

WYSOKOROZDZIELCZA KRYSTALOGRAFIA MAKROMOLEKUŁ

HIGH RESOLUTION CRYSTALLOGRAPHY OF MACROMOLECULES

Mirosław Gilski

*Zakład Krystalografii, Wydział Chemii UAM
ul. Grunwaldzka 6, 60-780 Poznań
Centrum Badań Biokrystalograficznych, Instytut Chemii Bioorganicznej PAN
ul. Noskowskiego 12/14, 61-704 Poznań
e-mail: mirek@amu.edu.pl*

Abstract

Wprowadzenie

1. Krystalizacja makromolekuł dla badań wysokorozdzielczych
2. Rejestracja i procesowanie danych dyfrakcyjnych
3. Rozwiązywanie struktur wysokorozdzielczych
 - 3.1. Metody bezpośrednie
4. Budowanie modelu struktury
5. Udokładnianie
 - 5.1. Modelowanie obszaru rozpuszczalnika
 - 5.2. Alternatywne konformacje
 - 5.3. Atomy wodoru w strukturach makromolekuł
 - 5.4. Standardowe parametry stereochemiczne
6. Deformacja gęstości elektronowej

Podsumowanie

Piśmiennictwo cytowane

Dr Mirosław Gilski jest absolwentem fizyki na Uniwersytecie im. A. Mickiewicza w Poznaniu. Po studiach rozpoczął pracę w Zakładzie Krystalografii na Wydziale Chemii Uniwersytetu im. A. Mickiewicza w Poznaniu, gdzie obecnie jest zatrudniony na stanowisku adiunkta. Oprócz tego pracuje, od momentu jego założenia, w Centrum Badań Biokrystalograficznych w Instytucie Chemii Bioorganicznej PAN w Poznaniu, które było pierwszym laboratorium krystalografii białek w Polsce. Jego zainteresowania naukowe koncentrują się na metodycznych aspektach wysokorozdzielczej krystalografii makromolekuł, metodach procesowania i automatycznej rejestracji synchrotronowych danych dyfrakcyjnych oraz strukturalnych badaniach białek i kwasów nukleinowych o ekstremalnie wysokiej rozdzielczości.

ABSTRACT

The recent developments and availability of modern third-generation synchrotron radiation facilities have a huge impact on macromolecular X-ray crystallography. In connection with a number of methodological improvements and new crystallographic software ranging from data processing to refinement, a unique opportunity has arisen to determine the macromolecular structures with unprecedented high resolution and quality, at a level traditionally reserved for small molecules.

At this resolution, individual atoms are clearly resolved and fine details of the structures become visible directly in the electron density maps.

The great importance of such structures is the possibility of having broader insights into macromolecule function. At very high resolution, hydrogen atoms can be seen in electron density maps and the detailed information about the protonation states of catalytically important residues can be studied, what often is critical for full understanding of enzymatic reactions.

Atomic resolution gives the opportunity for clear definition of multiple conformations, although the proportion of disordered residues is higher at higher resolution, and the disorder is seen as distinct alternative conformations.

Water in macromolecular crystal plays an important role in macromolecule's function and stabilization. Ultrahigh resolution data allows to refine water molecules with anisotropic displacement parameters and refine them with fractional occupancies. In this situation analyzing the subtle hydrogen bond network, involving precisely located water molecules, is possible.

Atomic resolution structures can be refined without or with only very weak stereochemical restraints. Macromolecular models refined at ultrahigh resolutions, for well ordered structures, can be used for validation and improvement of stereochemical restraint libraries, commonly used during refinement of macromolecular structures.

Keywords: macromolecular crystallography, atomic resolution structures, stereochemical restraints, charge density

Słowa kluczowe: krystalografia makromolekuł, struktury o rozdzielczości atomowej, więzy stereochemiczne, gęstość ładunku

WPROWADZENIE

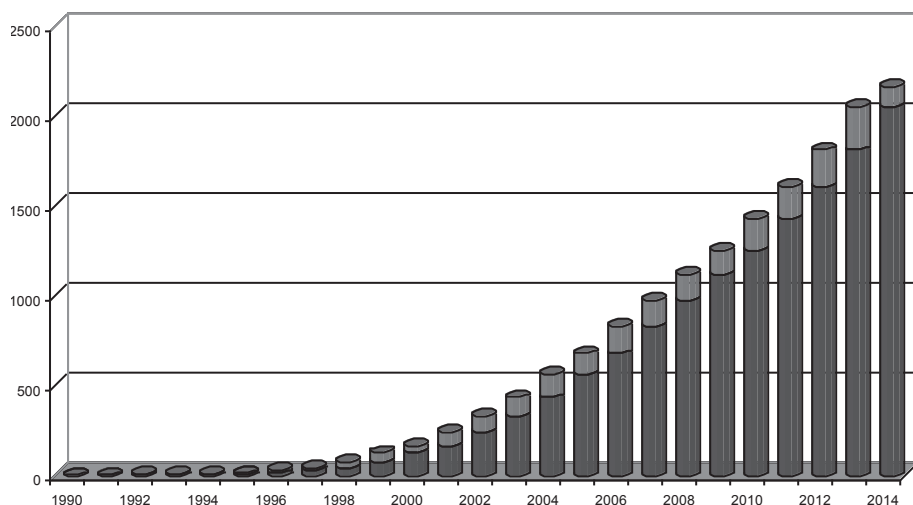
W krystalografii małych cząsteczek termin „rozdzielczość” jest używany dość rzadko. Najczęściej zakłada się, że dane dyfrakcyjne zostały zarejestrowane aż do maksymalnego kąta 2θ osiągalnego podczas pomiaru na dyfraktometrze z miedzianą lampą rentgenowską, co odpowiada rozdzielczości $\sim 0,8$ Å. Większość kryształów makromolekuł, takich jak białka czy kwasy nukleinowe, nie pozwala na uzyskanie tak dobrej dyfrakcji, głównie z powodu znacznie większych rozmiarów cząsteczek i wysokiej zawartości nieuporządkowanego rozpuszczalnika. W czasie ostatnich dwudziestu lat rozdzielczość i jakość krystalograficznych struktur makromolekularnych uległa znacznej poprawie, głównie z powodu różnych usprawnień technicznych, począwszy od lepszych technik ekspresji i krystalizacji, poprzez zastosowanie wysokowydajnych synchrotronowych źródeł promieniowania rentgenowskiego do pomiarów dyfrakcji a kończąc na znacznie ulepszonych metodach i algorytmach zaimplementowanych w programach komputerowych.

Stworzone zostały nowe możliwości określania struktur makrocząsteczek z niepotykaną dotąd wysoką rozdzielczością i precyzją, która tradycyjnie zarezerwowana jest dla krystalografii małych cząsteczek. Takie struktury są niezwykle wartościowe, ponieważ mogą się przyczynić do zrozumienia wielu subtelnych efektów strukturalnych i mogą zaoferować nowe spojrzenie na chemię strukturalną makrocząsteczek.

Rok 2014 jest szczególnie dla krystalografów nie tylko z powodu ogłoszonego przez ONZ Międzynarodowego Roku Krystalografii. Również w tym roku liczba struktur zdeponowanych w bazie Protein Data Bank (PDB) [1] gromadzącej informacje o strukturach makromolekuł przekroczy 100 tysięcy.

Termin „wysoka rozdzielczość” w krystalografii makromolekuł nie jest do końca jednoznaczny a struktury, które jeszcze kilkanaście lat temu były nazywane wysokorozdzielczymi dzisiaj możemy zaliczyć do średniorozdzielczych. W pracy tej określenie „wysokorozdzielczy” używane jest w stosunku do rozdzielczości atomowej, zdefiniowanej przez Sheldrick'a [2] jako rozdzielczość lepsza niż $1,2$ Å. Przy takiej rozdzielczości możliwe jest rozdzielenie poszczególnych atomów niewodorowych na mapach gęstości elektronowej. Natomiast termin „rozdzielczość subatomowa” odnosi się do struktur, dla których zarejestrowano dane o rozdzielczości $0,8$ Å lub lepszej, pozwalającej badać znacznie bardziej subtelne efekty związane np. z lokalizacją położeń atomów wodoru lub deformacją rozkładu gęstości elektronowej.

Wśród struktur zdeponowanych w bazie PDB tylko około 2% może być zaliczone do struktur o rozdzielczości atomowej, mniej niż 0,6% do struktur o rozdzielczości $< 1,0$ Å i tylko $\sim 0,05\%$ (47 struktur) do struktur o rozdzielczości subatomowej (Rys. 1).



Rysunek 1. Wzrost liczby struktur o rozdzielczości atomowej ($< 1,2 \text{ \AA}$) zdeponowanych w bazie PDB od roku 1990. W tym czasie obserwujemy eksponencjalny wzrost liczby takich struktur. Jaśniejszym kolorem zaznaczono roczne przyrosty liczby struktur o rozdzielczości atomowej

Figure 1. Growth of the number of the structures with atomic resolution ($< 1.2 \text{ \AA}$) deposited in the PDB database since year 1990. Over that time the deposition rate has become exponential. The lighter color represents yearly growth of number of atomic resolution structures

Planując eksperyment mający na celu otrzymanie modelu struktury krystalicznej dowolnej cząsteczki chemicznej zawsze dążymy do uzyskania wyników o najwyższej jakości i precyzji. W przypadku struktur makromolekularnych precyzja ta może być w znacznym stopniu ograniczona poprzez samą naturę makromolekuł. Cechą charakterystyczną biologicznych kryształów makromolekularnych jest duża zawartość rozpuszczalnika (30–70%, głównie wody). Każda cząsteczka białka lub kwasu nukleinowego jest otoczona przez cząsteczki wody co w znacznym stopniu ogranicza możliwości bezpośrednich kontaktów pomiędzy poszczególnymi makromolekułami i zwiększa labilność całej struktury oraz powoduje znacznie gorszą jakość i rozdzielczość obrazów dyfrakcyjnych niż w przypadku struktur małych cząsteczek (które z reguły nie zawierają rozpuszczalnika).

1. KRYSALIZACJA

Do rozpoczęcia badania struktury białka/kwasu nukleinowego metodą krystalografii strukturalnej konieczne jest dysponowanie kryształem pozwalającym uzyskać zadowalające dane dyfrakcyjne o zadawalającej jakości, a gdy celem jest uzyskanie struktury o wysokiej rozdzielczości, wymagania co do jakości kryształu oraz danych są bardzo wysokie. Przeznaczone do krystalizacji białko lub kwas nukleinowy uzyskane z materiału naturalnego, wytworzone za pomocą inżynierii gene-

tycznej lub syntetycznie musi być bardzo dobrze oczyszczone. W następnym kroku należy znaleźć odpowiednie warunki krystalizacyjne, poprzez zoptymalizowanie bardzo wielu możliwych parametrów, takich jak stężenie białka, rodzaj i stężenie odczynnika strącającego, rodzaj i pH buforu, temperatura, itd. [3] (w większości przypadków stosuje się znane protokoły krystalizacji lub metodę prób i błędów). Obecnie, zamiast żmudnej ręcznej pracy w laboratorium biochemicznym, coraz częściej używa się odpowiednio zaprogramowanych robotów krystalizacyjnych, które są w stanie automatycznie przygotować dużą liczbę prób krystalizacyjnych różniących się wybranymi parametrami.

Kryształami najlepiej nadającymi się do wysokorozdzielczych eksperymentów są kryształy charakteryzujące się dużym stopniem uporządkowania, gęstym upakowaniem, co najczęściej się zdarza w kryształach z niewielką zawartością rozpuszczalnika (z reguły poniżej 40%) i o niewielkich rozmiarach komórki elementarnej. Bardzo dobrze rozpuszczalne, o zwartej budowie (np. związanej z występowaniem mostków dwusiarczkowych) i monomeryczne cząsteczki mają znacznie większe tendencje do tworzenia uporządkowanych struktur i formowania dużych interfejsów międzycząsteczkowych. Przykładem może być krambina, dla której kryształów zarejestrowano dyfrakcję o najwyższej do tej pory rozdzielczości 0,48 Å [4].

Jednakże warunki pozwalające na utworzenie wysoce uporządkowanych struktur nie są łatwe do ustalenia. W przypadku struktury reduktazy aldozowej [5], której kryształy rozpraszały do rozdzielczości 0,66 Å, większość wyżej wymienionych warunków nie jest spełniona – znacznie większa komórka elementarna, brak mostków dwusiarczkowych, większa ilość rozpuszczalnika, brak szczególnie bliskich kontaktów międzycząsteczkowych.

Po uzyskaniu pierwszych kryształów dających dyfrakcję o dobrej rozdzielczości możliwe jest poprawienie jakości kryształów poprzez optymalizację warunków krystalizacyjnych. Dobre rezultaty można otrzymać poprzez zmianę takich czynników jak objętość kropli, temperatura, pH oraz użycie zarodków krystalizacji lub ko-kryształizację z ligandami [5].

2. REJESTRACJA I PROCESOWANIE DANYCH DYFRAKCYJNYCH

Rejestracja obrazów dyfrakcyjnych, która ma na celu pomiar intensywności oraz położenia poszczególnych refleksów, jest ostatnim eksperymentalnym etapem w procesie wyznaczania struktury metodą rentgenografii strukturalnej. Dlatego też zawsze należy zaplanować ją i wykonać w sposób bardzo precyzyjny. Jest to szczególnie istotne w przypadku, gdy chcemy uzyskać wysoką rozdzielczość. Zdecydowaną większość pomiarów dyfrakcyjnych kryształów makromolekularnych wykonuje się przy użyciu promieniowania synchrotronowego udostępnianego na dedykowanych stacjach pomiarowych w dużych ośrodkach synchrotronowych [6]. Aby zapobiec zbyt szybkiemu uszkodzeniu radiacyjnemu kryształu, konieczne jest monitorowanie całkowitej dawki promieniowania, którym naświetlamy kryształ.

Można to zrobić np. z pomocą programu RADDOS [7], który służy do określenia dawki promieniowania rentgenowskiego zaabsorbowanego przez kryształ w czasie pomiaru. Przy wyborze właściwej strategii pomiaru bardzo pomocne mogą być również dedykowane programy takie jak BEST [8] lub odpowiednie moduły programów do procesowania danych dyfrakcyjnych takich jak HKL2000, MOSFLM [9, 10], pomocne przy ustalaniu optymalnych zakresów kątowych i kąta oscylacji kryształu użytych podczas pomiaru. Bardzo istotnym elementem pomiaru wysokorozdzielczego jest typ detektora zamontowanego na linii pomiarowej rejestrującego obrazy dyfrakcyjne. Dla detektorów z 16-bitowym przetwornikiem z reguły nie jest możliwe jednoczesne zarejestrowanie bardzo silnych refleksów niskorozdzielczych i bardzo słabych refleksów wysokorozdzielczych. Dlatego też pomiar wykonuje się w kilku etapach (od 2–4) rejestrując dyfrakcję w kilku zakresach rozdzielczości, zaczynając od najniższej rozdzielczej (np. ∞ –2,0; 2,4–1,0; 1,5–0,7 Å [11]). Najnowsze typy detektorów (np. typu PILATUS [12], które coraz częściej montowane są na synchrotronowych liniach pomiarowych) mają przetworniki 20-bitowe, co znacznie zwiększa zakres rejestrowanych intensywności, i umożliwia wykonanie całego pomiaru dyfrakcji wysokorozdzielczej w jednym przebiegu. Oprócz tego detektory te mogą pracować w trybie ciągłym, co oznacza, że w czasie potrzebnym do obrotu kryształu detektor jest w stanie zarejestrować i zapisać dane dyfrakcyjne. Najnowsze modele mogą zarejestrować od 25 nawet do kilkuset obrazów dyfrakcyjnych w czasie sekundy. Możliwe jest to dzięki zastosowaniu bardzo szybkiej i wydajnej elektroniki oraz niezwykle intensywnej i o dużej jasności wiązki promieniowania rentgenowskiego. Podczas pomiaru należy wybrać optymalny kąt oscylacji kryształu. Użycie kąta oscylacji $<1^\circ$ jest korzystniejsze w trakcie procesowania danych (metoda „*fine slicing*”, procesowanie 3D, [6, 13]), jednak pomimo istniejących technicznych możliwości kąt ten nie powinien być znacznie mniejszy niż mozaikowość kryształu.

Dane dyfrakcyjne o najwyższej rozdzielczości znajdują się w obszarach wysokich kątów rozpraszania i to właśnie one pozwalają na precyzyjne wyznaczenie położenia atomów w strukturze. Jednak w tych zewnętrznych rejonach przestrzeni dyfrakcyjnej intensywność dyfrakcji zanika bardzo szybko, zbliżając się do poziomu szumu pomiarowego i staje się statystycznie nieznacząca. Dlatego też musimy określić granicę rozdzielczości, powyżej której dodatkowe eksperymentalne dane dyfrakcyjne nie będą wносиły żadnej znaczącej informacji.

Niestety precyzyjne określenie tej granicy jest w praktyce bardzo trudne i nie zawsze zupełnie jednoznaczne [14]. Jednym z powszechnie stosowanych kryteriów w krysztalografii makromolekuł określających jakość danych jest parametr obrazujący średni stosunek sygnału do poziomu szumu $\langle I/\sigma(I) \rangle$ obliczany dla poszczególnych zakresów rozdzielczości [15]. Zakłada się, że wartość $\langle I/\sigma(I) \rangle$, w najwyższym zakresie rozdzielczości, powinna wynosić co najmniej 2,0. W praktyce oznacza to, że w ostatnim przedziale rozdzielczości, gdzie znajdują się najsłabsze, a zarazem najważniejsze refleksy z punktu widzenia precyzji struktury, statystycznie około połowa zmierzonych intensywności będzie spełniać warunek $I > 2\sigma(I)$. Dla więk-

szości strukturalnych badań makromolekuł powyższe kryterium wraz z kompletnością i parametrem R_{merge} określającym wewnętrzną spójność danych dyfrakcyjnych stanowią podstawę do ustalenia jakości danych dyfrakcyjnych.

Wraz ze wzrostem rozdzielczości nowych struktur takie podejście jest ostatnio coraz częściej uznawane za zbyt rygorystyczne, szczególnie w odniesieniu do struktur o najwyższej rozdzielczości. Należy pamiętać, że liczba danych dyfrakcyjnych (refleksów) jest proporcjonalna do trzeciej potęgi rozdzielczości, co oznacza, że przy rozdzielczości wynoszącej 1,0 Å liczba refleksów jest ośmiokrotnie większa niż dla rozdzielczości 2,0 Å.

Dla takich struktur znaczna część zarejestrowanych refleksów to refleksy wysokorozdzielcze o zmniejszającej się intensywności i obciążone zwiększającym się błędem. Jednak każdy refleks powinien być traktowany jako bardzo cenny, gdyż wnosi dodatkową informację o strukturze i zbyt rygorystyczne kryterium odrzucania danych eksperymentalnych może być powodem zaniżenia rozdzielczości oraz, jak wykazali Karplus i Diederichs [15], ostateczny model struktury nie będzie tak dobry jak mógłby być przy bardziej efektywnym wykorzystaniu informacji pochodzącej z eksperymentu. Pomimo powszechnego stosowania, wskaźnik R_{merge} , oparty na wewnętrznej spójności danych, nie jest najlepszym kryterium w takich przypadkach. Ponieważ błąd względny pomiaru intensywności rośnie wraz z rozdzielczością, parametr R_{merge} będzie dążył do nieskończoności. Dlatego też kryterium określania „rzeczywistej rozdzielczości” na podstawie wartości R_{merge} (lub pochodnych parametrów jak R_{meas} lub $R_{\text{p.i.m}}$ [16]) nie jest zawsze najlepszym rozwiązaniem, może jednak być bardzo użyteczne przy porównywaniu różnych zestawów danych. Lepszą i bardziej wiarygodną miarą wewnętrznej spójności danych dyfrakcyjnych wydaje się zaproponowany przez tych samych autorów [15] współczynnik korelacji $CC_{1/2}$ pomiędzy dwoma zestawami danych powstałymi poprzez losowy podział danych na dwie połowy. Najmniejsza akceptowalna wartość tego współczynnika może być rzędu 0,1–0,2 co odpowiada korelacji $CC_{\text{true}} \sim 0,5$ pomiędzy rzeczywistym zestawem danych a idealnym zestawem wolnym od błędów. Oszacowanie tej wartości, oznaczone jako CC^* , jest wyznaczane przez większość nowych wersji programów do procesowania danych dyfrakcyjnych [9, 17]. Procesowanie „surowych” danych dyfrakcyjnych należy przeprowadzić szczególnie starannie, zwracając uwagę na skalowanie poszczególnych podzbiorów danych, jakość i precyzję w poszczególnych zakresach rozdzielczości, wybierając najlepsze zakresy pozwalające uzyskać kompletny zestaw danych [6].

3. ROZWIĄZYWANIE STRUKTUR WYSOKOROZDZIELCZYCH

Wynikiem pomiaru dyfrakcji promieni rentgenowskich na kryształach i procesowania danych są intensywności poszczególnych refleksów wraz z ich wskaźnikami hkl , co niestety nie jest wystarczające do rozwiązania struktury kryształu ze względu na niedostępność faz tych refleksów. Jest to tzw. problem fazowy, który musi być

rozwiązany w celu uzyskania modelu struktury. We współczesnej kryształografii makromolekuł głównymi metodami rozwikłania problemu fazowego są podstawienie cząsteczkowe [18] (ang. *Molecular Replacement*) oraz metoda polegająca na wprowadzeniu do struktury atomów o specjalnych właściwościach dyfrakcyjnych. W uproszczeniu, podstawienie cząsteczkowe polega na użyciu, jako przybliżonego modelu, makrocząsteczki o podobnej strukturze do badanej, np. jej mutantu lub homologu, oraz porównania dyfrakcji obserwowanej eksperymentalnie z teoretycznie obliczoną [19]. Druga metoda polega na wprowadzeniu do kryształu (np. poprzez nasączenie, lub modyfikację niektórych aminokwasów), bez zaburzenia jego pozostałej struktury, atomu o dużej liczbie elektronów dającego silną dyfrakcję lub takiego, dla którego można zarejestrować dyfrakcję anomalną i na tej podstawie wyznaczyć wstępne fazy. Obecnie najczęściej używanym wariantem tej metody jest SAD (ang. *Single-wavelength Anomalous Diffraction*), w której dokonuje się bardzo precyzyjnych pomiarów dyfrakcji anomalnej tylko przy jednej długości fali, przy zastosowaniu promieniowania synchrotronowego z możliwością dostrojenia długości fali do wybranej krawędzi absorpcji, co nie jest możliwe przy klasycznych źródłach promieniowania rentgenowskiego.

Gdy rozdzielczość i jakość danych dyfrakcyjnych jest wysoka, możliwe jest również wykorzystanie słabych sygnałów anomalnych pochodzących np. od atomów siarki [20, 21] lub fosforu [22], które są typowymi pierwiastkami występującymi w strukturach białek i kwasów nukleinowych.

3.1. METODY BEZPOŚREDNIE

Dysponując danymi o rozdzielczości atomowej możliwe jest również rozwiązanie problemu fazowego za pomocą tzw. metod bezpośrednich, które są rutynowo stosowane w kryształografii małych cząsteczek zawierających nie więcej niż 100 atomów niewodorowych. Metody te używają zaawansowanej teorii prawdopodobieństwa i nie wymagają żadnych innych informacji oprócz danych o symetrii i składzie chemicznym oraz zestawu eksperymentalnie zmierzonych intensywności, które i tak są niezbędne podczas udokładniania struktury. Korzystając z faktu, że dane dyfrakcyjne o atomowej rozdzielczości zawierają znacznie więcej zmierzonych intensywności niż wynosi liczba parametrów koniecznych do opisanie modelu atomowego i założenia, że atomy są dobrze rozdzielone, możliwe jest oszacowanie faz refleksów bezpośrednio z intensywności zmierzonych refleksów [23].

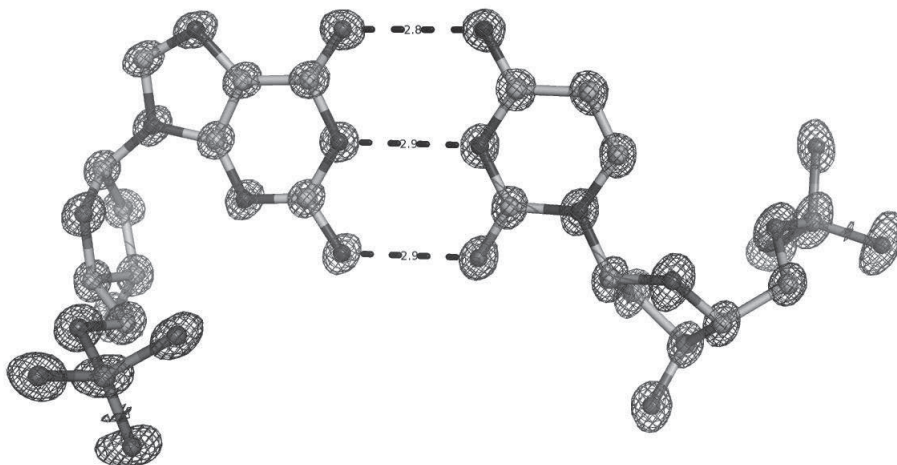
Ostatnio często do metod bezpośrednich zaliczane są też metody *ab initio*, które do rozwiązania problemu fazowego wykorzystują tylko natywne dane dyfrakcyjne bez używania informacji o wstępnych fazach uzyskanych na bazie pochodnych izomorficznych lub rozpraszania anomalnego. Jedną z metod *ab initio* w czystej formie jest metoda „odwracania ładunku” (ang. *charge-flipping*) [24], która odróżnia się od klasycznych metod bezpośrednich tym, że nie jest oparta na probabilistycznych zależnościach faz i do rozwiązania struktury nie wymaga żadnych innych infor-

macji oprócz danych zawierających intensywności i wskaźniki *hkl* refleksów. Algorytm *odwracania ładunku* należy do grupy iteracyjnych algorytmów operujących zarówno na przestrzeni odwrotnej jak i na rzeczywistej (ang. *dual-space algorithms*), polegających na iteracyjnym przemiennym udokładnianiu próbných faz w przestrzeni odwrotnej oraz modyfikacji map gęstości elektronowej w sieci rzeczywistej z więzami poprawiającymi fazy. Do grupy tej należą również takie programy jak: Shake-and-Bake [25], ShelxD [26], SIR2004 [27].

4. BUDOWANIE MODELU

Ostatecznym wynikiem rentgenowskiej analizy strukturalnej jest mapa gęstości elektronowej. Wstępne mapy, obliczone na podstawie faz wyznaczonych na etapie rozwiązania struktury i intensywności otrzymanych z eksperymentu (a co za tym idzie, proporcjonalnych do nich amplitud czynników struktury), są z reguły niedokładne, lecz umożliwiają zbudowanie wstępnego modelu struktury.

Taki model, często niekompletny, musi być uzupełniony poprzez dodanie wszystkich brakujących fragmentów. Budowanie kompletnego modelu możliwe jest, przy prawidłowo określonych wstępnych fazach, poprzez interpretację map gęstości elektronowej, głównie map różnicowych. W przypadku niewielkich struktur, jakie spotyka się w krystalografii małych cząsteczek, budowanie kompletnego modelu możliwe jest poprzez analizę listy najwyższych pików, która generowana jest poprzez automatyczną interpretację mapy gęstości elektronowej. W krystalografii makromolekularnej, z powodu wielkości badanych obiektów i z reguły nie najlepszej rozdzielczości, analiza w pełni automatyczna nie jest możliwa nawet dla map wysokorozdzielczych. Często stosuje się zaawansowane programy komputerowe takie jak Buccaneer [28], ArpWarp [29], Phenix.Autobuild [30], wspomagające budowanie i sprawdzanie poprawności modelu oraz jego wstępne dopasowanie do mapy gęstości elektronowej. Wyniki takich interpretacji zawsze wymagają sprawdzenia za pomocą programów wykorzystujących grafikę komputerową (takich jak Coot [31]), umożliwiających korektę niektórych fragmentów modelu. W przypadku struktur wysokorozdzielczych, mapy gęstości elektronowej są bardzo wysokiej jakości, pozwalającej rozróżnić poszczególne pojedyncze maksima reprezentujące atomy (Rys. 2).



Rysunek 2. Fragment mapy gęstości elektronowej struktury dupletu heksameru Z-DNA d(CGCGCG) o wyjątkowo wysokiej rozdzielczości [38]. Widoczna para zasad Cyt3-Gua10 z odpowiadającą mapą $2F_{\text{obs}} - F_{\text{calc}}$ przy konturze na poziomie $1,5\sigma$, rozdzielczość $0,53 \text{ \AA}$.

Figure 2. Fragment of electron density map of extremely high resolution structure of Z-DNA hexamer duplex d(CGCGCG) [38]. Cyt3-Gua10 base pair with the corresponding $2F_{\text{obs}} - F_{\text{calc}}$ map at 1.5σ contour level, resolution 0.53 \AA

Taka sytuacja znacznie ułatwia prawidłową interpretację mapy i często możliwe jest zbudowanie wstępnego modelu manualnie z użyciem programów do graficznej wizualizacji map gęstości elektronowej [31]. Jest to możliwe również dlatego, iż masa cząsteczkowa poznanych do tej pory makromolekuł, dla których określono struktury z najwyższą rozdzielczością, jest raczej niewielka. Dla większości struktur makromolekuł o rozdzielczości poniżej $1,0 \text{ \AA}$ masa cząsteczkowa jest mniejsza od 20 kD (średnia masa cząsteczkowa struktur zdeponowanych w ostatnich latach w bazie PDB zbliża się do 100 kD).

5. UDOKŁADNIANIE

Jednym z ostatnich kroków w procesie mającym na celu otrzymanie struktury rentgenograficznej makrocząsteczki jest udokładnienie modelu, prowadzone w celu maksymalizacji zgodności modelu z danymi eksperymentalnymi. W trakcie udokładniania optymalizowane są takie parametry jak współrzędne atomów, parametry przemieszczenia atomowego (ang. *Atomic Displacement Parameter*, ADP), czynniki obsadzenia i inne parametry opisujące bezpośrednio model.

Celem udokładniania jest znalezienie takiego zestawu parametrów modelu, dla którego obliczone na ich podstawie czynniki struktury $|F_{\text{calc}}|$ będą najbardziej zbliżone do amplitud zmierzonych eksperymentalnie $|F_{\text{obs}}|$. Ogólnie można powiedzieć, że udokładnianie struktur makromolekuł (jak również mniejszych struktur rentgenograficznych) jest optymalizacją funkcji zbioru obserwacji poprzez zmianę

parametrów modelu. W praktyce wykorzystywane są dwie funkcje: pierwsza oparta na zasadzie najmniejszych kwadratów (ang. *least squares*, LSQ), druga na największej wiarygodności (ang. *maximum likelihood*, ML). W metodzie najmniejszych kwadratów obserwacje mają ustalone wartości a parametry modelu zmieniane są tak, że obliczane na ich podstawie wartości zbliżają się w trakcie udokładniania do obserwacji tak blisko jak jest to możliwe. Natomiast główną ideą metody największej wiarygodności jest znalezienie takiego modelu, który będzie najbardziej zgodny z obserwacjami. Zgodność ta mierzona jest statystycznie i jeśli zmiana dokonana w modelu czyni obserwacje bardziej prawdopodobną to wiarygodność (ang. *likelihood*) rośnie, wskazując, że model jest lepszy. Metoda największej wiarygodności jest znacznie trudniejsza z punktu widzenia implementacji w programach komputerowych niż metoda najmniejszych kwadratów i autorzy programów muszą użyć wielu wstępnych założeń pozwalających zakończyć obliczenia w rozsądnym czasie.

Większość obecnie używanych programów do udokładniania struktur makrocząsteczek (np. REFMAC [32], Phenix.refine [33]) oparta jest na metodzie ML, głównie dlatego, że metoda ta daje znacznie lepsze wyniki dla struktur, w których wstępny model jest niekompletny, a fazy użyte do udokładniania nie są zbyt dokładne. Sytuacja taka jest typowa w badaniach strukturalnych makromolekuł, szczególnie tych bardzo dużych, dla których jakość danych i rozdzielczość jest niska. Programy te mogą być również z powodzeniem stosowane do udokładniania struktur wysokorozdzielczych, lecz należy je stosować z ostrożnością, zwracając szczególną uwagę na domyślne ustawienia parametrów udokładniania.

Warto zwrócić uwagę, że metoda najmniejszych kwadratów jest w zasadzie szczególnym przypadkiem bardziej ogólnej metody największej wiarygodności. W przypadku, gdy rozkład błędów obserwacji jest zgodny z rozkładem Gaussa można wykazać, że metoda największej wiarygodności daje bardzo podobne rezultaty jak metoda najmniejszych kwadratów.

Ponieważ wstępny model struktur o rozdzielczości atomowej uzyskany w wyniku rozwiązania problemu fazowego z reguły jest kompletny, a co za tym idzie wstępne fazy są dość precyzyjne, możliwe jest udokładnianie takiej struktury przy pomocy znacznie prostszego w implementacji algorytmu opartego na LSQ. Wśród programów wykorzystujących tę metodę jednym z najbardziej popularnych programów, szczególnie w krytalografii małych cząsteczek, jest SHELXL [34]. Pod koniec lat 90-tych program ten został dostosowany do udokładniania struktur białkowych i obecnie jest dominującym programem używanym do udokładniania struktur makromolekularnych o rozdzielczości $< 1,0 \text{ \AA}$.

Jedną z najbardziej wyróżniających cech programu SHELXL jest użycie konwencjonalnej metody obliczania czynników struktury, która jest bardziej precyzyjna niż metoda FFT (ang. *Fast Fourier Transform*). Taki sposób obliczeń poprawia zbieżność udokładniania, lecz program działa relatywnie wolniej. W ostatniej wersji programu zostało to w pewnym stopniu zrekomensowane poprzez efektywne użycie procesorów wielordzeniowych, powszechnie dostępnych we współczesnych

komputerach. Spośród innych zalet programu należy wymienić używanie podczas udokładniania kwadratów czynników struktury (F_{obs}^2 , proporcjonalnych do intensywności) zamiast ich amplitud $|F_{obs}|$ oraz możliwość oszacowania niepewności standardowych dla każdego parametru modelu poprzez inwersję macierzy najmniejszych kwadratów. Oprócz tego program oferuje bardzo elastyczny i wygodny sposób traktowania nieuporządkowania w strukturze.

W celu potwierdzenia poprawności nowych parametrów wprowadzanych do udokładniania powszechnie stosuje się kryterium R_{free} polegające na monitorowaniu wartości wskaźnika rozbieżności R_{free} , wyliczonego na podstawie modelu i losowo wybranego podzbioru około tysiąca refleksów, które nigdy nie były używane w udokładnianiu modelu. Jeśli zmiana dokonana w modelu była poprawna i zgodna z danymi eksperymentalnymi to wartości czynników R i R_{free} powinny się obniżyć. Należy jednak pamiętać, że różnica pomiędzy wartościami R i R_{free} nie powinna być większa niż 5–7%, gdyż w takim przypadku może to wskazywać na nadinterpretację lub inne poważne defekty modelu. Przykładem może być dodanie do struktury zbyt wielu cząsteczek wody, w położeniach niejednoznacznie potwierdzonych na mapie gęstości elektronowej, co spowoduje obniżenie wartości wskaźnika R , lecz nie R_{free} . Każda prawidłowa zmiana modelu skutkuje poprawą faz i na mapie gęstości elektronowej często ujawniają się nowe szczegóły modelu. Dlatego też, jak napisał Sheldrick [34], udokładniania wysokorozdzielczej struktury makromolekularnej nie można nigdy zakończyć, można je jedynie w pewnym momencie przerwać.

Przy rozdzielczościach rzędu 1,4 Å i wyższych w pełni uzasadnione jest wprowadzenie do udokładniania modelu z anizotropowymi atomowymi czynnikami temperaturowymi. Skutkuje to znaczącym poprawieniem zgodności modelu z danymi eksperymentalnymi, co wyraża się obniżeniem wskaźników rozbieżności R i R_{free} o 2–3% a przy rozdzielczościach rzędu 1,0 Å nawet o 5–6% [35].

5.1. MODELOWANIE OBSZARU ROZPUSZCZALNIKA

Jak już wspomniano, znaczącą część kryształu makromolekularnego zajmuje nieuporządkowany rozpuszczalnik i w większości są to cząsteczki wody. Dla kryształów białkowych średnia zawartość rozpuszczalnika wynosi ~50%, natomiast dla kwasów nukleinowych ~47% [36]. Zbudowanie precyzyjnego modelu tych fragmentów kryształu do tej pory jest dużym problemem krystalografii makromolekularnej. Wpływ nieuporządkowanego rozpuszczalnika na czynniki struktury jest znacznie większy w przypadku struktur o niskiej rozdzielczości, jednak również przy wysokich rozdzielczościach nie można go pominąć. Uporządkowane cząsteczki wody są integralnym składnikiem kryształów makromolekuł biologicznych i odgrywają kluczową rolę stabilizującą struktury wyższych rzędów, tworzone przez białka lub kwasy nukleinowe.

W zdecydowanej większości modeli struktur makromolekuł otrzymanych za pomocą dyfrakcji rentgenowskiej (o rozdzielczości lepszej niż 2,7 Å), na podstawie

interpretacji map gęstości elektronowej, możliwe jest zlokalizowanie pojedynczych cząsteczek rozpuszczalnika (głównie wody). Niestety w większości przypadków możliwe jest to tylko dla niektórych cząsteczek wody (10–20%), takich, które biorą udział w oddziaływaniach (głównie poprzez wiązania wodorowe) z makrocząsteczką lub innym związanym elementem struktury. Dzieje się tak dlatego, że krytalografia rentgenowska pozwala zaobserwować tylko obiekty, które powtarzają się w okresowy sposób w całej objętości kryształu. Pozostała część rozpuszczalnika znajduje się w nieuporządkowanych obszarach kryształu, dla których interpretacja mapy gęstości elektronowej nie jest możliwa. Ilość obserwowanych cząsteczek wody zależy również od rozdzielczości danych dyfrakcyjnych. W strukturach białek w prosty sposób można oszacować ile obserwowanych cząsteczek wody może znajdować się w strukturze: wystarczy pomnożyć liczbę aminokwasów przez liczbę $3-d$, gdzie d jest rozdzielczością [37]. Łatwo można zauważyć, że np. dla rozdzielczości $\sim 3,0$ Å praktycznie nie będziemy w stanie zidentyfikować cząsteczek wody na mapie gęstości elektronowej (rozdzielczość jest porównywalna z odległością donor-akceptor w wiązaniu wodorowym). Jednak w przypadku, gdy dysponujemy danymi wysokorozdzielczymi np. $\sim 1,0$ Å, liczba możliwych do zaobserwowania cząsteczek wody będzie około dwukrotnie większa niż liczba aminokwasów w badanym białku. W niektórych uporządkowanych strukturach krystalicznych kwasów nukleinowych o rozdzielczości subatomowej możliwa jest lokalizacja prawie wszystkich cząsteczek wody zawartych w kryształach. W takim przypadku modelowanie rozpraszania pochodzącego od nieuporządkowanego rozpuszczalnika nie prowadzi do znaczącej poprawy modelu struktury [38, 39]. W strukturach krystalicznych o rozdzielczości atomowej lub subatomowej duża ilość szczegółów możliwych do zinterpretowania na mapie gęstości elektronowej ma decydujący wpływ na wiarygodność modelu atomowego i zależy od ilości danych pomiarowych, czyli od ich rozdzielczości.

5.2. ALTERNATYWNE KONFORMACJE

Metoda dyfrakcji pozwala obserwować obiekty, które w okresowy sposób powtarzają się w badanym kryształach. Często jednak niektóre cząsteczki, lub ich fragmenty, występują w kryształach w kilku różnych konformacjach co może prowadzić do nieporządku w kryształach. W wyniku tego mapa gęstości elektronowej w takich miejscach jest mniej wyraźna i obciążona większym błędem. W strukturach o niskiej rozdzielczości w takim miejscu nieuporządkowania wprowadza się do modelu pojedynczą cząsteczkę z wyższymi czynnikami temperaturowymi w uśrednionym położeniu. Jednak w miarę zwiększającej się rozdzielczości możliwe jest rozszyfrowanie takiego nieporządku i wprowadzenie do modelu dwóch lub czasami nawet większej liczby alternatywnych konformacji, które są dobrze widoczne na mapie gęstości elektronowej (szczególnie przy bardzo wysokich rozdzielczościach). Ujawnienie wielokrotnych konformacji i precyzyjne określenie ich cząstkowych obsadzeń jest szczególnie ważne w okolicach miejsc aktywnych enzymów, ponie-

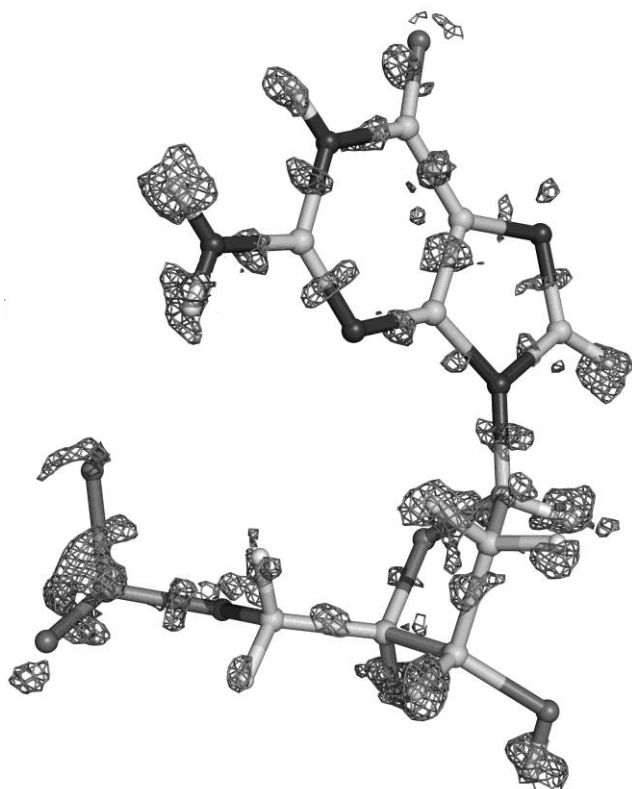
waż niektóre z tych konformacji mogą być ściśle powiązane z funkcją i aktywnością białka [40–42]. Średnia liczba reszt aminokwasowych, które zostały wprowadzone do modelu w wielokrotnych konformacjach (w większości podwójnych) w strukturach wysokorozdzielczych jest rzędu 20%, lecz niektóre interpretacje mapy gęstości elektronowej pozwoliły wyznaczyć znacznie więcej alternatywnych konformacji (np. 40% w krambinie [43], 25% w reduktazie aldozowej i synteninie [5, 44]).

5.3. ATOMY WODORU I WIĄZANIA WODOROWE

Powszechnie uważa się, że za pomocą dyfrakcji promieni rentgenowskich na kryształach makromolekularnych niemożliwa jest identyfikacja atomów wodoru, ponieważ pochodzący od nich sygnał jest zbyt słaby w porównaniu z dużym poziomem szumu pochodzącego np. od nieuporządkowanej części rozpuszczalnika.

W trakcie rejestrowania dyfrakcji promieniowanie rentgenowskie rozpraszane jest przez elektrony skupione wokół poszczególnych atomów a intensywność rozpraszania zależy bezpośrednio od ilości elektronów. Dlatego też nawet w strukturach małych cząsteczek, gęstość elektronowa dla elektronów obserwowana jest tylko w dobrze uporządkowanych fragmentach struktury. Należy zwrócić uwagę, że liczba atomów wodoru w makromolekułach jest znacząca, w białkach atomy te stanowią około 50% a w kwasach nukleinowych około 35% wszystkich atomów i odgrywają one kluczową rolę w mechanizmach reakcji enzymatycznych.

Większość struktur makromolekuł nie ma wystarczającej rozdzielczości aby określić położenie atomów wodoru, jednak fakt, że ich wkład do całkowitej mocy rozpraszania może być rzędu nawet 10%, nie pozwala na całkowite ich zignorowanie. Ponieważ stosunkowo łatwo można przewidzieć teoretyczną lokalizację atomów wodoru, w przypadku struktur średnio- i niskorozdzielczych atomy te dodawane są do modelu w procesie udokładniania jako jego uzupełnienie, jednak nie są deponowane w bazie danych PDB. Położenia atomów wodoru nie są udokładnianie a ich współrzędne są wyliczane na podstawie położenia atomu głównego (*angling model*), co nie powoduje zwiększenia liczby udokładnianych parametrów. Taki sposób postępowania poprawia zgodność obliczonych czynników struktury F_{calc} z danymi eksperymentalnymi co wyraża się lepszym parametrem zgodności R . Niestety położenia niektórych atomów wodoru bardzo trudno przewidzieć teoretycznie [np. atomy H w protonowanej grupie aminowej (NH_3^+), grupie metylowej (CH_3) lub grupie hydroksylowej OH] i atomy takie zawsze powinny być traktowane indywidualnie, a ich lokalizacja ustalona jest przeważnie na podstawie analizy sieci wiązań wodorowych oraz, jeśli rozdzielczość na to pozwala, mapy gęstości elektronowej (Rys. 3).



Rysunek 3. Fragment mapy gęstości elektronowej struktury duplexu heksameru Z-DNA d(CGCGCG) o wyjątkowo wysokiej rozdzielczości [38]. Widoczna mapa różnicowa $F_{\text{obs}} - F_{\text{calc}}$ z wyraźnymi pikami odpowiadającymi atomom wodoru oraz deformacyjna gęstość na środku wiązań pochodząca od elektronów wiążących. Kontur mapy (z pominiętymi atomami H) na poziomie $2,5\sigma$, rozdzielczość $0,53 \text{ \AA}$

Figure 3. Fragment of electron density map of extremely high resolution structure of Z-DNA hexamer duplex d(CGCGCG) [38]. Difference map $F_{\text{obs}} - F_{\text{calc}}$ with clearly visible peaks corresponding to hydrogen atoms and deformation density in the middle of bonds from lone pair electrons. Map (with omitted H atoms) was contoured at level of 2.5σ , resolution 0.53 \AA

W przypadku struktur wysokorozdzielczych ilość szczegółów możliwych do zaobserwowania na mapach gęstości elektronowej wzrasta wraz z rozdzielczością. Dla rozdzielczości atomowych można zlokalizować niektóre atomy wodoru, a dla bardzo wysokiej rozdzielczości $\sim 1,0 \text{ \AA}$ (porównywalnej z typową długością wiązania kowalencyjnego X–H), lub lepszej, możliwe jest zaobserwowanie dla większości atomów wodoru odpowiadających im pików na mapie gęstości elektronowej [38, 39]. Ponieważ stosunek liczby obserwacji do liczby udokładnianych parametrów przy takich rozdzielczościach jest wystarczająco duży, możliwe jest udokładnianie położenia niektórych atomów wodoru, szczególnie tych, których nie można jednoznacznie wyznaczyć geometrycznie. Przy najwyższych rozdzielczościach osiągniętych dla makromolekułów, które są rzędu $0,5 \text{ \AA}$, możliwe jest nawet włączenie do

udokładniania wszystkich atomów wodoru, jednak jak wykazał test przeprowadzony dla struktury Z-DNA o rozdzielczości 0,53 Å, poprawa zgodności modelu z danymi eksperymentalnymi była nieznacząca (parametr R obniżył się o kilka setnych procenta). Wobec tego, nawet przy najwyższych rozdzielczościach położenia większości atomów wodoru włączanych do modelu mogą być wyznaczone teoretycznie, a jedynie w niejednoznacznych przypadkach można zastosować pełne udokładnianie parametrów atomów wodoru z izotropowymi czynnikami ADP (ta grupa atomów wodoru zwykle ma kluczowe znaczenie dla struktury i funkcji makrocząsteczki).

Powyższe rozważania nie dotyczą z reguły atomów wodoru w cząsteczkach wody, która jest głównym składnikiem rozpuszczalnika. Ponieważ bardzo trudno przewidzieć położenie atomów wodoru w cząsteczce wody, atomy te dodawane są do modelu jedynie wtedy, gdy są wyraźnie widoczne na mapie gęstości elektronywej, a ich położenie można potwierdzić udziałem w wiązaniach wodorowych.

5.4. STANDARDOWE PARAMETRY STEREOCHEMICZNE

W kryształografii makromolekuł rozdzielczość mierzonych danych dyfrakcyjnych nie jest wysoka, szczególnie dla dużych białek lub kompleksów i nierzadko liczba zarejestrowanych refleksów jest porównywalna lub niewiele większa od liczby parametrów opisujących strukturę. Aby zwiększyć liczbę obserwacji przypadającą na jeden parametr korzysta się z informacji stereochemicznych o poszczególnych znanych elementach struktury, takich jak aminokwasy w białkach czy nukleotydy w kwasach nukleinowych [45, 46], traktując je jako dodatkowe obserwacje.

Znane wartości poszczególnych długości wiązań, kątów walencyjnych i niektórych innych parametrów (np. planarność pierścieni aromatycznych, chiralność niektórych atomów), dostarczają użytecznych więzów stosowanych w procesie udokładniania [46]. W przypadku, gdy dostępne są dane o wysokiej rozdzielczości, sytuacja zmienia się diametralnie. Liczba eksperymentalnych danych dyfrakcyjnych szybko zwiększa się wraz ze wzrostem rozdzielczości, co powoduje znaczną poprawę stosunku liczby refleksów przypadających na jeden parametr opisujący model. W takiej sytuacji możliwe jest udokładnianie bez użycia dodatkowych więzów lub z więzami bardzo rozluźnionymi.

Mimo nieustających dyskusji [47], większość struktur makrocząsteczek do tej pory jest udokładniania z więzami opracowanymi na początku lat 90. w oparciu o struktury małowcząsteczkowe [46]. Statystyczna analiza parametrów geometrycznych struktur wysokorozdzielczych powinna mieć olbrzymie znaczenie dla stworzenia biblioteki bardziej precyzyjnych więzów, które będą odzwierciedlały specyfikę makrocząsteczek. Takie więzy mogą być później wykorzystywane do udokładniania struktur o niskiej lub średniej rozdzielczości. Analiza danych geometrycznych [48] wykazała znaczące różnice pomiędzy obserwowanymi długościami wiązań i kątów pojedynczych aminokwasów a wartościami znajdującymi się w najczęściej używanej bibliotece [46]. Przykłady odstępstw od planarności wiązania peptydowego,

idealnych długości wiązania C-N i wartości kątów N-C α -C można znaleźć w wielu pracach, w których analizowane są parametry geometryczne struktur wysokorozdzielczych [49–51]. Analiza wybranych parametrów łańcucha głównego w uporządkowanych fragmentach struktur wysokorozdzielczych wykazała, że niektóre z obecnie używanych więzów geometrycznych powinny być uaktualnione lub stosowane z większą tolerancją. Szczególnie dotyczy to struktur o najwyższej rozdzielczości, gdzie dominującym członem powinny być dane dyfrakcyjne i w uporządkowanych obszarach struktury powinny być akceptowane większe odchylenia od idealnej geometrii. W pracy [47], która wywołała szeroką dyskusję [52–54], sugerowane jest wprowadzenie do programów służących do udokładniania zmiennych wag więzów, które zależne byłyby od czynników obsadzenia i temperaturowych poszczególnych atomów. Podobna metoda usprawnienia udokładniania struktur wysokorozdzielczych, polegająca na zastosowaniu biblioteki parametrów geometrycznych *CDL* (ang. *Conformation-Dependent backbone-geometry Library*) zależnych od konformacji, została zaproponowana przez Tronrunda i Karplusa [55].

6. ROZKŁAD GĘSTOŚCI ŁADUNKU I UDOKŁADNIANIE MULTIPÓLOWE

Przy rozdzielczościach $\sim 0,7$ Å lub lepszych możliwa jest bezpośrednia obserwacja deformacji gęstości elektronowej ukazującej odchylenia od powszechnie stosowanego podczas udokładniania modelu atomu sferycznego. Na mapach różnicowych można zaobserwować maksima dokładnie na środku niektórych wiązań oraz dodatkową gęstość, odpowiadającą wolnej parze elektronowej, w otoczeniu atomów tlenu (Rys. 3). Obecność takiej znaczącej gęstości elektronowej wskazuje, iż sferyczny model atomu, łącznie z anizotropowymi czynnikami temperaturowymi, daje zbyt zgrubne przybliżenie rzeczywistej gęstości elektronowej. Dlatego też coraz częściej stosuje się udokładnianie wykorzystujące niesferyczny model atomu. Podczas udokładniania multipolowego stosowany jest model, w którym elektrony atomu są podzielone na elektrony rdzenia, które można opisać modelem sferycznym oraz elektrony walencyjne, do których stosuje się funkcje multipolowe [56]. Udokładnianie multipolowe wymaga użycia znacznie większej liczby parametrów, multipolowy atom niewodorowy ma zwykle minimum 28 parametrów (18 multipolowych, 3 współrzędne, 6 parametrów temperaturowych oraz 1 czynnik obsadzenia). Udokładnianie multipolowe wymaga wyrafinowanego oprogramowania np. takiego jak program MoPro [57] i możliwe jest dla wybranej grupy przypadków z najwyższą rozdzielczością i jakością danych eksperymentalnych.

PODSUMOWANIE

Dynamiczny rozwój makromolekularnej krystalografii wysokorozdzielczej w ostatnich dwóch dekadach był możliwy głównie dzięki postępowi technologicz-

nemu i metodologicznemu. Należą do nich głównie powszechna dostępność niezwykle silnych i zogniskowanych źródeł promieniowania synchrotronowego, rutynowe kriogeniczne schładzanie kryształów w trakcie rejestracji danych dyfrakcyjnych oraz udoskonalone oprogramowanie krystalograficzne na wszystkich etapach od procesowania danych do udokładniania i sprawdzania poprawności struktury. Skutkuje to ciągłym wzrostem liczby struktur makrocząsteczek o rozdzielczości atomowej i subatomowej zdeponowanych w bazie PDB. Ilość szczegółów widocznych na mapach gęstości elektronowej takich struktur daje możliwość zupełnie nowego spojrzenia na funkcję białka i mechanizmy reakcji katalitycznych. Alternatywne konformacje łańcuchów bocznych jak również łańcuchów głównych oraz subtelne zmiany strukturalne mogą być często jednoznacznie zidentyfikowane, co jest szczególnie istotne w rejonach miejsc aktywnych makromolekuł. Model struktury krystalicznej charakteryzuje się dużą wiarygodnością a w przypadku najwyższych rozdzielczości w uporządkowanych rejonach można zlokalizować położenia atomów wodoru (nawet w niektórych cząsteczkach wody). Pozwala to na uzyskanie informacji o stanie protonacji katalitycznie istotnych reszt oraz sieci wiązań wodorowych. Ponieważ liczba zarejestrowanych refleksów jest znacznie większa od liczby parametrów opisujących strukturę, dobrze uporządkowane fragmenty struktury mogą być udokładniane bez więzów stereochemicznych. W takich przypadkach można zastosować pełnomacierzowe udokładnianie metodą najmniejszych kwadratów i oszacować standardowe niepewności parametrów geometrycznych. Niewiele kryształów makromolekularnych rozprasza promieniowanie rentgenowskie do rozdzielczości subatomowej (poza limit 0,8 Å), lecz są one szczególnie cenne z metodologicznego punktu widzenia, na przykład dla obiektywnego optymalizowania parametrów stereochemicznych, które są używane jako więzy przy udokładnianiu modeli przy niższych rozdzielczościach [11, 47]. Dysponując danymi o najwyższej rozdzielczości stają się możliwe badania deformacyjnej gęstości elektronowej oraz walencyjnych orbitali atomowych i molekularnych, które ukazują niezwykle wewnętrzny świat makrocząsteczek nie tylko na poziomie atomowym lecz również elektronowym.

PIŚMIENNICTWO CYTOWANE

- [1] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235.
- [2] G.M. Sheldrick, *Acta Cryst.*, 1990, **A46**, 467.
- [3] A. McPherson, Cold Spring Harbor Laboratory NY, Cold Spring Harbor Laboratory Press, 1999.
- [4] A. Schmidt, M. Teeter, E. Weckert, V.S. Lamzin, *Acta Cryst.*, 2011, **F67**, 424.
- [5] E.I. Howard, R. Sanishvili R, R.E. Cachau, A. Mitschler, B. Chevrier, P. Barth, V. Lamour, M. Van Zandt, E. Sibley, C. Bon, D. Moras, T.R. Schneider, A. Joachimiak, A. Podjarny, *Proteins*. 2004, **55**, 792.
- [6] M. Gilski, *Acta Phys. Polon.*, 2008, **A114**, 331.
- [7] K.S. Paithankar, R.L. Owen, E.F. Garman E.F., *J. Synch. Rad.*, 2009, **16**, 152.
- [8] A.N. Popov, G.P. Bourenkov, *Acta Crystallogr. D***59**, 2003, 1145.

- [9] Z. Otwinowski, W. Minor, *Methods Enzymol.*, 1997, **276**, 307.
- [10] A.G.W. Leslie, *Int CCP4/ESF-EACBM Newsl. Protein Crystallogr.*, 1992, **26**.
- [11] M. Jaskolski, *Advancing Methods for Biomolecular Crystallography*, R. Read, A.G. Urzhumtsev, V.Y. Lunin (Red.), Springer, 2013.
- [12] J. Aishima, R.L. Owen, D. Axford, E. Shepherd, G. Winter, K. Levik, P. Gibbons, et al., *Acta Cryst.*, 2010, **D66**, 1032.
- [13] M. Gilski, *Synchr. Radiat. Nat. Sci.*, 2007, **6**, 95.
- [14] P. Evans, *Science*, 2012, **336**, 986.
- [15] P.A. Karplus, K. Diederichs, *Science*, 2012, **336**, 1030.
- [16] M.S. Weiss, *J. Appl. Cryst.*, 2001, **34**, 130.
- [17] W. Kabsch, *Acta Cryst.*, 2010, **D66**, 125.
- [18] M.G. Rossmann & D.M. Blow, *Acta Cryst.*, 1962, **15**, 24.
- [19] Z. Dauter, M. Jaskolski, *Promieniowanie synchrotronowe w spektroskopii i badaniach strukturalnych. Wybrane zagadnienia*, B.J. Kowalski, W. Paszkowicz, E.A. Gorlich (Red.), Kraków, 2011, s. 301.
- [20] B.-C. Wang, *Methods Enzymol.*, 1985, **115**, 90.
- [21] Z. Dauter, M. Dauter, E. de La Fortelle, G. Bricogne, G.M. Sheldrick, *J. Mol. Biol.*, 1999, **289**, 83.
- [22] Z. Dauter, D.A. Adamiak, *Acta Cryst.*, 2001, **D57**, 990.
- [23] H. Hauptman, J. Karle, *Acta Cryst.*, 1955, **8**, 355.
- [24] C. Dumas, A. van der Lee, *Acta Cryst.*, **D64**, 2008, 864.
- [25] C.M. Weeks, G.T. DeTitta, H.A. Hauptman, P. Thuman, R. Miller, *Acta Cryst.*, 1994, **A50**, 210.
- [26] G.M. Sheldrick, R.O. Gould, *Acta Cryst.*, 1995, **B51**, 423.
- [27] M.C. Burla, R. Caliandro, M. Camalli, B. Carrozzini, G.L. Cascarano, L. De Caro, C. Giacovazzo, G. Polidori, G.R. Spagna, *J. Appl. Cryst.*, 2005, **38**, 381.
- [28] K. Cowtan, *Acta Cryst.*, 2006, **D62**, 1002.
- [29] G.G. Langer, S. Hazledine, T. Wiegels, C. Carolan, V.S. Lamzin, *Acta Cryst.*, 2013, **D69**, 635.
- [30] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, *Acta Cryst.*, 2010, **D66**, 213.
- [31] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, *Acta Cryst.*, 2010, **D66**, 486.
- [32] G.N. Murshudov, A.A. Vagin, E.J. Dodson, *Acta Cryst.*, 1997, **D53**, 240.
- [33] P.D. Adams, R.W. Grosse-Kunstleve, L.-W. Hung, T.R. Ioerger, A.J. McCoy, N.W. Moriarty, R.J. Read, J.C. Sacchettini, N.K. Sauter, T.C. Terwilliger, *Acta Cryst.*, 2002, **D58**, 1948.
- [34] G.M. Sheldrick, *Acta Cryst A*64, 2008, 112.
- [35] Z. Dauter, V.S. Lamzin, K.S. Wilson, *Curr. Opin. Struct. Biol.*, 1995, **5**, 784.
- [36] K.A. Kantardjieff, B. Rupp, *Protein Sci.*, 2003, **12**, 1865.
- [37] A. Wlodawer, W. Minor, Z. Dauter, M. Jaskolski, *FEBS J.*, 2008, **275**, 1.
- [38] K. Brzezinski, A. Brzuskiewicz, M. Dauter, M. Kubicki, M. Jaskolski, Z. Dauter, *Nucleic Acids Res.*, 2011, **39**, 6238.
- [39] P. Drodzdzal, M. Gilski, R. Kierzek, L. Lomozik, M. Jaskolski, *Acta Cryst.*, 2013, **D69**, 1180.
- [40] W.R. Rypniewski, P.R. Ostergaard, M. Nørregaard-Madsen, M. Dauter, K.S. Wilson, *Acta Cryst.* 2001, **D57**, 8.
- [41] I. Kursula, R.K. Wierenga, *J. Biol. Chem.*, 2003, **278**, 9544.
- [42] M. Bejger, B. Imiolczyk, D. Clavel, M. Gilski, A. Pajak, F. Marsolais, M. Jaskolski, *Acta Cryst. D*, 2014, in press.
- [43] C. Jelsch, M.M. Teeter, V. Lamzin, V. Pichon-Pesme, R.H. Blessing, C. Lecomte, *Proc. Natl. Acad. Sci. USA*, 2000, **97**, 3171.
- [44] B.S. Kang, Y. Devedjiev, U. Derewenda, Z.S. Derewenda, *J. Mol. Biol.*, 2004, **338**, 483.

-
- [45] J.H. Konnert, W.A. Hendrickson, *Acta Cryst.*, 1980, **A36**, 344.
[46] R.A. Engh, R. Huber, *Acta Cryst.*, 1991, **A47**, 392.
[47] M. Jaskolski, M. Gilski, Z. Dauter, A. Wlodawer, *Acta Cryst.*, 2007, **D63**, 611.
[48] J. Koepke, E.I. Scharff, C. Lucke, H. Ruterjans, G. Fritzsche, *Acta Cryst.*, 2003, **D59**, 1744.
[49] L. Esposito, L. Vitagliano, L. Mazzarella, *Protein Pept. Lett.* 2002, **9**, 95.
[50] A. Addlagatta, H. Czapinska, S. Krzywda, J. Otlewski, M. Jaskolski, *Acta Cryst.*, 2001, **D57**, 649.
[51] L. Esposito, L. Vitagliano, A. Zagari, L. Mazzarella, *Protein Eng.*, 2000, **13**, 825.
[52] B. Stec, *Acta Cryst.*, 2007, **D63**, 1113.
[53] I.J. Tickle, *Acta Cryst.*, 2007, **D63**, 1274.
[54] M. Jaskolski, M. Gilski, Z. Dauter, A. Wlodawer, *Acta Cryst.*, 2007, **D63**, 1282.
[55] D.E. Tronrud P.A. Karplus, *Acta Cryst.*, 2011, **D67**, 699.
[56] N.K. Hansen, P. Coppens, *Acta Cryst.*, 1978, **A34**, 909.
[57] C. Jelsch, B. Guillot, A. Lagoutte, C. Lecomte, *J. Appl. Cryst.*, 2005, **38**, 38.

Praca wpłynęła do Redakcji 16 maja 2014

