

**O KILKU OSOBLIWOŚCIACH
W ODDZIAŁYWANIACH MOLEKUŁ**

**ON SOME PECULARITIES OF INTERMOLECULAR
AND INTRAMOLECULAR INTERACTIONS**

Lucjan Piela

*Uniwersytet Warszawski, Pracownia Chemi Kwantowej
ul. Pasteura 1, 02-093 Warszawa
e-mail: piela@chem.uw.edu.pl*



Lucjan Piel jest profesorem zwyczajnym Uniwersytetu Warszawskiego, byłym kierownikiem Zakładu Chemii Teoretycznej i Krystalografii oraz Pracowni Chemii Kwantowej w Wydziale Chemii UW, członkiem Belgij-skiej Królewskiej Akademii Nauk i Europejskiej Akademii Nauk. Jego najważniejszymi dokonaniem naukowym są: rozwiązanie problemu sił dalekiego zasięgu w periodycznych polimerach, opis dalekozasięgowych wkładów wymiennych, koncepcja nowej izomerii (mne-mon), koncepcja optymalizacji globalnej poprzez deforma-cję funkcji optymalizowanej.

ABSTRACT

The ground state electronic energy represents a complicated function of the nuclear coordinates. Even for relatively small molecules this function may have many minima in the corresponding “energy landscape”, very often myriads of minima, each of them corresponding to a stable configuration of the nuclei. This is why predicting the lowest-energy conformation or configuration represents a formidable task. There were many attempts to solve this problem for protein molecules, for which it is believed their native conformation corresponds to the lowest free energy. The challenge to find this conformation from a given sequence of amino acids is known as a “second genetic code”. In fact all of these attempts based on some smoothing of the energy landscape. In the article some of these smoothing techniques are described, from a generic one to those, which finally turned out to be highly successful in finding native structures of globular proteins. When discussing the contributions to the conformational energy the importance of the hydrophobic effect as well as of the electrostatic interactions has been stressed. In particular it turned out that the dipole moments of the NH and of the CO bonds in proteins functioning in nature are oriented to good accuracy along the local intramolecular electric field. Thanks to enormous effort of the protein folding community it is possible to design such amino acid sequences, which fold to the desired protein 3D structure. A certain reliable theoretical technique of protein folding has been used to study a possibility of conformational autocatalysis. It turned out that a small protein of 32 amino acids, with carefully pre-designed amino acid sequence, exhibits indeed such an effect, which may be seen as a model of the prion disease propagation.

Keywords: global optimization, protein folding, dipole-dipole interaction, prion disease, autocatalysis

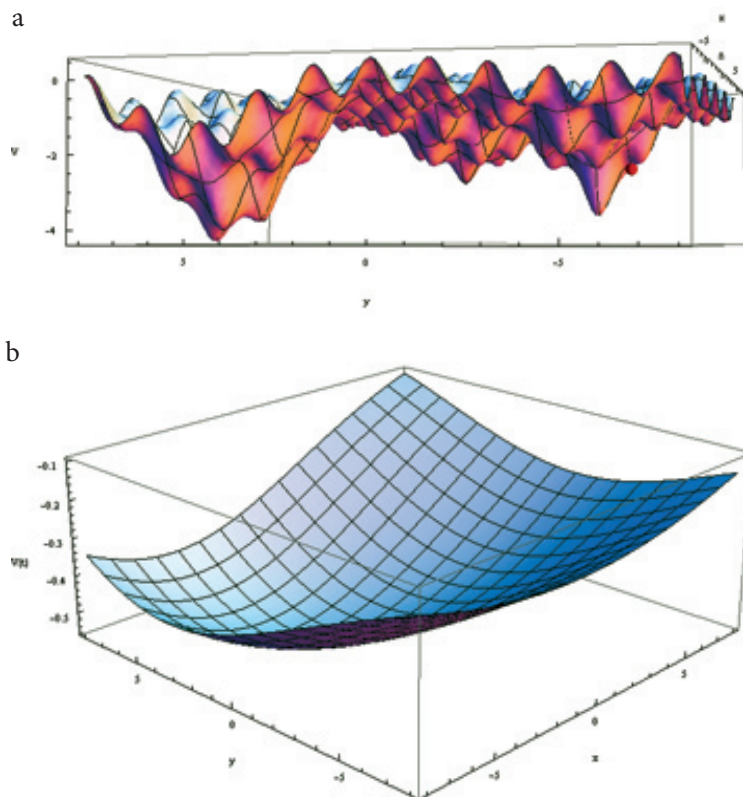
Słowa kluczowe: globalna optymalizacja, zwijanie białek, oddziaływanie dipol-dipol, choroba prionowa, autokataliza

WYKAZ STOSOWANYCH SKRÓTÓW

CASP	– ang. <i>Critical Assessment of Techniques for Protein Structure Prediction</i>
DEM	– ang. <i>Diffusion Equation Method</i>
EDMC	– ang. <i>Electrostatically Driven Monte Carlo</i>
PDB	– ang. <i>Protein Data Bank</i>
SCEF	– ang. <i>Self-Consistent Electrostatic Field</i>

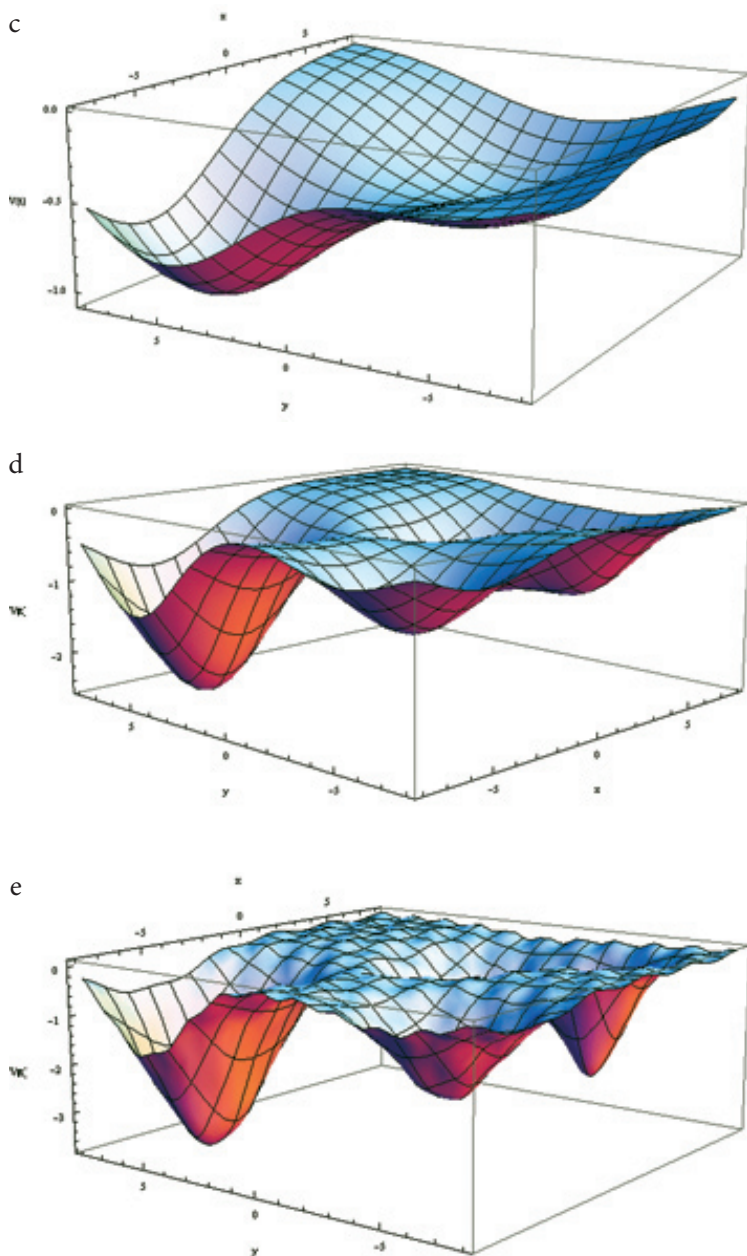
WPROWADZENIE

Zaproszenie redakcji „Wiadomości Chemicznych” traktuję jako zachętę do przystępnego opisu kilku zagadnień chemicznych z kręgu moich zainteresowań. Wybrałem kilka z nich, które w jakiś sposób wiążą się z optymalizacją energii konformacji makromolekuły.

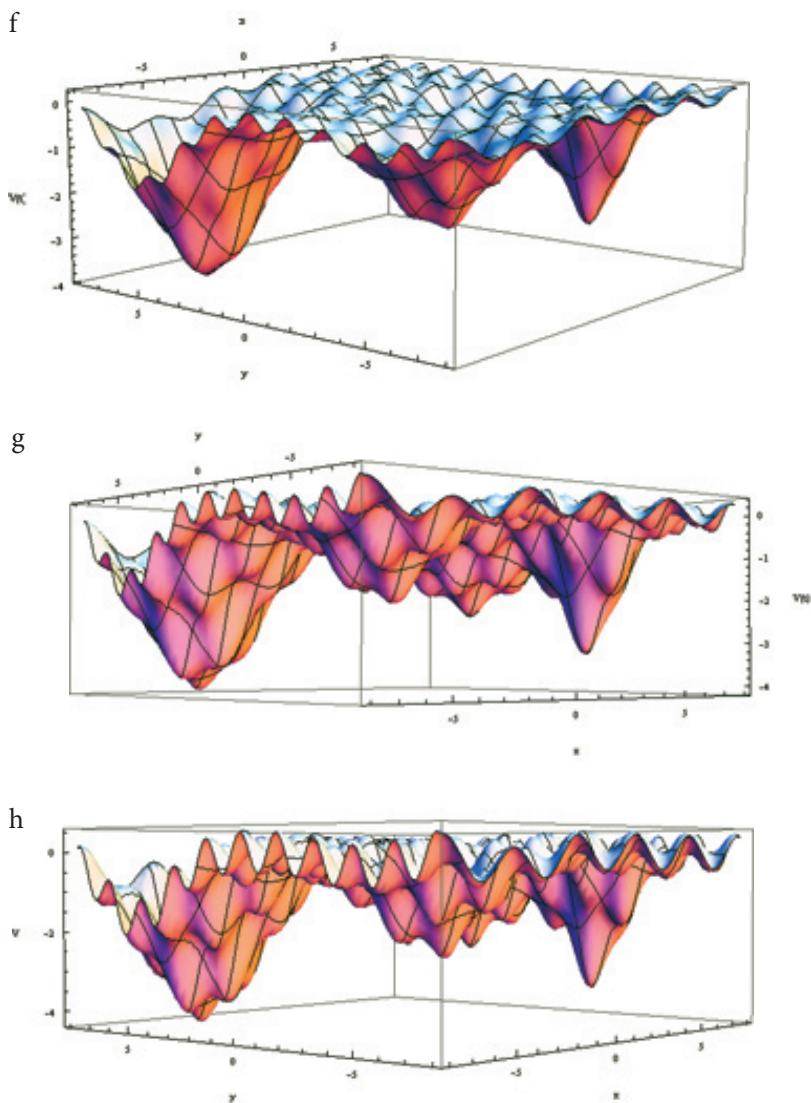


Rysunek 1. Zasada działania metody DEM znajdowania minimum globalnego pokazana na przykładzie funkcji $V(x,y) = -2\exp[-0,1(x-3)^2-0,2y^2]-3\exp[-0,5(x-5)^2-0,4(y+4)^2]-4\exp[-0,1(x+5)^2-0,2(y-5)^2]-2\exp[-0,2(x+5)^2-0,2(y+5)^2]-0,5\cos 2x\cos 4y$. Rysunek a pokazuje funkcję $V(x,y)$, czerwona kulka po prawej stronie rysunku wskazuje punkt startowy procedury optymalizacyjnej. Rysunek b odpowiada największej deformacji funkcji V . Optymalizacja wzdłuż minus gradientu funkcji zdeformowanej doprowadza do punktu jej jedyne minimum. Staje się on punktem startowym w optymalizacji funkcji pokazanej na Rys. c), itd. Ta procedura jest powtarzana i poprzez Rys. d,e,f,g doprowadza do znalezienia minimum globalnego funkcji V (Rys. h).

Figure 1. The principle of the DEM method designed for searching the global minimum of a function, in this example: $V(x,y) = -2\exp[-0,1(x-3)^2-0,2y^2]-3\exp[-0,5(x-5)^2-0,4(y+4)^2]-4\exp[-0,1(x+5)^2-0,2(y-5)^2]-2\exp[-0,2(x+5)^2-0,2(y+5)^2]-0,5\cos 2x\cos 4y$. Figure a shows $V(x,y)$, a sphere on the right-hand side of the picture indicates the starting point of the optimization procedure. Figure b corresponds to the largest deformation of V . Optimization along the minus gradient of the deformed function leads to the point of its single minimum. The point becomes the starting point in optimization of the function shown in Figure c), etc. This procedure is repeated and through Figs. d,e,f,g leads to the global minimum of V (Fig. h).



Rysunek 1. Ciąg dalszy
Figure 1. Continuation



Rysunek 1. Ciąg dalszy
Figure 1. Continuation

Problem optymalizacji energii konformacyjnej ma pewien istotny aspekt ogólny, skupimy się najpierw na optymalizacji (umownie: minimalizacji) funkcji danej wzorem matematycznym. Wtedy czasem udaje się znaleźć minimum globalne, nawet jeśli minimów lokalnych jest bardzo dużo, czasem astronomicznie wiele. Rysunek 1 ilustruje pewien nasz wkład w tym zakresie [1]. Pomysł jest bardzo prosty, pewnie zbyt prosty, aby być efektywnym w trudnych przypadkach, takich jak omawiany poniżej problem optymalnej struktury molekuly białka. Rysunek 1a przedstawia

funkcję o dość poważnej złożoności zależną (w celach ilustracyjnych) od dwóch zmiennych. Oczywiście, w podanym przypadku pokazany wycinek przestrzeni nie jest duży i nawet metoda przeszukiwania go za pomocą siatki punktów przyniosłaby tu sukces (znalezienie minimum globalnego). Jednak w przedstawianej idei godnym uwagi jest całkowicie inna strategia działania, która może być zastosowana w sytuacji, gdy takie właśnie „metody siłowe” przeszukiwania zawodzą (za dużo punktów na siatce). Polega ona na radykalnym uproszczeniu przez deformację (Rys. 1b) funkcji pierwotnej (Rys. 1a), znalezieniu minimum funkcji uproszczonej, a następnie stopniowym zmniejszaniu deformacji, przy każdym zmniejszeniu startując minimalizacją z pozycji minimum znalezionej dla poprzedniej, większej deformacji. W ten sposób posuwamy się poprzez Rys. 1c,d,e,f,g do Rys. 1h, co oznacza znalezienie minimum globalnego. Za tym opisem stoi, oczywiście, pewna matematyka, która określa wzór analityczny na funkcję zdeformowaną na każdym etapie deformacji (metoda DEM). Istota tej matematyki polega na wygładzaniu funkcji pierwotnej poprzez jej swoiste uśrednianie w otoczeniu każdego punktu przestrzeni i z różną szerokością uśredniania w zależności od parametru deformacji.

Jest rzeczą godną uwagi, że wygładzanie funkcji, choć inne niż wyżej opisane, stoi w istocie za wszystkimi opisanymi niżej ideami globalnej optymalizacji energii makromolekuł.

Asocjat kilkuatomowy ma zwykle jedną lub najwyżej kilka stabilnych konfiguracji jąder. W próbce materii poddanej jakiemś eksperymentowi niezwykle rzadko mamy do czynienia z małym asocjatem, zwykle mamy tam bardzo wiele oddziałujących atomów. W realnych sytuacjach to prowadzi do astronomicznej liczby stabilnych konfiguracji. Jedyną metodą teoretyczną rozsądzenia, która z tych konfiguracji może być tą obserwowaną w doświadczeniu, jest obliczenie energii wszystkich tych konfiguracji, a następnie zastosowanie rozkładu Boltzmanna, aby zobaczyć prawdopodobieństwo wystąpienia każdej z nich. Im niższa energia tym większe to prawdopodobieństwo. Nie ma wiarygodnego sposobu, aby przewidzieć przed obliczeniami, czy energia danej konfiguracji będzie wysoka czy niska. W ten sposób teoretyk zderza się z problemem wielu stabilnych konfiguracji/konformacji układu, który w naszym przypadku będzie uważany za molekułę („problem wielu minimów”).

1. WYRÓŻNIONA ROLA ELEKTROSTATYKI

Oddziaływanie międzycząsteczkowe w próżni jest wypadkową wielu wkładów, z których najistotniejsze to: oddziaływanie elektrostatyczne, odpychanie walencyjne, oddziaływanie indukcyjne i oddziaływanie dyspersyjne. Dwa ostatnie wkłady w praktyce zawsze istnieją, są znaczne, są zawsze przyciąganiem i przy odpowiednim wyborze środków układu współrzędnych są tylko w małym stopniu anizotropowe. Oddziaływanie elektrostatyczne jest inne. Jest niezerowe tylko, gdy oddziałujące układy mają niezerowe trwałe multipole (dipole, kwadrupole, *etc.*). W takim

przypadku jest często (choć nie zawsze) oddziaływaniem dużym i, co najważniejsze, bardzo anizotropowym. Np. odwrócenie jednej z molekuł o 180° może dać zamianę przyciągania na odpychanie. To powoduje, że nawet w sytuacjach, gdy jest ono małe, oddziaływanie elektrostatyczne decyduje o charakterze powstałej *struktury* międzymolekularnej, choć niekoniecznie o *stopniu jej trwałości* (to może zapewnić np. oddziaływanie dyspersyjne). Równie ważne, choć z całkiem innego powodu, jest odpychanie walencyjne, które wynika z zakazu Pauliego i rośnie gwałtownie, gdy jakaś część jednej molekuly wchodzi na jakąś część drugiej molekuly (kolizja w przestrzeni). Widać z tego ostatniego, że dopasowanie kształtu oddziałujących molekuł – jest to podstawa chemii supramolekularnej – może zapewnić nie tylko odpowiednią ich bliskość i styk na dużej powierzchni, ale także odpowiednią ich orientację. W tej wymuszonej przez odpychanie walencyjne orientacji, oddziaływanie elektrostatyczne może pomagać (wzmacniać kompleks) lub przeszkadzać (osłabiać kompleks). Najsilniejsze kompleksy powstają z kooperacyjnego kompromisu pomiędzy oddziaływaniem elektrostatycznym i odpychaniem walencyjnym.

Energia oddziaływania punktowych dipoli μ_A i μ_B połączonych wektorem \mathbf{R} opisana jest następującym wzorem, tu podanym w dość nietypowej, ale bardzo użytecznej formie

$$E_{\text{dip-dip}} = (\boldsymbol{\mu}_{\text{Aperp}} \cdot \boldsymbol{\mu}_{\text{Bperp}} - 2\boldsymbol{\mu}_{\text{Apar}} \cdot \boldsymbol{\mu}_{\text{Bpar}}) / R^3,$$

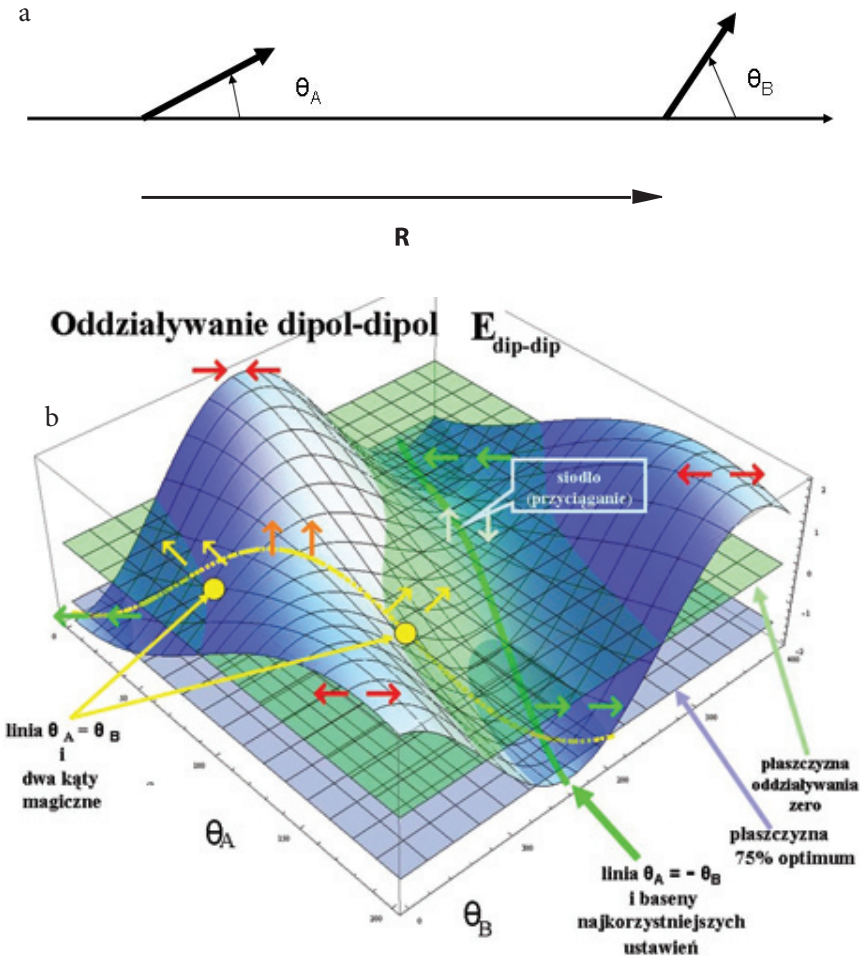
gdzie $\boldsymbol{\mu}_{\text{i perp}}$ oznacza składową wektora $\boldsymbol{\mu}_i$ prostopadłą do \mathbf{R} , a $\boldsymbol{\mu}_{\text{i par}}$ oznacza składową wektora $\boldsymbol{\mu}_i$ równoległą do \mathbf{R} .

Rysunek 2b przedstawia energię $E_{\text{dip-dip}}$ przy $\mu_A = \mu_B = \mu$ w jednostkach μ^2/R^3 (wtedy $|E_{\text{dip-dip}}| \leq 2$) jako funkcję kątów nachylenia dipoli do osi \mathbf{R} (θ_A i θ_B , Rys. 2a). Na pierwszy rzut oka powyższy wzór wygląda na banalnie prosty. Wydaje mi się jednak, a nie jestem w tym odosobniony [2], że wzór ten jednak informuje nas w sposób przejrzysty o pewnych ważnych, a czasem nawet nieoczekiwanych, cechach oddziaływania dipoli.

Na początku rzeczy dobrze znane. Widać, że oddziaływanie jest w wysokim stopniu anizotropowe, w zależności od orientacji dipoli mamy przyciąganie lub odpychanie. Najsilniejsze przyciąganie dwóch dipoli pojawia się przy ich współliniowości i zgodnym ustawieniu: $\rightarrow \rightarrow$ lub $\leftarrow \leftarrow$ ($E_{\text{dip-dip}} = -2$), najsilniejsze odpychanie zachodzi przy współliniowości i przeciwnym ustawieniu: $\rightarrow \leftarrow$ lub $\leftarrow \rightarrow$ ($E_{\text{dip-dip}} = 2$).

Mniej znane są dalsze wnioski z Rysunku 2b. Po pierwsze, ustawienie $\uparrow \downarrow$ lub $\downarrow \uparrow$ to nie minimum energii lecz punkt siodłowy (choć odpowiada mu przyciąganie $E_{\text{dip-dip}} = -1$). Najważniejsze jednak jest to, że wychodząc z konfiguracji $\rightarrow \rightarrow$ pewien ruch dipoli, a mianowicie pozwolenie dipolom zachowywać jednakowe, ale przeciwne odchylenia od łączącej je osi ($\theta_A = -\theta_B$) powoduje poruszanie się dnem kanionu o niskiej energii. Ten ruch zaczyna się od konfiguracji optymalnej $\rightarrow \rightarrow$ z $\theta_A = -\theta_B = 0$ poprzez konfiguracje typu $\nearrow \searrow$ do punktu siodłowego $\uparrow \downarrow$ lub $\downarrow \uparrow$. Ten kanion jest na początku (od strony optymalnych konfiguracji $\rightarrow \rightarrow$ oraz $\leftarrow \leftarrow$) *płaski*. Dlatego jakiegokolwiek dodatkowe (tzn. nieelektrostatyczne) oddziały-

wania, dla których optymalne ustawienia są inne niż $\rightarrow \rightarrow$ czy $\leftarrow \leftarrow$ (np. odpychanie walencyjne, dyspersyjne, indukcyjne, obniżenie energii wskutek superpozycji bazy funkcyjnej, itp.) z łatwością wpychają układ w ten kanion, *bez wielkiej straty na oddziaływaniu dipol-dipol*. Jest wobec tego zrozumiałe, że kąt między dipolami molekuł wody w dimerze wody to niemal kąt prosty ($\theta_A = -\theta_B = 45^\circ$), a mimo to jest to silne oddziaływanie elektrostatyczne dipol-dipol.



Rysunek 2. Energia oddziaływania dipol-dipol ($E_{\text{dip-dip}}$ w jednostkach μ^2/R^3) jako funkcja kątów θ_A i θ_B (Rys. a). Na Rysunku b widać, że zwiększanie przeciwnych odchyżeń od łączącej dipole osi ($\theta_A = -\theta_B$) powoduje poruszanie się dnem kanionu o bardzo niskiej energii. Z kolei zwiększanie takich wychyleń przy $\theta_A = \theta_B$ doprowadza do dwóch kątów magicznych, przy których $E_{\text{dip-dip}} = 0$.

Figure 2. Dipole-dipole interaction energy ($E_{\text{dip-dip}}$ in μ^2/R^3 units) as function of angles θ_A and θ_B (Fig. a). Figure b shows, that increasing the angular deviations in the opposite directions from the line connecting the two dipoles ($\theta_A = -\theta_B$) means moving along the bottom of a canyon of very low energy. On the other hand increasing the angular deviations with $\theta_A = \theta_B$ leads to two magic angles at which $E_{\text{dip-dip}} = 0$.

2. METODA SAMOUZGODNIONEGO POLA ELEKTROSTATYCZNEGO (SCEF)

Istotna część białek spotykanych w Przyrodzie ma określoną strukturę przestrzenną (konformacja natywna), którą białko uzyskuje na końcu czasem dość czasochłonnego procesu zwijania się w środowisku wodnym podczas i po jego syntezie w rybosomie. Czas zwijania białek ma ogromną rozpiętość (od milisekund do lat), a jest to zawsze bardzo długo w skali molekularnej i świadczy o złożonej kinetyce procesu zwijania. „Artykułem wiary” w tej dziedzinie jest, że konformacja natywna odpowiada minimum energii swobodnej molekuly białka, a ta jest określona przez sekwencję aminokwasów w tym białku („hipoteza termodynamiczna” Anfinsena [3]). W obecnym artykule skoncentrujemy się na przewidywaniu struktury przestrzennej białka z odpowiadającej mu sekwencji aminokwasów. Ten konkretny problem jest znany jako typowy problem wielu minimów, przy czym liczba tych minimów jest astronomiczna.

Weźmy np. oligomer alaniny $\text{CH}_3\text{CO}-(\text{Ala})_{19}-\text{NHCH}_3$. Minimum globalne dla tego układu jest znane: to α -helisa. Byłoby dobrze, aby minimalizacja energii startująca z dowolnej konformacji tej molekuly dawała na końcu minimum globalne. Niestety, jakkolwiek metoda minimalizacji posługująca się gradientem energii utknie w najbliższym lokalnym minimum, a tych minimów jest ok. 10^{20} . To bardzo duża liczba. Wydaje się, że nic nam nie pomaga w przeszukiwaniu przestrzeni konformacyjnej, bo łańcuch białka zachowuje się trochę jak drut kolczasty, który zahacza sam o siebie. Jednak pewna rzecz zwróciła uwagę. Otóż, minimum globalne powstaje tutaj wskutek możliwości tworzenia wiązań wodorowych. I choć wiązanie wodorowe proste nie jest (tu nie miałem na myśli jego geometrii) i ma pewne wkłady kowalencyjne, to jednak, gdy trochę donor protonu oddalić, to już praktycznie mamy tylko elektrostatykę. A w oddziaływaniu międzycząsteczkowym elektrostatyka decyduje. Niewiele da się zrobić wbrew elektrostatyce.

Teoria oddziaływań międzycząsteczkowych w próżni daje nam pouczające informacje dotyczące również oddziaływań *wewnątrz* pojedynczych molekuly. Np. we wspomnianym oligomerze alaniny w konfiguracji α -helisy można zainteresować się, jak układają się względem lokalnego wewnątrz molekularnego pola elektrycznego dipole NH i CO wchodzące w skład sztywnych i płaskich jednostek $-\text{HN}-\text{CO}-$ z wiązaniem peptydowym (z tych jednostek zbudowany jest rdzeń łańcucha białek). Okazało się [4], że pozycje dipoli wzdłuż łańcucha są bardzo bliskie kierunkowi lokalnego pola elektrycznego: straty energetyczne z powodu odchylenia wynosiły 0,02 kcal/mol dla CO i 0,12 kcal/mol dla NH. Nie było tak na końcach helisy, gdzie od strony *N*-końca ta strata wynosiła 2,24 dla CO, a dla *C*-końca strata dla CO wynosiła 0,75 kcal/mol. Jeśli w konformacji α -helisy zrobiono świadomie błędy łamiące lokalnie α -helisę (próby przeprowadzano do czterech błędów), były one widoczne przez bardzo silnie zlokalizowane straty energetyczne oddziaływania dipol-pole w miejscu wprowadzenia błędu. Taki „błąd ustawienia dipoli” łatwo więc w ten sposób zlokalizować, a także przedstawić można diagnozę, jaka zmiana

konformacyjna jest potrzebna, aby błąd usunąć. Pokazano, że taka procedura usuwania defektów prowadzi do minimum globalnego z pewnych punktów przestrzeni konformacyjnej odległych od minimum globalnego. Metoda SCEF została użyta do kierowania procedurą Monte Carlo w metodzie EDMC [5] przewidywania struktury białek. Do pewnego stopnia można to więc uważać za metodę zastępowania funkcji optymalizowanej przez funkcję prostszą, zdeformowaną (tu nawiązanie do poprzedniej metody) wynikającą z elektrostatyki. Niedawno wykonano analizę ustawienia dipoli wzdłuż lokalnego pola elektrycznego dla 112 białek występujących w Przyrodzie [6]. Badania te potwierdziły nasz wcześniejszy wynik, że dipole CO i NH orientują się w znacznym stopniu wzdłuż pola elektrycznego i to nie tylko w dwóch najważniejszych motywach struktury drugorzędowej, α -helisach i β -karkach, ale także w innych elementach strukturalnych białek (jak np. pętle).

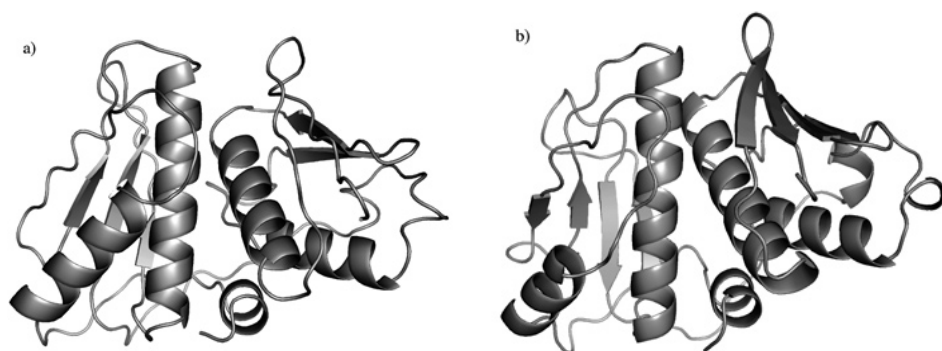
3. „DRUGI KOD GENETYCZNY”

Wielkim osiągnięciem współczesnej nauki jest praktyczne pokonanie problemu przewidywania kształtu przestrzennego białek globularnych (zagadnienie znane jako „drugi kod genetyczny”), przy ogromnym udziale uczonych polskich: głównie Krzysztofa Ginalskiego, Andrzeja Kolińskiego, Janusza Bujnickiego, Leszka Rychlewskiego. W roku 2004 w konkursie CASP z udziałem ponad 200 laboratoriów świata wyniki były następujące: 1. Ginalski (Polska), 2. Koliński-Bujnicki (Polska), 3. Baker (USA), 4. Skolnick-Zhang (USA), 5. Genesilico (Polska).

Wiele technik numerycznych prowadzi do tego sukcesu, a wśród najbardziej skutecznych są metody opierające się na porównywaniu sekwencji aminokwasów w białku, którego strukturę chcemy znaleźć, z białkami, których strukturę już wyznaczono eksperymentalnie (przeważnie w kryształach). Wadą takiego podejścia jest przedkładanie skuteczności nad zrozumienie mechanizmu fizycznego oddziaływań. Są jednak metody oparte w znacznie większym stopniu na fizyce i chemii oddziaływań. Zakładają one, że postać natywna białka odpowiada konformacji o najniższej energii i że rodzina tych konformacji jest oddzielona od innych przez szeroką przerwę energetyczną [7]. Tak więc, „krajobraz energetyczny” przedstawiający energię konformacyjną molekuly białka jest podobny do lejka, co ułatwia zwijanie się białek (zarówno w Przyrodzie jak i ...w komputerze). Gdyby stworzyć białko o przypadkowej sekwencji aminokwasów, najprawdopodobniej nie miałyby ono tej przerwy energetycznej, nie byłoby efektu lejka i nie byłoby sposobu na uzyskanie białka niezawodnie w jednej natywnej konformacji. To oznacza, że znajomość głównych mechanizmów zwijania się białek pozwoliłaby na zaprojektowanie białka, które miałyby się niezawodnie zwinąć do pewnej konformacji. Takie wymaganie w dużym stopniu spełnia model Kolińskiego [8]. Opiera się on na wprowadzeniu pola siłowego czyli uproszczonego wyrażenia na energię konformacji polipeptydu. W tym celu łańcuch polipeptydowy ulega uproszczeniu przez pominięcie na początku wielu szczegółów atomowych (odpowiada to swoistemu wygładzaniu powierzchni

energii). Energia łańcucha zawiera zarówno człony zapewniające odpowiedni stopień jego sztywności (potencjał krótkozasięgowy) jak i możliwe oddziaływania odległych wzdłuż łańcucha, ale bliskich w przestrzeni aminokwasów. Te ostatnie oddziaływania uwzględniane są przez energetyczne kary i nagrody zgodne z obserwowaną w światowej bazie danych statystyką kontaktów aminokwas-aminokwas, z uwzględnieniem typowych dla białek oddziaływań ich podstawowych elementów strukturalnych (α -helisy i β -kartki). To narzędzie teoretyczne jest stosowane wraz z metodą Monte Carlo, w której automatycznie wybierane są konformacje o energii pasującej do założonej w symulacji komputerowej temperatury. W praktyce stosuje się równoczesną symulację Monte Carlo wielu replik układu odpowiadającego różnym temperaturom, z dopuszczeniem wymiany replik między tymi równoległymi procesami numerycznymi (metoda Monte Carlo with Replica Exchange).

Rysunek 3a pokazuje przewidzianą teoretycznie przez Kolińskiego i Bujnickiego strukturę białka porównaną ze strukturą krystalograficzną (Rys. 3b) podaną po tej predykcji w ramach CASP6 w 2004 roku. Białko jest jednym z 1877 białek pewnej bakterii (*Thermotoga maritima*) znalezionej w pobliżu gorących wód podmorskich związanych z aktywnością wulkaniczną. Zastosowany model Kolińskiego i Bujnickiego był jednym z najefektywniejszych w CASP6.



Rysunek 3. Rysunek a przedstawia teoretycznie przewidywaną strukturę jednego z białek bakterii *Thermotoga maritima* podaną przez Kolińskiego i Bujnickiego, Rysunek b strukturę tego samego białka znaną metodami rentgenostrukturalnymi. Pozycje atomów obu struktur różnią się średnio zaledwie o 2,9 Å, co jest uważane za zgodność bardzo dobrą.

Figure 3. Figure a shows the 3D structure of a protein (of the bacterium *Thermotoga maritima*) predicted theoretically by Koliński and Bujnicki, while Figure b shows the corresponding structure obtained from X-ray analysis of a crystal. The corresponding atomic positions differ (r.m.s.) by about 2,9 Å, which means a very good agreement.

4. AUTOKATALIZA KONFORMACYJNA – JAKO MODEL CHOROBY PRIONOWEJ

Model teoretycznego zwiżania się białek może być zastosowany do sytuacji niezwykłych. Wyobraźmy sobie białko z sekwencją aminokwasów specjalnie zaprojektowaną tak, aby według opisanego modelu niezawodnie zwiżało się do pewnej konformacji typu A, która jest jego konformacją natywną, odpowiadającą najniższej energii. Jednocześnie sekwencja użyta ma jeszcze inną cechę, a mianowicie może tworzyć także konformację typu B, diametralnie różniącą się od A. Białko ma zwiżać się do pewnej konformacji natywnej, ale przy agregacji wielu molekuł białka może dojść do konkurencji dwóch rodzin konformacji: jednej podobnej do A a drugiej do B. Gdyby w oddziaływaniu międzymolekularnym niższą energię dawało oddziaływanie typu BB tak, że konformacja A w kontakcie z inną molekułą o konformacji B zmieniałaby się z AB na BB, to oznaczałoby to możliwość rozszerzania się „choroby konformacyjnej”.

Podjęto próbę takiego zaprojektowania sekwencji białka, aby opisana wyżej sytuacja się wydarzyła [9]. W tym celu w środku sekwencji aminokwasów umieszczone zostały dwie glicyny, aminokwasy znane z wyjątkowej giętkości (przyczyna: podstawnik w glicynie to atom wodoru). To powinno ułatwić tworzenie struktur zgiętych w połowie (typu „spinka do włosów”), pożądanych zarówno w planowanej strukturze α -helisy (ona odpowiadałaby wspomnianej konformacji A) jak i formy β (ta odpowiadałaby B). Dwie dalsze glicyny umieszczono w pozycjach 8 i 25, w miejscach ewentualnych zgieć w planowanej β -karcie. Kwasy glutaminowe i lizyny mają takie rozmieszczenie w sekwencji, które umożliwia stabilizację, o ile struktura jest α -helikalna (oddziałują one przez silne mostki solne $-\text{NH}_3^+ \dots \text{OOC}-$). Aby zwiększyć szansę utworzenia się struktur β , w sekwencji umieszczono charakterystyczny dla takich struktur układ walin i izoleucyn, które silnie się przyciągają poprzez efekt hydrofobowy związany ze strukturą wody. Struktura wody była tu uwzględniana tylko *implicite* poprzez potencjały statystyczne obliczone w obecności hydratacji białek z bazy danych PDB. Dokonując małych zmian w sekwencjach białek i następnie modelując zwiżanie się tych białek do konformacji typu A, B, AA, AB, BB wybrano ostatecznie jedną sekwencję:

GVEIAVKGAEEVAAKVG GVEIAVKAGEVAAKVG

(G = glicyna, V = walina, E = kwas glutaminowy, I = izoleucyna, A = alanina,
K = lizyna).

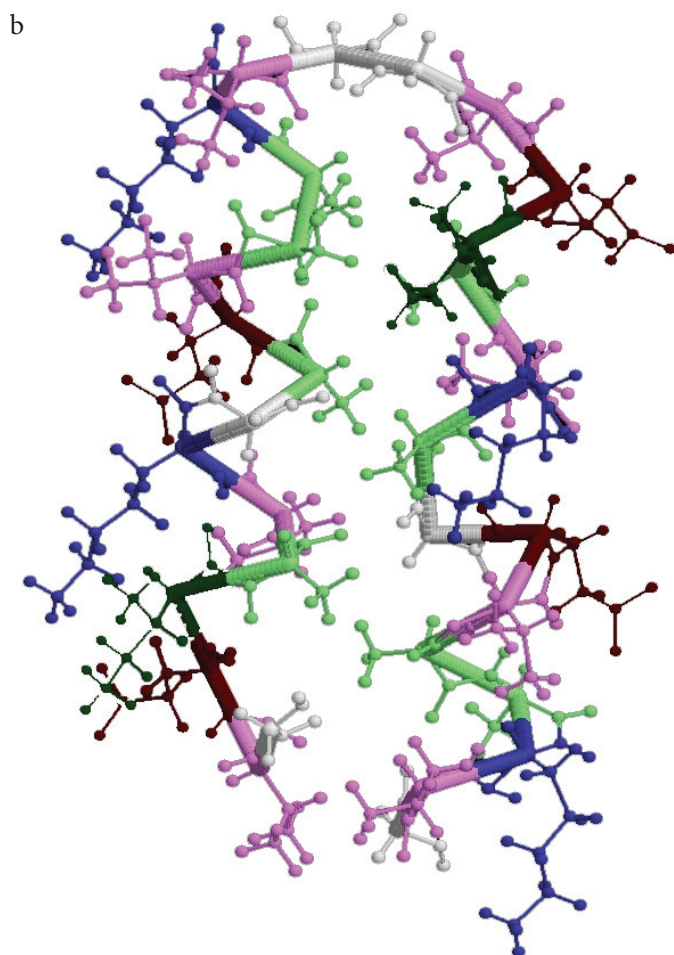
która jako jedyna spośród wielu zbadanych wykazała zjawisko autokatalizy konformacyjnej. Pojedyncza molekuła tego oligopeptydu w procedurze Monte Carlo osiąga globalne minimum energetyczne, któremu odpowiada dwuhelikalna wiązka (Rys. 4a,b), jeśli temperatura symulacji przewyższa pewną temperaturę minimalną. Poniżej tej temperatury, obok wspomnianej dwuhelikalnej wiązki, wykazuje stabilność czteroczłonowa baryłka β , Rys. 4c. Jeśli jednak mamy do czynienia z oddziaływaniem dwóch cząsteczek białka, i jedna z molekuł jest zamrożona w formie baryłki β

(z jakichkolwiek powodów, np. przez chemiczne lub fizyczne więzy, Rys.4d), a druga może się swobodnie zwijać, wtedy praktycznie zawsze zwija się ona do czterocząłowej baryłki β silnie oddziałującej międzycząsteczkowo z zamrożoną czterocząłową baryłką β (Rys. 4e). Dzieje się tak nawet wtedy, gdy ta druga molekula jest na starcie w swojej konformacji natywnej (Rys. 4d), wtedy się rozwinie i ponownie zwinie, ale do konformacji metastabilnej – baryłki β . Wykazano również, że trzecia molekula białka, w obecności już zwiniętej pary molekuł, zwija się podobnie, tworząc w ten sposób stos trzech czterocząłowych baryłek β , układ o niskiej energii.

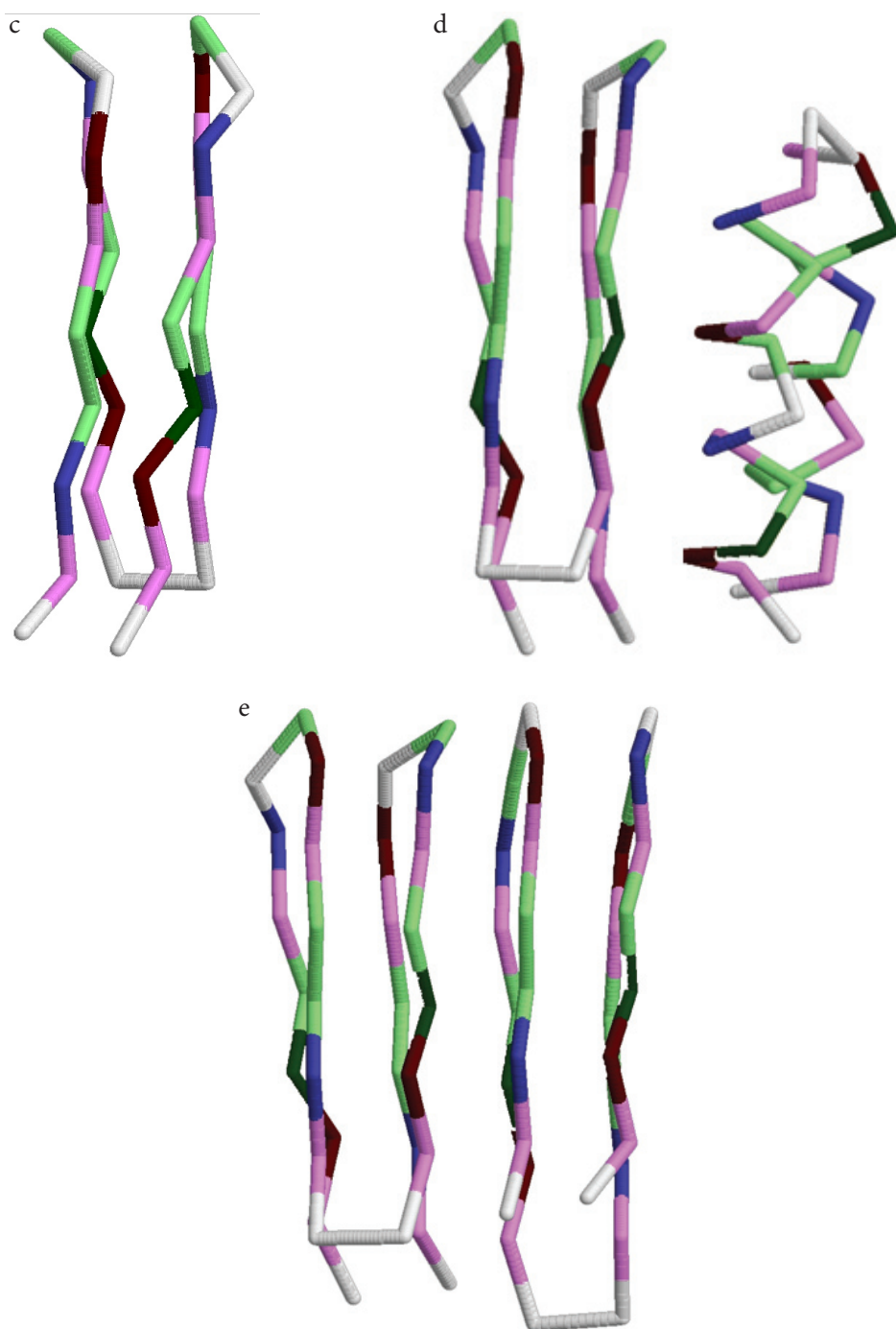


Rysunek 4. Autokataliza konformacyjna. Rysunek a przedstawia strukturę α -helikalnej wiązki białka o następującej sekwencji aminokwasów: GVEIAVKGAEVA AAKVGGVEI AVKAGEVA AAKVG (G = glicyna, V = walina, E = kwas glutaminowy, I = izoleucyna, A = alanina, K = lizyna) pokazaną w sposób uproszczony (każdy aminokwas jest jednym pseudoatomem). Ze struktur podobnych do a) da się odtworzyć reprezentację pełnoatomową (b). Jest to struktura odpowiadająca minimum globalnemu dla tego białka. Rysunek c pokazuje inną konformację tego samego białka, o energii wyższej niż minimum globalne, ale o sporej stabilności (stan metastabilny) wykazującą strukturę baryłki β . Rysunek d pokazuje konfigurację startową procedury Monte Carlo przy zamrożonej baryłce β , zaś Rysunek e wynik końcowy symulacji Monte Carlo.

Figure 4. Conformational autocatalysis. Figure a represents a two α -helix bundle of a protein with the following amino acid sequence: GVEIAVKGAEVA AAKVGGVEI AVKAGEVA AAKVG (G = glycine, V = valine, E = glutamic acid, I = isoleucine, A = alanine, K = lysine). This 3D structure is shown in a simplified way by representing a single amino acid by a sphere. The all atom representation of such structures can be recovered, as shown in Figure b for the two α -helix bundle. This structure corresponds to the global minimum of the protein molecule. In Figure c another conformation (metastable one) of the same protein is displayed, a β barrel, with a higher energy. In Figure d the starting configuration for the Monte Carlo procedure is given: the two α -helix bundle together with a frozen β barrel, while Figure e shows the final result: two interacting β barrels.



Rysunek 4. Ciąg dalszy
Figure 4. Continuation



Rysunek 4. Ciąg dalszy
Figure 4. Continuation

Wykazano w ten sposób tendencję wspomnianego oligomeru do autokatalizy konformacyjnej, czyli spontanicznego związania się molekuly, w obecności wyższej energetycznie konformacji innej molekuly tego białka, do takiej samej wyższej energetycznie konformacji, z zyskiem energetycznym wynikającym z oddziaływań międzymolekularnych. „Zła konformacja” (ta wyższa energetycznie) samorzutnie rozprzestrzenia się w układzie. Podobne zachowanie wykazują białka w groźnych chorobach prionowych.

UWAGI KOŃCOWE

Podane przykłady pokazują, że współczesne metody teoretyczne przewidywania struktury natywnej białek globularnych radzą sobie nawet w dość skomplikowanych sytuacjach. Główną przyczyną powodzenia jest swego rodzaju wygładzanie powierzchni energii potencjalnej (choć inne niż to omówione na początku artykułu) poprzez operowanie pseudoatomami reprezentującymi całe aminokwasy czy, w przypadku niektórych z nich, jeszcze dodatkowymi pseudoatomami. Do sukcesu tych metod przyczynia się także użycie metody Monte Carlo Replica Exchange i wykorzystanie potencjałów empirycznych opartych na częstości występowania par aminokwas-aminokwas w banku danych strukturalnych Protein Data Bank.

PIŚMIENNICTWO CYTOWANE

- [1] L. Piela, J. Kostrowicki, H.A. Scheraga, *J. Phys. Chem.*, 1989, **93**, 3339.
- [2] L. Kocbach, S. Lubbad, *Phys. Educ.*, 2010, **45**, 345.
- [3] Ch.B. Anfinsen, R.R. Redfield, W.L. Choate, J. Page, W.R. Carroll, *J. Biol. Chem.*, 1954, **207**, 201.
- [4] L. Piela, H.A. Scheraga, *Biopolymers*, 1987, **26**, S33.
- [5] D.R. Ripoll, H.A. Scheraga, *Biopolymers*, 1988, **27**, 1283.
- [6] D.R. Ripoll, J.A. Vila, H.A. Scheraga, *Proc. Natl Acad. Sci. (USA)*, 2005, **102**, 7559.
- [7] E.I. Shakanovich, A.M. Gutin, *Proc. Natl Acad. Sci. (USA)*, 1993, **90**, 7195.
- [8] M. Boniecki, P. Rotkiewicz, J. Skolnick, A. Koliński, *J. Comput. Aided Mol. Des.*, 2003, **17**, 725.
- [9] E. Małolepsza, M. Boniecki, A. Koliński, L. Piela, *Proc. Natl Acad. Sciences (USA)*, 2005, **102**, 7835.

Praca wpłynęła do Redakcji 14 listopada 2011