

PRACA DYPLOMOWA LICENCJACKA

Natalia Długołęcka

Nr albumu: 243262

podpis studenta:.....

Dokowanie symetryczne z utworzeniem kompleksu insuliny

Praca wykonana pod opieką prof. dr hab.
Rajmunda Kaźmierkiewicza w ramach
przedmiotu
Seminarium dyplomowe
prowadzonego przez prof. dr hab. Rajmunda
Kaźmierkiewicza

podpis opiekuna:

Gdańsk

OŚWIADCZENIE

Ja, niżej podpisana(y) oświadczam, iż przedłożona praca dyplomowa została wykonana przeze mnie samodzielnie, nie narusza praw autorskich, interesów prawnych i materialnych innych osób.

Ja, niżej podpisana(y) oświadczam, że praca dyplomowa przedstawiona w formie elektronicznej i wydrukowanej jest identyczna.

.....

..... data

.....

własnoręczny podpis studenta

Spis treści

Wstęp.....	4
Obiekt badań.....	4
Cel pracy.....	4
Metody.....	5
Program SAM.....	5
Model wyjściowy.....	6
Ekstrakcja monomeru.....	7
Dokowanie, wykorzystane narzędzia.....	8
Dokowanie dimeru.....	8
Przewidywanie struktury heksameru insuliny.....	9
Dokowanie sześciu monomerów.....	9
Dokowanie trzech dimerów.....	9
Wyniki.....	9
Dimer insuliny.....	9
Dimer 1.....	9
Dimer 2.....	11
Wybór dimeru.....	14
Heksamer insuliny.....	15
Przewidywanie struktury heksameru przy użyciu dimerów.....	15
Przewidywanie struktury heksameru przy użyciu monomerów.....	18
Interpretacja.....	24
Dyskusja.....	25
Bibliografia.....	26

Wstęp

Obiekt badań

W 1951 pojawiła się praca *The Amino-acid Sequence in the Phenylalanyl Chain of Insulin*, autorstwa Frederika Sangera i Hansa Tuppy'ego, w której przedstawiona została sekwencja jednego z łańcuchów insuliny - łańcucha B. Rok później znano już całą sekwencję insuliny, gdy Sanger wspólnie z Thompsonem sekwencjonowali łańcuch A¹.

Insulina jest niewielkim hormonem, którego gen INS koduje zaledwie 110 reszt aminokwasowych². Ludzka insulina zbudowana jest z łańcucha A i B, w których skład wchodzi odpowiednio 21 oraz 30 reszt aminokwasowych. Łańcuchy A i B są połączone dwoma mostkami disulfidowymi stabilizującymi strukturę trzeciorzędową insuliny. Oba łańcuchy przyjmują konformację α-helisy. W wyniku modyfikacji enzymatycznej nieaktywnej proinsuliny, z której zostaje usunięta środkowa część cząsteczki, następuje konwersja do dwułańcuchowej, aktywnej formy insuliny². W takiej postaci hormon magazynowany jest w komórkach β trzustki. W obecności kationów Zn²⁺ przy pH bliskim 7 dwie cząsteczki insuliny tworzą dimer. Co więcej, insulina może tworzyć również heksamery złożone z trzech dimerów, stabilizowane przez dwa dwuwartościowe kationy cynku⁴. Insulina w postaci heksameru wykazuje znaczną elastyczność struktury - dla porównania, nieaktywny heksamer proinsuliny jest stabilny⁴. Ta cecha insuliny objawia się przez występowanie trzech stanów konformacyjnych heksameru: T6, T3R5 oraz R62⁶. Stany te związane są ze zmianami zachodzącymi na n-końcu łańcucha B.

W stanie T, reszty 1-8 łańcucha B znajdują się w rozciągniętej konformacji (ang. *extended conformation*) natomiast jest zachowana α-helisa pomiędzy resztami B9-B19. Po przyjęciu przez insulinę konformacji R, helisa ta zostaje rozszerzona (ang. *fully extended*) także na reszty B1-B8. Podczas przejścia z konformacji T na R zmiany strukturalne są na tyle znaczące, że powstały R heksamer insuliny staje się przedmiotem badań nad zmianami allosterycznymi w białkach. Monomer insuliny również ulega przemianom konformacyjnym, jednak fizjologiczne znaczenie tego procesu nie zostało jeszcze do końca wyjaśnione⁶.

O ile od poznania sekwencji łańcucha B do uzyskania kompletnej sekwencji minął około rok, o tyle poznanie struktury trzeciorzędowej insuliny było dość problematyczne, ze względu na jej ruchliwość.

Cel pracy

Celem mojej pracy jest utworzenie modelu heksameru insuliny przy wykorzystaniu jednej z technik bioinformatycznych, dokowania symetrycznego. Moim zadaniem jest wykorzystanie struktury monomeron pochodzącego ze struktury natywnej do:

- uzyskania symetrycznego modelu dimeru insuliny, a następnie symetrycznego modelu heksameru, poprzez dokowanie trzech dimerów,

- uzyskanie modelu symetrycznego heksameru dokując bezpośrednio sześć monomerów.

Metody

Wykorzystanie technik bioinformatycznych takich jak dokowanie molekularne, pozwala, z jednej strony, na przyspieszenie badań nad strukturą kompleksów, umożliwia także prowadzenie poszukiwań bez konieczności odwoływania się do klasycznych eksperymentów. Algorytmy wykorzystywane w dokowaniu molekularnym umożliwiają znajdowanie najbardziej korzystnej konfiguracji układu białko-ligand - czyli takiej, dla której energia swobodna konformacji jest możliwie jak najniższa⁷. Dokowanie molekularne ma swoje zastosowanie głównie w procesie projektowania leków, pozwala także na odkrywanie struktur nieznanych białek⁷. Może się ono okazać skuteczne również w przypadku przewidywania struktury kompleksów niezbyt stabilnych w roztworze wodnym, to znaczy o dużej elastyczności, takich jak insulina.

W swojej pracy posłużyłam się oprogramowaniem komputerowym takim jak:

- *SAM 2.0.4* – program służący do dokowania symetrycznego,
- *UCSF Chimera 1.13, PyMOL 1.7.2.1* – programy do wizualizacji cząsteczek.

Program SAM

Szczególnym przypadkiem dokowania jest dokowanie symetryczne, w którym wykorzystuje się symetrię kompleksów białkowych i uzyskuje się dzięki temu struktury idealnie symetryczne. Jednym z programów komputerowych służących do dokowania jest program SAM (*Symmetry Assembler*), którego używałam w procesie dokowania symetrycznego podjednostek kompleksu insuliny⁵.

Program SAM został po raz pierwszy zaprezentowany w publikacji pod tytułem *Spherical polar Fourier assembly of protein complexes with arbitrary point group symmetry* autorstwa Davida W. Richiego i Sergeia Grudinina⁵ jako narzędzie do modelowania kompleksów charakteryzujących się dowolnym typem symetrii. Program SAM został opracowany przy wykorzystaniu programu HEX, którego algorytm oparty jest na szybkiej sferycznej transformacie Fouriera⁹. Odpowiednie modyfikacje tego algorytmu pozwoliły na otrzymanie kompleksów idealnie symetrycznych, o dowolnym typie symetrii punktowej. W doświadczeniu⁵, przetestowano program SAM dla 18 pobranych z bazy PDB (*Protein Data Bank*) struktur symetrycznych kompleksów białkowych. Każda struktura reprezentowała inny typ symetrii. W publikacji przeprowadzono dokowanie używając programu SAM bazując na monomerach badanych cząsteczek: 16 spośród 18 rekomendowanych modeli (posiadających najniższą energię konformacji) odpowiadało reprezentatywnej strukturze krystalicznej, a wskaźnik RMSD nie przekraczał wartości 3.0 Å.

Dzięki wykorzystaniu programu SAM wyniki dokowania uzyskiwałam stosunkowo szybko, sprecyzowanie interesującego mnie typu symetrii pozwoliło na utworzenie idealnie symetrycznych struktur⁵. Moim zdaniem jest warte podkreślenia to, iż program SAM traktuje cząsteczkę jako

chmurę elektronową posługując się funkcjami gęstości elektronowej, rezygnując z modelu „kulek i sprężynek”, co znacząco wpływa na precyzję wyników uzyskanych z użyciem tego programu⁵.

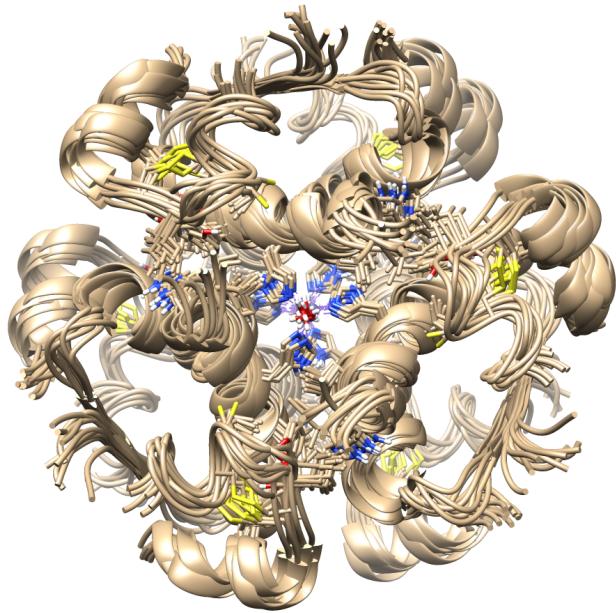
Aktualna wersja programu SAM¹⁰ działa tylko pod systemem operacyjnym Linux. Nie posiada on interfejsu graficznego. Jako skrypt napisany w języku programowania Python, uruchamiany jest w linii komend za pośrednictwem Terminala.

Model wyjściowy

W swojej pracy skupiłam się głównie na strukturze 1AIY pobranej z bazy PDB, zdeponowanej jako wynik opisany w pracy *Solution structures of the R6 human insulin hexamer* autorstwa Changi i Jorgesena⁶. W pracy tej określono przestrzenną strukturę heksameru ludzkiej insuliny stabilizowanego fenolem przy użyciu badań magnetycznego rezonansu jądrowego (NMR). Autorzy pracy zauważyl, że pierścień aromatyczny 25-tej reszty fenyloalaniny na łańcuchu B monomeru insuliny może przyjmować dwie różne orientacje: jest skierowany do wewnątrz cząsteczki heksameru (typ 1), lub na zewnątrz (typ 2)⁶. Autorzy wyróżnili dwie struktury heksameru:

- symetryczną – heksamer zbudowany jest z sześciu identycznych monomerów typu 1,
- niesymetryczną – heksamer zbudowany z pięciu identycznych monomerów typu 1 i jednego różniącego się od pozostałych typu 2.

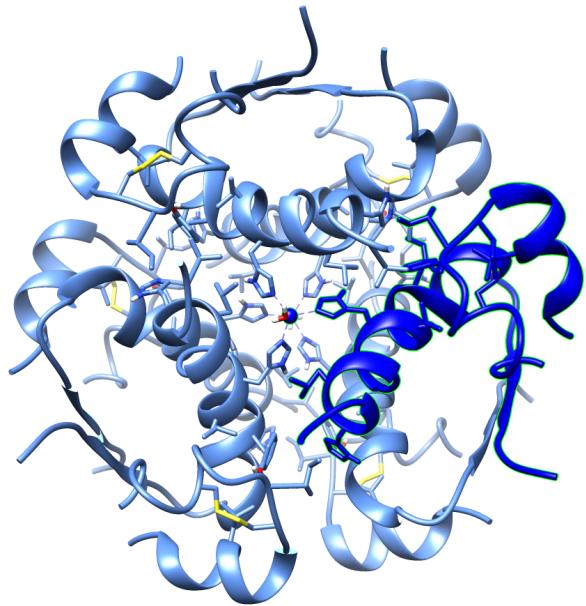
W wyniku dokowania symetrycznego uzyskałam strukturę idealnie symetryczną, dlatego spośród przedstawionych modeli wybrałam model nr 1, który przedstawia symetryczną postać heksameru insuliny. Wyekstrahowany z niego monomer wykorzystałam jako plik wejściowy programu SAM. Poniżej znajduje się rekord 1AIY przedstawiony w programie Chimera (Rysunek 1).



Rysunek 1: Rekord 1AIY zdeponowany w bazie PDB, przedstawiający 10 modeli heksameru insuliny. W każdym z nich można wyróżnić trzy podjednostki - dimery, które wspólnie tworzą heksamer. W centrum cząsteczki znajdują się dwuwartościowe jony cynku. Dodatkowo widoczna jest obecność pierścieni fenyloowych, które stabilizują strukturę.

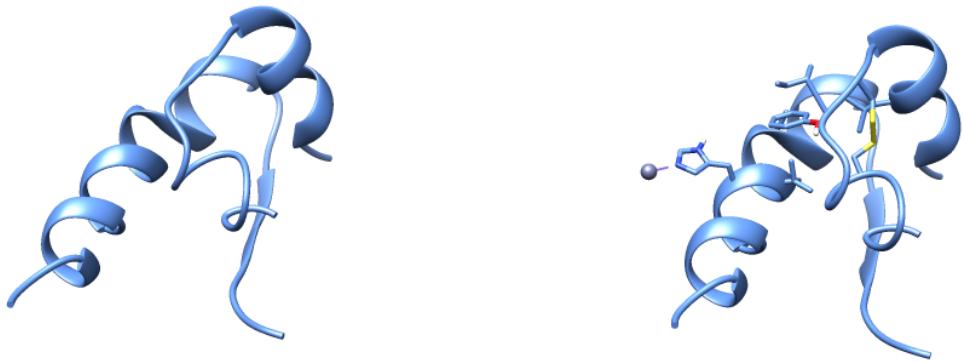
Ekstrakcja monomeru

W celu sprecyzowania, który monomer posłuży mi w dalszej części dokowania, skorzystałem z programu Chimera, który umożliwił mi obserwowanie graficznej reprezentacji modelu. Posłużyłem się wizualizacją pierwszej struktury kompleksu z pliku PDB, z której wyekstrahowałem monomer insuliny. Poniżej przedstawiłem model heksameru z zaznaczonym obszarem stanowiącym interesujący mnie monomer (Rysunek 2).



Rysunek 2: Pierwsza struktura rekordu 1AIY. Zaznaczona kolorem granatowym podjednostka stanowi wyekstrachowany monomer, w którego budowę wchodzą łańcuchy A i B przedstawionego heksameru. Wizualizację wykonałam w programie Chimera.

Obecność ligandów, które również są elementami struktury, może wpływać na postać kompleksów. Dokowanie cząsteczek podczas gdy ligand jest zachowany, może więc dać inne wyniki, niż dokowanie struktur jego pozbawionych. Z tego powodu utworzyłam dwie postacie monomeru. Cząsteczkę fenolu oraz kation cynku, traktowane jako ligandy, zachowałam w strukturze pierwszego monomeru (monomer 2). Drugi model cząsteczki jest ich pozbawiony (monomer 1) i przedstawia tylko strukturę łańcuchów A i B. Poniżej znajdują się uzyskane modele (Rysunek 3).



Rysunek 3: Modele monomerów insuliny uzyskanych z pierwszej struktury rekordu 1AIY: model pozbawiony ligandów (po lewej) oraz model, w którym zachowałam ligandy (po prawej).

Dokowanie, wykorzystane narzędzia

W procesie przewidywania struktur zarówno dimeru, jak i heksameru, wykorzystałem program SAM w wersji 2.0.4, uruchomiony pod kontrolą systemu operacyjnego Linux Ubuntu 16.04 LTS. Używany przeze mnie sprzęt komputerowy posiadał czterordzeniowy procesor Intel® Core™ i5-7300HQ 2.50GHz oraz kartę graficzną Intel HD Graphics 630 (Kaby Lake GT2).

Dokowanie dimeru

Posiadając modele monomerów insuliny, dla każdego z nich przeprowadziłem proces dokowania. Moim celem było przewidzenie struktury dimeru insuliny. Wybrany przeze mnie typ symetrii dimeru to C2. Program SAM przewidując struktury dimerów insuliny zawierające ligandy, znalazł 2 702 336 ich orientacji o podanym typie symetrii w czasie 14.56 sekund. Natomiast dla dimerów pozbawionych ligandów program przetestował taką samą liczbę orientacji w 15.39 sekund. Cały proces dokowania dla struktury wolnej od ligandów i zawierającej ligandy, trwał odpowiednio 2.27 minuty oraz 2.24 minuty.

Spośród wszystkich znalezionych konfiguracji, program SAM zapisał 20 najlepszych o najniższej energii w postaci plików PDB zarówno dla dimerów z ligandami, jak i bez. Łącznie otrzymałam 40 plików PDB.

Przewidywanie struktury heksameru insuliny

Dokowanie sześciu monomerów

Sześć wyekstrahowanych monomerów insuliny zadokowałem przy użyciu programu SAM. W celu utworzenia struktury symetrycznego heksameru insuliny z identycznych podjednostek, określiłem typ symetrii jako D3.

Dokowanie przeprowadziłem zarówno dla monomeru zawierającego ligandy, jak i dla struktury monomeru pozbawionej ligandów. Program SAM przeprowadził dokowanie monomerów w czasie 4 minut i 12 sekund (struktura wolna od ligandów) oraz w czasie 9.23 minut (struktura uwzględniająca ligandy).

Dokowanie trzech dimerów

Do przewidzenia kompleksu heksameru insuliny wykorzystałem także wcześniej uzyskane modele dimerów. Po wybraniu pliku PDB z najlepszą przewidzianą przez mnie strukturą dimeru insuliny, użyłem go jako plik wejściowy programu SAM. Typ symetrii, w przypadku dokowania trzech dimerów, określiłem jako C3.

Wyniki

Po przeprowadzeniu procesu dokowania przy użyciu programu SAM, wybrałem najlepsze modele przedstawiające heksamer insuliny. Po każdym dokowaniu, którego wynikiem była struktura dimeru jak i heksameru, zostało wygenerowane 20 plików PDB. Dodatkowo, w osobnym pliku, zostały zawarte pomiary RMSD pomiędzy strukturami wymodelowanymi przez program SAM, a natywną strukturą określoną eksperymentalnie. Jako strukturę natywną traktowałem zarówno sam rekord 1AIY, jak i wyekstrahowane podjednostki (dimery).

Dimer insuliny

Strukturę dimeru insuliny przewidziałam przy użyciu wcześniej wyekstrahowanych monomerów – monomer 1, w którym wykluczyłem obecność ligandów i model 2, w którym zostały one zachowane. Uzyskałem więc dwie struktury, które porównałem ze strukturą natywną dimeru insuliny. Pierwszy z nich nie posiada ligandów (**dimer 1**), drugi uwzględnia ich obecność (**dimer 2**).

Dimer 1

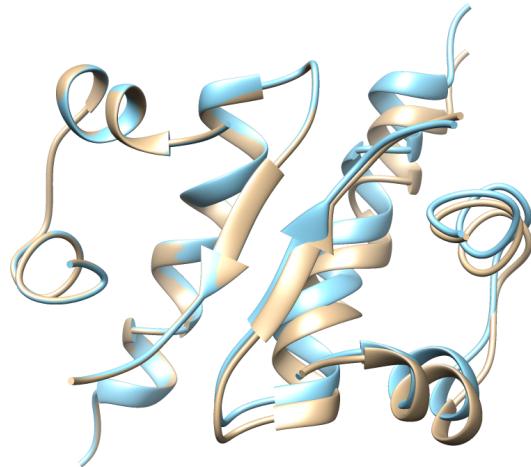
Pierwszy rozpatrywany przeze mnie przypadek to dimer pozbawiony ligandów. Przy użyciu programu SAM otrzymałem 20 plików PDB zawierających struktury przewidzianych dimerów insuliny, 5 pośród nich otrzymało wartości RMSD niższe niż 10.00 Å. Siódmy w kolejności model jest najbliższym strukturze natywnej z wartością RMSD równą 2.39 Å i energią konformacji równą -396.48 kJ/mol i nie jest to wartość najniższa – kompleks najbardziej zbliżony do natywnej

struktury dimeru insuliny nie jest kompleksem najkorzystniejszym energetycznie. W poniższej tabeli (Tabela 1) przedstawiam wartości energii konformacji oraz wskaźnika RMSD dla każdej z 20 struktur:

Numer struktury	Energia konformacji [kJ/mol]	RMSD [Å]
1	-539.12	29.28
2	-488.99	4.66
3	-431.69	18.94
4	-424.07	10.05
5	-417.16	8.40
6	-407.74	19.31
7	-396.48	2.39
8	-383.47	31.13
9	-375.45	12.22
10	-371.36	9.92
11	-367.69	22.86
12	-367.51	28.32
13	-366.30	8.26
14	-364.78	16.40
15	-358.00	27.43
16	-356.87	11.17
17	-355.43	18.18
18	-353.32	16.17
19	-352.41	17.82
20	-348.64	15.34

Tabela 1: Wyniki dokowania monomeru pozabawionego ligandów w celu uzyskania dimeru insuliny (dimer 1). Na 20 przewidzianych przez program SAM konfiguracji, model nr 7 posiadał najniższą wartość wskaźnika RMSD.

W programie PyMOL dokonałam pomiaru RMSD dla struktury numer 5 korzystając z komendy *align*. Otrzymany przez mnie wynik to 1.094 Å. Po znalezieniu struktury o najniższej wartości wskaźnika RMSD, za pomocą programu Chimera dokonałam nałożenia struktur, co umożliwiła mi funkcja *MatchMaker*. Poniżej znajduje się grafika ilustrująca dopasowanie struktury wymodelowanej do natywnej (Rysunek 4).



Rysunek 4: Nałożenie wymodelowanej struktury dimeru insuliny o najniższej wartości RMSD nieposiadającej ligandów ze strukturą natywną dimeru insuliny. Kolorem niebieskim zaznaczylem strukturę wymodelowaną, a kolorem beżowym natywną.

Dimer 2

Podczas przewidywania struktur dimeru w obecności ligandów, program SAM obliczył 20 konformacji, z czego 7 z nich miało wartości RMSD niższe niż 10.00 Å. Poniższa tabela (Tabela 2) przedstawia wyniki przeprowadzonego dokowania monomeru w obecności ligandów:

Numer struktury	Energia konformacji [kJ/mol]	RMSD [Å]
1	-529.57	4.67
2	-458.40	18.72
3	-430.91	19.33
4	-427.68	24.15
5	-417.08	9.99
6	-414.10	8.93
7	-411.26	2.40
8	-400.21	16.62
9	-389.66	8.23
10	-384.78	20.11
11	-380.26	16.22
12	-377.09	12.16
13	-375.79	19.41
14	-373.34	9.72
15	-372.79	25.01
16	-371.42	24.01
17	-368.66	19.56
18	-364.48	17.52
19	-361.48	1.51
20	-361.46	16.40

Tabela 2: Wyniki dokowania monomeru w obecności ligandów w celu uzyskania dimeru insuliny (dimer 2). Na 20 przewidzianych przez program SAM konfiguracji, model nr 19 posiadał najniższą wartość wskaźnika RMSD.

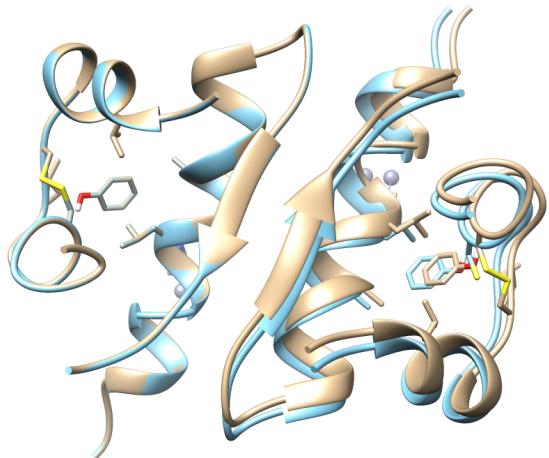
Najlepszą (najniższą) wartość RMSD uzyskała struktura o numerze 19 i wynosiła ona 1.51 Å. Energia konformacji w tym przypadku to -361.48 kJ/mol. Przewidziany model nie charakteryzuje się najniższą energią konformacji, która wynosi -529.57 kJ/mol. Kompleks o najwyższej wartości wskaźnika RMSD, a tym samym najbardziej podobny do struktury krystalograficznej dimeru insuliny, nie jest najkorzystniejszym energetycznie wariantem.

Dodatkowy pomiar RMSD wykonałam przy użyciu programu PyMOL. Ta sama struktura otrzymała wartość RMSD równą 0.594 Å. Podobnie jak w przypadku dimeru 1, nałożyłam strukturę

o najkorzystniejszej wartości RMSD (model 19, Rysunek 5) ze strukturą natywną w programie Chimera, czego rezultat znajduje się poniżej na Rysunku 6.



Rysunek 5: Otrzymana struktura dimeru insuliny w wyniku dokowania monomeru w obecności ligandów. Struktura ta charakteryzowała się najniższą wartością wskaźnika RMSD pośród wszystkich przewidzianych konfiguracji i jest modelem najbliższy zblizonym do postaci natywnej dimeru insuliny.

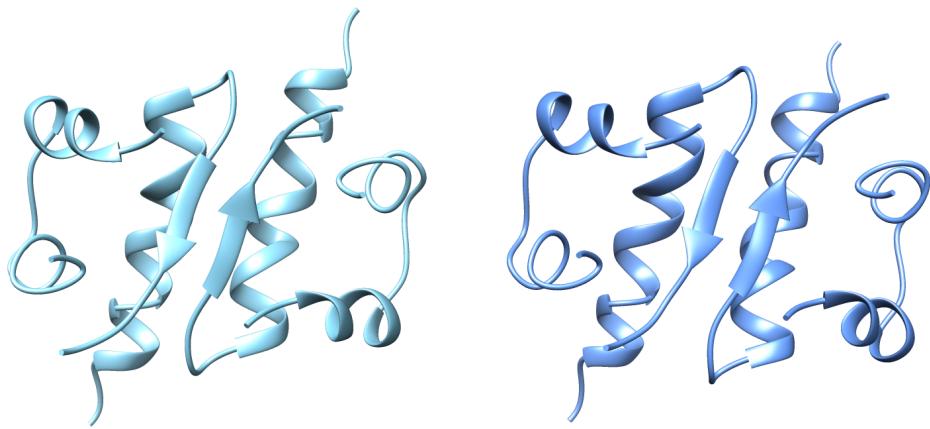


Rysunek 6: Nałożenie wymodelowanej struktury dimeru insuliny niepozbawionego ligandów z natywną strukturą dimeru insuliny. Kolorem beżowym zaznaczona jest struktura natywna, kolor błękitny odpowiada strukturze wymodelowanej.

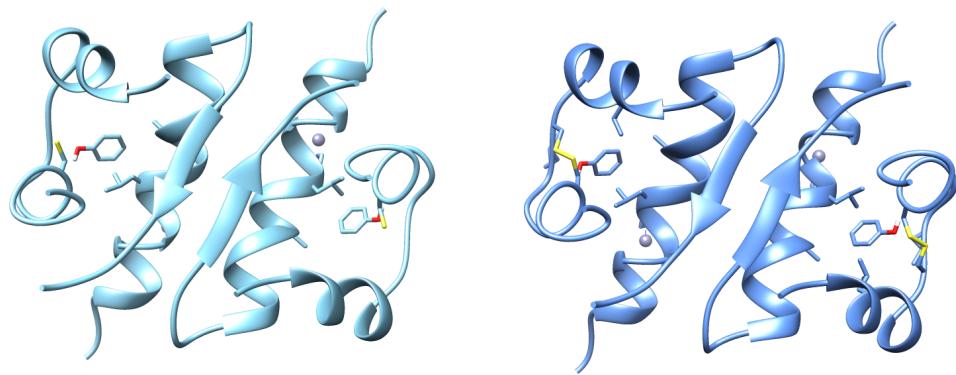
Wybór dimeru

Otrzymawszy strukturę dimeru zarówno posiadającego ligandy jak i pozbawionego ligandów, porównałem wartości RMSD dla struktur reprezentujących najlepszy wynik przeprowadzonych procesów dokowania. Struktura dimeru posiadającego ligandy o numerze 19 jest najbliższa strukturze natywnej i posłużyła mi jako podjednostka do przewidzenia struktury kompleksu heksameru insuliny. Założyłem, że otrzymany przeze mnie model heksameru insuliny będzie tym samym bardziej zbliżony do natywnej struktury, gdy wykorzystam dimer charakteryzujący się najniższą wartością wskaźnika RMSD.

Porównanie wymodelowanej przeze mnie struktury dimeru insuliny pozbawionego ligandów, jak i struktury dimeru insuliny przy obecności ligandów, ze strukturą natywną znajduje się odpowiednio na Rysunku 7 i Rysunku 8 poniżej.



Rysunek 7: Porównanie modeli dimeru insuliny pozabawionych ligandów: monomer naturalny (po lewej) i monomer uzyskany w procesie dokowania (po prawej)



Rysunek 8: Porównanie modeli dimeru insuliny posiadających ligandy: dimer naturalny (po prawej) i dimer wymodelowany (po lewej).

Heksamer insuliny

Przewidywanie struktury heksameru przy użyciu dimerów

Do przewidywania najkorzystniejszej konformacji heksameru insuliny użyłam wymodelowanego dimeru niepozbawionego ligandów z wcześniejszego dokowania. Dimer ten charakteryzował się najlepszą wartością RMSD, dlatego też wybrałam go jako plik wejściowy programu SAM.

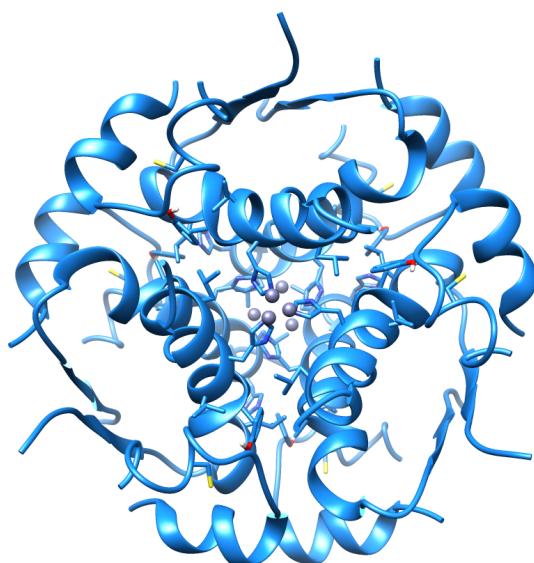
Symetrię kompleksu heksameru insuliny określiłam jako C3. W poniższej tabeli znajdują się wyniki dla 20 struktur heksameru przewidzianych przez program SAM:

Numer struktury	Energia konformacji [kJ/mol]	RMSD [\AA]
1	-410.37	31.50
2	-410.03	3.55
3	-398.88	31.50
4	-371.28	4.58
5	-361.88	32.49
6	-354.27	32.42
7	-345.80	5.74
8	-342.95	30.96
9	-335.34	32.63
10	-319.01	30.51
11	-319.01	33.85
12	-318.08	31.39
13	-316.32	17.51
14	-315.26	32.64
15	-314.10	34.79
16	-312.58	36.81
17	-305.95	34.80
18	-304.35	5.38
19	-302.80	31.74
20	-302.76	28.94

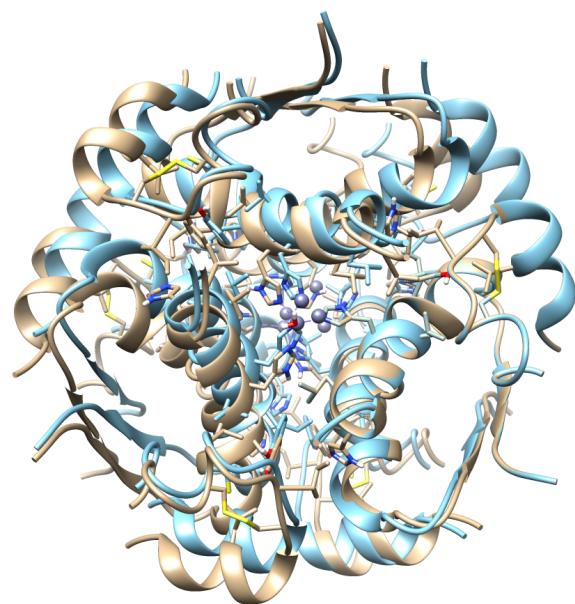
Tabela 3: Wyniki dokowania dimeru w obecności ligandów w celu uzyskania heksameru insuliny. Na 20 przewidzianych przez program SAM konfiguracji, model nr 2 posiadał najniższą wartość wskaźnika RMSD.

Kierując się wartościami wskaźników RMSD policzonymi przez program SAM, modelem najbardziej zbliżonym do struktury natywnej jest struktura numer 2, o wartości RMSD równej 3.55 Å i energii konformacji równej -410.03 kJ/mol. Wartość energii konformacji jest, zaraz po energii konformacji uzyskanej przez strukturę nr 1 (-410.37 kJ/mol), najkorzystniejszym wynikiem. Model nr 2 posiada więc najniższą wartość wskaźnika RMSD przy niskiej energii konformacji kompleksu.

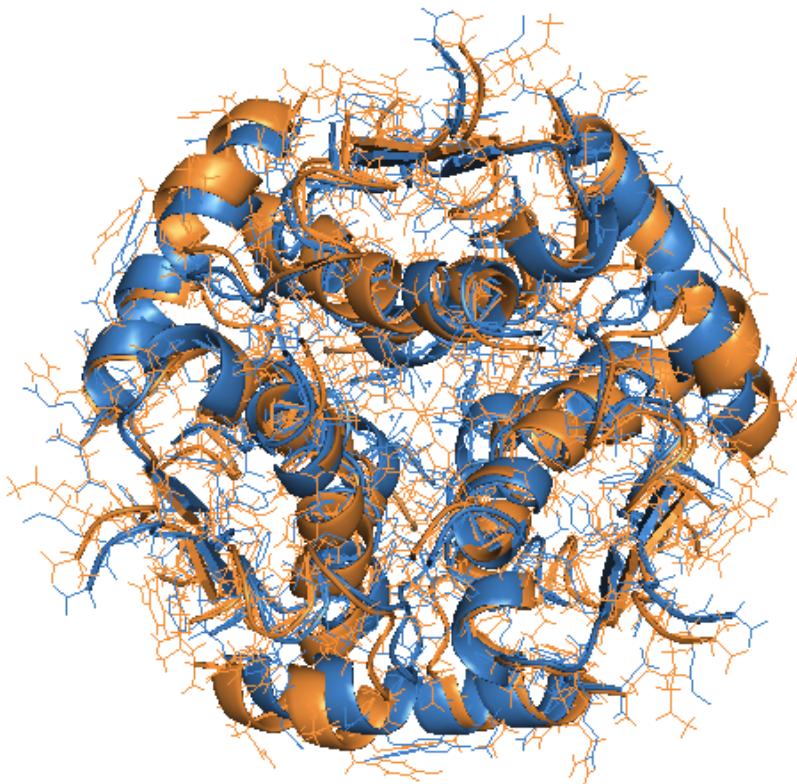
Powtarzając pomiar wartości RMSD w programie PyMOL, uzyskałam wynik 1.98 Å. Ilustracje struktury otrzymanego modelu heksameru insuliny wykonane przeze mnie w programie Chimera przedstawiam poniżej na Rysunkach 8, 10 i 11.



Rysunek 9: Przewidziany model heksameru insuliny, uzyskany w wyniku dokowania trzech dimerów przy symetrii C3.



Rysunek 10: Nałożenie struktury przewidzianego heksameru insuliny (kolor błękitny) z modelem natywnym (kolor beżowy) w programie Chimera.



Rysunek 11: Nałożenie struktury otrzymanego heksameru insuliny (kolor błękitny) z natywnym modelem (kolor pomarańczowy) w programie PyMOL.

Przewidywanie struktury heksameru przy użyciu monomerów

Drugim sposobem na uzyskanie struktury symetrycznego heksameru insuliny jest bezpośrednie dokowanie sześciu monomerów insuliny. Wykorzystałam podczas dokowania wcześniej uzyskane przeze mnie modele (Rysunek 3): monomer 1, w którym usunęłam ligandy oraz monomer 2, w którym ligandy zostały zachowane. Dokując odpowiednie monomery, uzyskałam dwie struktury heksamerów insuliny: heksamer posiadający ligandy (heksamer 2) oraz pozbawiony ligandów, (heksamer 1), podobnie jak w przypadku przewidywania struktury dimeru. Typ symetrii, zarówno przy dokowaniu monomeru 1, jak i monomeru 2, określiłam jako D3.

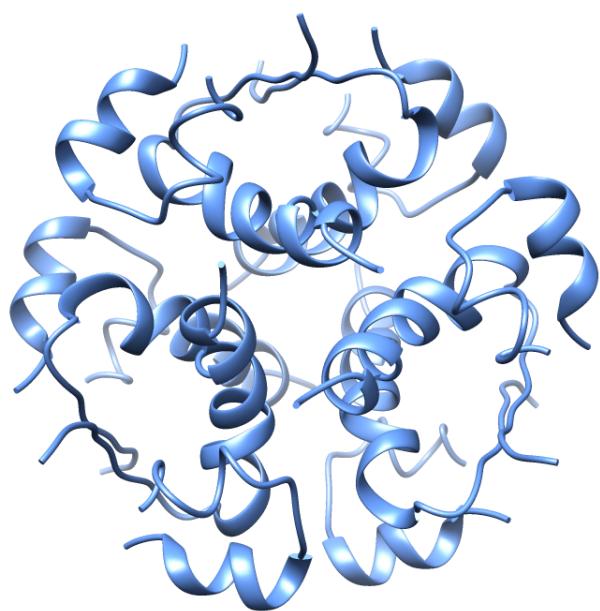
Heksamer 1

Strukturę heksameru 1 uzyskałam dokując monomery (monomer 1) pozbawione ligandów. W poniżej tabeli (Tabela 4) przedstawiam wyniki dla 20 struktur przewidzianych przez program SAM:

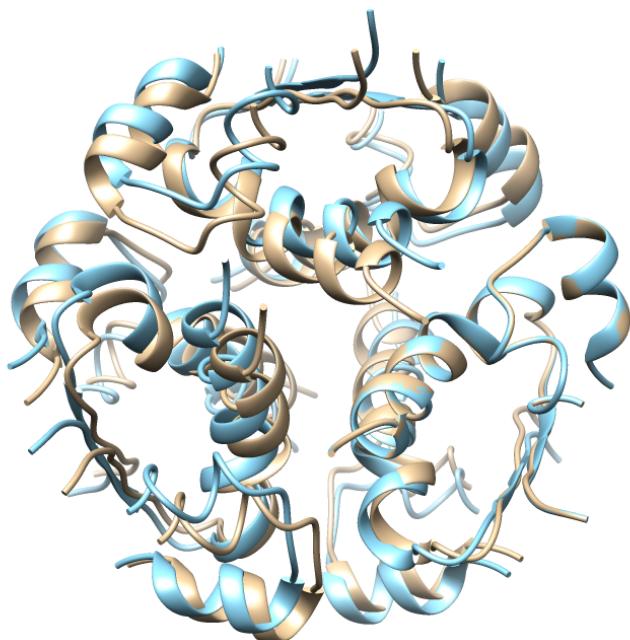
Numer struktury	Energia konformacji [kJ/mol]	RMSD [Å]
1	-606.30	28.86
2	-616.61	24.51
3	-624.90	5.09
4	-593.32	22.49
5	-592.17	16.18
6	-592.28	2.32
7	-565.75	28.16
8	-571.69	28.95
9	-559.23	23.60
10	-541.52	25.21
11	-545.25	32.24
12	-522.42	28.34
13	-529.20	22.84
14	-497.38	20.67
15	-473.13	23.61
16	-456.14	22.26
17	-448.31	29.34
18	-484.33	27.87
19	-409.93	19.81
20	-461.59	28.25

Tabela 4: Wyniki dokowania monomerów pozabawionych ligandów w celu uzyskania heksameru insuliny. Na 20 przewidzianych przez program SAM konfiguracji, model nr 6 posiadał najniższą wartość wskaźnika RMSD.

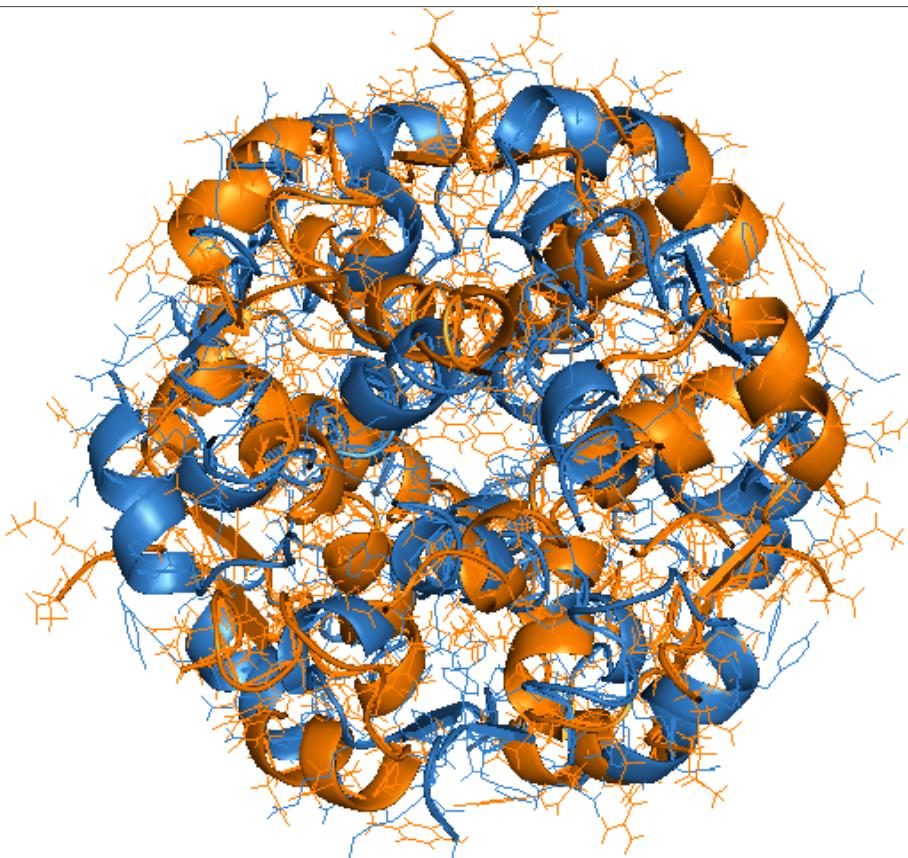
Biorąc pod uwagę wartości RMSD modeli wygenerowanych przez program SAM, najlepszą przewidzianą strukturą jest struktura nr 6, z wynikiem RMSD 2.32 Å. Podobnie jak dla pozostałych modeli, obliczyłem wartość RMSD w programie PyMOL. W odróżnieniu od pomiarów RMSD dla wcześniej uzyskanych przeze mnie struktur, tym razem otrzymałem wartość wskaźnika RMSD na poziomie 17.32 Å. Poniżej przedstawiam nałożenie struktury nr 6 z natywnym modelem heksameru insuliny (Rysunek 12, 13) oraz sam uzyskany model (Rysunek 11):



Rysunek 12: Struktura heksameru insuliny uzyskana w wyniku dokowania monomerów pozbawionych ligandów przy określonym typie symetrii D3.



Rysunek 13: Nałożenie struktury w programie Chimera o najniższym wskaźniku RMSD uzyskanej przy dokowaniu monomerów w obecności ligandów przy wybranej symetrii D3. Kolorem beżowym zaznaczona jest struktura natywna heksameru insuliny, kolorem niebieskim - uzyskany model.



Rysunek 14: Nałożenie struktury natywnej (kolor pomarańczowy) i struktury wymodelowanej przy wykorzystaniu monomerów pozbawionych ligandów (kolor niebieski), wykonane przeze mnie w programie PyMOL. Wynik ten stanowczo różni się od wcześniej wykonanych porównań natywnego heksameru insuliny z wymodelowaną strukturą – modele pokrywają się w niewielkim stopniu pomimo uzyskania wyniku RMSD równym 2.35 \AA obliczonego przez program SAM.

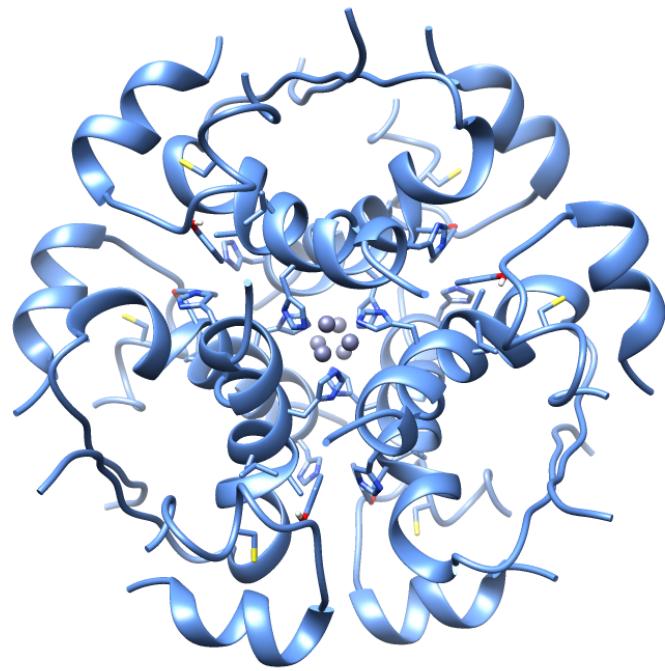
Heksamer 2

Aby uzyskać kolejny model, użyłam monomeru 2, w którego strukturze zachowałam obecność ligandów. Program SAM wygenerował 20 struktur, a wyniki przedstawione są poniżej w tabeli (Tabela 5):

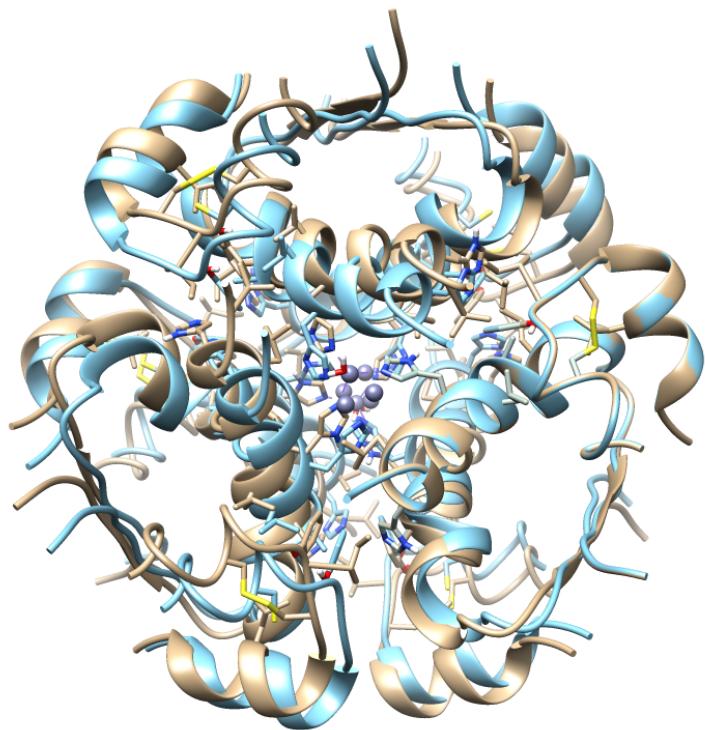
Numer struktury	Energia konformacji [kJ/mol]	RMSD [\AA]
1	-669.12	23.51
2	-646.78	17.24
3	-623.02	2.25
4	-594.54	4.44
5	-535.95	28.24
6	-540.58	22.88
7	-535.66	28.09
8	-527.75	32.20
9	-505.94	23.90
10	-442.47	26.77
11	-440.03	24.01
12	-494.38	24.01
13	-497.68	20.66
14	-428.71	27.16
15	-484.89	29.66
16	-475.41	22.19
17	-477.44	28.40
18	-456.33	26.15
19	-454.96	30.36
20	-435.77	19.61

Tabela 5: Wyniki dokowania monomeru w obecności ligandów w celu uzyskania heksameru insuliny. Na 20 przewidzianych przez program SAM konfiguracji, model nr 3 posiadał najniższą wartość wskaźnika RMSD.

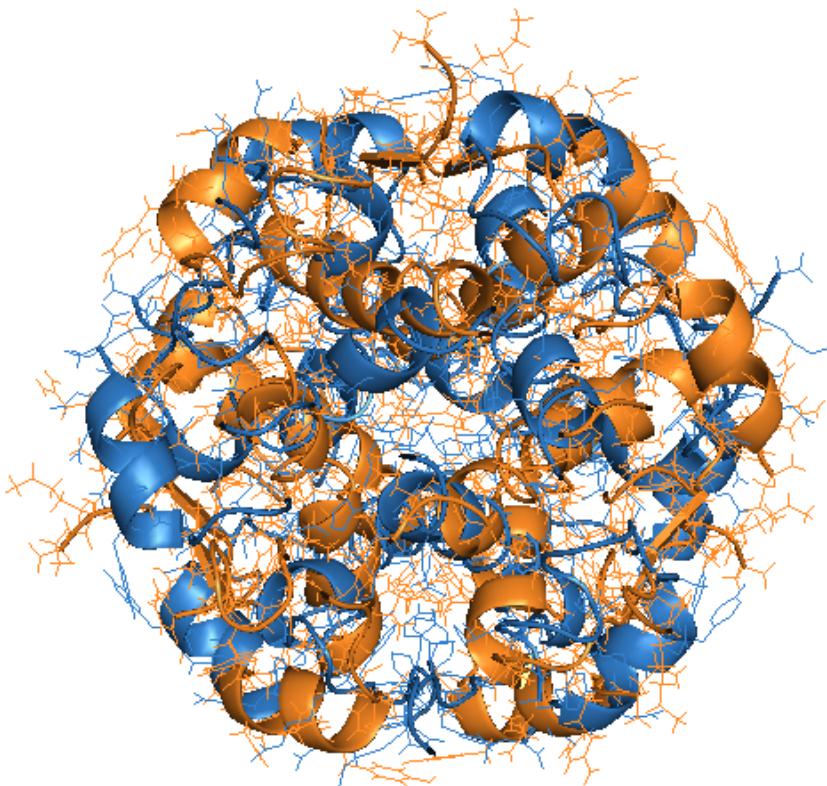
Struktura nr 3 charakteryzowała się najniższą wartością wskaźnika RMSD równą 2.25 \AA . Większość modeli przewidzianych przez program SAM nie pokrywała się w znacznym stopniu ze strukturą uzyskaną na drodze krystalografii (natywną). Tylko 2 na 20 modeli osiągnęło wartość wskaźnika RMSD niższą niż 10.00 \AA : model nr 3 z wynikiem 2.25 \AA i model nr 4 z wynikiem 4.44 \AA . Korzystając w funkcji programu PyMOL, obliczyłam wartość wskaźnika RMSD dla struktury nr 3 i otrzymałam wskaźnik równy 17.26 \AA . Wybrana przeze mnie struktura nr 3 (Rysunek 14) i porównanie jej z natywnym modelem (Rysunki 15, 16 i 17) przedstawione jest poniżej.



Rysunek 15: Model heksameru insuliny uzyskany w wyniku dokowania monomeru w obecności ligandów.



Rysunek 16: Nałożenie struktury wymodelowanej (kolor niebieski) z natywną strukturą heksameru insuliny (kolor beżowy). Porównanie zostało wykonane w programie Chimera.



Rysunek 17: Nałożenie struktury heksameru insuliny wymodelowanej przy wykorzystaniu monomeru niepozbawionego ligandów (kolor niebieski) ze strukturą natywną (kolor pomarańczowy). Ilustracja pokazuje, że pomimo uzyskanego wcześniej wyniku RMSD równego 2.25 \AA , w programie PyMOL (w którym zostało wykonane porównanie) struktury te nie pokrywają się ze sobą w dużym stopniu.

Interpretacja

Analizując otrzymane wyniki zarówno w przypadku przewidywania struktury dimeru, jak i całego heksameru insuliny, wybrałem dwa najlepsze modele:

- dimer – model nr 19 (Rysunek 5) uzyskany w wyniku dokowania monomeru, gdzie obecność ligandów została zachowana, z wynikiem RMSD równym 0.594 \AA ,
- heksamer – model nr 2 (Rysunek 8) uzyskany w wyniku dokowania wcześniejszego dimeru (model nr 19), z wynikiem RMSD równym 1.98 \AA .

Podczas przewidywania struktury symetrycznego heksameru insuliny z wykorzystaniem monomerów, otrzymałem dwa modele charakteryzujące się wartościami RMSD (obliczonymi przez program SAM) równymi 2.25 \AA oraz 2.32 \AA , dokując odpowiednio monomer zawierający ligandy (monomer 2) oraz monomer ich pozbawiony (monomer 1). Pomimo, że powyższe wyniki RMSD są niższe od wartości uzyskanych przez struktury otrzymane przez dokowanie dimerów, odrzuciłem je z powodu zaskakująco wysokich wyników RMSD w programie PyMOL, za pomocą funkcji *align*. Przy obliczaniu RMSD przez program PyMOL dla struktur heksamerów

przewidzianych w wyniku dokowania dimerów jak i samych dimerów insuliny, wartości otrzymane przeze mnie były niższe. W poniższej tabeli (Tabela 8) przedstawiłem otrzymane wartości RMSD dla wybranych struktur:

Model	RMSD [Å]	
	SAM	PyMOL
Dimer, symetria C2		
struktura 7, bez ligandów	2.39	1.094
struktura 19, ligandy	1.51	0.594
Heksamer		
• przewidywanie przy użyciu dimerów, symetria C3		
struktura 2, ligandy	3.55	1.98
• przewidywanie przy użyciu monomerów, symetria D3		
struktura 6, bez ligandów	2.32	17.32
struktura 3, ligandy	2.25	17.26

Tabela 6: Porównanie wartości RMSD wybranych, najlepszych modeli otrzymanych przeze mnie w wyniku dokowania symetrycznego, w zależności od użytego programu.

Podsumowując otrzymane w procesie dokowania symetrycznego wyniki, udało mi się przewidzieć strukturę heksameru insuliny bardzo zbliżoną do swojej postaci uzyskanej na drodze eksperymentalnej (Rysunek 2). Wartość RMSD równa 1.96 Å w przypadku dość dużego kompleksu jakim jest heksamer, jest bardzo dobrym wynikiem. Poprawność uzyskanej struktury potwierdza nie tylko sam wynik RMSD, ale i wizualne porównanie struktur (Rysunek 11). Szukając najlepszej konfiguracji dimeru insuliny, uzyskałem model różniący się od struktury natywnej zaledwie o 0.594 Å, co świadczy nie tylko o prawidłowo określonej symetrii kompleksu, ale i o tym, że modele te są *niemalże identyczne*.

Dyskusja

Niniejsza praca miała na celu przewidzenie struktury heksameru ludzkiej insuliny przy pomocy dokowania symetrycznego – techniki bioinformatycznej modelowania molekularnego. Korzystając tylko z narzędzi komputerowych, udało mi się przewidzieć strukturę heksameru.

Wykorzystany przeze mnie program SAM, uwzględniający symetrię kompleksów, policzył dla każdej z przewidzianych konfiguracji wartość wskaźnika RMSD. Dodatkowo, po wybraniu struktury o najwyższej jego wartości, potwórzyłem pomiar w programie PyMOL. Przy dokowaniu kompleksów z wykorzystaniem monomerów – pozbawionych ligandów oraz takich, gdzie obecność ligandów została zachowana – wartości RMSD liczone przez program PyMOL były znacznie wyższe. Taka rozbieżność wyników jest zastanawiająca. Aby rozwiązać problem, należałooby upewnić się czy postać heksameru jaką przewidziałał program SAM rozpoznawana jest w

odpowiedni sposób przez program PyMOL, tzn. czy poszczególne podjednostki kompleksu porównywane są ze swoimi „odpowiednikami” struktury natywnej, być może programy stosując różną kolejność monomerów, na przykład uwzględniają (lub nie) kolejność czytania monomerów z plików PDB. Program SAM utworzony niezależnie od programu PyMOL, może posiadać inne funkcje oceniające, których wynik uzależniony jest od innych parametrów przewidzianych kompleksów, programy te mogą stosować także różne algorytmy liczenia wartości RMSD. Program SAM może zakładać większą „swobodę” kompleksów przewidzianych przez niego podczas porównywania ich z postacią natywną wynikającą ze sposobu traktowania przez ów program cząsteczek jako chmur elektronowych.

Jednak sytuacja taka wystąpiła tylko w momencie dokowania monomerów. Przy próbie przewidzenia kompleksu ludzkiej insuliny dokując wcześniej uzyskane przeze mnie dimery, program PyMOL znacznie polepszył wynik RMSD. Sugeruje to, że tylko kompleksy heksamerów, których formowanie poprzedziła agregacja monomerów w dimery, są prawidłowo zadokowanymi (zbudowanymi) kompleksami lub kolejność podjednostek kompleksy została w tym wypadku zachowana w relacji do struktury natywnej. Aby rozwiązać ten problem, należałoby przeprowadzić dokowanie za pomocą innych programów dokujących lub skorzystać z innego oprogramowania wyliczającego wartość wskaźnika RMSD. Należy również rozważyć próbę porównywania struktury natywnej i przewidzianego modelu heksameru, analizując odpowiadające im składniki kompleksu i otrzymywane wówczas wartości wskaźnika RMSD (nakładając oddzielnie wybrane dimery lub trimery stanowiące części składowe przewidzianej struktury heksameru insuliny).

Przy próbie przewidzenia postaci heksameru insuliny poprzez dokowanie dimerów, wykorzystałam ten, w którym występowały ligandy (dwuwartościowy kation cynku, który ponadto oddziaływa z pozostałymi dwoma pierścieniami aromatycznymi reszty histydyny dłuższej helisy) oraz fenol (wykorzystany w celu stabilizacji cząsteczki). Dimer pozbawiony ligandów, posiadał mniejszą wartość RMSD – obecność ligandów w strukturze dimeru jak i heksameru przewidzianych modeli wpływa na dopasowanie do struktury natywnej, czego dowodem jest niższa wartość wskaźnika RMSD osiągnięta właśnie przez te modele. Wydaje się, że szczególną rolę w formowaniu się kompleksu ma kation cynku, znajdujący się w centralnych częściach cząsteczki, gdyż fenol zastosowany został w warunkach laboratoryjnych tylko w celu ograniczenia ruchliwości cząsteczki.

Stosowana przeze mnie metoda bioinformatyczna – dokowanie symetryczne – pozwoliła na przewidzenie idealnie symetrycznego modelu heksameru ludzkiej insuliny, którego typ symetrii określony jest jako D3. Bez konieczności sięgania po metody eksperymentalne, takie jak magnetyczny rezonans jądrowy (NMR) czy krystalografia, udało mi się przewidzieć jego strukturę z bardzo dobrym wynikiem RMSD. W celu kontynuacji badań należałoby skorzystać z innych, dostępnych programów służących do dokowania symetrycznego i porównać otrzymane struktury symetrycznego heksameru ludzkiej insuliny.

Bibiografia

- ¹ Sanger i inni, *The Amino-acid Sequence in the Phenylalanyl Chain of Insulin*, 1952
- ² Protein Data Bank, Insulin: *An early structure with sweet success*
- ³ P. Sieber i inni, *Total synthesis of human insulin under directed formation of the disulfide bonds*, 1974
- ⁴ Melinda Roys i inni, *Spectroscopic Signatures of the T to R Conformational Transition in the Insulin Hexamer*, 1989
- ⁵ David W. Ritchie i Sergei Grudinin, *Spherical polar Fourier assembly of protein complexes with arbitrary point group symmetry*, 2016
- ⁶ X. Chang i inni, *Solution Structures of the R6 Human Insulin Hexamer*, 1997
- ⁷ S. Pagadala i inni, *Software for molecular docking: a review*, 2017
- ⁹ David W. Ritchie, V. Venkatraman, *Ultra-fast FFT protein docking on graphics processors*, 2010
- ¹⁰ Sam Protein Structure Symmetry Assembler, www.sam.loria.fr