

CO NABÍZÍ SOUČASNÁ RTG STRUKTURNÍ ANALÝZA?

**BOHUMIL KRATOCHVÍL^a, MICHAL HUŠÁK^a,
JIŘÍ BRYNDA^{b,c} a JURAJ SEDLÁČEK^b**

^a Ústav chemie pevných látek, Vysoká škola chemicko-technologická v Praze, Technická 5, 166 28 Praha 6,

^b Laboratoř strukturní biologie, Ústav molekulární genetiky AV ČR, v.v.i., Flemingovo nám. 2, 166 37 Praha 6,

^c Tým strukturní biologie, Ústav organické chemie a biochemie AV ČR, v.v.i., Flemingovo nám. 2, 166 10 Praha 6
bohumi.kratochvil@vscht.cz, brynda@img.cas.cz

Došlo 4.5.08, přijato 3.7.08.

Klíčová slova: RTG difrakce, RTG strukturní analýza malých molekul, proteinová krystalografie

Věnováno profesoru Rudolfu Zahradníkovi k osmdesátým narozeninám.

Obsah

1. Historický úvod
2. Princip RTG strukturní analýzy
3. RTG difrakční experiment a jeho zpracování
 - 3.1. Hrubá výpočetní síla
4. RTG strukturní analýza malých molekul
 - 4.1. Tématické trendy v krystalografii malých molekul
5. Proteinová krystalografie
 - 5.1. Zvláštnosti postupů proteinové krystalografie
 - 5.2. Zuzitkování strukturní informace
6. Predikce
7. Závěr

1. Historický úvod

RTG difrakce je stará, bezmála stoletá dáma. O její objev se zasloužili Němci: Max von Laue, Walter Friedrich a Paul Knipping. V roce 1912 ozářili krystal modré skalice RTG svazkem a zjistili, že rozptýlená energie se od krystalu šíří pouze v určitých směrech, zatímco v jiných vyhasíná¹. Tento významný objev nedestruktivní interakce RTG záření s krystalickými materiály vedl ke vzniku nové analytické techniky – RTG strukturní krystalografie (našimi organiky nazývané „ix rej“).

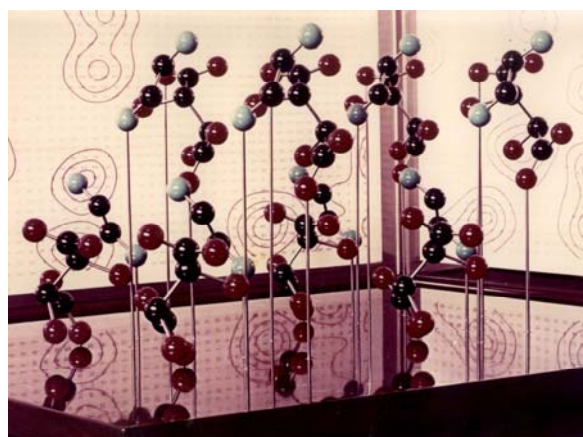
Fakt, že v RTG difrakčním obraze krystalu musí být zakódována informace o jeho vnitřní struktuře si poprvé uvědomil William L. Bragg, který ke svým pokusům začal

používat monochromatický RTG svazek. V letech 1912 až 1922 určil², za přispění svého otce Williama H. Bragga, krystalové struktury prvků (diamantu, grafitu, mědi), jednoduchých anorganických sloučenin a minerálů (chloridu sodného a draselného, sfaleritu, fluoritu, pyritu, kalcitu, korundu, spinelu, ledu) a první organické látky (naftalenu).

U nás se počátky RTG strukturní analýzy datují mnohem později. Okolo roku 1950 začal fyzik Allan Líněk řešit strukturu ethylendiamonium-tartarátu, aby vysvětlil mechanismus jeho piezoelektrických vlastností. Řešení dosáhl až v roce 1957 (cit.³) a model vyřešené struktury byl vystaven na EXPO1958 v Bruselu (obr. 1).

S počátky RTG strukturní analýzy na VŠCHT Praha je spojena i úsměvná historka, o které před lety vyprávěl prof. Jaroslav Bauer z tehdejší katedry mineralogie. Jakmile se koryfej české organické chemie prof. Rudolf Lukeš v 50. letech minulého století dozvěděl o možnosti určit krystalovou strukturu z difrakce RTG záření, okamžitě přinesl dva krystalické vzorky a žádal asistenta Bauera o jejich změření. Hned druhý den požadoval vyřešené struktury, ale chudák Bauer mu mohl ukázat pouze dva „šmouhogramy“ (práškové rentgenogramy pořízené na film metodou Debyeho-Scherrerovou). Při pohledu na ně ztratil prof. Lukeš o RTG strukturní analýzu zájem a zřejmě o tom informoval kolegy, protože dlouhá desetiletí potom čeští organici pohlíželi na „rentgenáře“ nedůvěřivě.

Dalším mezníkem světového vývoje bylo vyřešení prvních struktur biologických makromolekul. V roce 1962 obdrželi Nobelovu cenu za chemii John C. Kendrew a Max F. Perutz za výzkum struktury globulárních proteinů a za lékařství Francis H. C. Crick, James D. Watson a Maurice H. F. Wilkins za určení molekulární struktury nukleových kyselin pomocí RTG paprsků. Od této doby



Obr. 1. Model první vyřešené krystalové struktury u nás (ethylendiamonium-tartarát). Snímek laskavě poskytl Dr. Ctirad Novák

se RTG strukturní analýza postupně rozdělila na malé molekuly (zhruba do 1000 nevodíkových atomů) a velké molekuly (proteinová krystalografie, ale netýkající se pouze proteinů).

Je pravda, což dokumentuje i tento článek, že „tématické nůžky“ mezi malou a velkou krystalografií se stále více rozevírají (využití výsledků, metody krystalizace). Na druhé straně se zdokonalováním RTG difrakčního experimentu se „proteináři“ začínají zajímat i o věci z malé krystalografie (dvojčatění, korekce na absorpci, „disorder“, modulace). Na velkých krystalografických konferencích^{4,5} je krystalografie malých a velkých molekul stále prezentována „pod jednou střechou“.

Hlavní brzdou širší aplikace RTG strukturní analýzy v chemii byla v začátcích právě její relativní pomalost. Většinu chemiků nezajímala mnohaletá řešení krystalových struktur, ve kterých se utápěli RTG strukturní analytici. Pokrok přineslo jak výrazné zdokonalení techniky RTG difrakčních experimentů a metod pěstování monokrystalů, tak vývoj matematických metod řešení struktur a samozřejmě i obrovský skok ve vývoji počítačů. Za bezmála stoletou existenci obdrželi krystalografové za objev a zdokonalení metody RTG strukturní analýzy a za její aplikace v přírodních vědách dvanáctkrát Nobelovu cenu.

2. Princip RTG strukturní analýzy

Pokud pomineme defekty reálné krystalové struktury, je ideální krystal uvnitř (na atomové úrovni), uspořádán trojrozměrně periodicky, tzn. opakuje se určitý atomový motiv. V krystalech elementárních kovů je tímto motivem jeden atom, v krystalech biologických objektů složených z různých makromolekul to může být obrovský počet nevodíkových atomů (například u větší podjednotky ribosomu přes 64 tisíc⁶). Toto uspořádání, které nazýváme krystalová struktura, se bohužel nepřesně, ale velmi vžitě, označuje jako krystalová mřížka⁷. Aproximace atomového motivu bodem vede k pojmu prostorová mřížka. Pokud přes fyzikální realitu – krystalovou strukturu, přeložíme naši aproximaci – prostorovou mřížku, vymežíme v krystalové struktuře rovnoběžnostěny (elementární buňky), které jsou identické co do rozměrů i hmotné náplně a jejich orientace v prostoru.

Řešení (stanovení) ideální krystalové struktury znamená určení a upřesnění souřadnic a parametrů teplotního pohybu všech atomů v elementární buňce. Přesněji řečeno pouze v její asymetrické části, protože každá krystalová struktura je symetrická, jak popisuje její prostorová grupa. Prostorových grup je pouze omezený počet (230), což souvisí s faktem, že prostorová symetrie krystalů musí být konzistentní s jejich vnitřní periodicitou. Pozice všech atomů v krystalu (crystal packing) mohou být generovány příslušnými operacemi symetrie dané prostorové grupy. Kromě atomových parametrů jsou nezbytnými dalšími parametry vyřešené krystalové struktury rozměry elementární buňky (mřížkové parametry) a prostorová grupa symetrie.

K určení krystalové struktury musíme krystal ozářit monochromatickým RTG svazkem a tak získat jeho pozorovatelný difrakční obraz. Dopadající (primární) RTG svazek se pružně rozptyluje na elektronech měřeného krystalu (vzniká sekundární neboli difraktované záření). Důsledky tohoto jevu lze jednoduše popsat, podle Braggovy šikovní a fungující interpretace, jako reflexe primárního svazku na rovinách, které lze prokládat krystalovou strukturou. Difrakce se proto také označují jako reflexe. Reflektující roviny v krystalu se rozlišují hodnotami Millerových indexů *hkl*.

Difrakční obraz krystalu však není mikroskopickým obrazem jeho vnitřní struktury. Analyticky využitelnými veličinami v difrakčním obraze krystalu jsou intenzity a polohy (úhly) jednotlivých difrakcí. Kromě toho mohou některé difrakce systematicky vyhasínat (mít nulovou intenzitu). Z intenzit difrakcí stanovíme upřesněné pozice atomů a jejich teplotně-vibrační parametry, z poloh difrakcí rozměry elementární buňky a ze systematického vyhasínání prostorovou grupu.

Pro určení pozic atomů je třeba vypočítat mapu distribuce elektronové hustoty v asymetrické části elementární buňky. Maxima této mapy (těžiště elektronových obalů atomů), dobře koincidují (s výjimkou atomů vodíku) s pozicemi jader izolovaných atomů. Závěrem je nutné upřesnit pozice atomů, odečtené z map elektronových hustot, a jejich teplotně-vibrační parametry. Upřesňujeme proto, že máme k dispozici více pozorování (měřených reflexí) než parametrů (pozice atomů a jejich teplotní vibrace). Každá reflexe je nezávislý experiment s vlastní chybou.

Z hlediska chemie podává rutinní RTG strukturní analýza velmi užitečnou, ale na druhé straně statickou informaci. Krystalová struktura v ozařovaném objemu krystalu je prostorově zprůměrována do jedné elementární buňky, resp. její asymetrické části, a časově do délky trvání difrakčního experimentu. Tímto zprůměrováním se přichází o informace o defektech reálné krystalové struktury, což je ale řešitelné jinak, např. metodou tzv. difrakčního kontrastu.

Velkou výzvou RTG strukturní analýzy je sledovat měnící se strukturu látky během chemické reakce *in situ* (časově rozlišená RTG krystalografie).

3. RTG difrakční experiment a jeho zpracování

Výchozím materiálem pro RTG difrakční experiment je buď monokrystal nebo polykrystalický materiál (u malých molekul). Pokud je stanovení struktury provedeno z monokrystalových dat, hovoříme o monokrystalové RTG strukturní analýze, pokud řešíme strukturu z polykrystalických (práškových) dat, hovoříme o práškové RTG strukturní analýze. Progresivní metodiku práškové RTG strukturní analýzy jsme v Chemických listech již podrobně popsali⁸.

Experimentální zařízení pro monokrystalovou RTG strukturní analýzu se nazývá čtyřkruhový difraktometr,



Obr. 2. Laboratorní monokrystalový RTG difraktometr pro řešení struktury malých molekul. Zleva doprava: kontejner s kapalným dusíkem, čtyřkruhový goniometr s měřeným monokrystalem a pod ním zdroj RTG záření, chlazení RTG lampy, chlazení detektoru difraktovaných paprsků

který je v laboratorním provedení zhruba tak velký jako šatní skříň (obr. 2). K rutinní strukturní analýze malých molekul stačí monokrystalický vzorek o rozměrech 10^{-1} mm, přičemž v jednom rozměru to může být i 10^{-2} mm (např. monokrystalické jehličky nebo lístečky). Existují však i vícekruhové monokrystalové difraktometry, které jsou připojeny na synchrotronové záření v RTG oblasti. Zde je možné měřit i monokrystaly, které mají řádově všechny rozměry 10^{-2} mm. Zmenšováním rozměrů monokrystalu samozřejmě vzrůstají nároky na manipulační techniku.

Pro proteinovou krystalografii jsou nutné zdroje intenzivního RTG záření (vzhledem k velkým elementárním buňkám a slabým difrakcím). Generátory, nejčastěji s měděnou rotační anodou, se používají u laboratorního (in



Obr. 3. Proteinový difraktometr u RTG svazku ID14-1 (beam line) synchrotronu ESRF v Grenoblu¹⁰. Toto zařízení je samozřejmě podstatně větší než laboratorní difraktometr na obr. 2

house) sběru difrakčních dat (u nás slouží jeden – na pracovišti třetího a čtvrtého autora, díky Grantové agentuře ČR). Intenzivní a koncentrovaný synchrotronový RTG svazek vybudí i v nepatrném monokrystalu proteinu (10^{-2} mm) daleko více měřitelných reflexí než laboratorní generátor. Získané soubory difrakčních dat jsou pak kvalitnější a poskytují vyšší rozlišení. Jemná laditelnost vlnové délky synchrotronového RTG záření je nezbytná pro použití některých experimentálních technik určení fázi (např. MAD, Multiple Anomalous Dispersion). Statistika⁹ ukazuje, že v roce 2005 se počet proteinových struktur vyřešených synchrotronovými RTG svazky už přiblížil počtu struktur vyřešených laboratorními RTG generátory. Domácí proteinový krystalograf je zatím odkázán na zahraniční synchrotronová zařízení. Ta jsou ale dobře dostupná, pokud stručná žádost prokáže smysluplnost požadavku na měření.

Dosažený stupeň rozlišení struktury samozřejmě závisí na kvalitě měřeného monokrystalu. Počet difrakcí nutných k řešení krystalové struktury souvisí se složitostí měřeného krystalu (od několika desítek u krystalů kovů až k několika milionům u krystalů virů). Doba sběru dat kolísá od několika hodin k několika dnům. Měření se běžně provádí za nízké teploty (100–150 K), nejčastěji v proudě chladného dusíku. Důvodů je několik: potlačení teplotně-vibračního pohybu atomů, který řešení struktury komplikuje, a u proteinů hlavně ochrana před radiačním poškozením, které by rychle nastalo pod RTG svazkem za pokojové teploty.

Nezbytnost dostatečně kvalitního vstupního materiálu je pro monokrystalovou RTG strukturní analýzu vážným omezením. Často se nedaří připravit vhodný monokrystal, někdy to z principu není možné. U reakcí v pevném stavu získáme vzájemným „semletím“ dvou práškových reaktantů novou, ale opět práškovou fázi. V proteinové krystalografii je situace ještě komplikovanější a krystalizace se zde stala již samostatným specializovaným oborem¹¹. Preparát bílkoviny pro krystalizaci musí vykazovat uniformitu v kovalentní a konformační struktuře, možnost zahuštění ke koncentraci až 10 mg ml^{-1} , a s výhodou také testovatelnou biologickou aktivitu a stálost.

3.1. Hrubá výpočetní síla

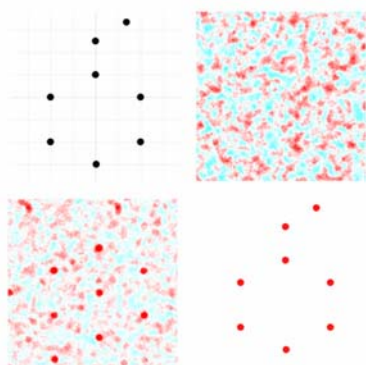
Pro výpočet map distribuce elektronové hustoty jsou kromě intenzit difrakcí zapotřebí i údaje o fázových úhlech (fázích) difraktovaných paprsků. Tyto údaje nejsou dostupné z RTG difrakčního experimentu. I při náhodných intenzitách, ale správných fázích bychom získali v zásadě správný model struktury, ale bohužel obráceně to neplatí. To proto, že průběh funkce elektronové hustoty výrazněji závisí na fázových úhlech než na intenzitách. Úvodní neznalost hodnot fází se ve strukturní krystalografii nazývá fázový problém. Tento problém je však v zásadě řešitelný, protože hodnoty fází lze extrahovat ze souboru naměřených intenzit tzv. přímými metodami – aplikací vztahů založených na nerovnostech, statistice a počtu pravděpodobností (Sayreho vztah, strukturní invarianty a semi-

invarianty, symbolická adiční procedura, vztah Σ_2 , tangen-
tová formule aj.). Na základě těchto vztahů se vybere nej-
více pravděpodobné fázování reflexí, a to se pak využije
pro výpočet mapy elektronové hustoty a určení pozic ato-
mů. Použití přímých metod vyžaduje vstupní informace
o symetrii a geometrii krystalové struktury, sumárním
chemickém složení a je předpokládáno i určité typické
rozložení elektronové hustoty v prostoru.

Přímé metody dosáhly časem tak vysokého stupně
propracovanosti a univerzálnosti, že by už nikdo nepřed-
pokládal jejich podstatný rozvoj. Přesto byla v roce 2004
(cit.¹²) popsána nová metoda, nazvaná „charge flipping“,
která RTG strukturní analýzu malých molekul řeší netra-
dičním způsobem. Tato metoda je z hlediska složitosti
přístupu krokem zpět ve vývoji algoritmů. Místo výše
uvedených sofistikovaných vztahů přímých metod se vyu-
žívá hrubé síly současné výpočetní techniky. Základem
použití je následující princip: elektronová hustota musí
mít v krystalu všude kladnou hodnotu. Pokud ji má zápor-
nou, musíme ji opravit. Jak algoritmus „charge flipping“
v praxi funguje? Postup lze rozdělit na následující kroky:

1. Vstupním datům (intenzitám naměřených difrakcí)
přiřadíme náhodné fáze.
2. Z intenzit a fází vypočteme mapu elektronové hustoty.
Ze začátku bude výsledek velice vzdálený skutečné
elektronové hustotě a v řadě míst bude mít mapa zá-
pornou hodnotu.
3. V mapě najdeme místa, jejichž velikost je menší než
určitá hranice, a hustotu v těchto místech násobíme -1 .
4. Z upravené mapy elektronové hustoty spočítáme fáze.
5. Nové fáze přiřadíme k experimentálně naměřeným
intenzitám a vracíme se do bodu 2.

Výpočet považujeme za dokončený, pokud je elektro-
nová hustota ve všech bodech výsledné mapy vyšší než
určený limit nebo pokud se v několika po sobě jdoucích
krocích výsledek nemění. Celý algoritmus je velice primi-
tivní, výpočetně náročný, ale je snadno implementovatelný
a u dostatečně kvalitních vstupních dat konverguje překva-



Obr. 4. Výsledek simulovaného řešení struktury metodou
„charge flipping“ (kladná elektronová hustota červeně, zá-
porná modře). Zleva doprava: ideální struktura, výsledek po 1.
cyklu iterace, po 5. cyklu, po 50. cyklu¹³

pivě rychle ke správnému řešení (obr. 4).

Jaké jsou zásadní výhody daného postupu?

Není nutný žádný předběžný předpoklad o prostorové
grupě látky. Problém lze řešit v nejméně symetrické grupě
 $P1$ a skutečnou symetrii zjišťovat v dalších fázích řešení.

Algoritmus při výpočtu nepoužívá žádné předběžné
předpoklady o rozmístění atomů. Dá se proto snadno mo-
difikovat pro řešení nestandardních problémů, jako jsou
modulované struktury nebo kvazikrystaly.

Časová náročnost algoritmu není díky možnostem
současné výpočetní techniky zásadním problémem.

Algoritmus „charge flipping“ je implementován
ve volně dostupných programech, zejména v softwaru
SUPERFLIP¹⁴. V současné době probíhá jeho testování
nejen pro řešení struktur malých molekul z monokrystalu,
ale i pro řešení struktury z práškových dat a struktur pro-
teinů. Zda se tato metoda stane plnohodnotnou náhradou
nebo jen doplňkem přímých metod, ukáže čas.

V souvislosti s řešením struktury z práškových dat
dochází k obnovení zájmu o rozvoj algoritmů pro indexaci
práškových difraktogramů. Správná indexace je prvním
důležitým krokem při řešení struktury. Kromě rozvoje
standardních algoritmů založených na statistice nebo ná-
hodném přiřazování Millerových indexů hkl jednotlivým
difrakcím (programy DICVOL, ITO, TREOR)⁸, se obje-
vují i nové přístupy založené na současných výpočetních
možnostech. Příkladem je program McMaille¹⁵. Tento
program řeší zadání buď metodou simulovaného žíhání
(hledání optimální buňky z náhodných nástřelů) nebo do-
konce systematickým prohledáváním všech možných bu-
něk. Oproti standardním metodám je algoritmus velmi
výpočetně náročný, ovšem praktické testy ukázaly¹⁶, že
jeho úspěšnost je ze všech existujících postupů nejvyšší.
V publikovaných testech byl software McMaille schopný
správně indexovat nejen data zatížená velkou experimen-
tální chybou, ale i jednodušší případy směsí. Extrémní
výpočetní náročnost programu (až několik dnů pro větší
buňky) se daří kompenzovat pomocí verzí optimalizova-
ných pro moderní vícejaderné procesory – výpočtem zalo-
ženým na paralelním řešení úlohy na více jádrech.

V proteinové krystalografii existují tři základní způ-
soby jak získat informace o fázích. Do první skupiny patří
metody založené na porovnávání Pattersonovy funkce
(funkce meziatomových vektorů) pro aktuální naměřená
data a pro podobný, homologický protein s již vyřešenou
strukturou. Mapu Pattersonovy funkce lze získat bez zna-
losti hodnot fází a její maxima odpovídají meziatomovým
vektorům. Její použití je zvláště výhodné, když protein
obsahuje nebo váže jeden nebo několik atomů s výrazně
vyššími protonovými čísly (např. selen inkorporovaný do
bílkovinného řetězce jako selenomethionin nebo xenon či
jód vnesený do krystalu).

Druhou skupinu tvoří metody, které se někdy označu-
jí jako experimentální. Je to např. metoda využívající ano-
málního rozptylu a metoda izomorfní záměny.

Do třetí skupiny patří přímé metody stanovení fází
založené na statistických metodách (např.: SHELXD¹⁷
nebo Shake-and-Bake¹⁸). Tyto metody mají v proteinové

krystalografii omezené použití vzhledem k velkému objemu dat a jejich nízkému rozlišení. S úspěchem jsou však používány pro řešení substruktur, což je nezbytný postupný krok druhé skupiny určení fází.

Pro metody první skupiny se vžil název „molekulové nahrazení“ (Molecular Replacement) a to proto, že pomocí známého strukturálního modelu pro podobnou molekulu hledáme vhodnou polohu a vhodné umístění studované molekuly v elementární buňce krystalu, který jsme měřili. Používané programy pro řešení problému fází metodou molekulového nahrazení s využitím Pattersonovy funkce v klasickém dvoukrokovém uspořádání jsou např. MOLREP¹⁹ a AMORE²⁰, které obsahuje soubor krystalografických programů CCP4 (Collaborative Computational Project, number 4.) Vzhledem k tlaku na robustní výpočetní metody jsou v současnosti ve větší míře využívány metody molekulového nahrazení, které prohledávají jednoduše šest parametrů umístění molekul v elementární buňce (tři translačně-polohové a tři rotačně-úhlové). Výhodou těchto metod je větší citlivost pro nalezení správného řešení, jelikož se netříští signál na příspěvek translační a rotační. Příkladem je program EPMR²¹, který využívá evoluční mechanismus k redukci výpočetního času a program BRUTE²², který opravdu systematicky prohledává N-rozměrný prostor se zvoleným krokem.

4. RTG strukturální analýza malých molekul

Výsledkem řešení krystalové struktury je seznam upřesněných pozic všech atomů v asymetrické části elementární buňky spolu s jejich upřesněnými teplotně-vibračními parametry. Přesněji řečeno kromě atomů vodíku, protože jejich pozice se většinou dopočítávají z očekávané geometrie (např. fenylové kruhy apod.).

Ukazatelem věrohodnosti vyřešené struktury je hodnota tzv. R-faktoru (reliability factor), což je v zásadě sumární relativní chyba mezi experimentálně zjištěnými intenzitami difrakcí a intenzitami vypočtenými z vyřešeného modelu struktury. U správně vyřešené struktury z monokrystalu se hodnota R-faktoru pohybuje okolo 5 % a méně. U struktur stanovených z RTG práškových dat se běžně setkáváme s hodnotami přes 10 %.

Výsledky RTG strukturální analýzy se dnes standardně ukládají do souboru cif (Crystallographic Information File²³), z kterého lze vypočítat libovolné údaje o geometrii molekulové i krystalové struktury (vzdálenosti, úhly, prokládání rovin a přímek atd.). Užitečný je také výpočet tvaru a objemu kavit (prázdných míst v krystalové struktuře). Tato informace nám může např. posloužit při předpovědi solvatace farmaceutických substancí, tzn. jak velké molekuly solventu a jakého tvaru se do krystalové struktury substance mohou vejít²⁴.

Za zvláštní zmínku stojí stanovení absolutní chiralitu z RTG difrakčních dat. To je možné pouze u enantiomeru, protože ten krystaluje v polární prostorové grupě (nemá střed symetrie ani jiné operace symetrie, které jej zahrnují). Z celkového počtu 230 prostorových grup je pouze 65

polárních. U racemátu je možné určit pouze relativní chiralitu (krystaluje v centrosymetrické prostorové grupě). Stanovení absolutní chiralitu je založeno na anomálním rozptylu primárního RTG svazku na elektronech měřené látky. Míra anomálního rozptylu závisí na velikosti atomu a použité vlnové délce dopadajícího RTG svazku. Předpokladem úspěchu při stanovení absolutní chiralitu je přítomnost těžkého atomu ve struktuře (např. Br a výše). Pokud struktura obsahuje pouze lehké atomy (C, N, O, H), jsme schopni stanovit pouze relativní chiralitu.

Kromě toho může RTG strukturální analýza sloužit jako východisko pro teoretické modely jiných metodik, příp. je korigovat (IČ, NMR).

Experimentálně vyřešené krystalové struktury se ukládají do databází, nejdůležitější jsou: CSD²⁵ (Cambridge Structural Database) – asi 424 tis. organických a organometalických struktur; ICSD²⁶ (Inorganic Crystal Structure Database – asi 92 tis. struktur prvků a anorganických sloučenin; CRYSTMET²⁶ – asi 110 tis. struktur kovů, slitin a intermetalických fází.

4.1. Tématické trendy v krystalografii malých molekul

Představu o tom, co řeší současná RTG strukturální krystalografie, lze získat z programu velkých krystalografických konferencí (24th European Crystallographic Meeting, ECM 24, Marakéš, Maroko 2007 a XXI Congress and General Assembly of the International Union of Crystallography, IUCr2008, Osaka, Japonsko 2008)^{4,5}.

Z hlediska složení se studují struktury anorganické, mineralogické, organické, organokovové a polymerní. Z hlediska aplikace struktury farmaceutické, struktury pro ukládání jaderného odpadu, struktury typu hostitel-host, struktury aperiodické, struktury potravinářských ingrediencí, struktury uhlíkatých nanotrubiček a dalších materiálů, struktury defektní, struktury amorfních a semikrystalických fází, struktury povrchů, struktury monitorované pro zachování a konzervaci kulturního dědictví a jiné. Vedle pevné fáze se studují i strukturální jevy v kapalinách. Hledají se cesty, jak řídit krystalizaci a budovat struktury elektronové, krystalové, mikro- a nanostruktury s předem vytipovanými vlastnostmi. Zkoumají se krystalové stavební jednotky a jejich vztah k samouspořádávacím procesům vedoucím k supramolekulám. Velká pozornost je věnována strukturálnímu studiu fázových přechodů a reakcím typu monokrystal → monokrystal a monokrystal → prášek. Zdokonaluje se RTG difrakční experiment za normálních a extrémních podmínek a jeho časově-rozlišené verze, vyvíjejí se RTG zdroje (laboratorní i synchrotronové) a detektory RTG záření a v neposlední řadě i matematické metody zpracování difrakčních dat. RTG strukturální analýza se kombinuje s ostatními strukturálními metodikami: elektronovou a neutronovou difrakcí a vysoce-rozlišenou elektronovou mikroskopií, NMR, hmotnostní spektroskopií atd.

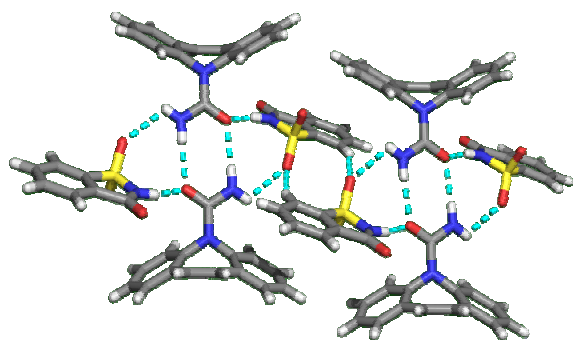
Rychleji se rozvíjí stanovení struktury z práškových dat před daty z monokrystalu. Práškový materiál je daleko

častěji k dispozici než monokrystal, jehož vypěstování se vůbec nemusí povést. Na druhé straně však z práškového materiálu nikdy nedostaneme tak velký soubor 3D difrakcí jako z monokrystalu. Tento rozdíl dohání hrubá výpočetní síla. Kromě toho jsou intenzity difrakcí z práškových dat, změřené na laboratorních difraktometrech, skoro vždy, více či méně, postiženy přednostní orientací (texturou) krystalitů, jejíž korekce je složitá. Základem úspěšného řešení struktury z prášku je proto měření na synchrotronu. Zde se práškový vzorek měří v kapiláře, kdy jednotlivé reflexe jsou dostatečně intenzivní a maximálně separované. Opatření dat ze synchrotronu je dnes komerčně dostupné zásilkovým způsobem. To však neznamená, že řešení struktury z prášku i ze synchrotronových dat je ve všech případech úspěšné. Zvláště u vyšších hodnot finálního R-faktoru si nemůžeme být jisti správností přijatého strukturálního modelu a z práškových dat se zatím nedají řešit molekuly velké a vnitřně velmi flexibilní⁸. Nicméně jsou prováděny i první pokusy řešit z práškových dat i struktury proteinů.

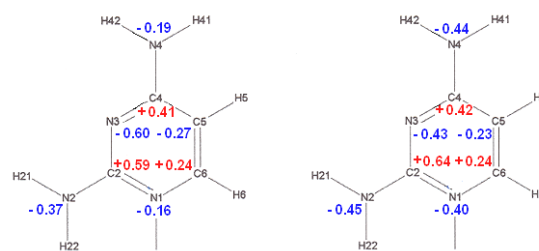
Velká pozornost je věnována aplikaci RTG práškové strukturální analýzy ve výzkumu léčivých látek. Velmi nadějně jsou tzv. molekulární kokrystaly, které modifikují farmakokinetické vlastnosti léčivých látek (především jejich rozpouštěcí rychlosti). Mechanochemickou syntézou dvou výchozích prášků, např. karbamazepinu (léčivá látka) a sacharinu (host), vzniká práškový produkt typu hostitel-host (kokrystal²⁷), kde molekulární poměr karbamazepin : sacharin = 1:1. Obě složky se v kokryystalu spojí vodíkovými můstky (přerušované čáry, obr. 5).

Pravidla pro vytváření kokryсталů s farmaceutickou aplikací jsou přísná. V současnosti přichází v úvahu asi 230 potenciálních hostů, kteří jsou pro kokrystaly farmaceuticky akceptovatelní. To spolu s asi 5000 léčivými látkami (hostiteli) představuje obrovský kombinační potenciál. Při vytváření kokryсталů se využívají postupy HT krystalizace (High-Throughput) a synthonový přístup.

Pomocí časově rozlišeného RTG difrakčního experimentu se studují strukturální změny u vybraných chemických a fyzikálních dějů *in situ*. Jedná se např. o fotoche-



Obr. 5. Molekulární kokrystal karbamazepin : sacharin (1:1), jehož struktura byla vyřešena z RTG práškových dat (obrázek laskavě poskytl prof. M. Zaworotko, University of South Florida)



Obr. 6. Srovnání experimentálních a teoretických nábojových hustot v jednotkách e. Vlevo výsledky z RTG difrakčního experimentu, vpravo teoretický výpočet (B3LYP/6-311G**). Obrázek laskavě poskytl Dr. M. Šlouf

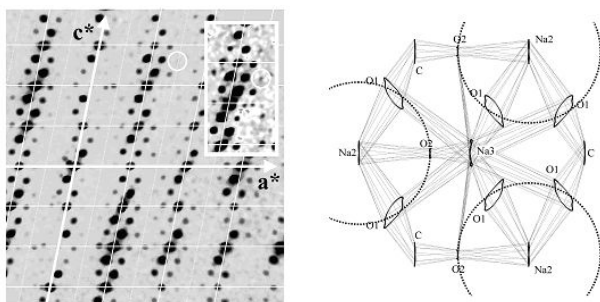
mické procesy, přechody nízkospinové \leftrightarrow vysokospinové fáze v komplexech, děje v solárních člancích, děje v polovodičích, mřížkové deformace v monokrystalech, chemické reakce v kapalinách, spalovací děje atd. Tyto procesy probíhají řádově v ms až fs a proto musí být např. u fotochemických reakcí laserový pulz (iniciující reakci) spřažen se synchrotronovým pulzem (sběr difrakčních dat). Tak byl např. studován přechod mezi nízkospinovým a vysokospinovým stavem komplexu Fe^{II} tris(1,10-fenantrolin)₃ (cit.²⁸), kdy synchrotronový pulz trval 50 ps.

Je známo, že chemickou reakci lze zpomalit a tak strukturálně charakterizovat nestálé intermediáty výrazným ochlazením. Tohoto principu využívá kryokrytalografie. Pro časově rozlišené RTG strukturální studie jsou vyvíjeny nové typy velmi výkonných a citlivých detektorů, např. detektor PILATUS 6M²⁹.

Pečlivé měření RTG difrakcí umožňuje stanovení deformačních a nábojových hustot³⁰. Tyto údaje lze potom porovnávat s teoretickými výpočty (obr. 6). Někdy se tak velmi překvapivě revidují modely klasické strukturální chemie. Studium nábojových hustot a jejich okamžitých změn je významné pro přesný popis elektronového stavu jakéhokoliv systému.

Významnou studovanou skupinou jsou aperiodické krystaly, mezi které řadíme kvazikrystaly, nesouměřitelné (incommensurate) a kompozitní struktury. Tyto látky nevykazují klasickou 3D periodicitu. Odchytky od 3D periodicity, tzv. modulace struktury, se projeví v difrakčním obraze vznikem dodatečných, tzv. satelitních reflexí, které jsou pravidelně uspořádány a lze je popsat, přidáme-li k elementární buňce jeden nebo více vektorů. To vede k zavedení vícedimenzionálního popisu symetrie aperiodických krystalů v tzv. superprostoru, (3+d)D. Superprostorová teorie umožňuje zobecnit základní krystalografické pojmy a vytvořit metody pro řešení a upřesňování aperiodických struktur. Atomy jsou v modulovaných strukturách vyjádřeny jako *d*-rozměrné atomové domény, charakterizující vychýlení atomů ze základních 3D poloh.

První modulované struktury byly nalézány zejména mezi anorganickými (obr. 7) a mineralogickými krystaly, brzy se však ukázalo, že tento fenomén je rozšířen i mezi strukturami organokovovými a organickými a v posled-



Obr. 7. Modulovaná struktura fáze γ - Na_2CO_3 . Vlevo jsou typické satelitní reflexe, které leží mimo reflexe základní trojrozměrné struktury, tj. mimo síť bílých čar. Vpravo jsou charakteristické výchylky atomů ze základních poloh (obrázek laskavě poskytl Dr. Michal Dušek)

ních letech se objevily i informace o satelitních reflexích pozorovaných u krystalů proteinů. Postupně se tak z okrajového oboru krystalografie stává nezbytný nástroj materiálového výzkumu a studia fyzikálních vlastností látek. Je potěšitelné, že práce laboratoře vedené Dr. Václavem Petříčkem ve Fyzikálním ústavu AV ČR představují v oboru modulovaných struktur světovou špičku a zde vyvíjený program Jana³¹ je základním výpočetním nástrojem pro studium aperiodických krystalů.

Aktuální RTG strukturní studie za extrémních podmínek (tlak 0,1 až 100 GPa a teplota 1 až 3000 K) jsou prováděny ve speciálních diamantových celách s využitím 3. generace výkonných synchrotronových zdrojů. Studují se vysokotlaké a vysokoteplotní fáze a jejich přechody buď z hlediska materiálového nebo geofyzikálního. Z hlediska materiálů byly za extrémních podmínek zkoumány např. materiály typu „nanocage“ (klecové či pastové)³², kam patří např. interkalované fulereny, křemíkové klatráty aj., nebo vysokotlaké fáze křemíku, elementárních kovů, vysokotlaké kovové fáze molekulárních krystalů aj. V souvislosti s geofyzikálními problémy je studována zejména dynamika zemského jádra – kapalné železo pod vysokými tlaky (až 10^2 GPa)³³.

Velmi důležité jsou strukturní studie u molekulárních materiálů, které vykazují zajímavé elektrické, magnetické nebo optické vlastnosti. Pro praktické aplikace je důležité, aby bylo možné tyto multifunkční vlastnosti skokově přepínat vnějším vlivem ($h\nu$, p , T). Příkladem je piezoelektrický materiál $\text{K}_{0,4}\text{Fe}^{\text{II}}_4[\text{Cr}^{\text{III}}(\text{CN})_6]_{2,8}$, kde přiloženým vnějším tlakem dojde k rotaci skupin CN a tím ke změně magnetického momentu³⁴.

5. Proteinová krystalografie

Pohled na užitečnost vyřešených bílkovinných struktur se od doby počátků proteinové krystalografie hodně změnil. Vyřešené struktury jsou dnes vnímány především jako klíč k odpovědím na otázky, jak vlastně mohou bílkovinné molekuly vstupovat do interakcí samy se sebou nebo

s jinými biologickými makromolekulami či s nízkomolekulárními ligandy. Trochu rigorózněji formulované otázky zní asi takto: kde je strukturní základ rozeznávání biologických molekul a co podmiňuje sílu a specifitu jejich interakcí? Dále: vodíkové vazby, elektrostatické a hydrofobní interakce stabilizují do jisté míry vnitřní strukturu bílkovinné molekuly, ale dovolují také určité konformační přechody: jak se tím mění tvar a povrchový náboj molekuly a jaká je zde spojitost s biologickou funkcí? V těchto směrech pak současně příklady uplatnění strukturních informací už zahrnou nejen konvenční problematiku tvorby komplexu enzymu se substrátem a následné katalýzy, ale třeba i proteinové inženýrství protilátek směrem k posilování jejich užitečných vlastností (a potlačování nežádoucích)³⁵ nebo mechanismus procesu, při němž stovky stejných bílkovinných molekul kondenzují do centrálního objektu, kapsidu, v retrovirové částici HIV³⁶. Asi teď už nikoho nepřekvapí, že proteinová krystalografie má zásadní význam také pro odhalování molekulárních základů mnoha chorob, včetně genetických poruch, a pro související vývoj úspěšných léčiv.

Údaje o primární struktuře bílkovin, aminokyselinových sekvencích, se dnes ponejvíc odvozují z přečtené genetické informace (tj. z nukleotidových sekvencí kódujících nukleových kyselin). K identifikaci kvartérních struktur bílkovin ve větších biologických objektech výborně slouží techniky elektronové mikroskopie. Sekundární a terciární struktura se však týká prostorové výstavby hlavního řetězce a poloh bočních skupin – jinými slovy prostorového uspořádání atomů v bílkovinné molekule. Zde je ústřední metodou proteinová krystalografie, jež v zásadě umožňuje získání informací až k atomovému rozlišení 1,0 Å.

„Strukturní biologií“ se rozumí obecněji pojatý fundamentální aspekt soudobé molekulární biologie, zaměřený na určení trojrozměrných struktur biologických molekul, obzvláště bílkovin, a poskytující pohled do funkce těchto molekul v živém organismu. Síla proteinové krystalografie (nebo strukturní biologie) spočívá kromě jiného v tom, že umožňuje identifikovat aminokyselinové zbytky podstatné pro funkci. Další sledování struktur se pak může zaměřit i na bílkovinné molekuly konstruované tak, aby jejich vybrané aminokyseliny byly zaměněny („cíleně mutovány“) nebo aby byly odstraněny nebo nově vneseny celé úseky polypeptidového řetězce, přirozené či umělé.

5.1. Zvláštnosti postupů proteinové krystalografie

Proteinová krystalografie je dnes multidisciplinárním oborem, v němž se protínají zájmy fyziků, biochemiků, biologů a dalších profesí, o komerčních zájmech zde zatím ještě nemluvě. Průnik do biomedicíny je patrný i ze zařazování kurzů proteinové krystalografie do výuky lékařů (např.³⁷). Praktické řešení 3D struktury bílkoviny má přitom samozřejmě stále stejný princip, RTG difrakci na monokrystalu, ale v průběhu posledních dvaceti let se jednotlivé postupné kroky vedoucí k řešení staly podstatně ro-

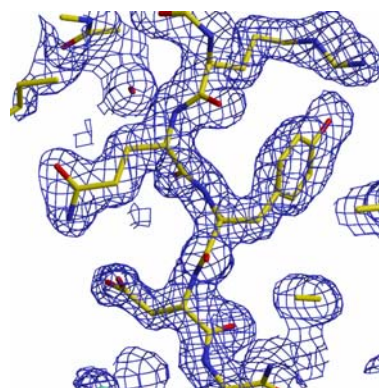
bustnějšími, rychlejšími a dostupnějšími. Některých aspektů tohoto vývoje jsme se dotkli již shora, jiné si však zasluhují samostatný výklad (přehled³⁸).

Produkce bílkovin pro proteinovou krystalografii se už nerealizuje izolací z přírodního materiálu (až na výjimky, jednu uvedeme níže), ale izolací ze systémů rekombinantní exprese (cizorodá bílkovina se produkuje například v bakteriích *E. coli* nebo v kvasinkách za kódování příslušnou nukleovou kyselinou, předem získanou molekulárním klonováním). Rekombinantní exprese je mnohostranně použitelný nástroj: větší bílkoviny mívají kromě pevně organizovaných domén i hodně dalších, inherentně neuspořádaných nebo příliš hydrofobních (např. transmembránových) součástí polypeptidového řetězce. Pokud takoveto bílkoviny nekystalizují, východiskem z nouze je exprese „osekaných“ (truncated) bílkovin nebo jednotlivých izolovaných domén; vyřešené parciální struktury pak případně poskytnou určitou „skládačku“. Zásadně se změnila i dostupnost bílkovin pro proteinovou krystalografii. Expresi v *E. coli* a následnou purifikaci lze dnes už poměrně pohodlně získat například retrovirovou proteasou HIV nesoucí konkrétní bodovou mutaci spojenou s rezistencí k léčivu³⁹. Konvenční preparace by zde schůdná nebyla, už proto, že by vyžadovala izolaci ze směsi s příliš podobnou normální bílkovinou a s jen velmi málo odlišnými bílkoviny nesusoucími jiné mutace.

Pro experimentální určení fází, jež jsou nezbytné k vyřešení struktury (podkap. 3.1.), se využívají postupná měření s RTG zářením o různých vlnových délkách, protože některé atomy krystalu mění své rozptylové vlastnosti v závislosti na energii záření (Multiple Anomalous Dispersion, MAD). Standardním postupem zavedení anomálně rozptylujících atomů do bílkovinné molekuly je rekombinantní exprese bílkoviny v obohaceném minimálním kulturačním médiu s obsahem selenomethioninu.

Hodně empirie stále zůstává v postupech nalézání krystalizačních podmínek (crystallization trials). Za skutečnost, že dnes zpravidla vystačíme „jen“ s několikamiligramovým množstvím preparátu bílkoviny, vděčíme mikrokystalizačním technikám, a za relativní experimentátorský komfort pak komerčně dostupným soupravám desítek až stovek krystalizačních roztoků (crystallization screens) a různých typů plastových destiček s difúzními komůrkami pro visící nebo sedící kapky o typickém objemu menším než deset μl . Krystalizace komplexů bílkovina/ligand se realizuje buď přítomností ligandu v krystalizačním médiu nebo namáčením bílkovinných krystalů do roztoku ligandu. Novějšími technikami použitelnými jak pro metodu MAD, tak pro metodu izomorfni modifikace k určování fází, je tlakové zavádění atomů xenonu do zmražených krystalů bílkoviny nebo namáčení krystalů do koncentrovaných roztoků jodidu.

Krystalizace zůstává stále úzkým hrdlem řešení bílkovinných struktur a stává se tedy předmětem snah o sestavení alternativních systémů (krystalizace v gelech nebo na rozhraních olej/vodný roztok) nebo snah o robotizaci a automatizaci standardních postupů¹¹. Příslušná instrumentace by zde ideálně měla „sama“ a opakovaně po delší

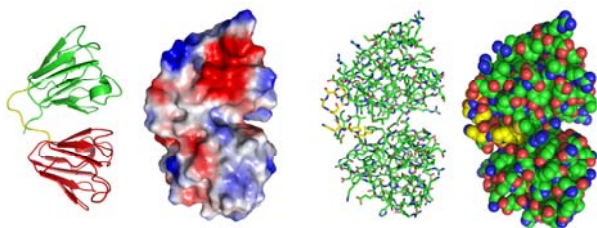


Obr. 8. Výřez modelu HIV proteasy v mapě elektronové hustoty³⁹, PDB kód 1NH0. Bílkovinná molekula je znázorněna tyčovým modelem, bodové atomy jsou pospojovány ve směru kovalentních vazeb, tyče nesou zbarvení podle druhu atomu, ze kterého vycházejí (žlutá uhlík, modrá dusík a červená kyslík). Mapa elektronové hustoty (modrá síť) je konturována na úrovni 1σ , zde cca 0,8 elektronu na \AA^3

dobu analyzovat a zaznamenávat mikroskopické obrazy kapek, rozeznávat v nich amorfni nebo olejovité sraženiny nebo i kýžené, ale třeba maličké, srostlé nebo deformované krystaly a poskytovat tak vodítka pro jemné vyladování krystalizačních podmínek. Nověji se robotizace uplatňuje u výměny a montáže krystalů do optimálních pozic vzhledem k synchrotronovému RTG svazku⁴⁰. Automatizace může dnes zasahovat i do výběru krystalů s nejlepší difrakční kvalitou, do sběru a zpracování difrakčních dat, popřípadě až po nekonvenční metody výstavby a upřesňování modelu⁴¹.

V běžném postupu se model bílkovinné struktury buduje do map elektronových hustot vypočtených na základě měřených intenzit difrakcí a jejich fází získaných výše zmíněnými metodami. Proces upřesňování prvotního modelu spočívá v postupném doladování jeho parametrů, konkrétně souboru souřadnic x_i, y_i, z_i , teplotního faktoru B_i a někdy též okupačního faktoru o_i každého i -tého atomu modelu tak, aby bylo dosaženo maximální možné shody s experimentálními daty (obr. 8). V proteinové krystalografii se za uspokojivě vyřešenou považuje struktura s R-faktorem pod konvenční hranici 25 % (i když se publikují i struktury s R-faktorem vyšším, např. v Nature v roce 2007 zhruba každá šestá).

Krystalografie biologických makromolekul není fundamentálně odlišná od krystalografie nízkomolekulárních sloučenin, ale rozhodně je složitější kvůli velikosti molekulárního systému. Pokud je struktura tvořena několika atomy, je na každý upřesňovaný parametr k dispozici dostatek experimentálně stanovených intenzit, takže upřesnění lze provést iterativním výpočtem, a do jehož průběhu není většinou nutné manuálně zasahovat. Typická makromolekula je však tvořena několika tisíci atomy a krystalizuje v elementárních buňkách o objemu milionů \AA^3 , takže poměr experimentálně zjištěných intenzit k počtu parametrů, které je třeba stanovit, se stává velmi nepříznivou



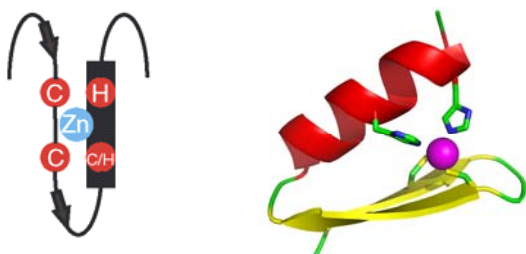
Obr. 9. Různé způsoby grafického zobrazení proteinů: molekula lidského Arg36Ser γ D-krystalinu⁴⁸, PDB kód 2G98, program pyMOL⁴³. Zleva: sekundární struktura s barevným rozlišením domén a spojky; povrch vybarven podle elektrostatického potenciálu; atomy pospojovány tyčemi; atomy reprezentované koulemi van der Waalsových poloměrů

vým. Komplexnost této úlohy vyžaduje kombinování iterativních upřesňovacích (rafinačních) metod využívajících matematických principů a manuálních zásahů řešitele při budování, resp. přebudování, modelu ve 3D prostoru (přidávání nebo odebrání konkrétních atomů, případně změna jejich polohy editací souřadnic).

Exponenciálně rostoucí počet vyřešených bílkovinných struktur si vyžádal vytvoření samostatných elektronických databází. V dominantní RCSB Protein Data Bank⁴² (dále jen „PDB“) je dnes deponováno přes 50 tisíc struktur, identifikovaných čtyřznakovými kódy. PDB je spravována konzorciem několika amerických institucí (Research Collaboratory for Structural Bioinformatics, RCSB) a slouží uživatelům zcela přátelským způsobem. Výstupy z PDB poskytují kromě vlastních struktur i literární odkazy a provázání k jiným databázím. Vizualizaci bílkovinných struktur na mnoho způsobů (s možnostmi vybarvení povrchu podle van der Waalsových parametrů nebo podle náboje, s vyznačením některých vzdáleností apod.) slouží programy počítačové grafiky, např. populární pyMOL⁴³, viz obr. 9.

5.2. Zužitkování strukturní informace

Co všechno lze ze strukturní informace vyžít a jak

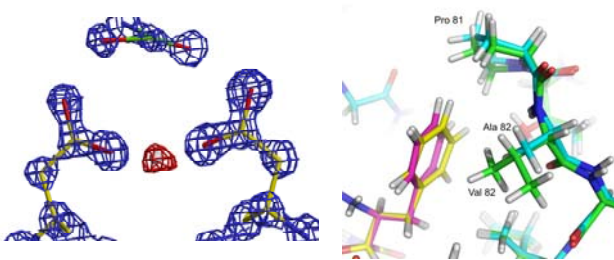


Obr. 10. Příklad topologie (vlevo) a sekundární struktury (vpravo) modulu zinkového prstu; C označuje cysteinové zbytky, H histidinové. Adaptováno z cit.⁴⁴

se orientovat v nových přínosech proteinové krystalografie? Ilustrativní může být krátký výběr z více než stovky „molekul měsíce“, vyhlášených a komentovaných v PDB. Pozpátku podle abecedy nalezneme „zinkové prsty“ (březen 2007). Zinkové „lepkavé“ prsty jsou tvořeny specifickými kompaktními moduly, jejichž prostřednictvím bílkovina rozeznává strukturní motivy v DNA, RNA nebo v jiných bílkovinách⁴⁴. Klasická topologie modulu zinkového prstu zahrnuje dva β -listy (poskytující k tetragonální vazbě zinkového atomu dva cysteiny) a jeden α -helix (poskytující k vazbě téhož atomu dva histidiny); část z ostatních asi 25 aminokyselinových zbytků modulu tvoří malé hydrofobní jádro modulu (leuciny a fenylalaniny) a zbývající „hledí ven“ a nacházejí partnery k pevným funkčním interakcím (obr. 10). V PDB je uloženo přes tisíc (!) struktur obsahujících zinkové prsty v kontextu různých bílkovin – fungují např. ve „čtení“ nukleotidové sekvence specifických segmentů DNA, jež regulují transkripci (přepis do RNA).

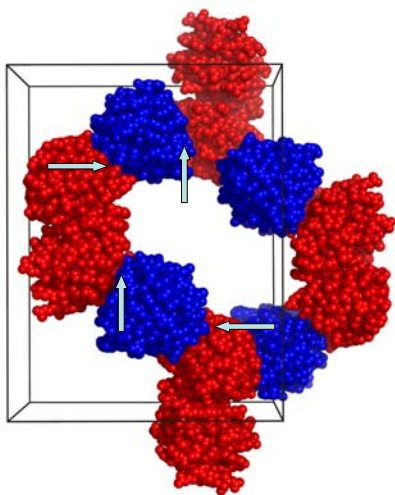
Strukturu transkripčního faktoru má i bílkovina p53 potlačující maligní růst (p53 Tumor Suppressor), jež byla vyhlášena molekulou měsíce července roku 2002. Funguje v případech, kdy jsou normální prvky regulace transkripcie v DNA narušené (zářením, kancerogeny apod.) a to zhruba tak, že ve své tetramerní formě „obejme“ DNA v relevantních místech genomu a v souhře s dalšími bílkovinami umožní nápravu (nebo řízenou buněčnou smrt, apoptózu). Bílkovinu p53 však inaktivují mutace, jež provázejí asi 50 procent lidských nádorů. Doménové struktury vyřešené do roku 2002 už umožnily pochopit některé onkogenní mutace p53. Například mutace jednoho aminokyselinového zbytku argininu představuje ztrátu „prstu“, který zapadá do malého žlábků šroubovice DNA. Skutečně fascinující je však nedávný vývoj na tomto poli^{45,46}. V důsledku jiných onkogenních mutací jsou destabilizovány lokální struktury v oblasti vazby p53 k DNA (zasazené struktury roztají již při tělesné teplotě nebo pod ní). Pro vyřešení kompletní struktury p53 s potřebným rozlišením byla napřed proteinově-inženýrsky zkonstruována praktičtější, ale nadále biologicky aktivní varianta p53 a na tomto základě byly řešeny struktury p53 nesoucí destabilizující onkogenní mutace. Ve vyřešených strukturách byly nalezeny dutiny, pro něž byly úspěšně navrženy malé syntetické molekuly zpevňující strukturu, tedy prototypy sloučenin k novému směru protinádorové intervence.

Dnes už klasickým příkladem intenzivního zapojení proteinové krystalografie do racionálního designu léčiv je vývoj inhibitorů retrovirové proteasy HIV (PDB molekuly měsíce června roku 2000, jež svého času doznala i časté prezentace v kalendářích a na titulních stranách firemních katalogů). Inhibitory proteasy přerušují infekční cyklus retroviru tím, že zabraňují proteolytickému zpracování bílkovinných prekurzorů na funkční stavební a enzymové bílkoviny. Prázkladem struktur inhibitorů se staly struktury přirozeně štěpitelných míst (např. oktapeptidů s náhradou štěpené peptidové vazby neštěpitelnou skupinou). Vodítkem pro syntetickou organickou chemii pak byly stovky struktur komplexů HIV proteasy s inhibitory



Obr. 11. Detaily struktury komplexu HIV proteasy s fenylnorstatinovým inhibítor, rozlišení 1,03 Å³⁹. Detail vlevo: Katalytické místo; červeně konturovaná mapa diferenční elektronové hustoty uprostřed přísluší vodíkovému atomu; červeně vyznačené kyslíkové atomy karboxylových skupin proteasy a hydroxylové a karbonylové skupiny inhibítora jsou kruhově propojeny vodíkovými vazbami. Detail napravo: Část inhibítora v „podmístě P1“ vazebné dutiny proteasy; při vazbě inhibítora (vybarven fialově) k normální 82Val protease (zeleně) je úhel napojení boční fenylové skupiny k hlavnímu řetězci inhibítora deformovaný, ale při vazbě ke mutované 82Ala protease je normální (žlutě a tyrkysově)

pocházejícími z různých fází strukturního designu⁴⁷. Struktura komplexu z našeho pracoviště³⁹ byla svého času strukturou s nejvyšším publikovaným rozlišením, 1,03 Å. Toto rozlišení umožnilo např. stanovit donory a akceptory vodíkových vazeb v katalytickém místě a určit polohu vodíkového atomu drženého dvěma katalytickými Asp zbytky (obr. 11). V jiném směru struktura vysvětlovala schopnost



Obr. 12. Elementární buňka krystalu lidského Arg36Ser γ D-kryštalínu⁴⁸. Vodorovné šipky směřují ke kontaktním zónám serinu 36 z červeně vybarvených molekul, svislé ke kontaktním zónám serinu 36 z modře vybarvených molekul

nové fenylnorstatinové sloučeniny inhibovat též HIV proteasu, jež mutací získala rezistenci k antiretrovirotickým léčivům, např. k Ritonaviru. Pro uvedenou rezistenci je podstatná mutace proteasového zbytku 82 (valinu na alanin) a pro její překonání prodloužená délka hlavního řetězce inhibítora v místě náhrady peptidové vazby. Jak ukazuje obr. 11, inhibitor dobře vyplňuje zvětšenou vazebnou dutinu mutovaného enzymu (jež je ale pro vazbu Ritonaviru jaksi příliš „volná“). Skutečnost, že mutovaná proteasa je novou sloučeninou inhibována dokonce poněkud lépe než proteasa divokého typu, lze vysvětlit překážkou, kterou pro inhibitor v „normální“ vazebné dutině představují dvě methylové skupiny bočního řetězce valinu 82.

Další zajímavý patogenní proces⁴⁸ můžeme vysvětlit i na jiných molekulách, než na prominentních „PDB molekulách měsíce“. Bílkovina γ D-kryštalín se ve vysoké koncentraci nachází v oční čočce a zajišťuje její světlostlonočnost. Některé mutace povrchových zbytků molekuly γ D-kryštalínu vedou ke vzniku kalného depozitu, což se projeví onemocněním kataraktou (lidově „šedým zákalem“). Místo oční čočky s poškozenou funkcí lze postiženým chirurgicky poskytnout umělou náhradu. Všímavý molekulární patolog zjistil, že na rozdíl od běžných případů amorfního bílkovinného depozitu má depozit v jednom chirurgicky odstraněném materiálu krystalický charakter. Po zjištění genetického základu onemocnění, potvrzeného i stanovením mutace v bílkovinné molekule jako Arg36Ser, předal drobné krystaly patogenní bílkoviny proteinově-kryštalografickému pracovišti. Krystaly neměly ideální kvalitu, ale vykazovaly celkem uspokojivou difrakci a umožnily vyřešení struktury s rozlišením 2,25 Å. Jaký strukturní prvek podmiňuje krystalizaci patogenní bílkoviny Arg36Ser γ D-kryštalínu v čočce lidského oka? Obr. 12 ukazuje, jak těsně se molekuly Arg36Ser γ D-kryštalínu v elementární buňce krystalu dotýkají. Do zóny každého kontaktu se ještě vejde boční řetězec aminokyselinového zbytku 36Ser, ale přítomnost 36Arg v normálním γ D-kryštalínu těsné kontakty znemožňuje (arginin má boční řetězec mnohem větší než serin). Vyřešená struktura byla první publikovanou strukturou lidského γ D-kryštalínu: normální bílkovina totiž navzdory svému pojmenování „tvrdohlavě“ odolává krystalizačním pokusům a setrvává v roztoku či gelu i při vysokých koncentracích, jak to ostatně odpovídá i její fyziologické funkci.

6. Predikce

V roce 1988 John Maddox* uštěpačně prohlásil, že: „je skandální, že teoretická fyzika stále neumí předpovědět krystalovou strukturu látky pouze z jejího vzorce“. O to více je skandální, že to jednoznačně neumí ani v roce 2008.

* Sir John Maddox (*1925), teoretický fyzik a chemik, editor časopisu Nature v letech 1966-73 a 1980-95

Problém má několik rovin. U malých molekul je teoretický výpočet dnes časově srovnatelný s experimentem. Teoretický výsledek je ovšem nejistý, tak proč teoretizovat? Na druhé straně je však teoretická predikce krystalové struktury stále prestižní a fundamentální otázkou. Predikci však výrazně komplikuje velmi rozšířený fenomén polymorfismu chemických entit. V závislosti na podmínkách krystalizace může jedna molekula vykristalovat v několika strukturách – polymorfech. Není snadné určit, která minima na hyperpotenciálové ploše odpovídají reálným polymorfům a která nikoliv. Právě predikce polymorfismu farmaceutických substancí by byla velmi užitečná např. pro originální farmaceutické firmy, které by tak blokovaly konkurenci generických firem. Bohužel však u patentových sporů farmaceutických firem soudy predikci jako důkaz neuznávají.

Ačkoliv se predikční algoritmy vyvíjejí stále k větší dokonalosti, čas od času provede Cambridge Crystallographic Data Centre „slepý“ test⁴⁹, jehož podstatou je predikovat *ab initio* neznámé a nepříliš složité krystalové struktury malých molekul. Tento test je pak zadán vybranému vzorku teoretických laboratoří disponujících různým programovým vybavením. Výsledek je však již několik let tentýž: zpravidla jsou zadány 4 sloučeniny a správnou strukturu všech čtyř nepředpoví žádná laboratoř. Většinou laboratoře predikují jednu nebo dvě správné struktury, ale často jiné kombinace. *Suma summarum*: predikce krystalové struktury malých molekul je možná, ale zcela nejednoznačná.

Jsou však i dílčí úspěchy. Příkladem je předpověď existence třetího polymorfu paracetamolu a jeho následné experimentální potvrzení⁵⁰. Na zdokonalování predikčních algoritmů však teoretici usilovně pracují.

Pokud je žádoucí nebo nutné získat informace o bílkovinné 3D struktuře, jež dosud nebyla vyřešena experimentálně, je v zásadě možné snažit se ji vypočítat postupem, jemuž se říká modelování. Homologní modelování⁵¹ spočívá v odhadu neznámé „cílové“ 3D struktury na základě struktury „vzorové“ (template) bílkoviny, získané zpravidla RTG strukturální analýzou a uložené v PDB, s tím, že uvažovaná „vzorová“ bílkovina je s „cílovou“ homologní co do aminokyselinové sekvence. Na stupni sekvenční homologie pak závisí i přesnost vypočteného modelu: Uprostřed spektra přesností leží modely založené na sekvenční identitě 30 až 50 %, u nichž až 85 % modelovaných Ca atomů leží ve vzdálenosti do 3,5 Å od správných poloh. Model, i když není absolutně dokonalý, může být velice užitečný. Může například napomoci návrhu cílených mutací k testování hypotéz o funkci bílkoviny, sloužit k identifikaci a vylepšování ligandů, k pochopení substrátové specifity a k racionalizaci experimentálních pozorování.

Popis metod homologního modelování by přesáhl rámeček tohoto článku. Jedná se, ve značné stručnosti řečeno, o následující kroky: Nalezení „vzorové“ struktury; porovnání (alignment) aminokyselinových sekvencí včetně potřebných korekcí; generování hlavního řetězce; na datech založené generování „kanonických“ smyček; genero-

vání poloh bočních řetězců a následná optimalizace; výstavba smyček založená na *ab initio* výpočtech; celková optimalizace modelu založená na výpočtech energetických minim a případné iterace předešlých kroků k odstranění chyb. Populárními programy pro homologní modelování jsou např. MODELLER, SWIS-MOD a WHAT IF umístěné na veřejně dostupných serverech. Program SWIS-MOD má odkaz na seznam REPOSITORY, tj. knihovnu modelovaných struktur vytvořených na základě dostatečně homologních aminokyselinových sekvencí nalezených v PDB.

Rozsah používání postupů modelování struktur biologických makromolekul kolísá podle objektů zájmu: U struktur smyček v řetězci variabilních oblastí imunoglobulinů je běžný³⁵, zatímco vyladování struktur bílkovinných ligandů jako potenciálních léčiv vyžaduje zpravidla vysoké rozlišení poskytované jen experimentálními krystalovými strukturami komplexů protein/ligand^{46,47}. Vztah proteinové krystalografie a modelování je obousměrný. Na jedné straně jsou základem porovnávacího modelování vyřešené krystalové struktury, na druhé straně může modelování ulehčovat postup molekulárního nahrazování v proteinové krystalografii a napomáhat designu bílkovinných molekul s kompaktní strukturou (bez dlouhých smyček a exponovaných hydrofobních oblastí) pro lepší šance úspěšné krystalizace.

7. Závěr

Současná RTG strukturální analýza představuje nejdůležitější analytickou techniku pro stanovení molekulové a krystalové struktury látek. Pokud je k dispozici monokrystal (u malých i velkých molekul), pak se s velkou pravděpodobností podaří vyřešit jeho strukturu. Omezení monokrystalové RTG strukturální analýzy (malých i velkých molekul) představuje právě vypěstování vhodného monokrystalu. U látek, které jsou složeny z malých a nepříliš vnitřně flexibilních molekul, a nedaří se je připravit ve formě monokrystalu, je šance vyřešit strukturu z práškových dat. Rychlost provedení rutinní RTG strukturální analýzy malých molekul (několik hodin) se vyrovná strukturálním spektroskopickým technikám, především NMR, se kterou se vhodně doplňuje.

Trendy v krystalografii malých molekul se přesouvají od statického stanovení struktury ke strukturálnímu sledování chemických reakcí *in situ* a ke stále podrobnějšímu a jemnějšímu popisu struktury (aperiodické krystaly). Velkou výzvou je predikce krystalové struktury z výpočtů *ab initio*, podobně jako nalezení fundamentálních příčin existence polymorfismu molekulárních krystalů. RTG krystalografie je stále více při řešení určitých problémů vhodně kombinována s ostatními metodami stanovení struktury látek – NMR, elektronovou mikroskopií, elektronovou diffrakcí aj.

Jak v současné „postgenomické“ éře načrtnout trendy proteinové krystalografie? Zřetelně se etabluje „strukturální genomika“, jejíž strategií je určování bílkovinných struktur ve velkých, byť zatím asi nutně neúplných, souborech

genových produktů kódovaných jednotlivými geny. Strukturální genomika se přednostně zaměřuje na lidský genom a genomy patogenních organismů, včetně parazitů. Akademický výzkum hledá (a nalézá) souhru genových produktů, např. v regulaci jejich tvorby, buněčném transportu a uplatnění v normálních a patologických procesech. Důležité je, že vhodné kandidáty na selektivní terapeutický zásah lze hledat (a nalézat) mezi bílkoviny se zvláštní nebo dokonce ojedinělou strukturou (v jistém slova smyslu mezi „molekulárními jednorozci“). Z druhé strany se proteinová krystalografie nezanedbatelně uplatňuje v proteinovém inženýrství bílkovin, jež mají do procesů probíhajících v živém organismu zasahovat: Ve třídě „biofarmaceutických“ bílkovin jsou stále nově konstruovány terapeutické protilátky, imunogeny a mnoho dalších perspektivních prostředků léčby nebo prevence. Shora zmíněné komerční zájmy se zde týkají ročního objemu světového trhu v řádech desítek miliard dolarů. V současných trendech, jistě ohromujících svou novostí a záběrem, může však pozorný proteinový krystalograf nebo strukturální biolog rozeznat i prvek kontinuity: kardinální místo pořád zaujímá prastaré téma vztahu struktury a funkce bílkovin. A tento zobecněný vztah struktura/vlastnosti platí i pro malé molekuly.

Práce byla podpořena výzkumným záměrem MŠMT ČR č. 604613730, projektem GA ČR 203/07/0040, výzkumným záměrem AV ČR AV0Z50520514 a výzkumným centrem MŠMT ČR č. M0505.

LITERATURA

- Friedrich W., Knipping P., Laue M.: *Sitzungsber. Bayer. Akad. Wiss.* 1912, 303.
- http://nobelprize.org/nobel_prizes/physics/laureates/1915/wl-bragg-lecture.pdf, staženo 14.3.2008.
- Novák C.: Podklady pro 250. jubilejní Rozhovory o aktuálních otázkách v rentgenové a neutronové strukturální analýzy. Ústav makromolekulární chemie AV ČR Praha 2001. http://www.xray.cz/xray/csca/r250_w.htm, staženo 14.3.2008.
- 24th European Crystallographic Meeting, Marrakech / Morocco, 22–27 August 2007. Abstracts.
- www.iucr2008.jp, staženo 14.4.2008.
- Ban N., Nissen P., Hansen J., Moore P. B., Steitz T. A.: *Science* 289, 905 (2000).
- Kratochvíl B.: *Chem. Listy* 99, 258 (2005).
- Hušák M., Rohlíček J., Čejka J., Kratochvíl B.: *Chem. Listy* 101, 697 (2007).
- <http://www.rigaku.com/generators/>, staženo 24.4.2008.
- http://www.esrf.eu/UsersAndScience/Experiments/MX/About_our_beamlines/ID14-1, staženo 24.4.2008.
- FEBS Advanced Course. Advanced methods in macromolecular crystallization III. Academic and University Center at Nové Hrady, Czech Republic October 03–10, 2008.
- Oszlányi G., Sütö A.: *Acta Crystallogr.* A60, 134 (2004).
- Chapius G., Schoeni N.: eCrystallography course, <http://escher.epfl.ch/eCrystallography>, staženo 10.12.2007.
- Palatinus L., Chapuis G.: *J. Appl. Crystallogr.* 40, 786 (2007).
- Le Bail A.: *Powder Diffraction* 19, 249 (2004).
- Bergmann J., Le Bail A., Shirley R., Zlokazov V.: *Z. Kristallogr.* 219, 783 (2004).
- Schneider T. R., Sheldrick G. M.: *Acta Crystallogr.* D58, 1772 (2002).
- Xu H., Hauptman H. A.: *Acta Crystallogr.* D62, 897 (2006).
- Vagin A., Teplyakov A.: *J. Appl. Crystallogr.* 30, 1022 (1997).
- Navaza G.: *Acta Crystallogr.* A50, 157 (1994).
- Kissinger C. R., Gehlhaar D. K., Fogel D. B.: *Acta Crystallogr.* D55, 484 (1999).
- Fujinaga M., Read R. J.: *J. Appl. Crystallogr.* 20, 517 (1987).
- <http://bioportal.weizmann.ac.il/iucr-top/cif/index.html>, staženo 10.2.2008.
- Jegorov A., Hušák M., Kratochvíl B., Cisařová I.: *Crystal Growth & Design* 3(4), 441 (2003).
- Hašek J.: *Materials Structure* 14, 97 (2007).
- Kužel R., Daniš S.: *Materials Structure* 14, 89 (2007).
- Vishweshwar P., McMahon J. A., Bis J. A., Zawortko M. J.: *J. Pharm. Sci.* 95, 499 (2006).
- Sato T., Nozawa S., Ichianagi K., Tomita A., Ichikawa H., Chollet M., Fujii H., Adachi S., Koshihara S.: *Acta Crystallogr.* A 64, C204 (2008).
- Broennimann C., Eikenberry E. F., Henrich B., Horisberger R., Huelsen G., Pohl E., Schmitt B., Schulze-Briese C., Suzuki M., Tomizaki T., Toyokawa H., Wagner A.: *J. Synchrotron Rad.* 13, 120 (2006).
- Šlouf M.: *Chem. Listy* 96, 3 (2002).
- Petříček V., Dušek M., Palatinus L.: Abstract KN02. ECM24, Marrakech 2007.
- San-Miguel A., Poloni R., Rey N., Toulemonde P., Le Floch S., Machon D., Pishedda V.: Abstract MS22 O2. ECM24, Marrakech 2007.
- Lundegaard F. L., McMahon M. I.: Abstract MS22 O4. ECM24, Marrakech 2007.
- Coronado E., Clemente-Leon M., Lopez-Jorda M., Romero F. M.: *Acta Crystallogr.* A 64, C49 (2008).
- Tsurushita N., Hinton P. R., Kumar S.: *Methods* 36, 69 (2005).
- Kelly B. N., Howard B. R., Wang H., Robinson H., Sundquist W. I., Hill C. P.: *Biochemistry* 45, 11257 (2006).
- <http://www.bcm.edu/catalog/gsbs-coursecatalog.html>, staženo 24.4.2008.
- Tickle I., Sharff A., Vinković M., Yon J., Jhoti H.: *Chem. Soc. Rev.* 33, 558 (2004).
- Brynda J., Řezáčová P., Fábry M., Hořejší M., Stouráčová R., Sedláček J., Souček M., Hradílek M., Lepšík

- M., Konvalinka J.: *J. Med. Chem.* 47, 2030 (2004).
40. Snell G., Cork C., Nordmeyer R., Cornell E., Meigs G., Yegian D., Jaklevic J., Jin J., Stevens R. C., Earnest T.: *Structure* 12, 537 (2004).
 41. Morris R. J., Zwart P. H., Cohen S., Fernandez F. J., Kakaris M., Kirillova O., Vornheim C., Perrakis A., Lamzin V. S.: *J. Synchrotron Rad.* 11, 56 (2004).
 42. Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N., Bourne P. E.: *Nucleic Acids Research* 28, 235 (2000).
 43. DeLano W.L.: The PyMOL Molecular Graphics System (2002) on World Wide Web, <http://www.pymol.org>.
 44. Gamsjaeger R., Liew C. K., Loughlin F. E., Crossley M., Mackay J. P.: *Trends Biochem. Sci.* 32, 63 (2007).
 45. Joerger A. C., Fersht A. R.: *Oncogene* 26, 2226 (2007).
 46. Joerger A. C., Ang H. C., Fersht A. R.: *Proc. Natl. Acad. Sci. U.S.A.* 103, 15056 (2006).
 47. Wlodawer A., Vondrášek J.: *Annu. Rev. Biophys. Biomol. Struct.* 27, 249 (1998).
 48. Kmoch S., Brynda J., Bekefadu A., Bezouška K., Novák P., Řezáčová P., Ondrová L., Filipec M., Sedláček J., Elleder M.: *Human Molecular Genetics* 9, 1779 (2000).
 49. Day M. G.: Abstract MS14O1. ECM24, Marrakech 2007.
 50. Beyer T., Day G. M., Price S. L.: *J. Am. Chem. Soc.* 123, 5086 (2001).
 51. Martí-Renom M. A., Stuart A. C., Fiser A., Sánchez R., Melo F., Šali A.: *Annu. Rev. Biophys. Biomol. Struct.* 29, 291 (2000).

B. Kratochvíl^a, M. Hušák^a, J. Brynda^{b,c}, and J. Sedláček^b (^a*Department of Solid State Chemistry, Institute of Chemical Technology, Prague, Czech Republic*; ^b*Laboratory of Structural Biology, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Prague*; ^c*Team of Structural Biology, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Prague*): **What Can the Current X-Ray Structure Analysis Offer?**

X-ray diffraction analysis is still the most important method for crystal structure determination. It is used in crystallography of small molecules and proteins. Although principles of both applications are the same, they differ in methodology of crystallization and utilization. The main limiting factor of progress is the preparation of single crystals. This article reviews current methods and trends, as presented at the last European Crystallographic Meeting and Congress of the International Union of Crystallography. The crystallography of small molecules tends towards solving structures from powder diffraction data, studying structural changes in chemical reactions in situ and describing detailed structure (aperiodic crystals). The advanced “charge-flipping” method for the phase problem solution is mentioned. In protein crystallography, solving crystal structures is easier with novel procedures for protein supply, crystallization tests, and data collection using common and synchrotron x-ray sources as well as advanced computational methods of phase determination and model refinement. The solved protein structures are exemplified with several “molecules of the month” in the Protein Data Bank, emphasis being put on structural basis of pathological mechanisms and the use of crystal structures in design of drugs and biopharmaceuticals. *Ab initio* predictions and simulations of crystal structures are discussed.