

Difrakční limit v makromolekulární krystalografii

Martin Malý^{1,2}, Jarmila Dušková², Lars Henrik Østergaard³, Jan Stránský^{1,2},
Leona Švecová^{1,2}, Petr Kolenko^{1,2}, Jan Dohnálek²

¹Katedra inženýrství pevných látek, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze, Břehová 7, 115 19, Praha 1, Česká republika

²Biotechnologický ústav AV ČR, v. v. i., BIOCEV, Průmyslová 595, 252 50, Vestec, Česká republika

³Novozymes A/S, Brudelysvej 26, DK-2880 Bagsvaerd, Dánsko
malymar9@fjfi.cvut.cz

Abstrakt

Difrakční limit pro vysoké rozlišení je jedním z důležitých parametrů, které je třeba vhodně zvolit při řešení makromolekulární struktury. Na testovací soubor difrakčních dat z proteinového krystalu byla aplikována různá kritéria volby difrakčního limitu. Následně byl zanalyzován jejich dopad na mapy elektronové hustoty, výslednou strukturu a její statistiky. Důraz byl kladen především na metodiku tzv. párového upřesňování – analýzu statistik upřesňování R_{work} and R_{free} .

Klíčová slova: Rentgenová strukturní analýza; Difrakce; Difrakční limit.

Úvod

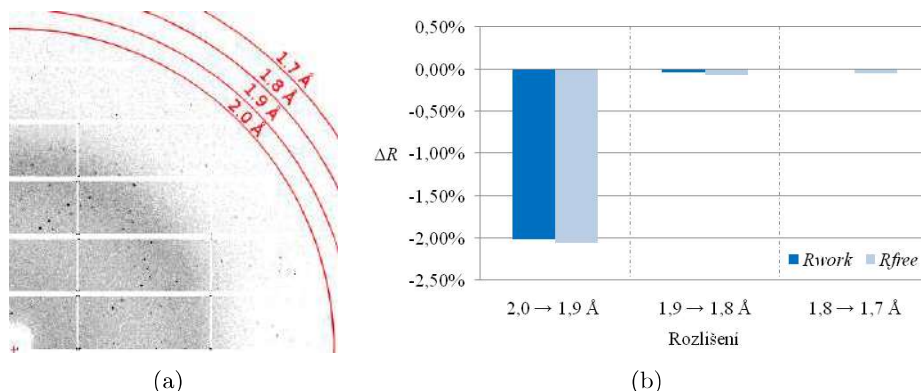
Intenzita reflexí pozorovaných na difrakčních snímcích obecně klesá s rostoucím rozlišením. V oblasti vysokého rozlišení proto data obvykle obsahují už jen velmi málo užitečného signálu a naopak jsou výrazně zatížena šumem. Při zpracování difrakčních dat je tedy nutné stanovit difrakční limit pro vysoké rozlišení – hodnotu rozlišení, do kterého budou data zpracována; zbylá data nebudou použita.

Pro určení difrakčního limitu je třeba kvalitu dat kvantifikovat. Mezi nejpoužívanější statistiky patří R -faktory R_{merge} a R_{meas} , poměr signálu ku šumu I/σ a korelační koeficient $CC_{1/2}$. Tyto indikátory kvality jsou však bez vztahu ke strukturnímu modelu. Dále je diskutabilní, které jejich hodnoty jsou ještě přípustné pro použitelná data.

Po zpracování dat a vyřešení fázového problému je možné ověřit volbu difrakčního limitu pomocí analýzy R -faktorů R_{work} a R_{free} v průběhu procesu upřesňování struktury [1]. Tyto statistiky jsou vztaženy k aktuálnímu strukturnímu modelu a jejich současný pokles (případně pokles pouze R_{free}) značí vylepšení modelu. Tento postup bude demonstrován na testovacích datech struktury enzymu s ligandem.

Experiment

Difrakční data byla získána na synchrotronu Petra III v Hamburku (stanoviště P13). Vzdálenost vzorek-detektor byla 419,67 mm, další parametry experimentu byly: $\Delta\phi = 0,1^\circ$; $\Delta t = 0,15$ s; $\phi_{\text{total}} = 200^\circ$. Difrakční data byla zpracována pomocí programu XDS [2].



Obrázek 1: (a) Vybraná část difrakčního snímku krystalu enzymu s vyznačenými slupkami s vysokým rozlišením. (b) Pokles hodnot R -faktorů R_{work} a R_{free} příslušných k zvoleným difrakčním limitům. ΔR je hodnota rozdílu hodnot R -faktoru nižšího a vyššího rozlišení.

Výsledky

Reflexe byly při vizuální inspekci spatřeny do rozlišení 2,2 Å. Při stanovení parametru průměrného $I/\sigma \geq 2$ pro vrstvu reflexí s nejvyšším rozlišením byl difrakční limit stanoven na 2,0 Å.

Fázový problém byl vyřešen metodou molekulárního nahrazení za použití struktury téhož enzymu – PDB kód 3RJ8 [3]. Model byl upřesněn pomocí programu Refmac5 [4].

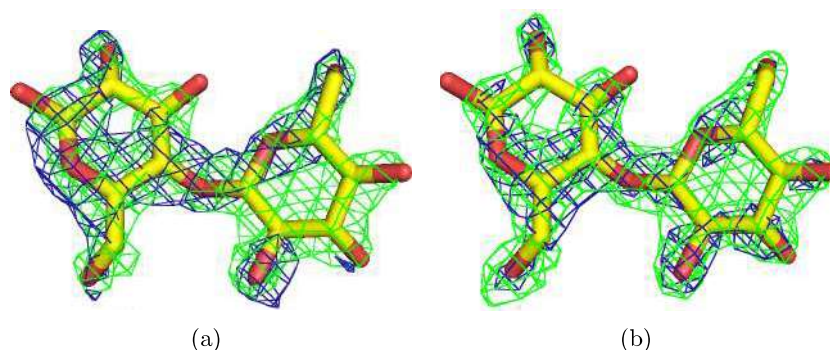
Před modelováním ligandů byla do procesu upřesňování struktury postupně přidávána data ze slupek 2,0–1,9 Å; 1,9–1,8 Å a 1,8–1,7 Å (Obr. 1(a)). Po přidání každé slupky byl strukturní model upřesněn (20 cyklů) a současně byly porovnány R -faktory na stejné hladině rozlišení (Obr. 1(b)). K významnému snížení hodnot obou R -faktorů došlo pouze po přidání dat ze slupky 2,0–1,9 Å. Difrakční limit pro vysoké rozlišení byl tedy na základě této analýzy změněn na 1,9 Å.

Tabulka 1: Statistické výsledky zpracování a upřesňování. Hodnoty v závorkách odpovídají slupce s nejvyšším rozlišením.

Vlnová délka [Å]	0,9201	Počet upřesňovaných atomů	4 350
Prostorová grupa	$C121$	Počet upřesňovaných molekul vody	521
	$a = 132,77; b = 56,99;$	Průměrný ADP	31
Elem. buňka [Å, °]	$c = 86,88; \beta = 95,66;$	$R_{\text{ms bond}}$ [Å]	0,010
	$\alpha = \gamma = 90,00$	$R_{\text{ms angle}}$ [°]	1,415
Difrakční limit [Å]	50,00–1,90 (2,01–1,90)	R_{free} (5 %)	0,213
Počet reflexí	135 014 (18 160)	R_{work} (95 %)	0,173
Počet nez. reflexí	48 682 (7 325)	Ramach.: povolené konformace [%]	100
R_{merge}	0,101 (1,013)	Ramach.: zakázané konformace [%]	0
R_{meas}	0,122 (1,251)		
$CC_{1/2}$	0,995 (0,354)		
I/σ [I]	7,9 (1,3)		
Redundance	2,8 (2,5)		
Kompletnost [%]	94,9 (89,5)		

Diskuze

Díky analýze R -faktorů byla použita i data ze slupky o rozlišení 2,0–1,9 Å, která byla původně podle konzervativního kritéria pro I/σ vyloučena. Tato slupka však obsahuje 15 % nezávislých reflexí (Tab. 1). Data z vysokého rozlišení navíc obsahují informace o strukturních detailech. V důsledku použití dat z této slupky došlo k vylepšení strukturního modelu, což bylo zejména patrné na elektronové hustotě molekuly ligandu (Obr. 2).



Obrázek 2: Mapy elektronové hustoty ligandu v aktivním místě enzymu. Modrá barva přísluší zpracování s difrakčním limitem 2,0 Å, zelená odpovídá optimalizovanému zpracování s difrakčním limitem 1,9 Å. (a) $2F_O - F_C$ mapy na hladině elektronové hustoty $1,5\sigma$. (b) $F_O - F_C$ (diferenční) „omit“ mapy na hladině elektronové hustoty $3,8\sigma$.

Závěr

Difrakční limit je důležitý parametr zpracování difrakčních dat. Na modelovém příkladě byl demonstrován postup, jak zvolit difrakční limit pro vysoké rozlišení tak, aby byla použita všechna data obsahující užitečný signál. Byl pozorován pozitivní vliv této volby na podobu map elektronové hustoty, a tedy přímo na kvalitu strukturního modelu.

Reference

- [1] P. A. Karplus, K. Diederichs. Linking Crystallographic Model and Data Quality. *Science* **336**: 1030–1033, 2012.
- [2] W. Kabsch. XDS. *Acta Cryst. D***66**: 125–132, 2010.
- [3] J. Dušková, T. Skálová, A. Štěpánková, T. Kovař, J. Hašek, L. H. Østergaard, C. C. Fuglsang, P. Kolenko, J. Dohnálek. PDB ID: 3RJ8. [2017-04-15]. <http://www.rcsb.org/pdb/explore/explore.do?structureId=3RJ8>. 2011.
- [4] G. N. Murshudov, A. A. Vagin, E. J. Dodson. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst. D***53**: 240–255, 1997.

Poděkování

Tato práce byla podpořena Interní grantovou agenturou ČVUT (projekt SGS16/246/OHK4/3T/14), MŠMT (projekty NPUII - LQ1604, LM2015043 a LG14009), ERDF (projekt BIOCEV - CZ.1.05/1.1.00/02.0109) a institucionální podporou BTÚ AV ČR, v.v.i. (RVO: 86652036).