近傍における熱水環境を模倣した反応炉（進化フローリアクター）を用い、そこで起こり得る化学反応を模倣的に実現する。本装置は高温環境と低温環境とが接する開放型非平衡反応炉である。目的とするのはタンパク質や核酸の前駆体としてのオリゴペプチド、オリゴヌクレオチドの生成である。グリシンを出発反応物質とした時に、ジケトピペラジン、グリシン 2～6 量体が生成され、反応の初期段階では反応生成物の生成量が時間の経過に対して指数関数的に増大することを既に確認した。次の段階としてグリシン以外に L-アラニンを加えた 2 種類のアミノ酸を出発反応物質として、生成される重合物の特定を行った。

グリシン (Gly)、L-アラニン (L-Ala) を出発反応物質としたときの反応生成物は多種に渡り、Gly や L-Ala の 2 量体、3 量体の他に、Ala-Gly、Gly-Ala や Gly-L-Ala-L-Ala、L-Ala-Gly-Gly などのオリゴマーが認められた。L-Ala 3 量体、4 量体の生成は L-Ala 2 量体の増大後に遅れて開始され、Gly-L-Ala、L-Ala-Gly の生成量の減少時期とも一致した。このことから Gly は Gly-L-Ala または L-Ala-Gly の生成を経て、L-Ala 3、4 量体の誘導体として機能する可能性がある。ヘテロなアミノ酸からなる重合物、たとえば Ala-Gly、Gly-Ala の生成量を比較すると、Gly-Ala の生成量が Ala-Gly のそれに比べて 10～100 倍ほど多くなった。さらに pH が低い場合や温度などの熱水環境条件にも依存し、それぞれの生成物において最適な条件が存在することが判明した。このことは海底熱水噴出孔のさまざまな熱水環境に応じて、アミノ酸の重合が選択的に行われることを示すものと考えている。

E. Imai, H. Honda, K. Hatori, K. Matsuno : Oligomerization of L-Alanine and Glycine in a Simulated Submarine Hydrothermal Environment