Přírodovědecká fakulta Univerzity Karlovy v Praze
Katedra fyzikální a makromolekulární chemie

Studium variantních forem lidské glutamátkarboxypeptidasy II kombinací experimentálních a teoretických metod

Bc. Václav Navrátil

Diplomová práce

Vedoucí diplomové práce:
Doc. RNDr. Jan Konvalinka, CSc.
Mgr. Lubomír Rulíšek, CSc.

Praha 2009

Rád bych poděkoval Janu Konvalinkovi, Lubomíru Rulíškovi a Vojtěchovi Klusákovi za laskavé a trpělivé vedení během vypracování mé diplomové práce. Dále bych rád poděkoval všem ostatním členům naší laboratoře za vytvoření příjemné pracovní atmosféry. Rodičům děkuji za podporu, kterou mi poskytovali během celého studia.

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením školitelů Doc. RNDr. Jana Konvalinky, CSc. a Mgr. Lubomíra Rulíška, CSc. a všechny použité prameny jsem řádně citoval.

V Praze dne 04. 09. 2009 .....

Obsah

Seznam zkratek 6

1. Předmluva 9

2. Úvod 11

2.1. Metody kvantové chemie....11
2.1.1. Hartreeho-Fockova metoda....11
2.1.2. Korelační energie, „post-SCF“ metody....12
2.1.3. Teorie funkcionálu hustoty....14
2.1.4. Partiční funkce, Gibbsova volná energie....17
2.1.5. Teorie tranzitního stavu a reakční rychlost....18
2.2. Enzymová kinetika a tvorba modelu....19
2.2.1. Enzymová kinetika....19
2.2.2. Přístupy k modelování enzymů....22
2.3. Lidská glutamátkarboxypeptidasa II (GCPII)....26
2.3.1. Reakční mechanismus GCPII....30
2.3.2. Biologická funkce a využití lidské GCPII v medicíně....32

3. Materiály a metody 35

3.1. Experimentální metody....35
3.2. Materiály....38
3.2.1. Seznam chemikálií, enzymů, kitů a jiného materiálu....38
3.2.2. Seznam přístrojů....39
3.3. Podrobnosti k výpočtům....40
3.3.1. Kombinované QM/MM modely lidské GCPII a jejích mutantních forem....40
3.3.2. Klastrové modely aktivního místa GCPII....42
3.3.3. Minimalistické klastrové modely GCPII a thermolysinu....45

4. Výsledky 46

4.1. Alternativní sestřih GCPII mRNA....46
4.2. Modelování záměn aminokyselin v aktivním místě GCPII pomocí QM/MM….51
4.2.1. Geometrie a reakční koordináta....51
4.2.2. Reakční bariéry....55
4.3. Modelování reakčního mechanismu pomocí klastrových modelů....57
4.3.1. Geometrie....58
4.3.2. Reakční koordináta....62
4.3.3. Reakční bariéra....62
4.3.4. Alternativní reakční mechanismus....65
4.4. Porovnání DFT a „post-SCF“ metod na malých klastrových modelech....69
4.4.1. Glutamátkarboxypeptidasa II....69

4.4.2. Aminopeptidasa z *Aeromonas proteolytica* (AAP).....74
4.4.3. Thermolysin.....76

5. Diskuse **78**

6. Závěr **84**

7. Literatura **85**

8. Seznam publikovaných a nepublikovaných prací **93**

Seznam zkratek

AAP	aminopeptidasa z *Aeromonas proteolytica*
B3LYP	Beckeho-3 hybridní korelačně-výměnný funkcionál
CCSD(T)	metoda vázaných klastrů zahrnující mono- a biexcitace, triexcitace jsou zahrnuty přibližně z poruchové metody
cDNA	„complementary“ DNA, DNA vznikající reverzní transkripcí mRNA
CI	„Configuration interaction“, konfigurační interakce
CISD	„Configuration interaction with single and double excitations“, konfigurační interakce se zahrnutím mono- a biexcitací
COSMO	„conductor-like screening model“, solvatační model polarizovatelného kontinua
DFT	„Density functional theory“, teorie funkcionálu hustoty
dNTPs	směs deoxynukleotid trifosfátů
EDTA	„ethylene diamine tetraacetic acid“, ethylendiamintetraoctová kyselina
EST	„expressed sequence tag“, označení pro (částečné) sekvence cDNA zaklonované a osekvenované z cDNA knihoven
FDA	„Food and Drug Administration“, americký úřad pro kontrolu potravin a léčiv
FOLH1	„Folate hydrolase 1“, název genu kódujícího GCPII
GC	glutaminylcyklasa
GCPII	glutamátkarboxypeptidasa II
GCPIII	glutamátkarboxypeptidasa III
GGA	„generalized gradient approximation“, třída korelačně-výměnných funkcionálů závisejících na elektronové hustotě a její prostorové první derivaci
HF metoda	Hartreeho-Fockova metoda
IMDM	„Iscove's Modified Dulbecco's Medium“, živné medium pro tkáňové kultury
INT	intermediát reakce
LB medium	Luria-Bertani medium
LB agar	Luria-Bertani agar
LNCaP	buněčná linie odvozená od metastáz karcinomu prostaty v lymfatických uzlinách
LYP	funkcionál „Lee-Yang-Parr“
MC	Michaelisův komplex
MHH MED1	buněčná linie odvozená od meduloblastomu
MM energie	molekulově mechanická energie příslušného systému

MM systém	systém uvažovaný na úrovni molekulové mechaniky (popisován empirickým silovým polem)
mRNA	„messenger“ RNA, zralá forma polymerní RNA kódující proteiny
PAP	„Prostate acidic phosphatase“, prostatická kyselá fosfatasa
PBE	funkcionál „Perdew-Burke-Ernzerhof“
PBS	„phosphate buffered saline“, fosfátový pufr
PCR	„polymerase chain reaction“, polymerasová řetězová reakce
PSA	„Prostate specific antigen“, prostatický specifický antigen
PSMA	„Prostate specific membrane antigen“, prostatický specifický membránový antigen
RICC2	„Resolution of identity - approximate 2nd order coupled cluster“, přibližná metoda vázaných klastrů druhého řádu využívající RI aproximaci
RIMP2	„Resolution of identity - Møller-Plesset 2nd order perturbation theory“, Møllerova-Plessetova poruchová metoda druhého řádu využívající RI aproximaci
r.m.s.d.	„root mean square deviation“, odmocnina střední kvadratické odchylky
QM/MM	„Quantum mechanics/molecular mechanics“, kombinovaná kvantově mechanická/molekulově mechanická metoda
QM energie	energie systému vypočtená kvantovou mechanikou
QM systém	systém popisovaný kvantovou mechanikou
qRT PCR	„quantitative reverse-transcriptase polymerase chain reaction“, kvantitativní polymerasová řetězová reakce se zpětným přepisem
QZVP	báze atomových orbitalů „quadruple zeta“ kvality, v programu Turbomole označovaná jako def2-QZVP
QZVPP	báze atomových orbitalů „quadruple zeta“ kvality, v programu Turbomole označovaná jako def2-QZVPP
RI	„Resolution of identity“, způsob nahrazení čtyřcentrických elektronových integrálů rozvojem třícentrických integrálů pomocí tzv. jednotkového rozvoje
RI-J	nahrazení čtyřcentrických coulombických integrálů v metodě DFT rozvojem třícentrických pomocí tzv. jednotkového rozvoje
RT PCR	„reverse-transcriptase polymerase chain reaction“, polymerasová řetězová reakce se zpětným přepisem
SV(P)	báze atomových orbitalů „double zeta“ kvality, v programu Turbomole označovaná jako def2-SVP

SVP	báze atomových orbitalů „double zeta“ kvality, v programu Turbomole označovaná jako def2-SV(P)
systém S	malý („small“) klastrový model
systém M	střední („medium“) klastrový model
systém L	velký („large“) klastrový model
TBE	„tris-borate-EDTA“, elektrodový pufr obsahující tris, kyselinu boritou a EDTA
TPSS	funkcionál „Tao-Perdew-Staroverov-Scuseria“
TS1	tranzitní stav 1
TS2	tranzitní stav 2
TZVP	báze atomových orbitalů „triple zeta“ kvality, v programu Turbomole označovaná jako def2-TZVP
VWN	funkcionál „Vosko-Wilk-Nusair“
ZPVE	„Zero point vibrational energy“, vibrační energie nulového bodu

1. Předmluva

Lidská glutamátkarboxypeptidasa II (GCPII, E.C. 3.4.17.21) je membránový glykoprotein druhého typu. Řadí se mezi dinukleární zinkové metalopeptidasy. Významná exprese GCPII je pozorována v prostatě, menší množství GCPII se vyskytuje také v játrech, mozku, ledvinách, tenkém střevě a některých dalších tkáních.

V mozku hydrolyzuje GCPII neurotransmitter *N*-acetyl-α-L-aspartyl-L-glutamát (*N*-Ac-Asp-Glu) za vzniku L-glutamátu a podílí se tak na modulaci synaptického přenosu, ve střevě pak odštěpováním C-terminálních glutamátů z folylpoly-γ-L-glutamátu umožňuje vstřebávání folátu ze stravy. V prostatě není její funkce známa.

Zvýšená exprese v karcinomu prostaty a ve velké většině odvozených metastáz činí z GCPII vhodný cíl pro diagnostiku této nemoci. Jako první se do klinické praxe prosadila protilátka s obchodním názvem ProstaScint značená isotopem ^{111}In, kterou schválila FDA již v roce 1996. Dále se ukazuje, že GCPII exprimují endoteliální buňky nově tvořených krevních vlásečnic mnoha pevných nádorů, což spolu se schopností vstoupit do buňky po navázání protilátky činí z GCPII velmi nadějný cíl nejen pro vizualizaci těchto nádorů, ale i pro případnou imunoterapii.

Mnoho lidských genů podléhá při transkripci genomové informace do mRNA procesu zvanému alternativní sestřih, který vede k tomu, že z jediného genu vzniká několik různých forem mRNA lišících se navzájem různými inzercemi či delecemi. GCPII není výjimkou a studium její mRNA vedlo k objevu celé řady alternativně sestřižených forem. Na základě několika studií pomocí PCR v reálném čase na malých klinických vzorcích se předpokládá, že množství některých variant se významně liší mezi zdravou prostatou, benigní hyperplazií prostaty a karcinomem prostaty. Kdyby se tyto výsledky potvrdily na větším vzorku tkání, pak by kvantifikace mRNA jednotlivých forem v bioptických vzorcích či v krvi mohla představovat významné doplnění a vylepšení dnešních diagnostických metod rakoviny prostaty, neboť tyto metody se stále potýkají s velkým množstvím falešně positivních a falešně negativních nálezů. Problémem dosavadních studií je, že se každá z nich zabývala pouze několika málo variantami a neexistuje žádná studie porovnávající všechny dnes známé varianty najednou, proto jsme se rozhodli takový test vyvinout. Aby měly jeho výsledky co největší vypovídací hodnotu, tak jsme nejprve detailně prozkoumali alternativní sestřih genu pro GCPII ve dvou nádorových buněčných liniích odvozených od karcinomu prostaty a meduloblastomu a exprimujících GCPII. Výsledky této studie tvoří experimentální část diplomové práce.

Doplňující přístup ke studování proteinů je představován metodami teoretické chemie. Za

předpokladu, že dnešní metody kvantové chemie jsou schopné poskytnout výsledky uspokojivé přesnosti, lze jejich využitím ušetřit mnoho cenného experimentálního času i nemalé prostředky. Právě modelování reakčního mechanismu GCPII pomocí různých metod kvantové chemie tvoří značnou část teoretické části práce.

Lze vyjít z nedávno publikované práce, ve které byl pomocí kombinované metody QM/MM s kvantově-mechanickým popisem na úrovni teorie funkcionálu hustoty (DFT) studován reakční mechanismus GCPII. Stejnými metodami a na stejném modelu jsem se pokusil popsat vliv záměny postranních řetězců některých aminokyselin v aktivním místě na reakční bariéru. Studované proteiny byly již dříve rekombinantně připraveny a byly stanoveny jejich kinetické charakteristiky. To umožňuje přímé srovnání teoretických dat s experimentálními.

Je nutné podotknout (a tím vysvětlit motivaci pro další kapitolu teoretické části diplomové práce), že teoretické studium variantních proteinů je provázeno mnohými úskalími. Přesnost, kterou požadujeme od teoretických metod, by měla být někde na úrovni 1-2 kcal.mol^{-1} (například rozdíl 1,42 kcal.mol^{-1} představuje rozdíl celého řádu v rychlostní konstantě k_{cat} při $t = 37$ °C). Je tedy nanejvýš vhodné přezkoumat přesnost používaných výpočetních metod a správnost používaných modelů, obzvláště uvážíme-li, že již zmíněný QM/MM model přecenil reakční bariéru o více než 10 kcal.mol^{-1}.

Poměrně často bývá pro osvětlení reakčních mechanismů enzymů používán model aktivního místa obklopený pouze polarizovatelným kontinuem nebo vakuem (v textu označován jako klastrový model). Stejný model byl proto vytvořen i pro GCPII a výsledky následně porovnány s původní metodou QM/MM. Pro klastrový model byla studována také konvergence získaných reakčních bariér v závislosti na velikosti modelu a namodelován alternativní reakční mechanismus, při kterém je na začátku katalytického cyklu karboxyskupina glutamátu sloužícího jako tzv. „proton shuttle“ deprotonována (na rozdíl od dříve publikovaného mechanismu).

Nakonec byla porovnána metoda DFT s metodou RICC2. Srovnání bylo provedeno na minimalistických klastrových modelech reaktantů, produktů, a tranzitních stavů aktivního místa GCPII a dalších dvou zinkových peptidas. Na dvou modelových reakcích bez přítomnosti kovu jsme se také pokusili porovnat metodu DFT s referenční metodou CCSD(T).

2. Úvod

2.1. Metody kvantové chemie

Velká část kvantové chemie se zabývá řešením Schrödingerovy vlnové rovnice pro mnohačásticové systémy (atomy a molekuly). Nedochází-li například k interakci s časově proměnným elektromagnetickým polem (tj. hamiltonián systému není časově závislý), je možné se omezit na řešení stacionární Schrödingerovy rovnice ve tvaru $\hat{H}\Psi(R_A, r_i) = E\Psi(R_A, r_i)$, kde $\hat{H}$ je hamiltonián plně definující systém. Vlnová funkce Ψ závisí na spinových a prostorových souřadnicích jader (*R*) a elektronů (*r*), *E* je pak energie systému. Hamiltonián systému ve vakuu obsahující *N* jader a *n* elektronů lze pak v atomových jednotkách psát:

$$\hat{H} = \hat{T}_n + \hat{T}_e + \hat{V}_{nn} + \hat{V}_{ne} + \hat{V}_{ee} = -\sum_{A=1}^{N} \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^{n} \frac{1}{2} \nabla_i^2 + \sum_{A=1}^{N} \sum_{B>A}^{N} \frac{Z_A Z_B}{R_{AB}} - \sum_{A=1}^{N} \sum_{i=1}^{n} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{n} \sum_{j>i}^{n} \frac{1}{r_{ij}},$$

kde *M* je jaderná hmotnost a *Z* jaderný náboj. V důsledku velkého rozdílu mezi hmotností jader a elektronů lze většinou separovat jejich pohyb a řešit tak elektronový problém pro danou geometrii jader samostatně. Toto zjednodušení se nazývá Bornova-Oppenheimerova aproximace. Z této aproximace vyplývá existence tzv. hyperplochy potenciální energie skládající se z elektronové energie (která je výsledkem řešení elektronového problému) a potenciální energie jader V_{nn}.

Ve zjednodušené Schrödingerově rovnici tedy vlnová funkce $\Psi(R_A, r_i)$ závisí na poloze jader pouze parametricky, elektronový hamiltonián pak neobsahuje členy $\hat{T}_n + \hat{V}_{nn}$. Přesto není kvůli poslednímu dvouelektronovému coulombickému členu $\hat{V}_{ee}$ možné tuto rovnici řešit analyticky a je proto nutné zavést další aproximace, které budou popsány v následujících odstavcích. Podrobnou diskusi metod kvantové chemie lze nalézt v řadě učebnic [Szabo 1989, Levine 1991, Yarkony 1995].

2.1.1. Hartreeho-Fockova metoda

První aproximací, která umožní separaci elektronových proměnných, je zápis vlnové funkce ve tvaru antisymetrizovaného součinu jednoelektronových funkcí (prostorových orbitalů), tzv. Slaterova determinantu. Dále zavedeme tzv. model nezávislých elektronů, kde každý elektron interaguje pouze se zprůměrovaným polem ostatních elektronů a definujeme pro něj jednoelektronový Fockův operátor (pro systém s uzavřenými slupkami):

$$\hat{F}(1) = \hat{h}_{(1)}^{core} + \sum_{j=1}^{n/2} [2\hat{J}_j(1) - \hat{K}_j(1)], \text{ kde } \hat{h}_{(1)}^{core} = -\frac{1}{2}\nabla_{(1)}^2 - \sum_{A=1}^{N} \frac{Z_A}{r_{1A}}.$$

Coulombický operátor $\hat{J}_j$ a výměnný operátor $\hat{K}_j$ jsou přitom definovány jako:

$$\hat{J}(1)f(1)=f(1)\int|\phi_j(2)|^2\frac{1}{r_{12}}dv_2 \text{ a } \hat{K}_j(1)f(1)=\phi_j(1)\int\frac{\phi_j^*(2)f(2)}{r_{12}}dv_2,$$

kde $\phi_j(1)$ je molekulový orbital, $f(1)$ je libovolná funkce a integrace probíhá přes celý prostor. Zavedení Fockova operátoru při řešení variačního problému $\langle\Phi|\hat{H}|\Phi\rangle\geqslant E$ (kde Φ představuje vlnovou funkci ve tvaru Slaterova determinantu) vede k jednoelektronovým Hartreeho-Fockovým (HF) rovnicím ve tvaru:

$$\hat{F}(1)\phi_i(1)=\varepsilon_i\phi(1).$$

Molekulové orbitaly ϕ_i jsou v praxi vyjádřeny lineární kombinací atomových orbitalů zpravidla tabelovaných v tzv. bázích atomových orbitalů, $\phi_i=\sum_\mu c_{\mu i}\chi_\mu(r)$. Dosazení tohoto rozvoje do HF rovnic (pro jednoduchost se dále omezíme na tzv. „restricted“ HF metodu, kdy dvojice elektronů s různým spinem sdílejí stejnou prostorovou část orbitalu) vede k maticové rovnici (zobecněnému vlastnímu problému):

$$\mathbf{FC} = \mathbf{SCE},$$

kde **C** je matice vektorů rozvojových koeficientů $c_{\mu i}$, **S** je překryvová matice s prvky $S_{\mu\nu}=\langle\chi_\mu|\chi_\nu\rangle$ a **F** je Fockova matice s prvky:

$$F_{\mu\nu}=\langle\chi_\mu|\hat{h}|\chi_\nu\rangle+\sum_j^{n/2}\sum_\rho\sum_\sigma c_{\rho j}c_{\sigma j}[2\langle\chi_\mu\chi_\rho|\frac{1}{r_{12}}|\chi_\sigma\chi_\nu\rangle-\langle\chi_\mu\chi_\rho|\frac{1}{r_{12}}|\chi_\nu\chi_\sigma\rangle].$$

Jelikož Fockův operátor závisí na tvaru molekulových orbitalů, je nutné tuto rovnici řešit iterativně. Řešení spočívá v diagonalizaci Fockovy matice **F** s okrajovou podmínkou $\mathbf{C}^\dagger\mathbf{SC} = \mathbf{1}$ (podmínka orthonormality), konstrukci nové Fockovy matice a její opětovné diagonalizaci až po dosažení tzv. self-konzistence, kdy se ve dvou po sobě následujících iteracích změní celková energie (a většinou i redukovaná matice hustoty) o méně než předem stanovenou mez. Prvky diagonální matice **E** pak odpovídají jednotlivým orbitálním energiím tzv. kanonických HF orbitalů, celková Hartreeho-Fockova energie je definována jako střední hodnota přesného hamiltoniánu:

$$E^{HF}=\langle\Phi|\hat{H}|\Phi\rangle=\frac{1}{2}\sum_i(\varepsilon_i+\langle\phi_i|\hat{h}|\phi_i\rangle),$$

kde sčítání probíhá přes obsazené orbitaly. HF energie tedy není pouhým součtem orbitálních energií.

2.1.2. Korelační energie, „post-SCF“ metody

Protože Fockův operátor popisuje jednoelektronový potenciál pomocí zprůměrované pole ostatních elektronů, nezahrnuje tzv. dynamickou elektronovou korelaci. Zatímco pohyb elektronů se stejným spinem je popisován v důsledku antisymetričnosti vlnové funkce (a z ní

vyplývající existence výměnných integrálů) poměrně správně (tzv. Fermiho korelace), pohyb elektronů s opačným spinem je popisován zcela nekorelovaně. Tato chyba narůstá s rostoucím počtem elektronů v systému a proto musí být HF vlnová funkce dále vylepšena, což je cílem tzv. „post-SCF“ metod. Podstata několika z nich bude stručně shrnuta v následujících odstavcích.

2.1.2.1. Møllerova-Plessetova poruchová metoda

Podrobnou diskusi poruchové metody lze nalézt v řadě klasických učebnic, proto se zde omezíme pouze na výsledky Møllerovy-Plessetovy poruchové metody bez uvedení odvození.

V této metodě je jako neporušený systém brán model nezávislých elektronů popsaný Fockovým operátorem, porucha je pak definována jako rozdíl mezi přesným hamiltoniánem a Fockovým operátorem. V korekci prvního řádu k energii dostaneme proto pouze původní HF energii, v důsledku použití kanonických HF orbitalů se podstatně zjednodušší korekce k energii druhého řádu, započtená v nejpoužívanější formě této metody MP2:

$$E_0^{(2)} = \sum_{b=a+1}^{\infty} \sum_{a=n+1}^{\infty} \sum_{i=j+1}^{n} \sum_{j=1}^{n-1} \frac{\langle \phi_a \phi_b | r_{12}^{-1} | \phi_i \phi_j \rangle - \langle \phi_a \phi_b | r_{12}^{-1} | \phi_j \phi_i \rangle}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} ,$$

kde ϕ_i a ϕ_j jsou obsazené molekulové spinorbitaly, ϕ_a a ϕ_b jsou neobsazené spinorbitaly (do kterých se „excitují“ elektrony z orbitalů ϕ_i a ϕ_j) a ε_i jsou příslušné orbitální energie. Sumace probíhá přes všechny spinorbitaly (postupně se provádí všechny možné „biexcitace“). Podmínkou konvergence poruchové metody je malá hodnota této korekce, tzn. že dostaneme velmi rychlou konvergenci např. pro atom helia, kde je malý podíl elektronové korelace, zatímco pro velké molekuly bude Møllerova-Plessetova poruchová metoda konvergovat pomaleji.

2.1.2.2. Konfigurační interakce

Jedním ze způsobů, jak lze dosáhnout vylepšení HF vlnové funkce, je rozšíření variačního prostoru při minimalizaci střední hodnoty přesného hamiltoniánu:

$$\langle \Psi | \hat{H} | \Psi \rangle \geqslant E .$$

Vlnová funkce pak může být vyjádřena například pomocí rozvoje:

$$\Psi = c_0 \Phi_0 + \sum_{a=n+1}^{\infty} \sum_{i=1}^{n} c_i^a \Phi_i^a + \sum_{b=a+1}^{\infty} \sum_{a=n+1}^{\infty} \sum_{j=i+1}^{n} \sum_{i=1}^{n} c_{ij}^{ab} \Phi_{ij}^{ab} .$$

Φ_0 značí Slaterův determinant základního stavu, Φ_i^a a Φ_{ij}^{ab} značí monoexcitovaný respektive biexcitovaný Slaterův determinant. Tento konkrétní případ konfigurační interakce bývá nazýván CISD („Configuration Interaction with Single and Double Excitations“). Zahrnutí všech možných excitací je nazýváno jako „*full*-CI“ rozvoj a při použití nekonečné báze lze docílit přesné hodnoty elektronové energie v nerelativistické aproximaci. Je ale zřejmé, že pro větší počet elektronů je „*full*-CI“ metoda (i v běžně používaných konečných bázích) výpočetně příliš

náročná. Širšímu používání neúplných rozvojů typu CISD pak brání nestejný popis korelační energie s rostoucí velikostí systému, například u molekul s deseti elektrony popíše CISD více než 90% korelační energie zatímco u molekul s 50 elektrony méně než 80% (tzv. problém „size extensivity“) [Sasaki 1977].

2.1.2.3. Metody vázaných klastrů (Coupled Cluster)

Elegantní formulování stejného problému jako u konfigurační interakce (i když nikoliv variační) spočívá v definici vlnové funkce jako:

$$\Psi = e^{\hat{T}} \Phi_0 ,$$

kde $\hat{T}$ představuje tzv. excitační operátor. Exponenciální tvar operátoru v předchozí rovnici lze vyjádřit Taylorovým rozvojem:

$$e^{\hat{T}} \equiv 1 + \hat{T} + \frac{\hat{T}^2}{2!} + \frac{\hat{T}^3}{3!} + ... = \sum_{k=0}^{\infty} \frac{\hat{T}^k}{k!} .$$

Působení mocninou excitačního operátoru odpovídá několikanásobnému postupnému působení na vlnovou funkci. Již pro monoexcitační operátor tak formálně dostáváme rozvoj vlnové funkce ve tvaru lineární kombinace mono-, bi-, triexcitovaných atd. Slaterových determinantů. Z toho plyne zásadní rozdíl oproti konfigurační interakci, který spočívá v tom, že shodné excitace (např. z orbitalu ϕ_i do ϕ_a) obsažené např. v monoexcitovaných a biexcitovaných Slaterových determinantech jsou váženy se stejným rozvojovým koeficientem t_i^a (nazývaným amplitudou). V CCSD metodě je vlnová funkce definována jako $\Psi = e^{(\hat{T}_1 + \hat{T}_2)} \Phi_0$, kde $\hat{T}_1$ je monoexcitační operátor a $\hat{T}_2$ biexcitační operátor. Řešení Schrödingerovy rovnice pak vede na soustavu nelineárních rovnic pro amplitudy, které musí být řešeny iterativně. Propojení („coupling“) amplitud pro různě excitované Slaterovy determinanty vede k tomu, že tato metoda je narozdíl od konfigurační interakce „size extensive“. CCSD(T) metoda pak na přibližné úrovni zahrnuje také triexcitace získané z poruchové teorie. Metody vázaných klastrů jsou shrnuty například v [Bartlett 1995].

2.1.3. Teorie funkcionálu hustoty

Ve všech dosud uvedených metodách byla počítána vlnová funkce která je závislá na spinových a prostorových souřadnicích všech elektronů (tedy $4N$ proměnných). Alternativní způsob, který zde stručně popíšeme, spočívá v určení pouze elektronové hustoty, která je závislá pouze na třech prostorových a jedné spinové souřadnici. Podrobnou diskusi teorie funkcionálu hustoty lze nalézt v [Parr 1994, Koch 2001].

Teoretickým základem teorie funkcionálu hustoty („density functional theory“, DFT) jsou

dva Hohenbergovy-Kohnovy teorémy. První tvrdí, že tzv. externí potenciál $v_{ext}(r)$ (který v praxi většinou odpovídá elektrostatickému působení jader na elektrony) je v případě systému s nedegenerovaným základním elektronovým stavem jednoznačným funkcionálem elektronové hustoty. Jinými slovy, systém (poloha a náboj jader a celkový počet elektronů) je jednoznačně definován elektronovou hustotou a naopak. Druhý pak definuje funkcionál energie systému a ukazuje, že pro něj v základním elektronovém stavu platí variační teorém [Hohenberg 1964].

Pro jednoduché vyjádření kinetické energie se definuje model neinteragujících elektronů, jehož kinetická energie je rovna:

$$T_e^{\mathrm{ni}} = -\frac{1}{2}\sum_{i=1}^{N} \langle \varphi_i | \nabla^2 | \varphi_i \rangle ,$$

kde φ_i jsou jednoelektronové orbitaly. Od tohoto systému dále požadujeme, aby jeho elektronová hustota, definovaná jako:

$$\rho(r) = \sum_{i=1}^{N} |\varphi_i(r)|^2 ,$$

byla stejná jako elektronová hustota reálného interagujícího systému. Funkcionál energie je pak:

$$E[\rho(r)] = T_e^{\mathrm{ni}}[\rho(r)] + V_{ne}[\rho(r)] + V_{ee}[\rho(r)] + \Delta T_e[\rho(r)] + \Delta V_{ee}[\rho(r)] ,$$

kde $T_e^{\mathrm{ni}}[\rho(r)]$ je kinetická energie neinteragujících elektronů, $V_{ne}[\rho(r)]$ představuje elektrostatickou atrakci mezi elektrony a jádry, $V_{ee}[\rho(r)]$ elektrostatickou repulzi mezi elektrony. $\Delta T_e[\rho(r)]$ je korekce ke kinetické energii v důsledku elektronové interakce a $\Delta V_{ee}[\rho(r)]$ je neklasický příspěvek k elektronové interakci. Poslední dva členy jsou zahrnuty do jediného funkcionálu nazvaného korelačně-výměnný $E_{xc}[\rho(r)]$, jehož explicitní tvar neznáme. Energii systému lze pak rozepsat jako:

$$E[\rho(r)] = -\frac{1}{2}\sum_{i}^{N} \langle \varphi_i | \nabla^2 | \varphi_i \rangle + \frac{1}{2}\sum_{i}^{N}\sum_{j}^{N} \int\int |\varphi(r_1)|^2 \frac{1}{r_{12}} |\varphi(r_2)|^2 dr_1 dr_2 + E_{xc}[\rho(r)] - \sum_{i}^{N} \int \sum_{A}^{M} \frac{Z_A}{r_{1\mathrm{A}}}$$

Aplikací variačního teorému lze odvodit tzv. Kohnovy-Shamovy rovnice [Kohn 1965] (které jsou obdobou HF rovnic):

$$\left(-\frac{1}{2}\nabla^2 + \left[\int \frac{\rho(r_2)}{r_{12}} dr_2 + V_{xc}(r_1) - \sum_{A}^{M} \frac{Z_A}{r_{1\mathrm{A}}}\right]\right)\varphi_i = \hat{\mathrm{f}}^{\mathrm{KS}}\varphi_i = \varepsilon_i \varphi_i ,$$

kde $\hat{\mathrm{f}}^{\mathrm{KS}}$ je jednoelektronový Kohnův-Shamův operátor. Explicitní tvar $V_{xc}(r_1)$ neznáme, prozatím jej definujeme pouze jako:

$$V_{xc}(r_1) = \frac{\delta E_{xc}}{\delta \rho} .$$

Kromě korekce ke kinetické energii a neklasického příspěvku k elektronové repulzi

(korelačně-výměnný člen) pak tento výraz obsahuje (často však náhodně) opravu interakce elektronu se sebou samým („self interaction correction“, SIC), která se v důsledku absence přesné výměny v Kohnových-Shamových rovnicích narozdíl od HF rovnic nevyruší. Použití báze atomových orbitalů vede k formálně stejným maticovým rovnicím jako v případě HF metody. Jejich řešení je také iterativní, neboť Kohnův-Shamův operátor závisí na elektronové hustotě systému (stejně jako Fockův operátor na vlnové funkci), je proto nutné ho v každé iteraci znovu přepočítat.

2.1.3.1. Korelačně-výměnné funkcionály

Přesný tvar korelačně-výměnného funkcionálu neznáme, proto se používá celá řada přibližných funkcionálů, které jsou zpravidla získávány na základě *ab initio* výpočtů a fitováním několika parametrů na experimentální data. Správnost tohoto funkcionálu je klíčová pro úspěch či neúspěch celé teorie funkcionálu hustoty.

Tzv. funkcionály „Local density approximation“ (LDA) závisí pouze na elektronové hustotě a jejich aplikace je především na poli výpočtů elektronové struktury pevné fáze. Dalším typem jsou tzv. funkcionály „Generalized gradient approximation“ (GGA), které závisí kromě elektronové hustoty také na její první prostorové derivaci. Příkladem GGA funkcionálu je funkcionál „Perdew-Burke-Ernzerhof“ (PBE) [Perdew 1996]. Další vylepšení představují tzv. „meta-GGA“ funkcionály, které závisí na elektronové hustotě, jejich první a druhé prostorové derivaci a na hustotě kinetické energie

$$\tau = \frac{1}{2}\sum_i |\nabla \chi_i|^2 ,$$

příkladem takového funkcionálu je funkcionál „Tao-Perdew-Staroverov-Scuseria“ (TPSS) [Tao 2003]. Hybridní funkcionály pak kombinují nejčastěji GGA funkcionál s částí přesné HF výměny. Pro chemické aplikace je velmi používaným hybridním funkcionálem B3LYP, jehož výměnná část pochází z Beckeho výměnného funkcionálu, zatímco korelační část tvoří korelační funkcionály „Lee-Yang-Parr“ (LYP) [Lee 1988] a „Vosko-Wilk-Nusair“ (VWN) [Vosko 1980]. To vše se kombinuje s částí přesné výměnné energie získané z HF metody [Becke 1993]:

$$E_{xc}^{B3LYP} = (1-a)E_x^{Slater} + aE_x^{HF} + bE_x^{B88} + (1-c)E_c^{VWN} + cE_c^{LYP}$$

kde E_x^{Slater} je Diracův-Slaterův výměnný funkcionál odvozený od uniformního elektronového plynu, E_x^{HF} je výměnná energie získaná HF metodou, E_x^{B88} je gradientová korekce k výměnné energii [Becke 1988], E_c^{VWN} a E_c^{LYP} jsou již zmíněné funkcionály LYP a VWN. Proměnné *a*, *b* a *c* jsou pak empirické parametry.

2.1.3.2. Omezení teorie funkcionálu hustoty

V důsledku nahrazení nelokální HF výměny výrazy závisející pouze lokálně či semilokálně na elektronové hustotě nepopisují standardní korelačně-výměnné funkcionály správně disperzní energii. Vylepšení v tomto směru představují tzv. dvojitě hybridní funkcionály či započtení empirické korekce [Grimme 2006, Schwabe 2007].

Další problémy představují např. nepřesné vyrušení interakce elektronu se sebou samým či absence statické korelace v důsledku jednodeterminantové formulace DFT.

2.1.4. Partiční funkce, Gibbsova volná energie

Při zkoumání chemických reakcí klasickými metodami kvantové chemie získáváme údaje o energii systému. Avšak pro děje probíhající při nenulové teplotě není v důsledku nenulové translační energie měřítkem samovolnosti děje samotná energie, ale volná energie. Statistické vlastnosti systému pak popisuje partiční funkce, podrobný úvod do statistické termodynamiky lze nalézt například v [Chandler 1987, McQuarrie 2000].

Analytický tvar partiční funkce lze naneštěstí odvodit pouze pro některé jednoduché modely, při odvozování molekulární partiční funkce je proto nutné učinit celou řadu aproximací. Při použití aproximace tuhého rotoru (tzn. že při rotaci nedochází k žádné distorzi molekuly), harmonické aproximace (vibrační hladiny jsou ekvidistantně separovány) a předpokladu že molekula tvoří ideální plyn, lze pro Gibbsovu volnou energii psát:

$$G = E_{\mathrm{el}} + E_{\mathrm{ZPVE}} - RT \ln\left(q_{\mathrm{trans}}\, q_{\mathrm{rot}}\, q_{\mathrm{vib}}\right),$$

kde E_{el} je elektronová energie, E_{ZPVE} je energie nulového bodu a pro q_{vib} lze psát (vibrační partiční funkce je při výpočtech v této práci hlavním zdrojem chyb, proto je zde podrobněji diskutována, viz kapitola 3.3.1.):

$$q_{vib} = \prod_i \frac{1}{1 - e^{(-h\nu_i / k_B T)}}$$

kde ν_i je frekvence i-tého vibračního módu. Pro entalpii lze psát:

$$H = E_{\mathrm{ZPVE}} + \frac{5}{2} RT + \sum_i h\nu_i \frac{1 + e^{(-h\nu_i / k_B T)}}{2\left[1 - e^{(-h\nu_i / k_B T)}\right]},$$

příspěvek jednoho vibračního módu k entalpii je tedy roven příslušnému členu v sumaci. Příspěvek ke Gibbsově volné energii G_{ν_i} je přitom roven:

$$G_{\nu_i} = -RT \ln \frac{1}{1 - e^{(-h\nu_i / k_B T)}}.$$

Absolutní hodnoty příspěvků jednotlivých vibračních módů o různých frekvencích k energii nulového bodu, entalpii a Gibbsově volné energii shrnuje tabulka 1. Z této tabulky je zřejmé, že příspěvek měkkých vibračních módů k energii nulového bodu je zcela zanedbatelný. Naproti

tomu příspěvek k entalpii není možné zanedbat, avšak jeho hodnota vykazuje asymptotické chování pro limitu frekvence proti nule. Příspěvek módů k entalpii s rostoucí frekvencí pomalu stoupá, avšak pro měkké módy mezi 1 až 100 cm^{-1} je takřka konstantní ($\sim k_B T$). Největší příspěvek měkkých módů je pak ke Gibbsově energii, navíc se zmenšující frekvencí velmi výrazně mění. Proč je toto chování důležité pro výpočty v této práci, je uvedeno v kapitole 3.3.1.

Tabulka 1. Příspěvek jednoho vibračního módu k energii nulového bodu, entalpii a Gibbsově volné energii při teplotě *t*=37 °C.

Levý sloupec udává frekvenci daného vibračního módu, E_{ZPVE} značí energii nulového bodu, H entalpii a G Gibbsovu volnou energii.

Vibrace	E_{ZPVE} [kcal.mol^{-1}]	H [kcal.mol^{-1}]	G [kcal.mol^{-1}]
1 cm^{-1}	0,0014	0,62	-3,3
10 cm^{-1}	0,014	0,62	-1,9
100 cm^{-1}	0,14	0,63	-0,60
1000 cm^{-1}	1,4	1,6	-0,06

2.1.5. Teorie tranzitního stavu a reakční rychlost

Reakční rychlosti chemických reakcí vysvětluje teorie tranzitního stavu, odvozená Eyringem, Evansovou a Polanyim v roce 1935. Tato teorie předpokládá: (1) platnost Bornovy-Oppenheimerovy aproximace a tudíž existenci hyperplochy potenciální energie, (2) tato hyperplocha potenciální energie je rozdělena na oblast reaktantů a produktů, sedlový bod mezi oběma oblastmi odpovídá tranzitnímu stavu, (3) populovanost jednotlivých struktur vyhovuje Boltzmannově distribuci. Jelikož je pohyb jader po reakční koordinátě popisován klasickou mechanikou, zanedbává tedy např. možnost tunelového efektu. Eyringovu rovnici vyjadřující rychlostní konstantu reakce k lze pak psát:

$$k = \frac{k_B T}{h} e^{\frac{-\Delta G^{\neq}}{RT}},$$

kde k_B je Boltzmannova konstanta, R univerzální plynová konstanta, h Planckova konstanta, T teplota a $\Delta G^{\neq}$ rozdíl Gibbsovy volné energie mezi reaktantem a tranzitním stavem. Rychlostní konstanta 1 s^{-1} pak odpovídá při t=37 °C reakční bariéře 18,2 kcal.mol^{-1} [Truhlar 1996].

2.2. Enzymová kinetika a tvorba modelu

Řada životně důležitých chemických reakcí v našem těle by v důsledku vysoké reakční bariéry probíhala bez katalysy s rychlostní konstantou v řádu let či ještě více (například přeměna glukosy na oxid uhličitý), čili v časovém měřítku nedostačujícím pro udržení homeostázy tak, jak ji známe. Roli katalysátorů v živých organismech plní enzymy, což jsou biologické polymery kódované genomem a složené v drtivé většině z aminokyselin (proteiny) či výjimečně ribonukleotidů (ribozymy) [Kruger 1982]. Od běžných anorganických katalysátorů se enzymy liší nejen svou výjimečnou účinností ale také specifitou katalysované reakce. U řady enzymů stačí k úplné ztrátě aktivity velmi malá modifikace substrátu, například přidání methylové skupiny [McClain 1993, Casadesus 2006]. Toto je velmi důležitá vlastnost, zvláště když si uvědomíme, kolik vedlejších produktů vzniká při běžné synthese v organické chemii. Přitom obdobných synthetických reakcí probíhají v buňkách tisíce [Kanehisa 2008]. Další velmi důležitá vlastnost enzymů je jejich regulovatelnost. V závislosti na potřebách buňky může být enzym (přechodně) aktivován či inhibován, například kovalentní modifikací nebo vazbou nízkomolekulárního ligandu či jiného proteinu [Hicke 2001, Bozulic 2009, Shemarova 2009].

Detailní pochopení reakčního mechanismu enzymů je často experimentálně špatně přístupné, proto se nabízí alternativní přístup modelování enzymových reakcí na kvantově mechanické úrovni. Výpočetní výkon současných počítačů umožňuje studovat (například na úrovni DFT) pohodlně systémy obsahující 150 až 200 atomů. To je ale stále řádově méně než běžná velikost enzymu. Je proto nutné se omezit jen na aktivní místo, ve kterém probíhá samotná reakce, a zbytek proteinu popsat pouze přibližně, ať už pomocí molekulové mechaniky či solvatačních modelů vycházejících z polarizovatelného continua [Senn 2007, Siegbahn 2009].

V následujících odstavcích je stručně diskutována enzymová kinetika, princip enzymové katalysy a způsoby modelování aktivního místa glutamátkarboxypeptidasy II a dalších zinkových metalopeptidas, které jsou použity v této práci.

2.2.1. Enzymová kinetika

Jako modelová reakce poslouží jednoduchá jednosubstrátová reakce typu:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} EP \xrightarrow{k_3} E + P$$

kde E značí volný enzym, S volný substrát, ES komplex enzym-substrát, EP komplex enzym-produkt, P volný produkt a k_1, k_{-1}, k_2, k_3 jsou rychlostní konstanty jednotlivých kroků. Štěpení dipeptidu *N*-Ac-Asp-Glu glutamátkarboxypeptidasou II je příkladem takové reakce.

V předchozí rovnici neuvažujeme zpětnou reakci $EP \rightarrow ES$, neboť předpokládáme, že na počátku reakce je koncentrace produktu velmi malá a proto je rychlost této reakce zanedbatelná.

Vzhledem k tomu, že se reakční rychlosti experimentálně stanovují metodou počátečních rychlostí, tzn. při malých konverzích substrátu na produkt, nepředstavuje tento předpoklad významnou chybu.

Dále předpokládáme, že se po určité počáteční fázi reakce ustanoví rovnovážná koncentrace komplexu *ES*, tzn. že rychlost tvorby a zániku tohoto komplexu je stejná:

$$k_1[E][S]=(k_{-1}+k_2)[ES]\,.$$

Dále zavedeme celkovou koncentraci enzymu (volného i vázaného v *ES*):

$$[E_t]=[E]+[ES]$$

a zkombinováním posledních dvou výrazů a vyjádřením pro ES dostáváme:

$$[ES]=\frac{[E_t][S]}{[S]+(k_{-1}+k_2)/k_1}\,.$$

Člen $(k_{-1}+k_2)/k_1$ se nazývá Michaelisova konstanta - K_m. Dosadíme-li za *ES* do výrazu pro rychlost reakce ($V_0=k_2[ES]$), dostaneme:

$$V_0=\frac{k_2[E_t][S]}{K_m+[S]}\,.$$

Jelikož maximální rychlost reakce nabývá svého maxima tehdy, je-li enzym plně saturovaný substrátem, lze V_{max} definovat jako $V_{max}=k_2[E_t]$. Po dosazení do předchozího vztahu dostáváme rovnici Michaelise a Mentenové, popisující jednosubstrátovou enzymově katalysovanou reakci:

$$V_0=\frac{V_{max}[S]}{K_m+[S]}\,.$$

Michaelisova konstanta K_m má tedy rozměr koncentrace, a je-li koncentrace substrátu rovna K_m, pak rychlost reakce odpovídá polovině maximální rychlosti V_{max}, což je velmi praktická definice. Za předpokladu, že druhý krok celkové reakce (charakterizovaný rychlostní konstantou k_2) je krokem určujícím rychlost reakce, to jest $k_{-1}\gg k_2$, pak lze K_m interpretovat jako míru afinity enzymu k substrátu, neboť výraz $(k_{-1}+k_2)/k_1$ se zjednoduší na $(k_{-1})/k_1$, což odpovídá rovnovážné konstantě popisující vznik komplexu *ES*.

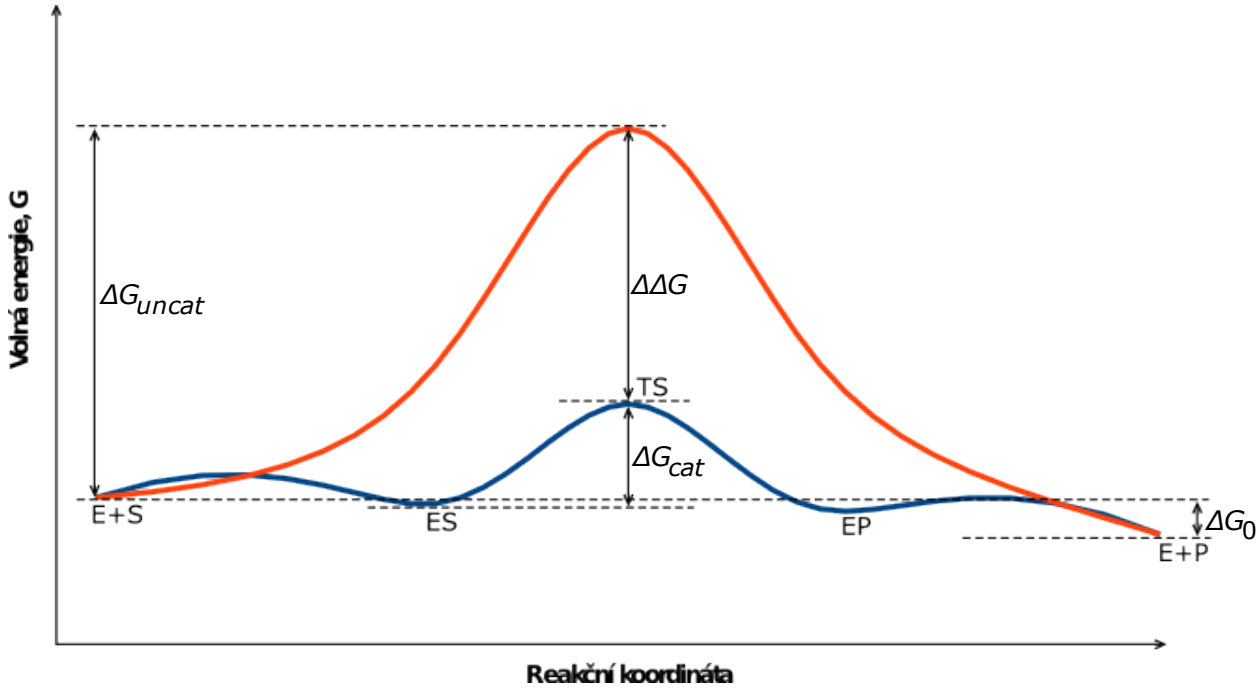
Ne všechny jednosubstrátové reakce se řídí výše uvedenou kinetikou, jako příklad uveďme inhibici enzymu substrátem či produktem. V případě inhibice substrátem se rychlost reakce se stoupající koncentrací substrátu již nepřibližuje asymptoticky k V_{max} tak jako u Michaelise-Mentenové kinetiky, nýbrž po dosažení určitého maxima opět klesá.

Místo k_2 se používá označení k_{cat}, které obecně odpovídá rychlostní konstantě kroku určujícímu rychlost reakce. Je-li například nejpomalejším krokem reakce disociace produktu

z aktivního místa enzymu, pak k_{cat} odpovídá k_3 ve vzorové reakci uvedené na začátku tohoto oddílu. Experimentálně se hodnota k_{cat} určuje ze vztahu [Copeland 2000].

$$k_{cat} = \frac{V_{max}}{[E_t]} .$$

Urychlení reakce katalýzou je důsledkem snížení celkové bariéry volné Gibbsovy energie dané reakce v důsledku stabilizace energeticky nejvýše položených tranzitních stavů. Obrázek 1 schematicky znázorňuje profily volné energie katalysované i nekatalysovaná reakce:



Obrázek 1. Schematické znázornění profilů volné energie katalysované a nekatalysované reakce.
ΔG_{uncat} odpovídá bariéře volné energie nekatalysované reakce v roztoku, ΔG_{cat} pak bariéře katalysované reakce a $\Delta\Delta G$ stabilizaci tranzitního stavu enzymem. ΔG_0 značí rozdíl volné energie celé reakce, jeho hodnota nemá vliv na rychlost reakce ale ovlivňuje rovnovážný poměr koncentrací obou látek v roztoku.

Je tedy zřejmé, že aktivní místo musí být uzpůsobeno tak, aby nejsilněji vázalo a tím stabilizovalo jinak energeticky nestabilní tranzitní stavy. Naproti tomu velká afinita enzymu k substrátu by vedla ke zvýšení reakční bariéry, jak je vidět na obrázku 1.

Některé enzymy byly evolucí selektovány tak, aby jimi katalysované reakce probíhaly co nejrychleji (např. katalasa či karbonická anhydrasa), jiné zase vynikají substrátovou specifitou. Zde lze jako příklad uvést aminoacyl-tRNA synthetasy, které se podílejí na přepisu genetického kódu do aminokyselinového kódu proteinů. Tyto enzymy dokáží rozlišit i tak podobné aminokyseliny jako leucin, isoleucin, valin atp. a jsou charakterizovány velmi nízkými

hodnotami Michaelisovy konstanty K_m [Jakubowski 1996, Nelson 2004]. Předchozí příklady naznačují, že lze jen obtížně hledat zcela obecné principy enzymové katalysy a je nutné studovat reakční mechanismus každého enzymu jednotlivě, nejlépe kombinací experimentálních metod a pokročilých metod kvantové chemie, které správně popisují elektronovou strukturu. Potom není nutné se omezovat pouze na fenomenologický popis.

Je též důležité si uvědomit, že se substrát váže do aktivního místa v konformaci vhodné pro reakci, lze proto předpokládat že změna konformační entropie v průběhu reakce je poměrně malá. Podrobnější diskusi lze nalézt například v [Nelson 2004, Liao 2009a]. Toto je dobrá zpráva, neboť pomocí standardních kvantově chemických modelů lze jen velmi obtížně entropii správně popsat (viz kapitola 3.3.1). Podobné závěry nelze učinit o entropickém zabarvení samotné vazby substrátu do aktivního místa [Perozzo 2000, Kožíšek 2007].

2.2.2. Přístupy k modelování enzymů

Jak již bylo řečeno, výpočetní výkon současných počítačů zatím neumožňuje studium celých proteinů metodami DFT či *ab initio* kvantově chemickými metodami. Je proto nutné vytvořit malý model aktivního místa, kde probíhá samotná reakce, a zbytek proteinu uvažovat pouze na přibližné úrovni. Pro další zmenšení tohoto modelu je většinou nutné se omezit pouze na postranní zbytky důležitých aminokyselin, u metaloenzymů jsou typicky zahrnuty přechodné kovy nacházející se v aktivním místě, s nimi interagující postranní řetězce aminokyselin, substrát (nebo jeho část) a ideálně ještě postranní zbytky aminokyselin interagující se substrátem [Siegbahn 2009].

Při tvorbě modelu se nejčastěji vychází ze struktur proteinů získaných rentgenovou difrakcí na proteinovém krystalu (zřídka pak metodami nukleární magnetické rezonance, NMR). Tyto struktury jsou uložené a volně přístupné v databázi Protein Data Bank (PDB). Nejvhodnější jsou přitom struktury určené s rozlišením lepším než 2.0 Å. Jsou-li dostupné struktury enzymu s prázdným aktivním místem, s nehydrolysovatelným analogem substrátu, inhibitorem napodobujícím strukturu tranzitního stavu, či neaktivního mutantu se substrátem navázaným v aktivním místě, mohou poskytnout (vedle kinetických dat) další zpětnou vazbu modelovaného reakčního mechanismu s experimentálními poznatky.

Používanými metodami k efektivnímu zahrnutí zbytku proteinu jsou tzv. kombinované „Quantum Mechanics/Molecular Mechanics“ (QM/MM) metody [Senn 2007], ve kterých je molekula proteinu (mimo aktivní místo, které je popsané na kvantově mechanické úrovni) popisován méně přesným empirickým potenciálem (MM část). Často je tato MM část rozdělena na dvě podčásti. Atomy vnější podčásti (systém MM3, okraj proteinu) zůstávají fixované v

polohách známých z krystalové struktury proteinu, zatímco okolí aktivního místa je umožněno relaxovat v odpovědi na konformační změny uvnitř aktivního místa, které nastávají v průběhu reakce. Zafixováním okraje proteinu se zamezí nechtěným konformačním změnám velkého rozsahu, které by mohly být artefaktem výpočtu.

Alternativní přístup spočívá v nahrazení zbytku proteinu polarizovatelným kontinuem [Siegbahn 2009]. Tento přístup se někdy nazývá klastrový model. V této práci byly oba tyto přístupy použity a následně porovnány na příkladu reakčního mechanismu GCPII a jsou proto podrobněji diskutovány v následujících odstavcích.

2.2.2.1. Kombinovaná metoda QM/MM

Pro QM/MM výpočty byl v této práci použit program ComQum [Ryde 1996]. V následujících odstavcích proto bude popsáno pouze QM/MM schema implementované v tomto programu. Obecnější diskusi metod QM/MM lze nalézt např. v [Senn 2007].

ComQum ve standardní podobě spojuje program Turbomole [Ahlrichs 1989] pro část popisovanou pomocí kvantové mechaniky (dále označovanou jako QM část) a Amber [Case 2005] pro část popisovanou pomocí klasické mechaniky (dále označovanou jako MM část). Enzym je obvykle stejně jako v metodě ONIOM [Vreven 2006] rozdělen na tři části (viz obrázek 2). Centrální systém 1 zahrnuje aktivní místo enzymu a je optimalizován na QM úrovni. Systém 2 bývá definován uvnitř určitého poloměru okolo systému 1 a je optimalizován na MM úrovni. Zbytek enzymu a molekuly vody ležící od systému 1 dále než zvolený poloměr tvoří systém 3, který je sice zahrnut ve výpočtech energie, ale je ponechán geometrii známé z krystalu. Systémy 2 a 3 jsou při QM výpočtech systému 1 reprezentovány neceločíselnými bodovými náboji pro každý atom MM oblasti v molekulovém hamiltoniánu systému 1. Ve většině případů se jedná o parciální náboje získané přímo z jejich elektrostatické reprezentace (knihovny) použitého silového pole. Vzájemné elektrostatické působení mezi QM systémem a zbytkem proteinu je tedy zahrnuto v QM výpočtu (tzv. „electrostatic embedding“).

V případě existence kovalentní vazby spojující atom ležící v systému 1 s atomem ležícím v systému 2 je systém 1 utvářen tak, že atomy těchto vazeb ležící v systému 2 (tzv. junction atomy) jsou nahrazeny vodíky, přičemž jejich poloha je přímo úměrná poloze původních „junction atomů“ v neporušeném systému (viz obrázek 2). Aby se zabránilo přílišné (a arteficiální) polarizaci kvantového systému, jsou bodové náboje na atomech v systému 2 spojené vazbou s „junction atomy“ vynulovány a jejich náboj je rozpočítán na zbylé atomy aminokyselinového fragmentu, tak aby zůstal zachován celkový náboj.

Celková energie je pak rovna:

$$E_{tot} = E_{QM} + E_{MM123} - E_{MM1} ,$$

kde E_{QM} je QM energie systému 1 (s „junction atomy“ nahrazenými vodíky) zahrnující elektrostatickou interakci mezi systémem 1 a bodovými náboji reprezentujícími systémy 2 a 3 avšak bez interakce bodových nábojů se sebou samými. E_{MM1} je pak MM energie systému 1 bez jakékoliv elektrostatické interakce (s vynulovanými náboji) a E_{MM123} je MM energie celého proteinu, se všemi původními atomy. Aby se nezapočítávala podruhé elektrostatická interakce mezi systémem 1 a zbytkem proteinu, jsou při výpočtech E_{MM123} náboje na atomech náležejících systému 1 též vynulovány. Toto schema vede k tomu, že se chyby způsobené přerušením vazeb na okraji systému 1 obsažené v E_{QM} a E_{MM1} navzájem přibližně vyruší.

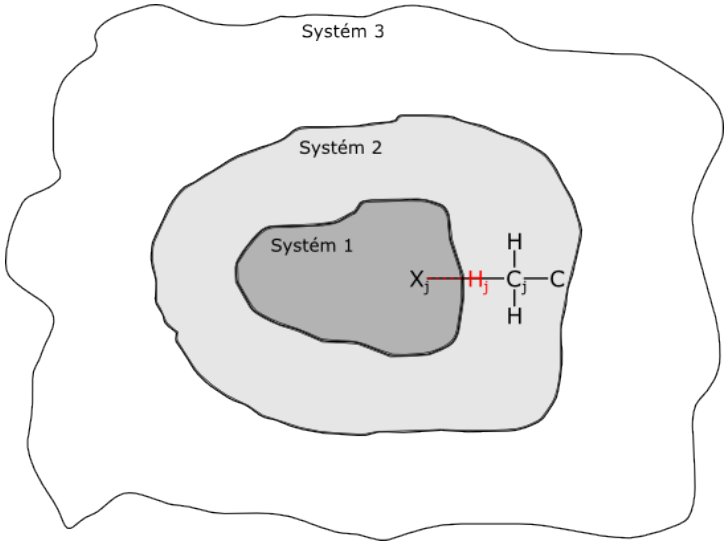
Vypočtené síly jsou pak zápornými gradienty příslušných energií:

$$-\nabla E_{tot}(r_1, r_{H_j}, r_{2-3}) = -\nabla E_{QM}(r_1, r_{H_j}) - \nabla E_{MM123}(r_1, r_{C_j}, r_{2-3}) + \nabla E_{MM1}(r_1, r_{H_j})$$

a pro polohové vektory různých reprezentací junction atomů použitých v jednotlivých systémech platí následující vztah:

$$r_{C_j} = r_{X_j} + k(r_{H_j} - r_{X_j}) ,$$

kde $k = (r_{C_j} - r_{X_j})/(r_{H_j} - r_{X_j})$. Symboly r_1 značí polohové vektory atomů systému 1, $r_{2\text{-}3}$ značí polohové vektory atomů systémů 2 a 3, r_{C_j}, r_{X_j} a r_{H_j} značí polohové vektory atomů C_j, X_j a H_j, tak jak jsou znázorněny na obrázku 2.



Obrázek 2. Schematické znázornění rozdělení enzymu do tří systémů v metodě QM/MM.

Systém 1 je uvažován pomocí kvantové mechaniky, systémy 2 a 3 pomocí klasické mechaniky.

Označení atomů C_j, X_j, a H_j odpovídá označení v předchozím textu. Dále jsou zobrazeny atomy systému 2 spojené vazbou s atomem C_j, náboje těchto atomů jsou pro QM výpočty vynulovány.

Výpočetní protokol při optimalizaci QM/MM lze pak shrnout v tomto schématu:

```
Vypočti QM vlnovou funkci
Opakuj
      Vypočti QM síly (S1)
      Vypočti MM síly (S1+2+3, S1)
      Sečti QM a MM síly
      Relaxuj geometrii S1 za použití těchto sil (1 krok)
      Změň prostorové souřadnice S1 v MM reprezentaci
      Má-li být optimalizován S2 pak
            Vypočti QM náboje na atomech S1
            Přepiš tyto náboje v MM reprezentaci S1
            Relaxuj S2 MM minimalizací při fixovaných S1 a S3 (více kroků,
                  až do konvergence)
            Změň prostorové souřadnice S2 v QM reprezentaci
      Vypočti QM vlnovou funkci a energii (S1)
      Vypočti MM energie (S1+2+3 a S1)
      Sečti QM a MM energie
až do splnění konvergenčních kritérií
```

2.2.2.2. Klastrový model aktivního místa

Jednodušší způsob modelování reakčních mechanismů spočívá v tom, že se vytvoří model aktivního místa obdobně jako u metody QM/MM systém 1, ale dále se uvažuje pouze tento systém mimikující reakční centrum a zbytek proteinu se nahradí modelem polarizovatelného kontinua. Konstrukce těchto modelů se oproti tvorbě systému 1 u metody QM/MM může lišit ve výběru zahrnutých zbytků aminokyselin, je totiž například nutné aby celkový náboj systému byl blízký nule. Na rozdíl od metody QM/MM nejsou v okolí žádné bodové náboje, které by jej kompenzovaly [Chen 2008, Liao 2009b, Siegbahn 2009].

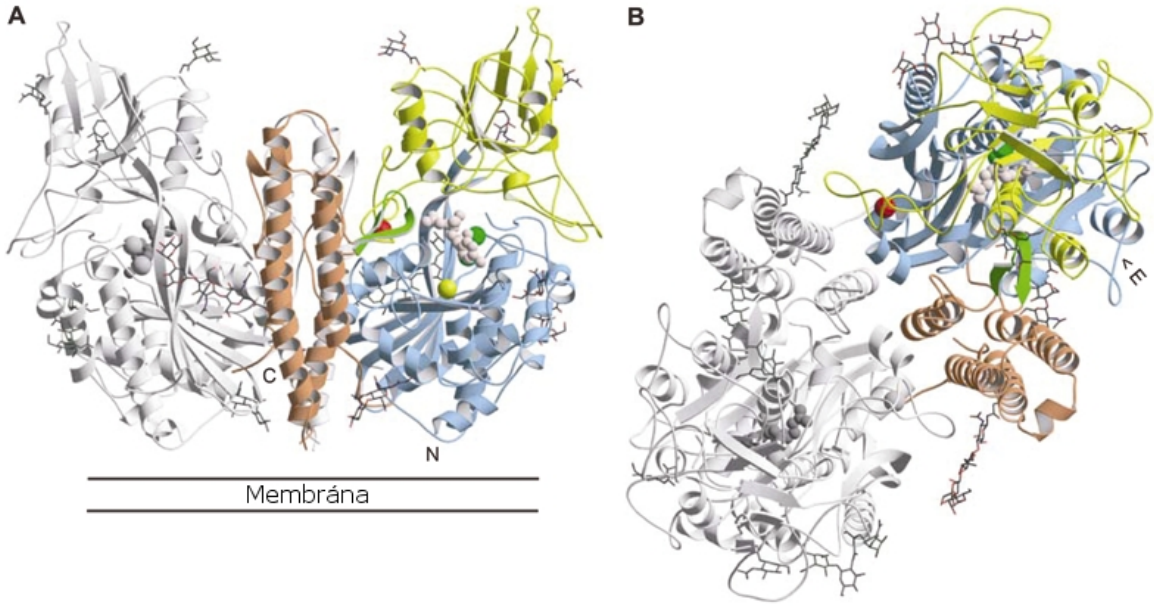
Jelikož proteinové okolí aktivního místa, které je zde ignorováno, v reálném proteinu stericky omezuje řadu rotací a translací skupin obsažených v modelu aktivního místa, jsou zpravidla u těchto modelů fixovány prostorové souřadnice těch krajních atomů, u kterých byla přerušena vazba s původním proteinovým okolím. Geometrické optimalizace jsou často prováděny ve vakuu, energie výsledných struktur pak bývají přepočítány za použití dielektrického kontinuálního solvatačního modelu, oblíbených Tomasiho modelů PCM („Polarized continuum model“) [Cossi 2003] či modelu COSMO („Conductor-like screening model“) [Klamt 1993], implemetovaného například v programu Turbomole [Schäfer 2000].

2.3. Lidská glutamátkarboxypeptidasa II (GCPII)

GCPII je integrální membránový protein II. typu o délce 750 aminokyselin. Jeho molekulová hmotnost (včetně deseti N-glykosylací) činí přibližně 110 kDa. Jeho deglykosylace vede ke ztrátě enzymové aktivity [Holmes 1996, Bařinka 2002, Ghosh 2003, Bařinka 2004]. Aktivní formou GCPII v roztoku je homodimer [Schulke 2003].

Známými přirozenými substráty GCPII jsou *N*-acetyl-α-L-aspartyl-L-glutamát (*N*-Ac-Asp-Glu) [Robinson 1987] a folylpoly-γ-L-glutamát [Pinto 1996]. Nejvyšší katalytické účinnosti pro štěpení *N*-Ac-Asp-Glu dosahuje GCPII okolo pH=8. Této závislosti dominuje pH závislost Michaelisovy konstanty K_m, zatímco katalytická konstanta k_{cat} se se změnou pH téměř nemění [Hlouchová 2007]. Slabými inhibitory GCPII jsou chelátory dvojmocných kationtů, polyvalentní anionty typu sulfátu či fosfátu či produkt štěpení *N*-Ac-Asp-Glu L-glutamát [Robinson 1987]. Deriváty kyseliny L-glutarové jako například kyselina 2-fosfonomethylpentandiová (2-PMPA) [Jackson 1996] či kyselina 2-(3-mercaptopropyl) pentandiová pak představují jedny z nejúčinnějších známých inhibitorů GCPII. Druhý jmenovaný inhibitor je navíc aplikovatelný perorálně a prochází přes mozkomíšní bariéru [Majer 2003].

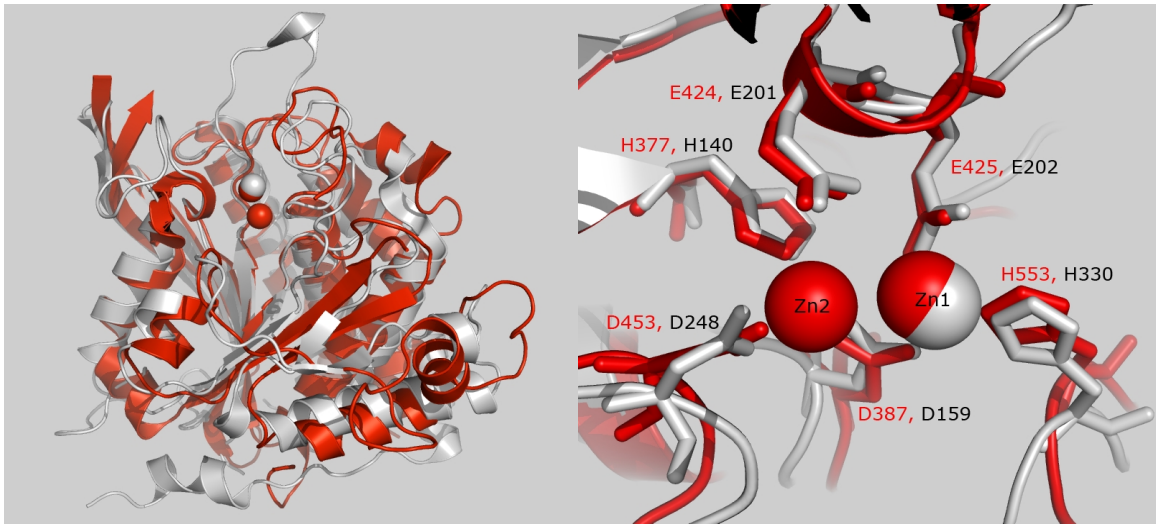
Krystalová struktura extracelulární části GCPII (znázorněná na obrázku 3) ukazuje tři strukturní domény: proteasovou, apikální a dimerizační [Davis 2005, Mesters 2006]. V proteasové doméně se nachází katalytické centrum se dvěma zinečnatými kationty. Helikální dimerizační doména tvoří velkou část kontaktní plochy obou podjednotek (spolu s proteasovou doménou), žádná konkrétní funkce apikální domény není známa. V současné době jsou vyřešeny struktury GCPII s prázdným aktivním místem (PDB kód 2OOT; rozlišení 1,64 Å) [Bařinka 2007b], s navázaným L-glutamátem v aktivním místě (2C6G; 2,2 Å), s navázaným fosfátem v aktivním místě (2C6P; 2,4 Å) [Mesters 2006] či s navázaným inhibitorem 2-PMPA (2PVW; 1,71 Å) [Bařinka 2007a] a řada dalších struktur s různými inhibitory. Existuje také krystalová struktura neaktivního mutantu GCPII, u kterého je zaměněn Glu424 za Ala424, s navázaným substrátem GCPII *N*-acetyl-α-L-aspartyl-L-glutamátem v aktivním místě (3BXM; 1,71 Å). Jelikož struktura aktivního místa zůstala oproti GCPII divokého typu zachována a přesto došlo ke ztrátě aktivity, je zřejmé že tento aminokyselinový postranní řetězec hraje klíčovou úlohu při štěpení peptidové vazby [Klusák 2009].



Obrázek 3. Krystalová struktura extracelulární části GCPII.

První podjednotka znázorněného dimeru je zbarvena šedě, zatímco druhá barevně podle doménového uspořádání (proteasová doména modře, apikální žlutě a dimerizační červeně). Dva zinečnaté ionty jsou znázorněny zelenými kuličkami, chloridový iont žlutou kuličkou a vápenatý iont červenou kuličkou. Glykosidické složky jsou znázorněny tyčinkovými modely. E značí vstup k aktivnímu místu, C- resp. N-konec jsou vyznačeny písmeny C resp N.
(A) Pohled na homodimer GCPII z povrchu membrány, (B) pohled z vrchu. Převzato z [Mesters 2006].

GCPII patří do rodiny dizinkových metalopeptidas M28 [Rawlings 2008] a její struktura ukázala shodný strukturní motiv (fold) proteasové domény s ostatními členy této rodiny, např. bakteriálními aminopeptidasami ze *Streptomyces griseus* [Greenblatt 1997] a *Aeromonas proteolytica* [Chevrier 1994]. S lidskými zástupci této rodiny, glutamátkarboxypeptidasou III (70% sekvenční identity s GCPII) [Hlouchová 2009] a transferrinovým receptorem I (25% sekvenční identity s GCPII, není katalyticky aktivní) [Lawrence 1999] sdílí GCPII stejný celkový fold všech tří domén. Celkový fold proteasové domény je stejný dokonce i u nejvzdálenějšího lidského homologu GCPII - glutaminylcyklasy (10% sekvenční identity s GCPII, katalysuje cyklizaci N-terminálního glutaminu a glutamátu různých peptidů) [Fischer 1987, Schilling 2004, Huang 2005]. Společný všem katalyticky aktivním členům této rodiny je aminokyselinový motiv HDEDH, který v aktivním místě enzymu váže dva zinečnaté ionty. Glutamátu podílejícímu se na vazbě zinečnatých iontů pak v primární struktuře bezprostředně předchází další glutamát, který je klíčový pro enzymatickou aktivitu. Zajímavé je, že celý tento motiv je zachován i u glutaminylcyklasy, která v aktivním místě váže pouze jeden zinečnatý iont. Porovnání celkových foldů proteasových domén GCPII a GC, stejně jako porovnání strukturního uspořádání okolí zinečnatých iontů obou enzymů je znázorněno na obrázku 4.



Obrázek 4. Porovnání struktur proteasových domén a zinkových center GCPII a GC.

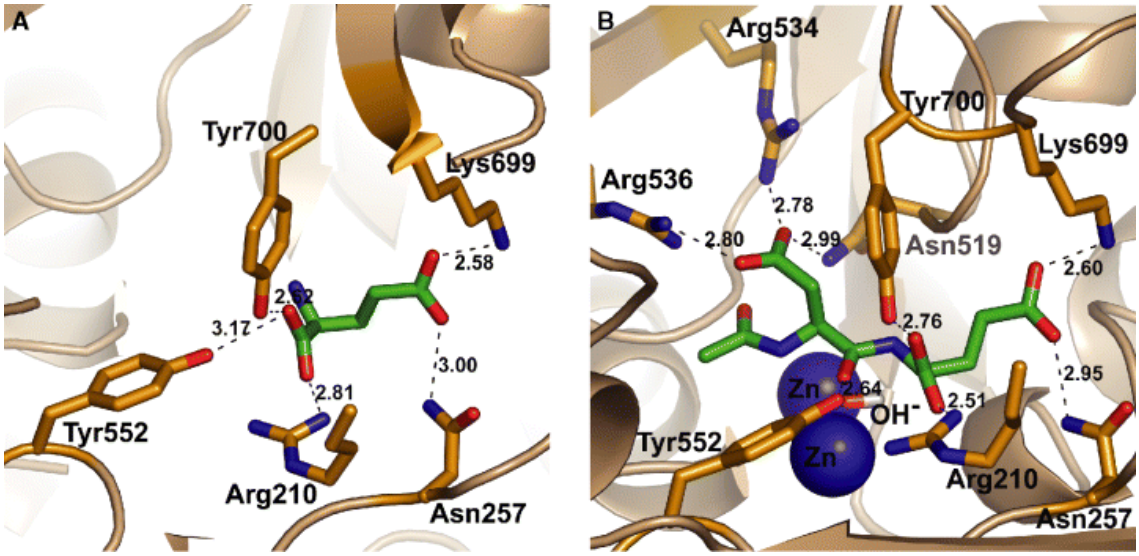
GCPII je znázorněna červeně, GC šedivě.

(A) Celkový pohled na proteasové domény obou enzymů, kuličky představují zinečnaté kationty. Přeložení struktur bylo vytvořeno s pomocí algoritmu DaliLite [Holm 2000] s r.m.s.d. 2,5 Å pro 256 C_α atomů z 329 porovnávaných.

(B) Detailní pohled na dinukleární zinkové centrum obou enzymů. Kuličky představují zinečnaté ionty, tyčinkovými modely jsou znázorněny aminokyseliny vážící zinečnaté ionty a katalytický glutamát (Glu424 u GCPII). Červené popisky odpovídají pořadí jednotlivých aminokyselin v primární struktuře GCPII, černé pak v primární struktuře GC. Převzato z [Navrátil 2007].

Ze struktury GCPII s navázaným L-glutamátem [Mesters 2006] lze jednoznačně určit aminokyseliny podílející se na jeho vazbě. Jsou to Arg210, Asn257, Gly518, Lys699 a Tyr700 (obrázek 5, panel A), tato vazebná kapsa se nazývá S1' (jedná se o standardní zápis, kde apostrof značí C-konec substrátu vzhledem k místu štěpení a číslo 1 značí, že se jedná o první aminokyselinu v pořadí vzhledem k místu štěpení). Aminokyseliny podílející se na vazbě zbytku substrátu, tvořící tzv. S1 vazebnou kapsu, se podařilo identifikovat pomocí QM/MM modelu [Mlčochová 2007] a později potvrdit pomocí krystalové struktury neaktivního mutantu GCPII Glu424Ala ($k_{cat} < 0{,}001\ s^{-1}$) s nerozštěpeným substrátem v aktivním místě (uspořádání aktivního místa není mutací dotčeno) [Klusák 2009]. Na vazbě aspartátového zbytku se podílí Asn519, Arg534, Arg536 a Tyr552, zatímco karbonylový kyslík acetylové skupiny substrátu interaguje s Arg536 (obrázek 5, panel B) [Mlčochová 2007].

Důležitost jednotlivých postranních řetězců aminokyselin vážících substrát byla podrobně studována pomocí cílené mutageneze následované rekombinantní expresí. Purifikované rekombinantní proteiny pak byly kineticky charakterizovány. Nejdůležitější pro zachování katalytické účinnosti se ukázaly být zbytky vážící glutamát. Naměřená kinetická data u jednotlivých záměn aminokyselin shrnuje tabulka 2 [Hlouchová 2007, Mlčochová 2007].



Obrázek 5. Vazba substrátu *N*-Ac-Asp-Glu do aktivního místa GCPII.

Uvedené vzdálenosti jsou v ångströmech, peptidová páteř je znázorněna pomocí stužkového modelu.

(A) L-glutamát navázaný v aktivním místě GCPII (krystalová struktura, PDB kód 2C6G [Mesters 2006]). Znázorněny jsou postranní řetězce aminokyselin interagující s L-glutamátem s výjimkou Gly518 interagujícího s aminoskupinou L-glutamátu (Gly518 viz obrázek 6).

(B) QM/MM struktura GCPII s navázaným *N*-Ac-Asp-Glu v aktivním místě. Znázorněny jsou aminokyselinové zbytky interagující se substrátem, opět s výjimkou Gly518 interagujícího s -NH skupinou peptidové vazby (Gly518 viz obrázek 6) [Klusák 2009]. Převzato z [Mlčochová 2007].

Tabulka 2. Kinetické parametry hydrolysy *N*-Ac-Asp-Glu enzymem GCPII a jeho mutantními formami.

Aminokyselina Asp520 se nepodílí na vazbě substrátu, nicméně se ve struktuře GCPII vyskytuje v blízkosti Glu424, který hraje klíčovou úlohu při štěpení peptidové vazby. Upraveno podle [Mlčochová 2007, Hlouchová 2007].

Mutace	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1}·s^{-1}$)
Wild-type	1.15 ± 0.57	1.1 ± 0.2	957
D520N	2.30 ± 0.180	0.007 ± 0.001	3.04
Aminokyseliny S1' vazebné kapsy			
R210A	294 ± 15	0.023 ± 0.001	0.08
R210K	801 ± 124	0.130 ± 0.020	0.16
N257D	68.10 ± 19.7	0.320 ± 0.080	4.70
K699S	40.50 ± 22.9	0.270 ± 0.060	6.67
Y700F	45.70 ± 6.6	0.075 ± 0.003	1.64
Aminokyseliny S1 vazebné kapsy			
G518P	2.20 ± 0.028	0.090 ± 0.020	40.9
N519D	27.60 ± 0.300	0.078 ± 0.005	2.83
N519S	0.10 ± 0.03	0.42 ± 0.02	41
N519V	0.67 ± 0.066	0.036 ± 0.001	53.7
R534L	0.14 ± 0.072	0.100 ± 0.040	714
R536L	0.18 ± 0.005	0.010 ± 0.005	55.6
Y552I	0.15 ± 0.036	0.014 ± 0.001	93.3

2.3.1. Reakční mechanismus GCPII

Reakční mechanismus GCPII byl nedávno podrobně popsán pomocí QM/MM modelu [Klusák 2009], jeho nejdůležitější kroky budou podrobněji rozebrány v následujících odstavcích. Geometrie QM/MM struktury Michaelisova komplexu enzymu jsou v této kapitole porovnávány s krystalovou strukturou mutantní formy GCPII Glu424Ala s navázaným substrátem v aktivním místě (PDB kód 3BXM [Klusák 2009]).

2.3.1.1. Dinukleární zinkové centrum

Mezi dvěma zinečnatými ionty je v krystalových strukturách přibližně symetricky vázán atom kyslíku (Zn...O vzdálenosti 2,2 Å v optimalizované QM/MM struktuře Michaelisova komplexu v porovnání s 2,0 Å v krystalové struktuře Glu424Ala; Zn...Zn vzdálenost je pak 3,3 Å v obou strukturách). Rentgenová difrakce při daném rozlišení bohužel nedává odpověď, jestli tento kyslík patří vodě nebo hydroxidovému aniontu. Porovnáním s krystalografickými vzdálenostmi pro Zn...OH^- a Zn...H_2O v Cambridge Structural Database a výpočty bylo ukázáno, že se jedná o hydroxidový aniont. Jako nejpravděpodobnější akceptorová skupina pro odstupující vodík z molekuly vody po navázání na zinkové centrum byla identifikována karboxylová skupina patřící Glu424. Druhý (neprotonovaný) kyslík karboxylové skupiny Glu424 pak slouží jako akceptor vodíkové vazby můstkového OH^- iontu, zatímco hydroxylová část tohoto karboxylu tvoří slabou vodíkovou vazbu s karbonylem peptidové páteře patřící Gly518 (viz obrázek 6) [Klusák 2009].

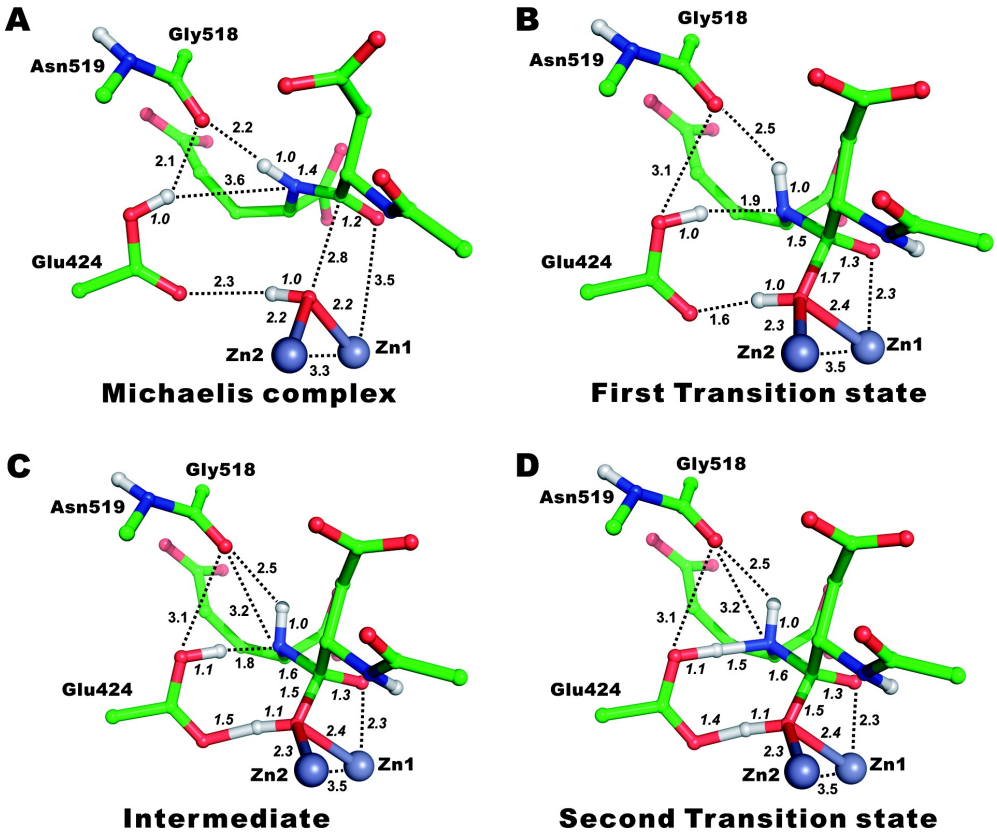
2.3.1.2. Interakce enzymu se štěpenou peptidovou vazbou

Substrát je v aktivním místě orientován tak, že -NH skupina peptidové vazby tvoří vodíkovou vazbu s karbonylem peptidové páteře patřící Gly518 (O...N vzdálenost 3,2 Å v QM/MM modelu versus 2,9 Å v krystalu). Karbonylový kyslík této peptidové vazby pak slouží jako akceptor vodíkové vazby OH skupiny Tyr552 a zároveň slabě interaguje s jedním ze zinků (3,5 Å v porovnání s 2,6 Å v krystalu, viz obrázek 6) [Klusák 2009].

2.3.1.3. Reakční koordináta

Peptidová vazba je v aktivním místě orientována tak, že karbonylový uhlík štěpené peptidové vazby je vzdálen pouze 2,8 Å od můstkového hydroxidového aniontu, který během reakce nukleofilně napadá právě tento uhlíkový atom. To je spojeno se změnou hybridizace z sp^2 na sp^3 nejen na tomto uhlíku, ale také na sousedním dusíku, který je v této konfiguraci stabilizován vodíkovou vazbou s hydroxylovou skupinou karboxylu Glu424, což vede ke vzniku stabilního tetraedrálního intermediátu (viz obrázek 6). Disociace peptidové vazby následuje po překonání bariéry pro přeskok protonu z Glu424 na pyramidalizovaný dusík této vazby.

Vypočtená reakční bariéra na úrovni QM(B3LYP/TZVP)/MM, $\Delta E^{\#}$ = 29 kcal.mol^{-1}, je přibližně o 11 kcal.mol^{-1} vyšší než hodnota plynoucí z experimentálně stanoveného $k_{cat} \approx 1$ s^{-1}. Relativní energie QM/MM struktur reakční koordináty vůči Michaelisovu komplexu jsou: $\Delta E^{\#}$(TS1) = 26,8 kcal.mol^{-1} , ΔE(INT) = 26,1 kcal.mol^{-1} a $\Delta E^{\#}$(TS2) = 29,0 kcal.mol^{-1}. Popsaný reakční mechanismus je velmi podobný mechanismu navrženému pro bakteriální aminopeptidasu z *Aeromonas Proteolytica* [Chen 2008].



Obrázek 6. Reakční mechanismus GCPII.

Znázorněny jsou pouze zinečnaté kationty Zn1 a Zn2, můstkový hydroxid, postranní řetězec Glu424, peptidová páteř Gly518-Asn519 a substrát (*N*-Ac-Asp-Glu). Koordinačně kovalentní vazby mezi hydroxidovým aniontem a zinečnatými ionty jsou také znázorněny. Udávané vzdálenosti jsou v ångströmech.

(A) Struktura Michaelisova komplexu. Nukleofilní útok můstkového hydroxidu na uhlíkový atom peptidové vazby za současné stabilizace pyramidálního dusíku vodíkovou vazbou od Glu424 vede přes (B) první tranzitní stav ke stabilnímu (C) intermediátu. (D) Po překonání nízké bariéry (řádově 3 kcal.mol^{-1}) pro přeskok protonu z Glu424 na pyramidální dusík dochází k disociaci peptidové vazby. Převzato z [Klusák 2009].

2.3.2. Biologická funkce a využití lidské GCPII v medicíně

GCPII je exprimována především v prostatě [Horoszewicz 1987, Kinoshita 2006], výrazně méně pak v mozku [Robinson 1987, Šácha 2007], játrech, ledvinách či tenkém střevě a řadě dalších tkání [Cunha 2006]. Ve střevě se GCPII odštěpováním C-terminálních glutamátů z folylpoly-γ-L-glutamátu podílí na vstřebávání folátu ze stravy [Heston 1997, Donelly 2001]. V centrální nervové soustavě se GCPII vyskytuje téměř výhradně na cytoplasmatické membráně astrocytů, které svými výběžky těsně obklopují neurosynaptické štěrbiny, ve které pak pravděpodobně GCPII štěpením neurotransmitteru *N*-Ac-Asp-Glu za uvolnění jiného neurotransmitteru L-glutamátu moduluje synaptický přenos [Robinson 1987, Šácha 2007]. Inhibice GCPII se ukázala být neuroprotektivní například během mozkové mrtvice u myší [Slusher 1999]. Role *N*-Ac-Asp-Glu v centrální nervové soustavě je shrnuta např. v [Neale 2000], role L-glutamátu v [Nedergaard 2002], úloha astrocytů pro nervový přenos v [Haydon 2001, Volterra 2005] a nakonec vliv inhibice GCPII např. v [Navrátil 2007].

Funkce GCPII v jiných tkáních není zatím zřejmá. Z lékařského hlediska je důležitá exprese GCPII v adenokarcinomu prostaty, v jejích metastázách a v endotelu neovaskulatury většiny pevných nádorů (v endotelu neovaskulatury zdravých tkání se nevyskytuje) [Liu 1997, Silver 1997, Chang 1999a, Chang 1999b]. Využití GCPII při diagnostice karcinomu prostaty bude podrobněji popsáno v následujících odstavcích.

2.3.2.1. Karcinom prostaty

Ač se rakovina prostaty vyvíjí velmi pomalu, je v pozdějších stádiích obtížně léčitelná a u mužů představuje jednu z nejčastějších příčin úmrtí na nádorová onemocnění. Pro terapii časných stadií (lokalizovaný karcinom bez metastáz) se s dobrými výsledky používá radikální prostatektomie (odebrání celé prostaty). Avšak tato metoda selhává u pacientů s metastazujícím karcinomem, navíc vzhledem k invazivitě metody není vhodné ji provádět u pacientů zbytečně. Dnešní diagnostické metody bohužel nedovedou včasně a spolehlivě rozeznat počínající karcinom prostaty a proto je velká poptávka po nových způsobech diagnosy, přičemž GCPII patří mezi nové a nadějné cíle [Schiffer 2007, Botchorishvili 2009].

Původně byla GCPII studována v prostatě a mozku nezávisle na sobě, teprve později se zjistilo, že se jedná o identický protein [Horoszewicz 1987, Robinson 1987, Carter 1996]. Odtud plyne její druhý, dodnes používaný název prostatický specifický membránový antigen (PSMA).

Kvantitativní PCR s reverzní transkripcí (qRT PCR) ukázala, že exprese GCPII v prostatě je více než desetkrát vyšší než v kterékoliv jiné tkáni, což představuje menší rozdíl než u současně používaných markerů karcinomu prostaty prostatickému specifickému antigenu

(PSA, [Wang 1981]) a kyselé prostatické fosfatase (PAP, [Chu 1978]) [Cunha 2006]. Exprese GCPII v prostatě je omezená na sekreční acinární epitel, přičemž 95% nádorů prostaty jsou právě adenokarcinomy exprimující GCPII. Pro vizualizaci těchto nádorů a jejich metastáz se v klinické praxi využívá zatím jediná protilátka proti GCPII s komerčním názvem ProstaScint [Howell 1999], novější protilátka označovaná J591 zatím prošla první fází klinického testování s velmi nadějnými výsledky [Liu 1997, Morris 2005]. Hlavní výhodu GCPII oproti nejhojněji používané PSA představuje zvýšení její exprese během antiandrogenní hormonální terapie používané během léčení rakoviny prostaty [Wright 1996, Perner 2007], zatímco exprese PSA je potlačena a nemůže tak sloužit jako prognostický marker [Murphy 1996]. Navíc některé tzv. 'high grade' karcinomy ztrácí oba jmenované markery (PSA i PAP), zatímco jsou intenzivně barveny protilátkami proti GCPII [Miller 2005].

Z exprese GCPII v endotelu neovaskulatur mnoha pevných nádorů plyne širší uplatnění GCPII jako nádorového markeru [Milowsky 2007]. Internalizace GCPII po navázání protilátky [Liu 1998] by bylo také pravděpodobně možné využít pro cílenou protinádorovou terapii, např. ve formě protilátky konjugované s toxinem, který by se specificky dostával do nádorových buněk [Fracasso 2002, Bander 2003].

2.3.2.2. Alternativně sestřižené varianty GCPII

Oproti používání značené rekombinantní protilátky jako kontrastní látky pro zviditelnění nádorů představuje přesné stanovení specifické mRNA pomocí qRT PCR levnější a citlivější metodu. Její citlivosti lze využít pro stanovení GCPII v krvi jako průvodního znaku cirkulujících metastáz karcinomu prostaty, které jsou pravděpodobně spojeny s horší prognózou pacienta [Miyake 2007, Zhang 2008].

Pomocí klasické RT PCR bylo v karcinomu prostaty či buněčných liniích od ní odvozených popsáno několik alternativně sestřižených variant mRNA GCPII. Varianta kódující celou formu GCPII obsahuje celkem 19 exonů a bývá nazývána stejně jako vlastní protein PSMA. První popsaná alternativně sestřižená varianta byla PSM', která místo 379 nukleotidů dlouhého prvního exonu obsahuje pouze prvních 113 nukleotidů tohoto exonu [Su 1995]. U další varianty označované PSMC následuje po tomto zkráceném prvním exonu nový alternativní exon čítající 132 nukleotidů. Varianta PSMD pak obsahuje stejný zkrácený první exon, který následuje ještě jiný alternativní exon než u PSMC, jeho délka je 97 nukleotidů [O'Keefe 2000]. Posledními variantami popsanými v literatuře jsou PSMF, která obsahuje celý první exon následovaný stejným alternativním exonem jako u PSMD, a PSME, která je stejná jako PSMF až na to, že neobsahuje 93 nukleotidů dlouhý osmnáctý exon [Cao 2007]. Všechny tyto varianty jsou

schematicky znázorněny na obrázku 7 v kapitole 4.1.

Translace PSM' mRNA vede ke vzniku 693 aminokyselin dlouhého proteinového produktu začínajícího Met58 (číslování odpovídá celé formě GCPII), který neobsahuje cytosolární a transmembránovou část, neobsahuje tedy žádnou sekvenci, které by jej směřovala do endoplasmického retikula a jedná se proto pravděpodobně o cytosolární, neglykosylovaný a enzymově neaktivní protein [Bařinka 2002]. V cytoplasmě se vyskytující glykosylovaná, enzymově aktivní forma GCPII nazývaná též PSM' a začínající Ala60 je s největší pravděpodobností produktem proteolytického štěpení kompletní formy GCPII [Mlčochová 2009]. PSMC obsahuje stejný nejdelší čtecí rámec jako PSM', avšak doposud nebylo demonstrováno zda nějaký protein z této mRNA vůbec vzniká. Naproti tomu translace variant PSMD a PSMF by vedla ke vzniku 735 aminokyselin dlouhého proteinového produktu, který by v porovnání s celou formou GCPII obsahoval alternativních 24 aminokyselin na N-konci následovaných 711 shodnými aminokyselinami. PSME protein by oproti PSMF neobsahoval 31 aminokyselin dlouhý úsek nedaleko C-konce. Z těchto variant byla zatím v transfekovaných buňkách demonstrována exprese pouze PSME varianty [O'Keefe 2000, Cao 2007].

Od objevení první varianty PSM' se spekulovalo o tom, že se poměr množství PSMA a PSM' variant liší mezi zdravou prostatou a karcinomem prostaty. Novější studie používající qRT PCR tuto domněnku potvrdila, avšak pro klinické využití není tato změna dostatečně statisticky významná. Z kvantitativních studií se zatím zdá, že množství PSMD varianty je statisticky významně vyšší u metastáz než u primárního karcinomu prostaty a že množství PSME se významně liší mezi zdravou prostatou, benigní hyperplazií prostaty, karcinomem prostaty a dokonce koreluje s Gleasonovým skórem nádoru. Množství PSMA varianty je sice vyšší u karcinomu prostaty než u zdravé prostaty, tato změna ale není zřejmě statisticky významná [Schmittgen 2003, Cao 2007].

3. Materiály a metody

3.1. Experimentální metody

Tkáňové kultury

Buněčné linie LNCaP (linie odvozená od metastáz karcinomu prostaty v míšních uzlinách, z angl. „Lymph Node Carcinoma of the Prostate“) a MHH MED1 (linie odvozená od meduloblastomu, což je mozkový nádor odvozený od neuroektodermu, který se typicky objevuje u dětí do 10 let věku) byly narůstány v Petriho miskách pro tkáňové kultury (TPP, Trasadingen, Švýcarsko) při 37°C a 5% CO_2 atmosféře v IMDM mediu (Sigma Aldrich, Buchs, Švýcarsko) s přídavkem 10% fetálního hovězího séra (Gibco, Carlsbad, USA) a 2% 200mM glutaminu. Z narostlých kultur bylo medium odebráno a buňky rozsuspendovány v PBS pufru, přibližně $1 \cdot 10^6$ jich bylo přeneseno do mikrozkumavky a stočeno 3 minuty při 1500 ot/min ve stolní centrifuze. Po odebrání supernatantu byly buňky skladovány buď samotné při -20°C nebo při 4°C po přidání 100µL roztoku RNAlater (Qiagen, Germantown, USA). Pro dlouhodobější uchovávání byly buňky v roztoku RNAlater po 1 dnu při 4°C dále uskladněny při -20°C.

Isolace RNA a reverzní transkripce

Celková RNA byla z buněčných kultur isolována buď pomocí reagencie TRIZOL (Invitrogen, Carlsbad, USA) či kitu RNeasy (Qiagen) v kombinaci s homogenizačními kolonkami QIAshredder (Qiagen). Čistota a množství RNA byla stanovena proměřením absorbčního spektra mezi 220nm a 350nm na přístroji Nanodrop ND-1000 (Thermo Scientific, Rochester, USA), její celistvost pak pomocí kapilární elektroforesy (2100 Bioanalyzer od firmy Agilent, Santa Clara, USA). 1 až 3µg RNA byly převedeny na cDNA pomocí kitu „Superscript III First-strand cDNA synthesis“ (Invitrogen) dle doporučení výrobce a s použitím směsi náhodných hexamerů (5ng/µL v reakci).

Polymerasová řetězová reakce

Použité páry primerů byly navrženy tak, aby specificky amplifikovaly oblast prvních tří konstitutivních exonů cDNA GCPII a zároveň aby byly odděleny alespoň jedním intronem. Případná kontaminace genomovou DNA by pak neměla ovlivnit výsledek PCR reakcí, sekvence všech primerů použitých pro amplifikaci a klonování variant jsou shrnuty v tabulce 3. Po třiceti cyklech PCR amplifikace (počáteční 1' denaturace při 94°C následovaná 30 cykly po 30" 94°C, 1' 60°C až 68°C dle použitých primerů a 1' 72°C; použit termální cyklér Biometra Tgradient od firmy Labrepco, Horsham, USA; koncentrace primerů v reakci byla 1mM, koncentrace dNTPs 200µM) s použitím *Taq* polymerasy (TopBio, Praha, Česko) byla namnožená DNA analyzována pomocí horizontální agarosové elektroforesy (hustota gelu byla 1% až 2% v závislosti na

velikosti analysovaných DNA fragmentů, jako elektrodový pufr byl používán pufr TBE; použita byla aparatura „EasyCast mini gel system“ od firmy Thermo Scientific). Gel byl barven chemikálií GelRed (Biotium, Hayward, USA) a DNA následně zviditelněna pomocí UV lampy. Proužky gelu obsahující DNA byly vyřezány z gelu skalpelem a DNA byla z gelu poté isolována pomocí kitu „QIAquick gel extraction“ (Qiagen). V případě, že výtěžek určitého fragmentu DNA byl příliš malý pro zaklonování do bakteriálního vektoru, byla jeho DNA znovu amplifikována pomocí PCR za použití původních primerů, analyzována na agarosovém gelu a po isolaci z gelu zaklonována.

Klonování cDNA fragmentů

DNA fragmenty byly zaligovány do vektoru pCRII-TOPO a vektoru pcDNA3.1/V5-His TOPO dle instrukcí výrobce (Invitrogen), ligační směsí transformované buňky TOP10 *E. coli* (Invitrogen) pak byly po hodinové inkubaci ve sterilizovaném LB mediu (Sigma Aldrich) natřeny na sterilizovaný LB Agar (Sigma Aldrich) s obsahem ampicilinu 100µg/ml (Biotica, Slovenská Ľupča, Slovensko) a přes noc inkubovány při 37°C. Rezistentní klony pak byly narůstány v 5ml sterilizovaného LB media s obsahem ampicilinu 100µg/ml a vektorová DNA z nich byla poté isolována pomocí kitu „QIAprep spin miniprep“ (Qiagen) s výtěžky okolo 15µg plasmidové DNA. Klony obsahující ligovaný fragment DNA byly vybrány pomocí restrikčního štěpení endonukleasou *EcoRI* (New England Biolabs, Ipswich, USA) v případě vektoru pCRII-TOPO nebo endonukleasami *KpnI* a *XhoI* (New England Biolabs) v případě vektoru pcDNA3.1/V5-His TOPO a následně osekvenovány.

Zpracování dat

Translace mRNA sekvencí na proteiny byly spolu s vzájemným porovnáním aminokyselinových sekvencí těchto proteinů prováděny v programu Vector NTI Suite 10.3 (Invitrogen). Přítomnost transmembránových helixů byla předpovídána programem TMHMM 2.0 [Krogh 2001], signální peptidy pak programem SignalP 3.0 [Bendtsen 2004]. Pro porovnání cDNA sekvencí s genomovou DNA a určení exon/intronové struktury byl použit program Spidey [Wheelan 2001]. Programy TMHMM, SignalP a Spidey jsou volně dostupné na internetu.

Složení pufrů

PBS: vodný roztok 8,0 g/L NaCl; 0,2 g/L KCl; 0,2 g/L KH_2PO_4; 1,15 g/L Na_2HPO_4; pH=7,4

TBE: vodný roztok 10,8 g Tris; 5,5 g/L H_3BO_3; 0,37 g/L Na_2H_2EDTA; pH=8,3

Tabulka 3. Nukleotidové sekvence primerů použitých při amplifikaci klonovaných cDNA fragmentů.

Primery Q157 a Q158 mají shodnou sekvenci se stejně pojmenovanými primery v [Schmittgen 2003].

Forwardní primer	Sekvence (5' - 3')	Nukleotidů
Q158	GCTGTGGTGGAGAAACTG	18
Var6Fmut	GGAGAAACTGGACCCCAGAGGC	22
FRTPSMA1	GCGCCGAGATGTGGAATCTCC	21
FRTPSMA2	GGAGCAAACCTCGGAGTCTTCCC	23
FRTPSMC1	GCAAGCTTGGGAACTGTGTGAACC	24
Reverzní primer	**Sekvence (5' - 3')**	**Nukleotidů**
Q157	GTTCCTGCTAAATGTGGTATCTGTGTA	27
Group4R	CCTGTGTGGTTCTGCTTTTCTTA	23
RRTPSM1b'	AACCACCCTGGTTCACACAGTTCC	24
RRTPSMA2b	CTGAGATCACGCCATTGCACTCC	23
RRTPSMA3	TGGGATTGAATTTGCTTTGCAAGC	24

3.2. Materiály

3.2.1. Seznam chemikálií, enzymů, kitů a jiného materiálu

Biotium, Hayward (USA)

GelRed

Generi Biotech, Hradec Králové (Česko)

oligonukleotidy

Gibco, Carlsbad (USA)

fetální bovinní sérum

Hirschmann Geräte, Eberstadt (Německo)

Pasteurovy pipety

Invitrogen, Carlsbad (USA)

agarosa, buňky TOP10 *E. coli,* kit „Superscript-III First-strand cDNA synthesis“, reagencie TRIZOL, vektor pcDNA3.1/V5-His TOPO, vektor pCRII-TOPO

New England Biolabs, Ipswich (USA)

BSA, EcoRI pufr, NEB1 pufr, restrikční endonukleasy *EcoRI, KpnI, XhoI*

Promega, Madison (USA)

tris

Qiagen, Germantown (USA)

kit „QIAprep spin miniprep“, kit „QIAquick gel extraction“, kit RNeasy, homogenizační kolonky QIAshredder, RNAlater, X-Gal

Sigma Aldrich, Buchs (Švýcarsko)

disodná sůl EDTA, IMDM medium, LB agar, LB medium, sacharosa

Serva, Heidelberg (Německo)

bromfenolová modř

TopBio, Praha (Česko)

dNTPs, PCR reakční pufr, roztok chloridu hořečnatého*, Taq* DNA polymerasa

TPP, Trasadingen (Švýcarsko)

Petriho misky pro tkáňové kultury

USB, Cleveland (USA)

kyselina boritá

ÚMG AV ČR, Praha (Česko)

PBS

3.2.2. Seznam přístrojů

Autokláv MLS-3020U (Sanyo, Japonsko)

Biometra Tgradient thermal cycler (Labrepco, USA)

CO_2 inkubátor MCO-17AI (Sanyo, Japonsko)

Hematocytometr (Sigma-Aldrich, USA)

Horizontální agarosová elektroforesa EasyCast mini gel system (Thermo Scientific, USA)

Inkubátor TCH 100 (Laboratorní přístroje Praha, Česko)

Kapilární elektroforesa 2100 Bioanalyzer (Agilent, USA)

Měřič pH Unicam 9450 (Unicam, UK)

Nanodrop ND-1000 (Thermo Scientific, USA)

Stolní mikrocentrifuga Biofuge pico (Heraeus Instrument, Německo)

Stolní mikrocentrifuga Eppendorf 5414R (Eppendorf, Německo)

UV Transilumitátor (UltraLum, USA)

Zdroj elektrického napětí MP-500V (Major Science, Tchai-wan)

3.3. Podrobnosti k výpočtům

Při kvantově-chemických výpočtech byly používány báze molekulových orbitalů implementované v programu Turbomole [Ahlrichs 1989] pod označením def2-xVP. Báze def2-SVP jsou tzv. „double-zeta“ kvality, tedy každý valenční orbital je reprezentovaný dvěma kontrahovanými gaussiány. V textu jsou označovány jako SV(P) a SVP. Pro výpočty energií na optimalizovaných strukturách byly pak použity báze „triple-zeta“ (označené jako TZVP) a „quadruple-zeta“ kvality (QZVP a QZVPP) [Schäfer 1992, Weigend 1998, Weigend 2003, Hattig 2005, Weigend 2005, Weigend 2006, Hellweg 2007].

Při optimalizacích molekulové geometrie, které byly prováděny pokaždé na úrovni DFT(PBE)/SVP (není-li řečeno výslovně jinak), byla použita konvergenční kriteria 1.10^{-6} atomové jednotky (a. u.) pro změnu energie a 1.10^{-3} a. u. pro normu gradientu energie. V textu je používáno standardní označení úrovně teorie jako M/B (metoda/báze atomových orbitalů).

Pro všechny výpočty s nehybridními funkcionály (PBE a TPSS) byla používána RI-J aproximace („resolution-of-the-identity“ někdy též nazývaná „density-fitting“).

V případech, kdy byla zahrnuta tzv. implicitní solvatace, byl použit model COSMO [Klamt 1993]. Jako solvatační poloměr atomu zinku byl zvolen jeho van der Waalsův poloměr, jako solvatační poloměry ostatních atomů byly použity optimalizované van der Waalsovy poloměry, tak jak jsou implementované v programu Turbomole 6.0 [Schäfer 2000]. Empirická korekce k disperzi v metodě DFT byla zahrnuta dle postupu Grimmeho [Grimme 2006], tak jak je implementována v programu Turbomole 6.0 (včetně použitých parametrů). Pro všechny výpočty (není-li řečeno jinak) uvedené v kapitolách 3.3.2. a 3.3.3. byl použit program Turbomole 6.0 [Ahlrichs 1989].

Všechny obrázky struktur byly v této práci vytvořeny v programu Pymol [DeLano 2002].

3.3.1. Kombinované QM/MM modely lidské GCPII a jejích mutantních forem

Pro všechny QM/MM výpočty byl použit program ComQum verze 9 [Ryde 1996], který používá program Turbomole 5.7 [Ahlrichs 1989] pro QM výpočty a program Amber 8 [Case 2005] se silovým polem parm99 [Wang 2000] pro MM výpočty.

Jako výchozí struktury pro QM/MM optimalizace byly použity výsledné (optimalizované) struktury Michaelisova komplexu (MC) a tetraedrálního intermediátu (INT) GCPII popisující reakční koordinátu hydrolysy *N*-acetyl-α-L-aspartyl-L-glutamátu převzaté z ref. [Klusák 2009]. Východiskem pro modelování těchto struktur byla krystalová struktura GCPII s navázaným inhibitorem GPI-18431 ((S)-2-(4-iodobenzylphosphonomethyl)-pentandiová kyselina) s PDB kódem 2C6C [Mesters 2006], další podrobnosti viz [Klusák 2009].

Bodové náboje na atomech systému 1 popisovaného kvantovou mechanikou (dále jen QM systém) byly pro potřeby MM minimalizace systému 2 vypočteny pomocí metody ESP („electrostatic potential fitting“) [Merz 1992], tedy tak, aby byly v souladu s náboji v silovém poli AMBER.

Postranní řetězce vybraných aminokyselin byly zaměněny tak, že z PDB souboru byly odstraněny přebytečné atomy původního aminokyselinového zbytku, přepsán název tohoto zbytku a chybějící atomy tohoto prvku pak byly doplněny pomocí skriptu *tleap* (Amber). Pro mutaci Glu424Ala byly navíc přidány dvě molekuly vody, polohové vektory atomů v nich obsažených kyslíků odpovídaly polohovým vektorům kyslíků v karboxylové skupině původního Glu424. Před optimalizací metodou QM/MM byla geometrie těchto vod zoptimalizována v modelu QM systému umístěného ve vakuu na úrovni funkcionálu PBE a báze SVP (všechny ostatní atomy QM systému byly fixovány v původní poloze).

Při geometrické optimalizaci obsahoval QM systém (systém 1) následující aminokyseliny, skupiny či atomy (popis se týká původního GCPII divokého typu, definice QM systémů mutovaných forem GCPII je popsán později v textu): substrát (*N*-Ac-Asp-Glu), dva zinečnaté ionty s navázaným hydroxidovým aniontem přemosťujícím oba kovy, guanidiniové skupiny Arg210 a Arg536, imidazolové kruhy His377 a His553, fenoly patřící Tyr552 a Tyr700, acetáty patřící Asp387, Glu424, Glu425 a Asp453, jednu molekulu vody interagující s karboxylovou skupinou Glu425, aminokyselinové zbytky Asn519 a Asp520, včetně peptidové páteře ukončené karbonylovou skupinou Gly518 na N-konci a uhlíkem C_α aminokyseliny Phe521 na C-konci. Celková velikost QM systému byla 175 atomů. Pro substituci Glu424Ala byl místo acetylové skupiny Glu424 v QM systému zahrnut methyl Ala424 a dvě molekuly vody, vyplňující prostor původní karboxylové skupiny. Pro substituci Tyr700Phe jednoduše ubyla hydroxidová skupina tohoto zbytku a pro substituci Asn520Asp byl amidový dusík s dvěma vodíky nahrazen atomem kyslíku. Celkový náboj QM systému byl mínus dva, pro substituci Asn520Asp pak mínus tři. Tento náboj byl kompenzován kladnými bodovými náboji Arg534 a Lys699 v MM systému 2.

Při každém optimalizačním kroku byl vypočten gradient energie pro QM systém a provedena úplná minimalizace MM systému 2, který zahrnoval všechny aminokyselinové zbytky, molekuly vody a ionty do vzdálenosti 6 Å od jakéhokoliv atomu QM systému (viz též kapitola 2.2.2.1). Aby se systém 2 nelišil u struktur MC a INT, byl pro obě struktury systém 2 nadefinován na základě MC. Na výsledných strukturách pak byly přepočítávány energie na úrovni funkcionálů PBE, TPSS a B3LYP s použitím TZVP báze. Pro tyto výpočty byl QM

systém rozšířen ještě o acetamidovou skupinu Asn257, methylaminovou skupinu z Lys699, guanidiniovou skupinu Arg534 a hydroxymethylovou skupinu Ser431 spolu se dvěma molekulami vody interagující s karboxylovou skupinou Asn520 (celkem 217 atomů). Celkový náboj tohoto systému byl nulový (pro mutaci Asn520Asp pak mínus jedna). Oba QM systémy byly nadefinovány naprosto shodně jako ve studii reakčního mechanismu GCPII divokého typu [Klusák 2009], přičemž rozšířený QM systém byl tvořen tak, aby byl nábojově neutrální a aby bylo zahrnuto pokud možno co nejvíce vodíkových vazeb stabilizujících polární skupiny v původním menším QM systému.

Pro přibližný odhad entalpické bariéry reakce byly výsledné struktury QM systému (menší varianty) dále optimalizovány ve vakuu. Krajní atomy uhlíku, pro které byla při tvorbě modelu QM systému přerušena vazba s atomy v MM systému 2, byly fixovány v původní poloze. Pro GCPII divokého typu a pro substituci Asp424Ala byla navíc zahrnuta methylaminová skupina Lys699. Na výsledných strukturách byl vypočítán hessián na úrovni PBE/SVP a pomocí skriptu *freeh* (Turbomole) byly vypočteny energie nulového bodu (*ZPVE*) a entalpie pro T=298,15 K. Použitá konvergenční kriteria vedla k několika imaginárním normálním vibračním módům (typicky menším než 30i cm^{-1}), které byly pro výpočty těchto energií ignorovány. V případě použití tvrdších konvergenčních kriterií by se z těchto vibrací staly nejpravděpodobněji měkké vibrace s frekvencí maximálně do 100 cm^{-1}. Jejich příspěvek k *ZPVE* je zcela minimální, avšak příspěvek k entalpii každého takového módu je přibližně 0,6 $kcal.mol^{-1}$. Tato hodnota byla k energiím jednotlivých struktur připočtena za každou imaginární vibraci. Vzhledem k velkému příspěvku měkkých módů ke Gibbsově volné energii a jeho velké změně s poměrně malou změnou frekvence nelze podobnou korekci pro Gibbsovu energii použít a je nutné se omezit pouze na výpočet entalpie (příspěvek jednotlivých vibračních módů k entalpii je diskutován v kapitole 2.1.4.).

3.3.2. Klastrové modely aktivního místa GCPII

Jako výchozí struktura pro tvorbu klastrových modelů aktivního místa GCPII byla použita krystalová struktura mutantní formy Glu424Ala s navázaným nerozštěpeným substrátem *N*-Ac-Asp-Glu v aktivním místě (PDB kód 3BXM [Klusák 2009]). Z této struktury byly odstraněny dvě molekuly vody vyplňující dutinu po chybějícím postranním řetězci Glu424, postranní řetězec Glu424 byl následně doplněn pomocí skriptu *tleap* (Amber) stejným postupem jako v případě zaměňování postranních řetězců aminokyselin v QM/MM modelech. Dihedrální úhly postranního řetězce Glu424 byly manuálně změněny tak, aby jeho karboxylová skupina interagovala s hydroxidem vázaným mezi zinečnatými ionty a karbonylovým kyslíkem

peptidové páteře Gly518. Glu424 byl primárně modelován protonovaný, paralelně byl vytvořen také jeden alternativní model s deprotonovaným Glu424 (popsán dále v textu).

Z takto upravené struktury GCPII byl připraven klastrový model aktivního místa čítající 297 atomů, tzv. systém L, jeho přesné složení je uvedeno na konci této kapitoly. Tam, kde bylo nutné při tvorbě modelu přerušit vazbu, byl chybějící atom nahrazen atomem vodíku. Celý tento systém byl vytvořen tak, aby jeho van der Waalsův povrch co nejvíce připomínal kouli, čímž jsme se snažili minimalizovat arteficiálnost systému (resp. minimalizovat množství atomů v kontaktu s vakuem). Tam, kde to bylo možné, jsme kromě postranních řetězců zahrnuli též peptidovou páteř, která je spojovala, čímž jsme do modelu vnesli alespoň malou míru sterických požadavků daných skutečným proteinem.

Geometrie tohoto modelu byla optimalizována ve vakuu, přičemž nejprve byly všechny atomy kromě Glu424 ponechány v původní poloze (tj. určené rentgenovou difrakcí). Po dosažení konvergence byly pro další optimalizaci fixovány polohy pouze některých vybraných atomů, konkrétně buď C_α atomů v případě aminokyselinových zbytků zahrnutých i s jejich peptidovou páteří nebo v případě zahrnutí pouze části postranního řetězce uhlíkový atom, jehož vazba ke zbytku aminokyseliny byla přerušena, tyto atomy jsou uvedeny u složení systému L na konci této kapitoly. Z takto získané struktury Michaelisova komplexu byl modelován reakční meziprodukt pyramidalizací dusíku a uhlíku štěpené peptidové vazby a zkracováním vzdálenosti mezi hydroxidovým kyslíkem a karbonylovým uhlíkem štěpené peptidové vazby.

Ze zkonvergovaných struktur Michaelisova komplexu a intermediátu byl opětovným zmenšením modelu vytvořen tzv. systém M, jehož struktury byly opětovně optimalizovány s fixovanými vybranými atomy, způsob jejich výběru byla stejná jakou u systému L. Složení systému M a atomy fixované během optimalizace jsou uvedeny na konci této kapitoly. Ze zkonvergovaných struktur systému M pak byl dalším zmenšením vytvořen nejměnší systém S, jehož složení včetně atomů fixovaných v průběhu optimalizace je rovněž uvedeno na konci kapitoly.

Zkonvergované struktury všech tří systémů byly na závěr optimalizovány ve vakuu bez jakýchkoliv omezení. Pro všechny takto získané struktury byly vypočteny energie na různých úrovních teorie, odhadnut vliv proteinového okolí pomocí modelu COSMO [Klamt 1993]. U systémů S a M byl vypočten příspěvek *ZPVE* a odhad příspěvku partiční funkce k entalpii, postup byl stejný jako u QM/MM modelů mutantních forem GCPII.

V následujících odstavcích je popsáno složení jednotlivých systémů, uhlíkové atomy fixované během geometrické optimalizace jsou uvedeny v závorce u příslušného zbytku.

Systém S

Dva zinečnaté kationty vázané acetáty patřící Asp387 (C_β), Glu425 (C_γ), Asp453 (C_β) a imidazoly patřící His377 (C_γ) a His553 (C_γ). Hydroxidový aniont vázaný oběma zinky, acetát patřící Glu424 (C_γ), substrát *N*-Ac-Asp-Glu zde byl nahrazen modelovým substrátem *N*-Ac-Ala-Ala, protože v tomto systému chybí postranní řetězce aminokyselin proteinu kompenzující záporné náboje karboxylových skupin aspartátu a glutamátu. Celková velikost byla 78 atomů.

Systém M:

Dva zinečnaté kationty vázané propionáty patřící Asp387 (C_α), Glu425 (C_β), Asp453 (C_α) a 4-methylimidazoly patřící His377 (C_β) a His553 (C_β). Hydroxidový aniont vázaný oběma zinky, propionát patřící Glu424 (C_β), substrát *N*-Ac-Asp-Glu. Guanidiny patřící Arg210 (C_δ) a Arg536 (C_δ), fenoly patřící Tyr552 (C_γ) a Tyr700 (C_γ), methylamin patřící Lys699 (C_ε). Fragment obsahující peptidovou páteř počínající na N-konci C_α atomem Gly518 (C_α), pokračující aminokyselinovým zbytkem Asn519 a zakončeného aminokyselinovým zbytkem Asp520 (C_α) bez C-terminálního karbonylu peptidové vazby. Molekula vody interagující s karboxylovým kyslíkem Asp520 a amidovým kyslíkem Asn519. Celková velikost byla 197 atomů.

Systém L:

Dva zinečnaté kationty s navázaným hydroxidovým aniontem, substrát *N*-Ac-Asp-Glu, methylbenzen Phe209 (C_β), guanidin Arg210 (C_δ), peptidový fragment složený z aminokyselinových zbytků Gly256 (C_α) bez N-terminálního peptidového dusíku a Asn257 (C_α) bez C-terminálního peptidového karbonylu. Dále 4-methylimidazol His377 (C_β), indol Trp381 (C_γ), propionát Asp387 (C_α), peptidový fragment začínající Glu424 a zkrácený o N-terminální peptidový dusík, pokračující kompletním Glu425 (C_α), peptidovou páteří s C_β atomem Phe426 (C_α), Gly427 (C_α) a Leu428 (C_α) s vynechaným C-koncovým karbonylem. Propionát Asp453 (C_α), peptidový fragment začínající C_α atomem Gly518 (C_α), pokračující Asn519 (C_α) a zakončený Asp520 (C_α) bez C-terminálního karbonylu, guanidin Arg536 (C_δ), fenol Tyr552 (C_γ), 4-methylimidazol His553 (C_β), methylamin Lys699 (C_ε) a fenol Tyr700 (C_γ). Molekula vody interagující s karboxylovou skupinou Glu425, molekula vody interagující s karboxylovým kyslíkem Asp520 a amidovým kyslíkem Asn519 a molekula vody interagující s α-karboxylovou kyselinou substrátového glutamátu a karbonylem peptidové páteře Glu425. Celková velikost byla 297 atomů.

3.3.3. Minimalistické klastrové modely GCPII a thermolysinu

GCPII

Všechny struktury byly odvozeny od výsledné struktury intermediátu systému S klastrových modelů GCPII (složení tohoto modelu je uvedeno v předchozí kapitole). Krajní atomy byly přitom během všech optimalizací fixovány.

Tranzitní stavy byly hledány sledováním imaginárního vibračního módu odpovídajícího reakční koordinátě. Během optimalizace byl každý čtvrtý krok přepočítán hessián. Dobrý odhad tranzitního stavu 2 se nepodařilo získat jednoduchou manipulací geometrie intermediátu, pro jeho získání byl proto použit algoritmus využívající pro hledání tranzitního stavu lokální hessián, implementovaný v programu Gaussian03 [Frisch 2004]. Vnitřní souřadnice použité pro výpočet hessiánu byly vzdálenosti mezi uhlíkem a dusíkem peptidové vazby a mezi dusíkem peptidové vazby a protonem na karboxylové skupině Glu424. Výsledná struktura byla dále optimalizována v programu Turbomole. Odhady změn entalpie v průběhu reakce byly vypočteny stejným způsobem jako v případě QM/MM metody.

Thermolysin

Jako výchozí struktura posloužila přibližná struktura tranzitního stavu 2 QM/MM modelu reakčního mechanismu z [Blumberger 2007] poskytnutá autory. Z ní byl vytvořen malý model skládající se ze zinečnatého kationtu, molekuly vody, imidazolových kruhů patřících His142 (C_γ) a His146 (C_γ), acetátů Glu143 (C_γ) a Glu166 (C_γ) a substrátu zmenšeného na N-methyl-acetamid. Atomy fixované během optimalizací jsou uvedeny v závorkách.

4. Výsledky

4.1. Alternativní sestřih GCPII mRNA

Isolovaná RNA běžně dosahovala hodnot RIN („RNA integrity number“) mezi 6 až 9, někdy i vyšších, přičemž obecně byla kvalitnější (méně štěpená) RNA získávána pomocí kitu RNeasy (Qiagen) při současném použití homogenizačních kolonek QIAshredder (Qiagen) než pomocí reagencie TRIZOL (Invitrogen). Vynechání homogenizace přes kolonky QIAshredder vedlo přibližně k desetinásobnému snížení výtěžku. Stejné zkušenosti jsme udělali při isolaci RNA z klinických vzorků tkání.

Zkoumaná oblast GCPII cDNA představovala přibližně prvních 8300 nukleotidů genu kódujícího GCPII (označovaného FOLH1). Všechny v literatuře dosud popsané alternativně sestřižené varianty v této oblasti mRNA byly nalezeny i v této studii. Kromě nich bylo objeveno dalších deset dosud nepopsaných variant, které obsahují dohromady šest nových alternativních exonů. Všechny tyto varianty včetně označení jednotlivých exonů používaných dále v textu jsou schematicky znázorněny na obrázku 7, délky jednotlivých exonů a jejich poloha v rámci genu FOLH1 pak v tabulce 4. Již známé varianty byly označovány zažitými jmény (PSM', PSMC, PSMD a PSMA, varianty PSME a PSMF, které se v této oblasti neliší, byly označeny jako jediná varianta PSME/F). Dosud nepopsané varianty byly označeny pořadovým číslem (Var0, Var1, Var4, Var5, Var7, Var10, Var11, Var12, Var13 a Var14). Stejně tak již známé alternativní exony byly označovány zažitými jmény (exon 1b, exon 1c, zkrácený první konstitutivní exon byl nazván podle PSM' varianty exon I'), nové alternativní exony 1b', 1b'' a 1b* byly pojmenovány podle podobnosti s exonem 1b, nový exon následující po druhém konstitutivním exonu byl nazván exon 2b a prodloužené formy prvního konstitutivního exonu byly označeny jednoduše podle počtu přidaných nukleotidů (exon I+68bps a exon I+68bps+70bps).

Části sekvencí variant Var7 (jeden fragment) a Var11 (tři fragmenty) jsou uloženy v databázi EST („expressed sequenced tag“, zaklonovaný a osekvenovaný fragment cDNA; často jsou tyto sekvence generovány při klonování celých knihoven cDNA) na serveru National Center for Biomedical Information (NCBI) [Sayers 2009], nejsou tedy námi popsány poprvé. Tyto sekvence obsahují celý exon 1b* obsažený ve Var7 a velkou část exonu I+68bps obsaženého ve Var11 (chybí pouze část na 3' konci exonu, která leží ještě před začátkem nejdelšího otevřeného čtecího rámce). Žádné další nové varianty jsme v této databázi nenalezli.

Naše varianty PSMC a PSMD se mírně liší sekvencemi alternativních exonů 1c a 1b oproti sekvencím publikovaným v [O'Keefe 1999, Schmittgen 2003], přesto byly pro přehlednost označovány stejnými jmény. Dříve popsaná sekvence PSMD obsahuje oproti námi nalezené

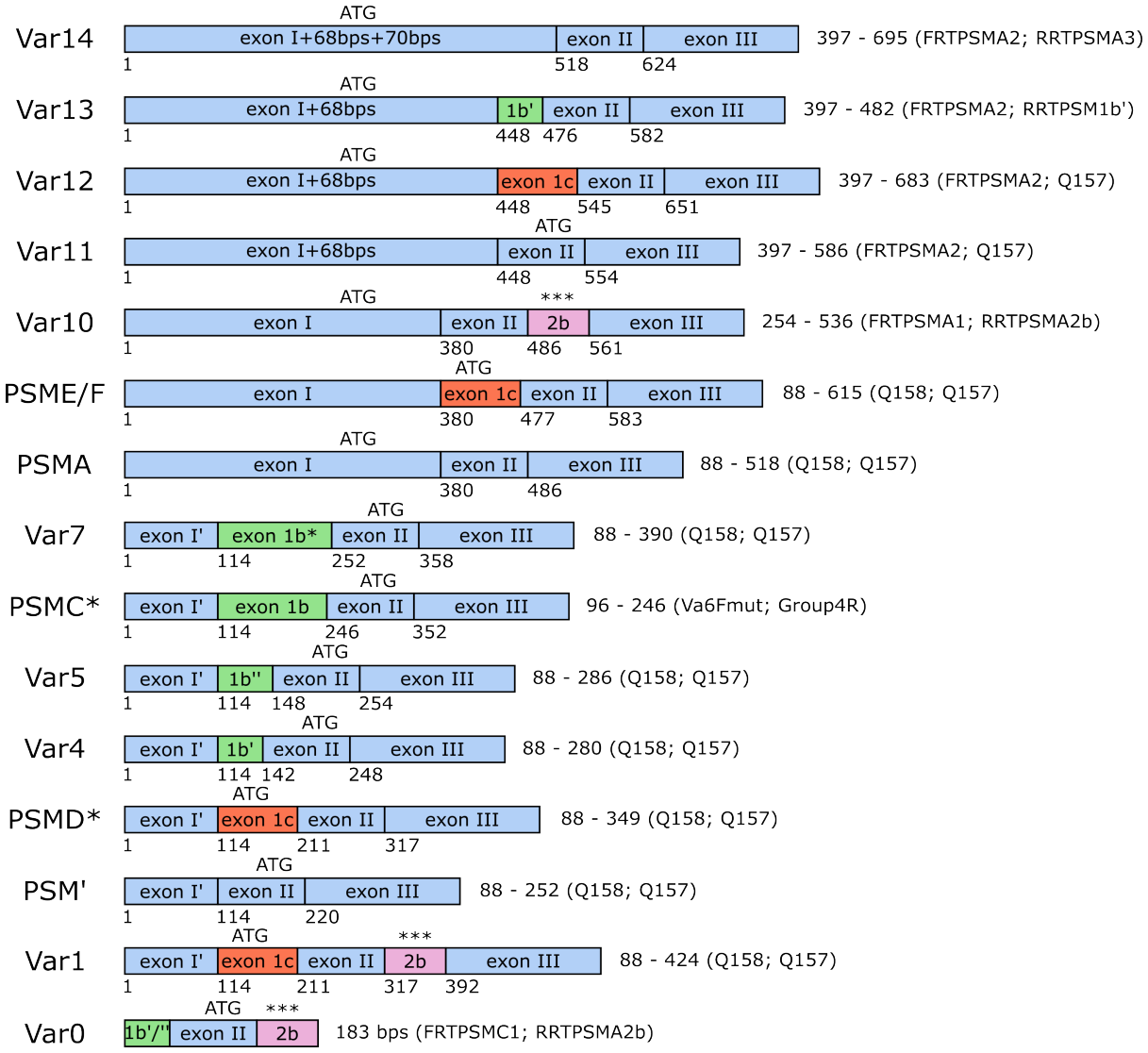
variantě na 5' konci exonu 1c čtyři nukleotidy genomové sekvence navíc. Takto prodloužený exon 1c se nám nepodařilo nalézt, ani když jsme osekvenovali celkem 4 klony varianty PSMD, 2 klony varianty PSME/F a 1 klon varianty Var12 (PSME/F a Var12 také obsahují exon 1c), naše sekvence se navíc shoduje se sekvencí tohoto exonu ve variantě PSME/F, tak jak je popsána v [Cao 2007]. Dále jsme v databasi EST [Sayers 2009] nalezli řadu sekvencí odpovídajících této kratší variantě exonu 1c, avšak ani jeden záznam odpovídající původně popsané formě. Původně popsaná sekvence PSMC pak obsahuje na 5' konci exonu 1b jeden nukleotid genomové sekvence navíc. V databasi EST jsme nalezli pouze jedinou sekvenci obsahující exon 1b, ta odpovídá námi popsané kratší variantě tohoto exonu. Pro tuto variantu svědčí také shodný 5' konec s exonem 1b' (stejně jako má exon 1b* shodný 5' konec s exonem 1b", viz také tabulka 4). Je však nutné podotknout, že variantu PSMC jsme v této podobě zaklonovali až pomocí primeru, který přesahoval z exonu I do exonu 1b a v podstatě jsme tak ani nemohli jiný 5' konec exonu 1b, než byl obsažen v tomto primeru, nalézt. Takto jsme postupovali proto, že v důsledku rozdílu v délce amplikonu variant PSMC a Var7 pouhých šesti nukleotidů není jednoduché tyto dvě varianty rozdělit na agarosovém gelu a proto byly do vektoru ligovány ve směsi a ze čtyř připravených klonů neobsahoval sekvenci PSMC ani jeden z nich. Pro přesné určení sekvence exonu 1b tak bude ještě nutné připravit a osekvenovat další klony. Nakonec budiž zmíněno, že drtivá většina lidských exonů je ohraničena na 5' konci nukleotidy AG. Zatímco námi nalezené sekvence exonů 1b a 1c toto splňují, dříve popsané sekvence nikoliv. Avšak stejně tak je většina exonů ohraničena na 3' konci nukleotidy GT a konkrétně exon 1b je zde ohraničen nukleotidy GC, což nám potvrzuje že nelze považovat sekvenci těchto nukleotidů za absolutní kriterium [Mount 1982].

Varianta PSMA byla zaklonována z knihoven cDNA obou použitých buněčných linií, všechny ostatní varianty kromě PSME/F byly zaklonovány z knihovny cDNA buněčné linie LNCaP, varianta PSME/F byly zaklonovány z knihovny cDNA buněčné linie MHH MED1. Pomocí primerů obsahujících přesahy mezi exony se nám dále podařilo z linie LNCaP zaklonovat dalších osm variant, plynoucích z dalších kombinací zde popsaných alternativních exonů. Je však nutné ještě ověřit specifitu těchto primerů a proto zde nejsou tyto varianty dále popisovány.

Za předpokladu, že by se nalezené varianty vyskytovaly ve formě kompletní mRNA (obsahovaly by všech 19 konstitutivních exonů), pak by většina z nich obsahovala nejdelší otevřený čtecí rámec shodný s variantou PSM' a kódovala by tak protein, který se od GCPII divokého typu (varianta PSMA) liší delecí prvních 57 aminokyselin a v důsledku absence

jakékoliv signální sekvence pro směřování do endoplasmického retikula představuje cytoplasmatický neglykosylovaný a enzymaticky neaktivní protein [Bařinka 2002]. Některé varianty by však kódovaly potenciálně zajímavé proteinové formy lišící se od GCPII divokého typu aminokyselinovými insercemi či substitucemi na N-konci. Porovnání aminokyselinových sekvencí N-konců těchto proteinů je znázorněno na obrázku 8.

Varianty PSMD a PSMF by vedly k proteinu, který by oproti GCPII divokého typu obsahoval místo prvních 39 aminokyselin (cytosolární a transmembránová část) alternativní sekvenci 24 aminokyselin, která podle předpovědi programu SignalP [Bendtsen 2004] obsahuje signální peptid. Tyto varianty tak představují potencionálně sekretovanou formu GCPII. Proteinový produkt PSME varianty by se oproti nim lišil delecí 31 aminokyselin a substitucí 1 aminokyseliny nedaleko C-konce peptidového řetězce. Celá kódující oblast variant PSME a PSMF již byla dříve zaklonována a dokonce byla prokázána tvorba proteinu rozeznávaného protilátkou proti GCPII v buňkách transfekovaných expresním vektorem s insertovanou kódující oblastí PSME varianty [Cao 2007]. Nově objevené varianty Var12, Var13 a Var14 by pak vedly k různě dlouhým insercím mezi transmembránovou a extracelulární část GCPII. Aminokyselinová inserce varianty Var12 by mimo jiné obsahovala stejných 24 aminokyselin jako N-konec proteinových produktů variant PSMD, PSME a PSMF. Jelikož exon 2b, obsažený ve variantách Var1 a Var10 obsahuje ve čtecím rámci stejném jako u GCPII terminační kodon, tak by translace těchto variant vedla ke vzniku krátkých peptidů. Konkrétně Var1 by kódovala 59 aminokyselin, obsahující stejný předpovězený signální peptid jako varianty PSMD, PSME a PSMF, zatímco Var10 peptid dlouhý 74 aminokyselin a skládající se z cytosolární, transmembránové a krátkého úseku extracelulární části GCPII divokého typu.



Obrázek 7. Schematické znázornění exonové struktury nalezených variant mRNA GCPII.

Na každém řádku jsou uvedeny (zleva): název varianty, schematické znázornění exonové struktury této varianty včetně pořadových čísel prvního nukleotidu každého exonu, pořadová čísla nukleotidů ohraničující v této studii zaklonovaný fragment cDNA příslušné varianty (u Var0 pouze délku tohoto fragmentu) a v závorce pár primerů použitý při amplifikaci dané varianty z knihovny cDNA.

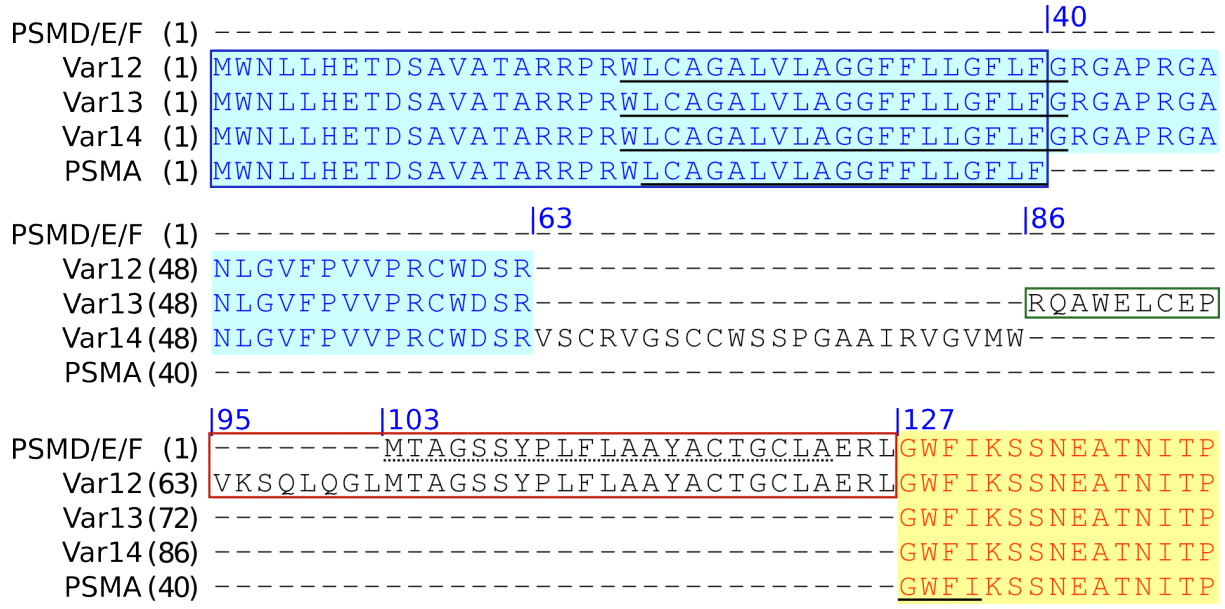
Jednotlivé exony jsou znázorněny jako barevné obdélníky, jejichž délka přesně odpovídá délce daného exonu (s výjimkou exonů 1b' a 1b'', které jsou vůči tomuto měřítku dvojnásobně zvětšené). Znázorněna je pouze část mRNA zkoumaná v této studii (tzn. oblast prvních tří konstitutivních exonů, mRNA GCPII divokého typu obsahuje celkem 19 exonů). Počátek prvního konstitutivního exonu odpovídá počátku mRNA GCPII divokého typu, tak jak je uložena v databasi Genbank (kód NM_004476). Protože amplikon varianty Var0 neobsahuje žádnou část prvního exonu a není tak zřejmé, jak tato varianta začíná, nejsou prvním nukleotidům jednotlivých exonů přiřazena žádná pořadová čísla.

ATG značí iniciační kodon nejdelšího otevřeného čtecího rámce začínajícího v zobrazené oblasti mRNA, *** značí terminační kodon tohoto čtecího rámce (Var0, Var1, Var10).

Tabulka 4. Délky jednotlivých exonů a jejich poloha v genu FOLH1.

Uvedeny jsou pouze exony v námi zkoumané části genu, tzn. v oblasti prvních 8300 nukleotidů.

název exonu	nukleotidů	gen FOLH1
I + 68 + 70	517	1 - 517
I + 68	447	1 - 447
I	379	1 - 379
I'	113	1 - 113
1b*	138	770 - 907
1b	132	776 - 907
1b''	34	770 - 803
1b'	28	776 - 803
1c	97	1797 - 1893
II	106	2499 - 2604
2b	75	3470 - 3544
III	187	8230 - 8416



Obrázek 8. Porovnání aminokyselinových sekvencí N-konců vybraných variant GCPII.

Názvy variant jsou vypsány vlevo, aminokyselinové sekvence na modrém pozadí ohraničené modrým obdélníkem pocházejí z translace exonu I, sekvence pouze na modrém pozadí pak z translace prodloužení exonu I o 68 nukleotidů (I+68bps), nijak nezvýrazněná sekvence z translace dalšího prodloužení tohoto exonu o 70 nukleotidů (exon I+68bps+70bps), sekvence v zeleném obdélníku z translace exonu 1b', sekvence v červeném obdélníku z translace exonu 1c a konečně sekvence červeným písmem na žlutém pozadí z pochází z translace počátku druhého konstitutivního exonu (viz také obrázek 7). U variant Var12, Var13 a Var14 se jedná pouze o hypotetické proteinové produkty, jež byly vytvořeny *in silico* translací sekvencí mRNA.

Sekvence předpokládaných transmembránových helixů jsou podtrženy plnou čarou, předpovězený signální peptid je podtržen tečkovaně. Modrá čísla značí pořadová čísla pozic aminokyselin v rámci tohoto porovnání.

4.2. Modelování záměn aminokyselin v aktivním místě GCPII pomocí QM/MM

Pomocí QM/MM modelu jsem se pokusil popsat několik aminokyselinových záměn v aktivním místě GCPII, čímž jsem chtěl ověřit robustnost a správnost dříve publikovaného modelu [Klusák 2009]. Pro zvolené mutace jsou stejně jako pro GCPII divokého typu dostupná kinetická data štěpení substrátu (*N*-Ac-Asp-Glu). Výsledky teoretických výpočtů lze tedy přímo porovnat s experimentem. Mutace jsem vybíral tak, aby jejich k_{cat} pro štěpení N-Ac-Asp-Glu bylo přinejmenším o řád menší než k_{cat} GCPII divokého typu (řádová změna v reakční rychlosti odpovídá při t=37 °C změně reakční bariéry pouhých 1,4 kcal.mol^{-1}) Dalším kriteriem bylo, aby se zaměňované aminokyseliny lišily v důležité funkční skupině, avšak zároveň byly stericky podobné a aby se nacházely uvnitř QM systému. Takto byly vybrány mutace Asn519Asp (Asn519 interaguje amidovým dusíkem s karbonylovým kyslíkem N-acetylové skupiny substrátu a zároveň s karboxylovou skupinou Asp453), Tyr700Phe (Tyr700 tvoří vodíkovou vazbu s kyslíkem C-terminální karboxylové skupiny substrátu) a zcela neaktivní Glu424Ala, pro který je dostupná krystalová struktura s navázaným substrátem v aktivním místě.

Na základě toho, že nejméně stabilní tranzitní stav 2 leží u QM/MM modelu reakčního mechanismu GCPII divokého typu jen o necelé 3 kcal.mol^{-1} výše než energie tetraedrálního intermediátu, byla aktivační bariéra aproximována relativní energií intermediátu vůči Michaelisově komplexu. Domníváme se, že pro pouhé porovnání změn reakční bariéry u daných mutací oproti GCPII divokého typu (nikoliv tedy jejich absolutních hodnot) se jedná o přijatelnou aproximaci. Jinými slovy předpokládáme, že energie tranzitních stavů bude vyšší oproti intermediátu u všech mutantů o přibližně stejnou hodnotu a neovlivní tak rozdíly mezi jednotlivými formami GCPII. Označení atomů používaných dále v textu odpovídá obrázku 10.

4.2.1. Geometrie a reakční koordináta

U všech tří mutací se podařilo nalézt stabilní intermediát, strukturně odpovídající intermediátu při štěpení *N*-Ac-Asp-Glu GCPII divokého typu. Ve formách GCPII obsahující Glu424 ve své sekvenci (tedy kromě Glu424Ala mutantu) byl na počátku reakce uvažován Glu424 protonovaný na $O_{\varepsilon 2}$ atomu.

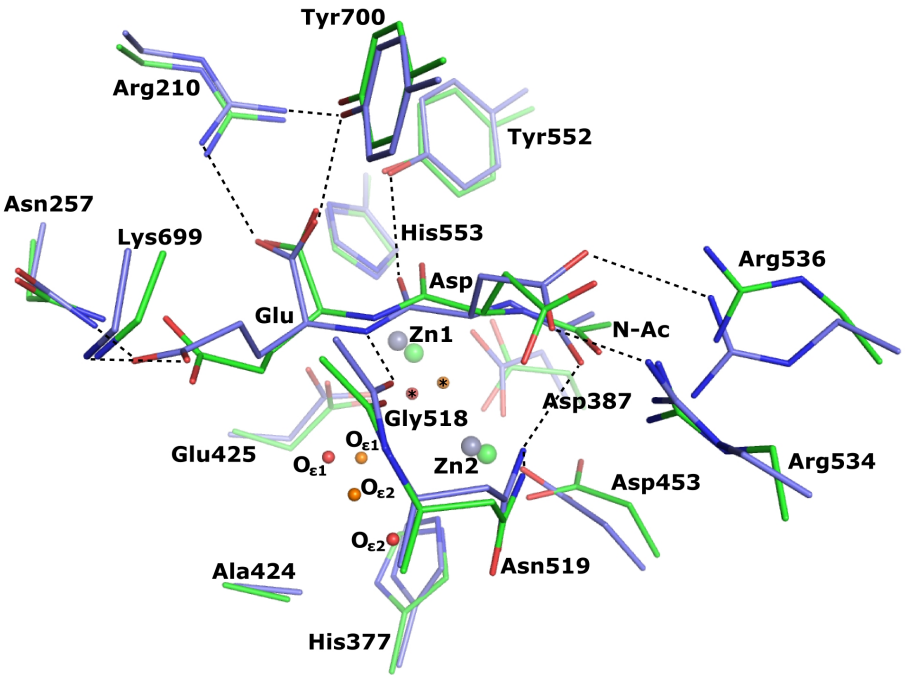
Záměna Glu424Ala nevedla k žádným zásadním změnám geometrie aktivního místa, což je ve shodě s krystalovými strukturami GCPII divokého typu (např. s navázaným analogem tranzitního stavu, PDB kód 3BHX [Bařinka 2008]) a zmíněné mutantní formy (PDB kód 3BXM [Klusák 2009]). Z porovnání QM/MM modelu Glu424Ala formy GCPII s její krystalovou strukturou (viz obrázek 9) je patrné, že QM/MM model velmi dobře reprodukuje experimentální stukturu. Největší rozdíl spočívá v poloze dinukleárního zinkového centra, které je v krystalu v

porovnání s modelem posunuté o přibližně 0,8 Å směrem k dutině vzniklé po chybějícím acetátu Glu424. Poloha obou molekul vody vyplňujících tuto dutinu je u těchto struktur také jiná (rozdíl v polohách jejich kyslíkových atomů činí 1,3 Å-1,4 Å při srovnání krystalové struktury s modelem). Vzdálenost mezi hydroxidovým aniontem a oběma zinečnatými kationty je v QM/MM modelu 2,2 Å, zatímco v krystalové struktuře pouze 2,0 Å. Toto je společné QM/MM modelům všech forem GCPII. Dalším společným jmenovatelem QM/MM struktur je, že nedostatečně popisují vzdálenost mezi karbonylovým kyslíkem štěpené peptidové vazby substrátu a zinečnatým iontem označovaným Zn_1 (2.6 Å v krystalové struktuře, přibližně 3,5 Å v QM/MM modelech). Orientace Glu424 se pak v Michaelisově komplexu QM/MM modelu GCPII divokého typu liší oproti krystalovým strukturám GCPII divokého typu s prázdným aktivním místem (PDB kód 2OOT [Bařinka 20007b]) a s navázaným glutamátem (PDB kód 2C6G [Mesters 2006]). Glu424 těsněji interaguje s karbonylovou skupinou Gly518 (vzdálenost $O_{\varepsilon 2}...O_{G518}$ je 2,9 Å v porovnání s 3,4 Å v krystalových strukturách) a zároveň se odklání od hydroxidu vázaného zinečnatými ionty (vzdálenost $O_{OH}...O_{\varepsilon 2}$ 4,8 Å v porovnání s 3,7 Å v 2OOT struktuře a 3,4 Å v 2C6G) a dusíku štěpené peptidové vazby (vzdálenost $O_{\varepsilon 2}...N_{pept}$ 4,5Å v porovnání s 2,7 Å ve struktuře 2C6G, kde je ale takto malá vzdálenost dána tvorbou silné vodíkové vazby mezi $O_{\varepsilon 2}$ a aminoskupinou vázaného glutamátu). Vybrané meziatomové vzdálenosti v aktivním místě krystalové struktury 3BXM [Klusák 2009], Michaelisova komplexu Glu424Ala formy GCPII a GCPII divokého typu jsou shrnuty v tabulce 5

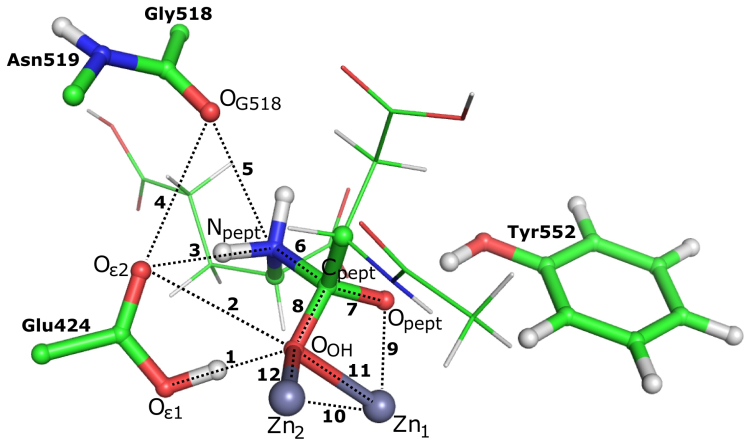
Jediná výraznější změna geometrie u mutantu Asn519Asp oproti QM/MM modelu GCPII divokého typu spočívala v rotaci karboxylové skupiny Asp519 směrem od karboxylové skupiny substrátového aspartátu (se kterou interagoval původní amidový dusík) k amidový dusíku Asn451. Mutací zavedený karboxylový atom kyslíku se posunul o 1,4 Å zatímco druhý (původní) pouze o 0,1 Å. Záměna Tyr700Phe nevedla prakticky k žádným strukturním změnám, nejvýraznější rozdíl spočíval v malé rotaci fenylových skupin Phe509 a Phe700 (maximální vzdálenost mezi stejnými atomy těchto postranních řetězců GCPII divokého typu a mutantní formy byla 0,4 Å).

V optimalizovaných strukturách Michaelisova komplexu a intermediátu mutantní formy Glu424Ala tvoří dvě molekuly vody nalézající se v aktivním místě v dutině po postranním řetězci Glu424 síť vodíkových vazeb, která se podobá interakci protonovaného karboxylu Glu424 se substrátem a proteinovým okolím v GCPII divokého typu. Ve struktuře Michaelisova komplexu slouží voda označená $O_{\varepsilon 2}$ jako akceptor vodíkové vazby můstkového hydroxidu a zároveň jako donor vodíkové vazby pro kyslík molekuly vody $O_{\varepsilon 1}$, která dále tvoří vodíkovou

vazbu s karbonylovým kyslíkem peptidové páteře Gly518. Tento pak tvoří další vodíkovou vazbu s -NH skupinou štěpené peptidové vazby substrátu. Po nukleofilní adici můstkového hydroxidu na uhlíkový atom peptidové vazby substrátu C_{pept} vzniká stabilní tetrahedrální intermediát, který je spojen se změnou hybridizace z sp^2 na sp^3 nejen u C_{pept} ale i u N_{pept}. Tento pyramidální dusík je stabilizován vodíkovou vazbou vody $O_{\varepsilon 1}$, obdobně jako v GCPII divokého typu protonovanou karboxylovou skupinou Glu424. Délka peptidové vazby se z původních 1,38 Å v Michaelisově komplexu prodlouží na 1,55 Å, což ukazuje na ztrátu částečně dvojného charakteru této vazby (detail aktivního místa Michaelisova komplexu a intermediátu je znázorněn na obrázku 11).



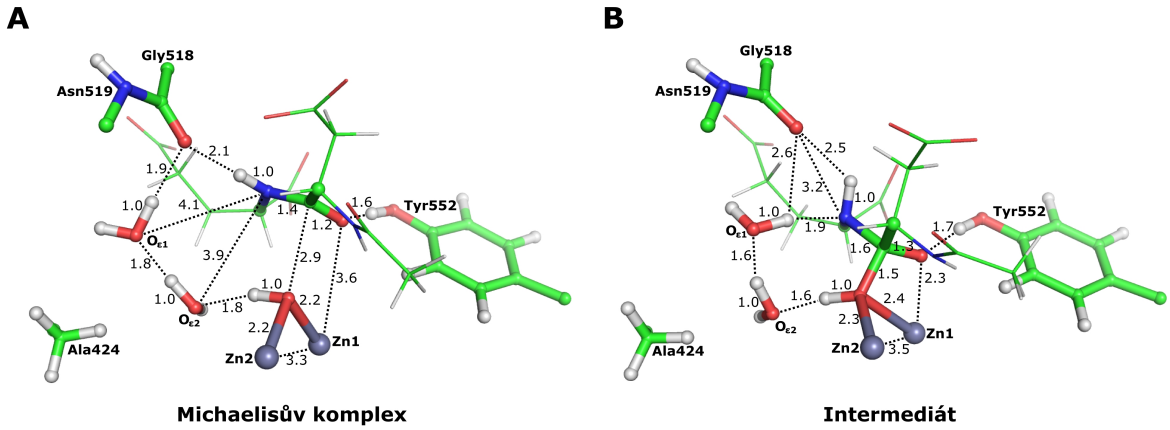
Obrázek 9. Porovnání aktivního místa krystalové struktury Glu424Ala GCPII a jejího QM/MM modelu.
Atomy uhlíku a zinečnaté ionty krystalové struktury Glu424Ala mutantu GCPII jsou znázorněny modře (PDB kód 3BXM [Klusák 2009]), molekuly vody červenou kuličkou a můstkový hydroxid vázaný na oba zinečnaté ionty červenou kuličkou s hvězdičkou. Atomy uhlíku a zinečnaté ionty QM/MM modelu jsou znázorněny zeleně, molekuly vody pak oranžovými kuličkami, můstkový hydroxid oranžovou kuličkou s hvězdičkou. Ostatní atomy jsou u obou molekul znázorněny shodnou barvou (červeně atomy kyslíku, modře atomy dusíku).
Znázorněny jsou aminokyseliny koordinující zinečnaté ionty (His377 a Asp453 koordinující Zn2, Glu425 a His553 koordinující Zn1 a Asp453 přemosťující oba kovy), postranní řetězec Ala424, substrát a postranní řetězce aminokyselin podílejících se na jeho vazbě. Důležité interakce jsou zvýrazněny tečkovanou čarou. Molekuly vody jsou pro rozlišení označeny jako $O_{\varepsilon 1}$ a $O_{\varepsilon 2}$.



Obrázek 10. Přehled porovnávaných vzdáleností mezi jednotlivými strukturami a modely aktivního místa. Označení atomů je stejné jako v textu a v tabulkách 5 a 11. Porovnávané vzdálenosti jsou zvýrazněny tečkovanou čarou a očíslovány. Toto číslování odpovídá číslování v tabulkách 5 a 11. Zinečnaté ionty jsou znázorněny šedými kuličkami. Tyčinkovým a kuličkovým modelem jsou pak znázorněny zbytky Glu424, Gly518, Asn519, Tyr552. Ze substrátu *N*-Ac-Asp-Glu je tyčinkovým a kuličkovým modelem znázorněna pouze štěpené peptidová vazba, zbytek je znázorněn pouze tenkými čárami.

Tabulka 5. Porovnání vybraných meziatomových vzdáleností v krystalové struktuře a QM/MM modelu. Vzdálenosti jsou udávány v ångströmech, první sloupec popisuje krystalovou strukturu mutantní formy GCPII Glu424Ala s PDB kódem 3BXM [Klusák 2009], další dva sloupce struktury Michaelisova komplexu QM/MM modelů stejného mutantu Glu424Ala a GCPII divokého typu. Značení atomů, stejně jako očíslování jednotlivých vzdáleností v levé části tabulky odpovídá značení na obrázku 10. Kurzívou jsou uvedeny vzdálenosti týkající se dvou molekul vody (viz obrázek 9 a 11) nahrazující Glu424 v Glu424Ala mutantní formě GCPII, tučně je pak uvedena vzdálenost, kterou QM/MM model popisuje odlišně.

		٣BXM	Glu٤٢٤Ala	GCPII div. typu
١	O_{OH}-$O_{\varepsilon 1}$	٢,٩	٢,٧٢	٣,٢٣
٢	O_{OH}-$O_{\varepsilon 2}$	٤,٤	٤,٥٠	٤,٨٠
٣	$O_{\varepsilon 2}$-N_{pept}	٤,٨	٤,١٠	٤,٤٦
٤	$O_{\varepsilon 2}$-O_{G518}	٤,٢	٢,٨٩	٢,٩٢
٥	O_{G518}-N_{pept}	٢,٩	٣,١١	٣,٢١
٦	**CN vazba**	١,٣	١,٣٨	١,٣٧
٧	**CO vazba**	١,٢	١,٢٤	١,٢٤
٨	**OH-C_{pept}**	٢,٧	٢,٨٧	٢,٨٣
٩	**Zn_1-O_{pept}**	**٢,٦**	**٣,٦٠**	**٣,٤٩**
١٠	**Zn_1-Zn_2**	٣,٣	٣,٣٣	٣,٣٦
١١	**Zn_1-O_{OH}**	٢,٠	٢,٢٠	٢,٢٠
١٢	**Zn_2-O_{OH}**	٢,٠	٢,١٩	٢,١٨



Obrázek 11. Struktury Michaelisova komplexu a intermediátu QM/MM modelu mutantní formy Glu424Ala GCPII.

Zinečnaté kationty jsou znázorněny šedými kuličkami. Aminokyselinové zbytky Ala424, Gly518, Asn519, Tyr552, dvě molekuly vody označené $O_{\varepsilon 1}$ a $O_{\varepsilon 2}$ a hydroxidový aniont vázaný zinečnatými kationty jsou znázorněny tyčinkovým a kuličkovým modelem. Ze substrátu *N*-Ac-Asp-Glu je tyčinkovým a kuličkovým modelem znázorněna pouze štěpená peptidová vazba, zbytek je znázorněn tenkými čárami. Vzdálenosti jsou uváděny v ångströmech.

4.2.2. Reakční bariéry

Na rozšířeném QM systému byly s pomocí různých funkcionálů vypočteny relativní energie intermediátu vůči Michaelisově komplexu, kterými je aproximována celková reakční bariéra. Byl také vypočten přibližný příspěvek energie nulového bodu (*ΔZPVE*) a příspěvek partiční funkce spolu s *ΔZPVE* k entalpii (*ΔH*). Výsledná entalpická bariéra pak byla součtem relativních DFT energií (QM energie), relativních MM energií, *ΔH* a korekce k *ΔH*, která aproximovala entalpický příspěvek několika malých imaginárních vibrací, které se objevily v důsledku příliš měkkých konvergenčních kriterií a které byly v původním výpočtu *ΔH* ignorovány. Jako nejlepší aproximace bariéry byly brány hodnoty získané pomocí hybridního B3LYP funkcionálu, protože se jedná o zřejmě nejpřesnější funkcionál ze třech použitých. Přibližná entalpická bariéra u GCPII divokého typu pak byla rovna 25,3 kcal.mol^{-1}, Tyr700Phe formy GCPII 27,7 kcal.mol^{-1}, Asn519Asp formy GCPII 27,8 kcal.mol^{-1} a Glu424Ala formy GCPII pak 28,8 kcal.mol^{-1}. Použití funkcionálů PBE a TPSS poskytlo u forem Tyr700Phe a Asn519Asp bariéru o přibližně 5 kcal.mol^{-1} nižší, u Glu424Ala formy o přibližně 4 kcal.mol^{-1}, zatímco u GCPII divokého typu o více než 7 kcal.mol^{-1}. Výsledné hodnoty reakčních bariér získané různými funkcionály jsou shrnuty v tabulce 6.

Pro přibližný odhad vlivu okolí byly s použitím funkcionálu PBE vypočteny energie rozšířeného QM systému v poli bodových nábojů stejném jako v původním QM/MM modelu (pro tento model byl vypočten také vliv empirické disperze), ve vakuu a v solvatačním modelu

COSMO (s dielektrickou konstantou zvolenou buď čtyři, což je obvyklá hodnota používaná pro prostředí proteinu, nebo nekonečno). Tyto energie jsou shrnuty v tabulce 7, z níž je zřejmé, že rozptyl energií v různých modelech okolí byl u jednotlivých forem mezi 2-3 kcal.mol^{-1}. Zajímavý je vliv empirické disperze, která stabilizuje v případě GCPII divokého typu intermediát na úkor Michaelisova komplexu o 13,4 kcal.mol^{-1}, stabilizace o více jak 10 kcal.mol^{-1} nastává i u mutantních forem GCPII.

Tabulka 6. Přibližná hodnota reakční bariéry štěpení *N*-Ac-Asp-Glu enzymem GCPII a jeho mutantními formami.

Levý sloupec udává výpočetní metodu a jednotlivé příspěvky k celkové energii, hodnoty reakčních bariér jsou udávány v kcal.mol^{-1}. Tyto reakční bariéry jsou aproximovány relativní energií intermediátu vůči Michaelisově komplexu. WT značí GCPII divokého typu, E424A mutantní formu Glu424Ala, N519D mutantní formu Asn519Asp a Y700F mutantní formu Tyr700Phe. *ΔH* značí příspěvek partiční funkce k entalpii (se započtením *ZPVE*), korekce k této hodnotě viz kapitola 3.3.1. Celková energie odpovídá vypočtené entalpické bariéře reakce, celkové energie vycházející z funkcionálu B3LYP jsou pak zvýrazněny tučným fontem.

† geometrie a vypočtená QM/MM energie GCPII divokého typu převzata z [Klusák 2009].

B3LYP/TZVP	**WT**†	**E424A**	**N519D**	**Y700F**
QM energie	28.3	27.9	26.5	22.8
MM energie	-2.23	1.28	2.35	5.64
ΔZPVE	1.7	1.2	0.5	0.5
ΔH	1.0	2.0	0.2	-0.2
QM/MM energie	26.1†	29.2	28.9	28.5
QM/MM energie + ZPVE	27.8	30.4	29.3	28.9
QM/MM energie + ΔH	27.1	31.2	29.0	28.3
oprava k ΔH	-1.8	-2.4	-1.2	-0.6
celková energie	**25.3**	**28.8**	**27.8**	**27.7**

TPSS/TZVP	**WT**	**E424A**	**N519D**	**Y700F**
QM energie	21.2	24.1	21.5	17.8
QM/MM energie	19.0	25.3	23.9	23.4
QM/MM energie + ZPVE	20.6	26.5	24.3	23.9
QM/MM energie + ΔH	20.0	27.4	24.0	23.2
celková energie	18.2	25.0	22.8	22.6

PBE/TZVP	**WT**	**E424A**	**N519D**	**Y700F**
QM energie	20.2	23.8	20.8	17.7
QM/MM energie	18.0	25.1	23.2	23.3
QM/MM energie + ZPVE	19.7	26.3	23.6	23.8
QM/MM energie + ΔH	19.0	27.1	23.3	23.2
celková energie	17.2	24.7	22.1	22.6

Tabulka 7. Přibližná hodnota reakční bariéry štěpení *N*-Ac-Asp-Glu enzymem GCPII a jeho mutantními formami v různých modelech okolí včetně vlivu disperze.

Hodnoty bariér uváděné v kcal.mol^{-1} jsou zde aproximovány relativní energií intermediátu vůči Michaelisově komplexu (diskuse viz text) QM systému umístěného v různých modelech okolí.

Levý sloupec udává model okolí QM systému, přičemž všechny hodnoty byly vypočteny na úrovni teorie PBE/TZVP. Pole bodových nábojů odpovídá původnímu QM/MM modelu, pro tento model byl vypočten také vliv empirické disperze.

† geometrie QM systému GCPII divokého typu převzata z [Klusák 2009].

PBE/TZVP	WT†	E424A	N519D	Y700F
bodové náboje	20.2	23.8	20.8	17.7
bodové náboje, disperze	6.8	13.1	10.1	6.1
vakuum	22.1	26.0	20.9	20.1
COSMO $\varepsilon = 4$	21.0	23.9	19.0	19.5
COSMO $\varepsilon = \infty$	20.4	22.9	17.9	19.2

4.3. Modelování reakčního mechanismu pomocí klastrových modelů

Pro modelování reakčního mechanismu byl vytvořen alternativní klastrový model aktivního místa GCPII umístěný ve vakuu. Pro posouzení spolehlivosti tohoto modelu v závislosti na velikosti systému, byly vytvořeny tři různě velké modely, systém S obsahoval 78 atomů, systém M 197 atomů a systém L 297 atomů. Nejmenší systém zahrnoval kromě dvou zinečnatých iontů a na ně vázaného hydroxidu pouze první koordinační sféru zinečnatých iontů, zmenšený model substrátu (*N*-Ac-Ala-Ala) a katalytický zbytek Glu424. Ve středním systému byl zahrnut celý substrát *N*-Ac-Asp-Glu a většina postranních řetězců s ním interagujících. Do největšího systému bylo zahrnuto několik dalších postranních řetězců interagujících s aminokyselinami zahrnutými ve středním systému a také části peptidové páteře, tak aby se zamezilo arteficiálnímu pohybu některých postranních řetězců během geometrické optimalizace.

Počáteční optimalizace probíhaly s fixovanými krajními atomy, takto byla přibližně zahrnuta sterická omezení daná zbytkem proteinu (který zde není explicitně uvažován) na pohybu těchto atomů. Pro posouzení sterického pnutí, vzniklého tímto postupem byly následně všechny struktury přeoptimalizovány bez jakýchkoliv omezení. Za nejsprávnější model byl považován největší systém optimalizovaný s fixovanými krajními atomy, ostatní systémy s ním byly porovnávány.

Při modelování reakční bariéry bylo stejně jako u QM/MM modelu předpokládáno, že reakční meziprodukt leží energeticky velmi blízko tranzitním stavům a proto byla celková reakční bariéra aproximována relativní energií intermediátu oproti Michaelisovu komplexu. Modelovány tak byly pouze Michaelisův komplex a intermediát, pouze pro nejmenší systém S

byly vyhledány i oba tranzitní stavy (popsáno v následující kapitole). Atom kyslíku $O_{\varepsilon 2}$ na postranním řetězci Glu424 byl opět uvažován protonovaný, pro největší systém L byla posléze namodelována přibližná reakční kooordináta bez protonu na Glu424 (kapitola alternativní reakční mechanismus). Označení atomů používaných zde a dále v textu odpovídá obrázku 10 v předchozí kapitole.

4.3.1. Geometrie

Struktury optimalizované s fixovanými krajními atomy

Optimalizovaná struktura Michaelisova komplexu systému L velmi dobře reprodukuje krystalovou strukturu GCPII (PDB kód 3BXM), ze které byl tento model vytvořen, největší rozdíl spočíval v rotaci postranního řetězce Asn257. Přeložení obou struktur je pak na obrázku 12. Oproti QM/MM modelu je velký rozdíl v poloze Glu424, který se přiklání k hydroxidu koordinovanému zinečnatými ionty a vytváří s ním dvě vodíkové vazby, čímž se poněkud vzdaluje od karbonylového kyslíku Gly518. Toto je charakterizováno vzdáleností $O_{OH}...O_{\varepsilon 2}$ pouhých 2,7 Å (oproti 3,7 Å v krystalové struktuře s prázdným aktivním místem 2OOT [Bařinka 2007] a 4,8 Å v QM/MM modelu) a kratší vzdáleností $O_{\varepsilon 2}...N_{pept}$ v porovnání s QM/MM (3,3 Å oproti 4,5 Å, avšak 2,7 Å v krystalové struktuře s navázaným glutamátem 2C6G [Mesters 2006]). Naopak vzdálenost $O_{\varepsilon 2}...O_{G518}$ se oproti QM/MM prodloužila z 2,9 Å na 3,7 Å (3,4 Å ve strukturách 2OOT a 2C6G). Vybrané meziatomové vzdálenosti tohoto modelu, modelu alternativního reakčního mechanismu s deprotonovaným Glu424 v Michaelisově komplexu, QM/MM modelu reakčního mechanismu a krystalové struktury neaktivní formy GCPII s navázaným substrátem v aktivním místě (PDB kód 3BXM) jsou uvedeny v tabulce 11 v kapitole 4.3.4.

Optimalizovaná struktura systému M se v geometrii aktivního centra od systému L lišila jen velmi málo, avšak geometrie nejmenšího systému S se lišila velmi výrazně v poloze substrátu, což lze přičíst především tomu, že zde chybí jakékoliv interakce proteinového okolí se substrátem (kromě dizinkového centra a Glu424). Rozdíly mezi jednotlivými strukturami nejlépe ilustrují vzdálenosti Glu424 $O_{\varepsilon 2}...N_{pept}$ (3,3 Å v systému L; 3,5 Å v systému M a 4,6 Å v systému S) a $O_{OH}...C_{pept}$ (2,9 Å v systému L; 2,9 Å v systému M a 3,6 Å v systému S).

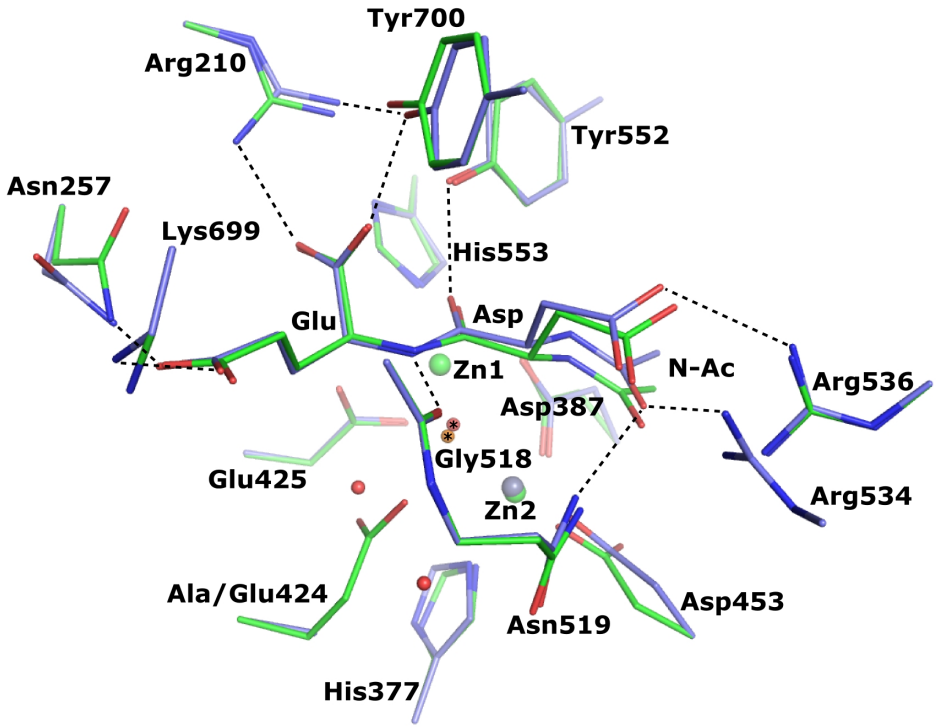
Struktury optimalizované bez fixování krajních atomů

Uvolnění krajních atomů při geometrické optimalizaci vedlo u všech systémů k výrazným strukturním změnám, které jsou u Michaelisova komplexu opět nejlépe ilustrovány vzdálenostmi Glu424 $O_{\varepsilon 2}...N_{pept}$ (3,4 Å v systému L; 3,3 Å v systému M a 4,2 Å v systému S) a $O_{OH}...C_{pept}$ (3,1 Å v systému L; 3,6 Å v systému M a 4,1 Å v systému S). Geometrie obou

zinečnatých iontů s navázaným hydroxidem zůstala v Michaelisově komplexu u všech tří systémů stejná, zatímco v intermediátu se zmenšujícím se systémem postupně z můstkového hydroxidu stával terminální, vázaný pouze atomem Zn_2. Je vidět jasná tendence zmenšování rozdílů mezi strukturami optimalizovanými oběma způsoby se zvyšujícím se počtem atomů modelu. U největšího systému L pak jediný výraznější rozdíl spočívá v odklonu Gly518, který leží na okraji tohoto systému, z aktivního místa (o 0,6 Å v případě vzdálenosti $O_{\varepsilon 2}...O_{G518}$).

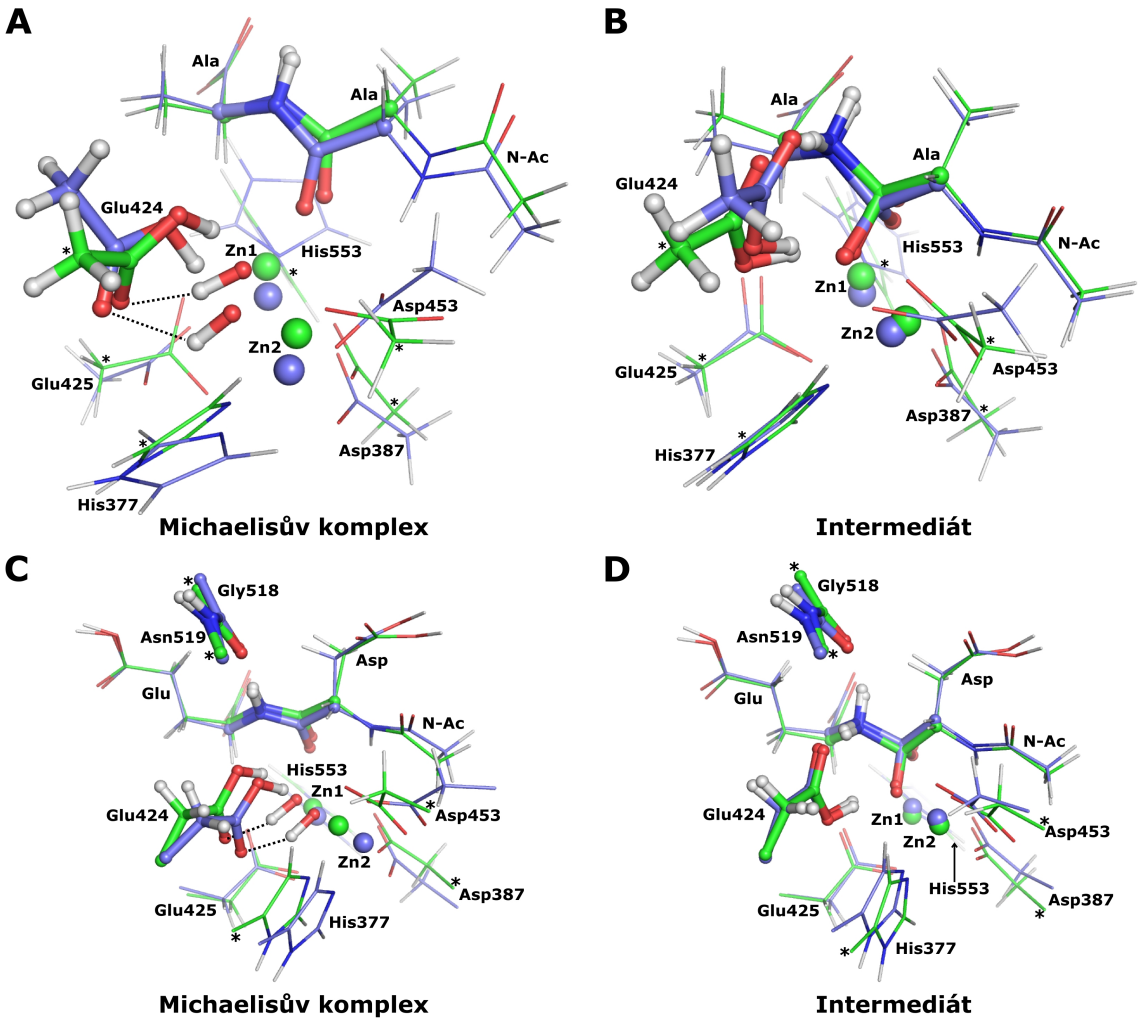
Porovnáme-li změny v polohách atomů, které byly v původní optimalizaci fixovány a následně uvolněny, vyjdou nám r.m.s.d. hodnoty 1,67 Å (Michaelisův komplex, MC) a 1,00 Å (intermediát, INT) pro systém S; 2,16 Å (MC) a 2.28 Å (INT) pro systém M; 1,40 Å (MC) a 1,12 Å (INT) pro systém L. Tyto hodnoty naznačují, že na okrajích systému L dochází k menším změnám než u systému M. Malé hodnoty r.m.s.d. u systému S lze vysvětlit tím, že se jednalo pouze o C_γ atom Glu424 a uhlíkové atomy postranních řetězců aminokyselin koordinujících zinečnaté ionty, které v důsledku interakce s kovy nemají přílišnou volnost pohybu, výrazný posun substrátu se v těchto hodnotách nepromítne, neboť u něj nebyly fixovány žádné atomy. Porovnání aktivních center systému S a systému L získaných omezenou či volnou optimalizací je znázorněno na obrázku 13.

Karboxylové skupiny postranních řetězců aspartátu a glutamátu substrátu byly ve výsledných strukturách protonovány (proton na ně přešel z guanidiniové skupiny Arg536 a aminoskupiny Lys699), což lze přičíst jejich poloze u okraje klastrového modelu, kde je původně polární okolí proteinu nahrazeno z velké části vakuem, ve kterém jsou separované náboje velmi nestabilní.



Obrázek 12. Porovnání krystalové struktury Glu424Ala GCPII a klastrového modelu GCPII divokého typu (systém L optimalizovaný s fixovanými krajními atomy).

Atomy uhlíku a zinku krystalové struktury Glu424Ala formy GCPII (PDB kód 3BXM [Klusák 2009]) jsou znázorněny modře, molekuly vody červenou kuličkou, můstkový hydroxid pak červenou kuličkou s hvězdičkou. Atomy uhlíku a zinku klastrového modelu systému L (s fixovanými krajními atomy) jsou znázorněny zeleně, můstkový hydroxid pak oranžovou kuličkou s hvězdičkou, Arg534 nebyl v tomto modelu zahrnut. Důležité interakce proteinu se substrátem jsou zvýrazněny tečkovaně, His377 a Asp453 koordinují zinečnatý iont Zn2, Glu425 a His553 zinečnatý iont Zn1 a Asp387 oba tyto ionty.



Obrázek 13. Porovnání geometrií aktivního místa systémů S a L optimalizovaných s fixovanými krajními atomy a zcela uvolněně (bez fixování atomů).

Zelenou barvou jsou znázorněny struktury optimalizované s fixovanými atomy, modrou barvou struktury optimalizované bez jakýchkoliv omezení. Ze substrátu je tyčinkovým a kuličkovým modelem znázorněna pouze štěpená peptidová vazba, zbytek je znázorněn pouze tenkými čárami, stejně jako postranní řetězce koordinující oba zinečnaté ionty (His377 a Asp453 koordinující Zn2, Glu425 a His553 koordinující Zn1 a Asp387 přemosťující oba kovy), které jsou znázorněny kuličkami. Atomy uhlíků, které byly během omezené optimalizace fixovány, jsou označeny hvězdičkou. Interakce mezi hydroxidem přemosťujícím zinečnaté ionty a karboxylovou skupinou Glu424 jsou znázorněny tečkovaně.

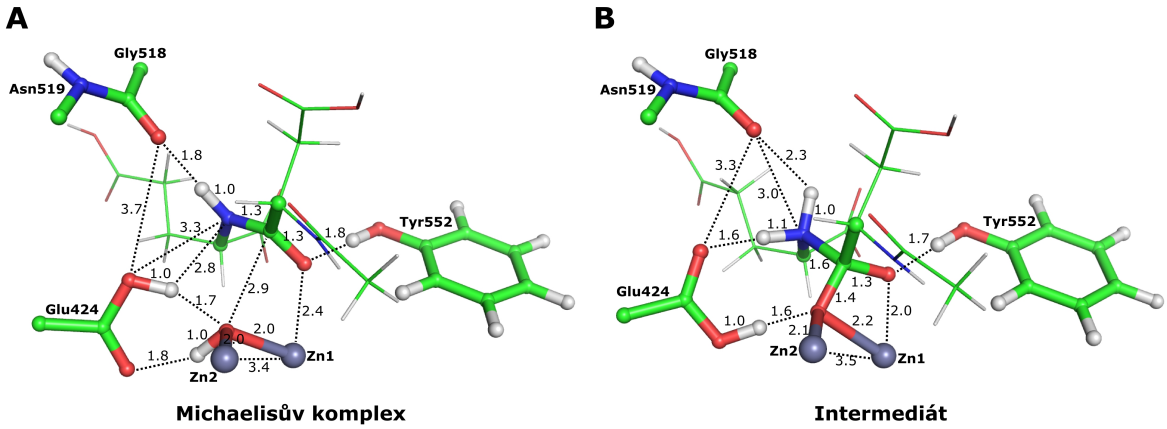
(A) U Michaelisova komplexu systému S byla poloha substrátu výrazně odlišná od větších systémů M a L již po optimalizaci s fixovanými krajními atomy. Po jejich uvolnění lze pozorovat velké geometrické změny u postranních řetězců Asp453 a His553. (B) Rozdíly mezi strukturami intermediátu systému S optimalizovanými oběma způsoby byly podstatně menší než v případě Michaelisova komplexu. (C) Rozdíly u Michaelisova komplexu systému L spočívaly především v posunu His377 a Asp453 a jimi koordinovaného zinečnatého iontu Zn2. (D) Intermediát systému L.

U obou struktur systémů L je zřetelné, že geometrie těch aminokyselin, které byly spojené peptidovou páteří (Glu424 a Glu425, Gly518 a Asn519 spojené dále s Asp520, který není znázorněn) se po uvolnění fixovaných atomů téměř nezměnila, na rozdíl od aminokyselin reprezentovaných pouze jejich postranními řetězci.

4.3.2. Reakční koordináta

Zde popisované struktury Michaelisova komplexu a intermediátu systému L vznikly optimalizací s fixovanými krajními atomy, značení atomů v textu odpovídá obrázku 10 v kapitole 4.2.

Ve struktuře Michaelisova komplexu tvoří -NH skupina štěpené peptidové vazby substrátu vodíkovou vazbu s karbonylem Gly518, zatímco Glu424 tvoří dvě vodíkové vazby s hydroxidovým aniontem, jehož kyslík se nachází ve vzdálenosti 2,9 Å od atomu uhlíku C_{pept}. Ke vzniku stabilního intermediátu vede nukleofilní útok hydroxidového kyslíku na zmíněný uhlíkový atom C_{pept}, za současného odklonění atomu kyslíku $O_{\varepsilon 2}$ patřící Glu424 a přenesení protonu nacházejícího se na tomto atomu na atom dusíku N_{pept}, který je tak stabilizován v novém pyramidálním uspořádání. Proton, který byl v Michaelisově komplexu součástí hydroxidového aniontu se přenese na atom $O_{\varepsilon 1}$ residua Glu424. Vzniklý stabilní intermediát je charakterizován sp^3 hybridizací na C_{pept} i N_{pept} a prodloužením peptidové vazby z 1,33 Å v Michaelisově komplexu na 1,59 Å v důsledku ztráty částečně dvojitého charakteru. Detaily aktivního místa obou struktur jsou spolu s vybranými meziatomovými vzdálenostmi znázorněny na obrázku 14.



Obrázek 14. Struktury Michaelisova komplexu a intermediátu klastrového modelu GCPII (systém L s fixovanými krajními atomy).

Zinečnaté kationty jsou znázorněny šedými kuličkami. Aminokyselinové zbytky Glu424, Gly518, Asn519, Tyr552 a hydroxidový aniont vázaný zinečnatými kationty jsou znázorněny tyčinkovým a kuličkovým modelem. Ze substrátu *N*-Ac-Asp-Glu je tyčinkovým a kuličkovým modelem znázorněna pouze štěpená peptidová vazba, zbytek je znázorněn tenkými čárami. Vzdálenosti jsou uváděny v ångströmech.

4.3.3. Reakční bariéra

Na výsledných strukturách všech zmíněných modelů byly vypočteny relativní energie intermediátu vůči Michaelisovu komplexu na úrovni PBE/TZVP, TPSS/TZVP a B3LYP/TZVP, kterými byla aproximována reakční bariéra. Hodnota entalpické bariéry byla získáná započtením

ZPVE a entalpického příspěvku partiční funkce k B3LYP/TZVP energii. Jelikož se obdobně jako u QM/MM modelu objevilo několik měkkých imaginárních vibračních módů (především u struktur optimalizovaných s fixovanými krajními atomy, ojediněle i u struktur optimalizovaných zcela uvolněně), byl jejich příspěvek k entalpii zahrnut stejně jako u QM/MM v korekci k ΔH (pro systém L byly vypočteny pouze energie, neboť vypočtení hessiánu pro tento systém bylo mimo naše výpočetní možnosti). Výsledné B3LYP bariéry u systémů s fixovanými krajními atomy vykazovaly asymptoticky klesající tendenci se zvětšujícím se modelem, hodnota reakční bariéry potenciální energie pro největší model byla 17,2 kcal.mol^{-1}, což bylo pouze o 3,7 kcal.mol^{-1} méně než v případě nejmenšího systému. Bariéry vypočtené PBE funkcionálem se s velikostí systému prakticky neměnily a zůstávaly na hodnotě okolo 13 kcal.mol^{-1} (shrnuto v tabulce 7). Naproti tomu nelze pozorovat žádnou tendenci velikosti reakční bariéry v závislosti na rozsahu modelu u systémů optimalizovaných zcela volně, přičemž bariéry vypočtené B3LYP se pohybovaly okolo 20 kcal.mol^{-1} (shrnuto v tabulce 8).

Vliv empirické disperze a vliv solvatace modelovaný pomocí COSMO (při zvolené dielektrické konstantě 4 napodobující zprůměrovaný vnitřek proteinu) byl započítán na PBE/TZVP úrovni. Všem systémům je společná velká stabilizace intermediátu oproti Michaelisově komplexu zohledněním empirické disperze (o 6-9 kcal.mol^{-1}). Stejně tak přítomnost modelového polarizovatelného kontinua v okolí klastrového modelu vede ke snížení relativní energie intermediátu, přičemž tento efekt se stává méně významným se zvětšující se velikostí systému (pro největší systém L a dielektrickou konstantu rovnou nekonečnu se jedná o 1,4 kcal/mol v případě omezené optimalizace a 5,4 kcal.mol^{-1} v případě uvolněné optimalizace). Jednotlivé hodnoty reakčních bariér shrnují tabulky 7 a 8.

Tabulka 8. Reakční bariéry pro různé velikosti klastrového modelu optimalizovaného s fixovanými krajními atomy.

Systém S značí nejmenší model, systém M střední model a systém L největší model. Hodnoty energií odpovídají relativním energiím intermediátu oproti Michaelisovu komplexu a jsou udávány v kcal.mol^{-1}. Použitý funkcionál je uveden v levém sloupci, báze atomových orbitalů byla vždy TZVP. Disperze značí započtení empirické disperze a u výpočtů se započtením solvatace modelu pomocí COSMO je uvedena hodnota dielektrické konstanty ε. *ΔZPVE* značí příspěvek vibrační energie nulového bodu, *ΔH* pak součet entalpického příspěvku partiční funkce a *ΔZPVE*. Oprava k *ΔH* přibližně kompenzuje chybu vzniklou přítomností několika malých imaginárních vibračních módů.

Omezená optimalizace	Systém S	Systém M	Systém L
PBE	12.8	13.3	13.2
PBE + disperze	3.07	8.12	7.20
PBE, COSMO, ε = 4	10.5	9.29	12.4
PBE, COSMO, ε = ∞	8.74	6.79	11.8
TPSS	12.9	12.0	12.5
B3LYP	20.9	18.7	17.2
ΔZPVE	-0.2	-0.2	-
ΔH	-1.0	-0.4	-
oprava k ΔH	0.0	-0.6	-
entalpie	**19.9**	**17.7**	**17.2**

Tabulka 9. Reakční bariéry pro různé velikosti klastrového modelu optimalizovaného zcela uvolněně.

Systém S značí nejmenší model, systém M střední model a systém L největší model. Hodnoty energií odpovídají relativním energiím intermediátu oproti Michaelisovu komplexu a jsou udávány v kcal.mol^{-1}. Použitý funkcionál je uveden v levém sloupci, báze atomových orbitalů byla vždy TZVP. Disperze značí započtení empirické disperze a u výpočtů se započtením solvatace modelu pomocí COSMO je uvedena hodnota dielektrické konstanty ε. *ΔZPVE* značí příspěvek vibrační energie nulového bodu, *ΔH* pak součet entalpického příspěvku partiční funkce a *ΔZPVE*. Oprava k *ΔH* přibližně kompenzuje chybu vzniklou přítomností několika malých imaginárních vibračních módů.

Uvolněná optimalizace	Systém S	Systém M	Systém L
PBE	١٩.٨	١٤.٥٦	١٨.١
PBE + disperze	١٢.١	٦.٨٦	٨.٧٤
PBE, COSMO, ε = 4	١٤.٤	٩.٩١	١٤.٩
PBE, COSMO, ε = ∞	١٠.٧	٦.٩١	١٢.٧
TPSS	٢٠.٥	١٢.٦	١٨.١
B٣LYP	٢٢.١	١٧.٧	٢١.٩
ΔZPVE	١.٩	٢.٩	-
ΔH	١.٨	١.٧	-
oprava k ΔH	-٠.٦	٠.٦	-
entalpie	**٢٣.٢**	**٢٠.١**	**٢١.٩**

4.3.4. Alternativní reakční mechanismus

Na základě struktur systému L (s fixovanými krajními atomy) jsem se pokusil namodelovat alternativní reakční mechanismus, lišící se protonačním stavem Glu424. Zatímco doposud bylo předpokládáno, že Glu424 je na počátku reakce protonován na $O_{\varepsilon 2}$ atomu, zde byl modelován v deprotonovaném stavu. Kromě Michaelisova komplexu byly nalezeny dva stabilní meziprodukty a byl učiněn pokus o popis přibližné reakční bariéry mezi těmito dvěma stavy. Používaná označení atomů dále v textu odpovídají obrázku 10 v kapitole 4.2.

Geometrie Michaelisova komplexu se oproti struktuře s protonovaným Glu424 liší především polohou Glu424, což je dáno tím, že po ztrátě protonu již nemůže karboxylová skupina Glu424 tvořit dvě vodíkové vazby s hydroxidem koordinovaným zinečnatými ionty. Místo toho slouží $O_{\varepsilon 2}$ této skupiny jako akceptor vodíkové vazby -NH skupiny štěpené peptidové vazby substrátu na úkor karbonylové skupiny Gly518. Vezmeme-li jako referenci krystalovou strukturu GCPII s navázaným glutamátem (i) (PDB kód 2C6G [Mesters 2006]), tak model alternativního reakčního mechanismu (ii) popisuje lépe než model s protonovaným Glu424 (iii) vzdálenosti $O_{OH}...O_{\varepsilon 2}$ (3,5 Å {i} versus 2,7 Å {ii}; 3,4 Å {iii} v 2C6G) a $O_{\varepsilon 2}...N_{pept}$ (2,7 Å versus 3,3 Å; 2,7 Å v 2C6G), zatímco vzdálenost $O_{G518}...N_{pept}$ (3,3 Å versus 2,8 Å; 2,9 Å v 2C6G) hůře. QM/MM model pak tyto vzdálenosti popisuje zcela odlišně, například vzdálenost $O_{OH}...O_{\varepsilon 2}$ popisuje jinak i v porovnání s dalšími krystalovými strukturami 2OOT [Bařinka 2007b] či 3BHX [Bařinka 2008]. Problém ale spočívá právě ve výběru referenční struktury, neboť není žádná krystalová struktura GCPII divokého typu s navázaným substrátem v aktivním místě a aminoskupina glutamátu vázaného ve struktuře 2C6G má mnohem větší možnosti tvorby vodíkových vazeb než -NH skupina planární peptidové vazby. Vybrané vzdálenosti všech tří modelů a krystalové struktury Glu424Ala formy GCPII s navázaným substrátem v aktivním místě (pdb kód 3BXM [Klusák 2009]) shrnuje tabulka 11.

4.3.4.1. Reakční koordináta a reakční bariéra

Ve struktuře Michaelisova komplexu (viz obrázek 15) je peptidová vazba orientována tak, že vzdálenost mezi atomem kyslíku O_{OH} a uhlíku C_{pept}, který je zmíněným atomem kyslíku nukleofilně napadán, je pouhých 2,7 Å. Adice atomu kyslíku O_{OH} na atom C_{pept} má za následek změnu hybridizace atomů C_{pept} i N_{pept} z sp^2 na sp^3, přičemž $O_{\varepsilon 1}$ atom karboxylové skupiny Glu424 extrahuje proton z tohoto kyslíku. Vodíkový atom -NH skupiny původně planární peptidové vazby se přiklání k $O_{\varepsilon 2}$ atomu karboxylové skupiny Glu424 a tvoří s ním vodíkovou vazbu (stabilní intermediát 1). Pro rozštěpení peptidové vazby je však ještě nutné přenést proton nacházející se na $O_{\varepsilon 1}$ kyslíku na N_{pept}. To je možné tak, že se vodíkový atom na N_{pept} odkloní od

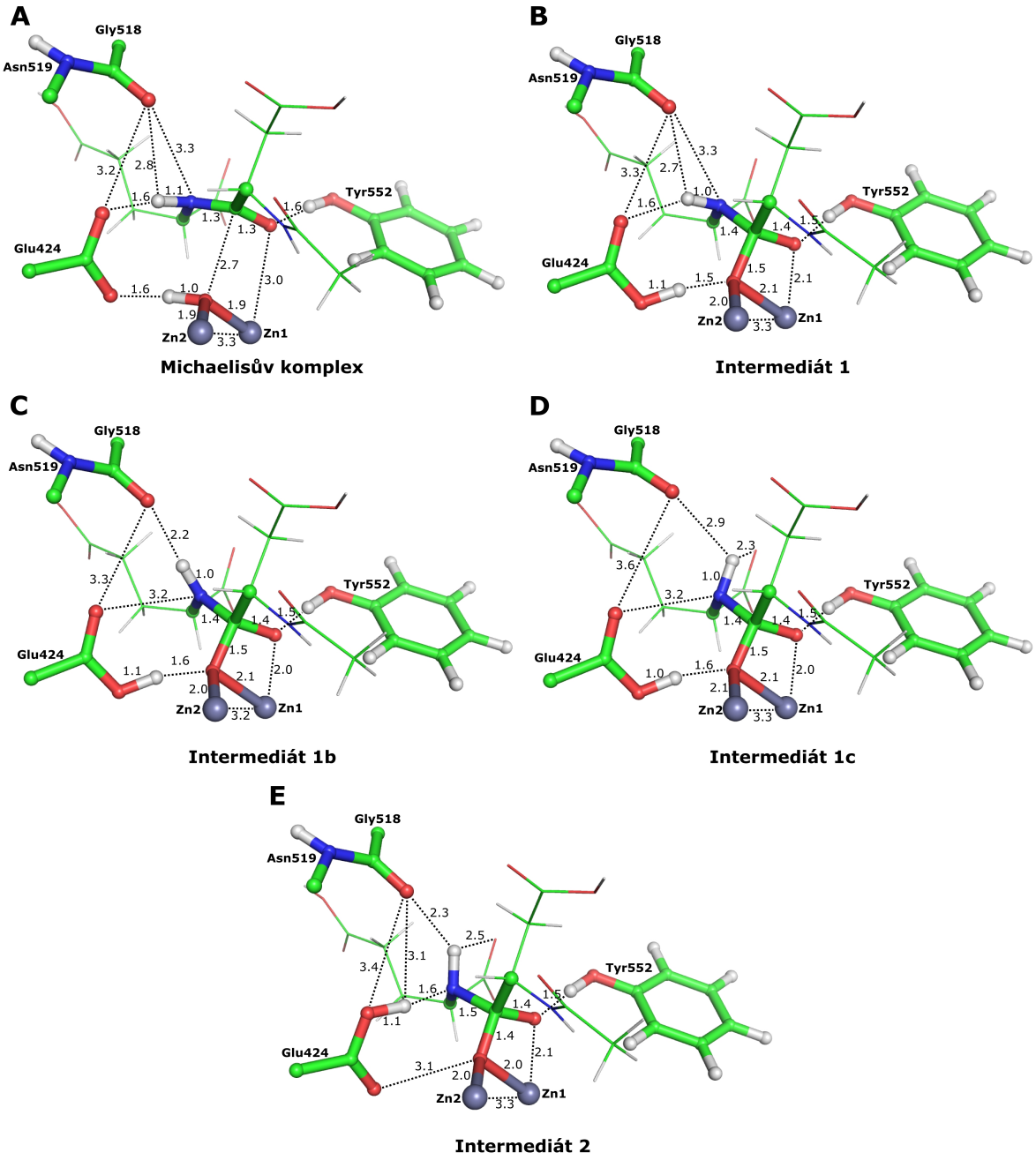
$O_{\varepsilon 2}$ atomu a zároveň dojde k rotaci karboxylové skupiny Glu424 tak, že se $O_{\varepsilon 1}$ atom i s vodíkem dostane do blízkosti N_{pept}, což vede ke vzniku dalšího stabilního intermediátu 2. Předpokládáme, že prodloužení peptidové vazby v této struktuře povede k přeskočení vodíku na peptidový dusík a k disociaci peptidové vazby (tento poslední krok reakce již nebyl modelován). Jelikož cesta od intermediátu 1 k intermediátu 2 je poměrně složitá, tak jsem se pokusil odhadnout energetickou bariéru pro překlopení vodíku na peptidovém dusíku. Při modelování různých orientací tohoto vodíku bylo zjištěno, že žádná bariéra pro tento jev není, pravděpodobně v důsledku tvorby vodíkové vazby s karbonylovým kyslíkem Gly518 v planární konformaci -NH skupiny. Celá rotace probíhá proti gradientu energie, pro ilustraci energetického profilu byla namodelována strukturu s planární -NH skupinou (intermediát 1b) a struktura s vodíkem zcela odkloněným od $O_{\varepsilon 2}$ atomu (intermediát 1c). Nejedná se tedy o struktury v minimu, ale o mezistavy, které vedou k intermediátu 2. Energetický profil rotace karboxylové skupiny Glu424 nebyl zatím prozkoumán. Z délky peptidové vazby 1,43 Å v intermediátu 1 a 1,51 Å v intermediátu 2 je vidět, že je tato vazba pevnější než v případě intermediátu reakčního mechanismu s protonovaným Glu424 (délka peptidové vazby 1,59 Å). Detaily aktivního místa jednotlivých struktur v průběhu reakce jsou znázorněny na obrázku 15.

Na výsledných strukturách byly vypočteny energie na úrovni PBE/TZVP, relativní energie oproti struktuře Michaelisova komplexu shrnuje tabulka 10. Výsledná reakční bariéra nejméně 27,7 kcal.mol^{-1} je o více než 14 kcal.mol^{-1} vyšší než v případě stejného modelu, lišícího se pouze přítomností protonu na karboxylové skupině Glu424.

Tabulka 10. Relativní energie geometrií vznikajících v průběhu reakčního mechanismu GCPII s deprotonovanou karboxylovou skupinou Glu424 v Michaelisově komplexu.

Energie jsou uváděny v kcal.mol^{-1}, PBE/TZVP značí úroveň teorie pro výpočet energií. MC značí Michaelisův komplex, INT1 intermediát 1, INT1b intermediát 1b, INT1c intermediát 1c a INT2 intermediát 2.

	MC	INT1	INT1b	INT1c	INT2
PBE/TZVP	0,0	18,6	24,4	27,7	15,0



Obrázek 15. Reakční koordináta klastrového modelu GCPII (systém L s fixovanými krajními atomy) s deprotonovanou karboxylovou skupinou Glu424 v Michaelisově komplexu.

Zinečnaté kationty jsou znázorněny šedými kuličkami. Aminokyselinové zbytky Glu424, Gly518, Asn519, Tyr552 a hydroxidový aniont vázaný zinečnatými kationty jsou znázorněny tyčinkovým a kuličkovým modelem. Ze substrátu *N*-Ac-Asp-Glu je tyčinkovým a kuličkovým modelem znázorněna pouze štěpená peptidová vazba, zbytek je znázorněn tenkými čárami. Vzdálenosti jsou uváděny v ångströmech.

Tabulka 11. Porovnání vybraných vzdáleností mezi QM/MM a klastrovými modely a krystalovou stukturou. Vzdálenosti jsou udávány v ångströmech, „3BXM“ je PDB kód krystalové struktury Glu424Ala formy GCPII s navázaným substrátem v aktivním místě [Klusák 2009], „systém L“ značí největší klastrový model optimalizovaný s fixovanými krajními atomy, „Alt. mechanismus“ stejný model s deprotonovaným Glu424 a „QM/MM“ pak QM/MM model GCPII divokého typu. Značení atomů, stejně jako očíslování jednotlivých vzdáleností v levé části tabulky odpovídá značení na obrázku 10 v kapitole 4.2. Kurzívou jsou uvedeny vzdálenosti týkající se dvou molekul vody nahrazující Glu424 v Glu424Ala mutantní formě GCPII, tučně jsou pak uvedeny vzdálenosti diskutované v textu.

		3BXM	Systém L	Alt. mechanismus	QM/MM
1	**O_{OH}-$O_{\varepsilon 1}$**	*2.9*	2.71	2.61	3.23
2	**O_{OH}-$O_{\varepsilon 2}$**	*4.4*	**2.66**	**3.53**	**4.80**
3	**$O_{\varepsilon 2}$-N_{pept}**	*4.8*	**3.27**	**2.69**	**4.46**
4	**$O_{\varepsilon 2}$-O_{G518}**	*4.2*	3.66	3.22	2.92
5	**O_{G518}-N_{pept}**	**2.9**	**2.84**	**3.31**	**3.21**
6	**CN vazba**	1.3	1.33	1.34	1.37
7	**CO vazba**	1.2	1.27	1.27	1.24
8	**OH-C_{pept}**	2.7	2.86	2.73	2.83
9	**Zn_1-O_{pept}**	2.6	2.36	3.01	3.49
10	**Zn_1-Zn_2**	3.3	3.39	3.28	3.36
11	**Zn_1-O_{OH}**	2.0	2.04	1.92	2.20
12	**Zn_2-O_{OH}**	2.0	2.02	1.95	2.18

4.4. Porovnání DFT a „post-SCF“ metod na malých klastrových modelech

V tomto oddíle byla pro nejmenší klastrový model GCPII domodelována kompletní reakční cesta včetně tranzitních stavů. Energie výsledných struktur byla vypočtena nejen pomocí DFT a různých funkcionálů, ale také pomocí Hartreeho-Fockovy metody a „post-SCF“ metod RIMP2 a RICC2 [Hattig 2000]. Cílem bylo nejenom určit hodnotu reakční bariéry GCPII v tomto modelu, ale také získat porovnání s referenční *ab initio* metodami, přičemž pro pokročilejší metody typu CCSD(T) je tento model bohužel stále příliš velký.

Pro srovnání DFT s referenčními výpočty na více modelech byly převzaty struktury malého klastrového modelu reakčního mechanismu dinukleární zinkové aminopeptidasy z *Aeromonas proteolytica* (AAP) patřící do stejné třídy peptidas M28 [Rawlings 2008] jako GCPII a vypočtena reakční bariéra stejnými metodami jako u modelu GCPII. Jako další referenční systém byl zvolen malý klastrový model reakčního mechanismu bakteriální monozinkové peptidasy thermolysinu, který byl vytvořen podle [Blumberger 2007].

4.4.1. Glutamátkarboxypeptidasa II

Výsledné struktury Michaelisova komplexu, intermediátu, produktu a obou tranzitních stavů jsou znázorněny na obrázku 16. Struktury tranzitních stavů nebyly dooptimalizovány až do úplné konvergence, neboť fixování krajních atomů modelu velmi výrazně zpomaluje konvergenci optimalizace do tranzitního stavu. Avšak změna energie při posledních optimalizačních krocích byla poměrně malá ($3 \cdot 10^{-5}$ Hartree) a také jsem ukázal, že manipulace nejdůležitější vnitřní souřadnice jedním nebo druhým směrem vede po následné optimalizaci ke strukturám nacházejícím se na odlišné straně bariéry. Z těchto důvodů jsem tyto struktury považoval za dobré přiblížení obou tranzitních stavů. Značení atomů použité v následujícím textu odpovídá značení na obrázku 10 v kapitole 4.2.

Prvním krokem reakce je nukleofilní útok atomu kyslíku hydroxidu (O_{OH}) vázaného zinečnatými ionty na atom uhlíku štěpení peptidové vazby (C_{pept}). To je spojeno s odkloněním protonovaného atomu kyslíku $O_{\varepsilon 2}$ Glu424 od zmíněného atomu O_{OH}, se kterým původně tvořil vodíkovou vazbu, směrem k dusíku štěpené peptidové vazby (N_{pept}). Tranzitní stav 1 je pak charakterizován normálním vibračním módem o frekvenci 408i cm^{-1}, který je spojen se změnou $O_{OH}...C_{pept}$ vzdálenosti a přeskokem protonu z $O_{\varepsilon 2}$ na N_{pept}. Geometrie N_{pept} odpovídá v této struktuře sp^3 hybridizaci, geometrie C_{pept} odpovídá přechodu mezi sp^2 a sp^3 hybridizací. Délka peptidové vazby je v této struktuře 1,44 Å oproti 1,34 Å ve struktuře Michaelisova komplexu, $O_{OH}...C_{pept}$ vzdálenost je pak 2,00 Å. Prodloužení této vzdálenosti o 0,1 Å vedlo po minimalizaci geometrie ke struktuře Michaelisova komplexu, zatímco její zkrácení o 0,1 Å vedlo ke struktuře

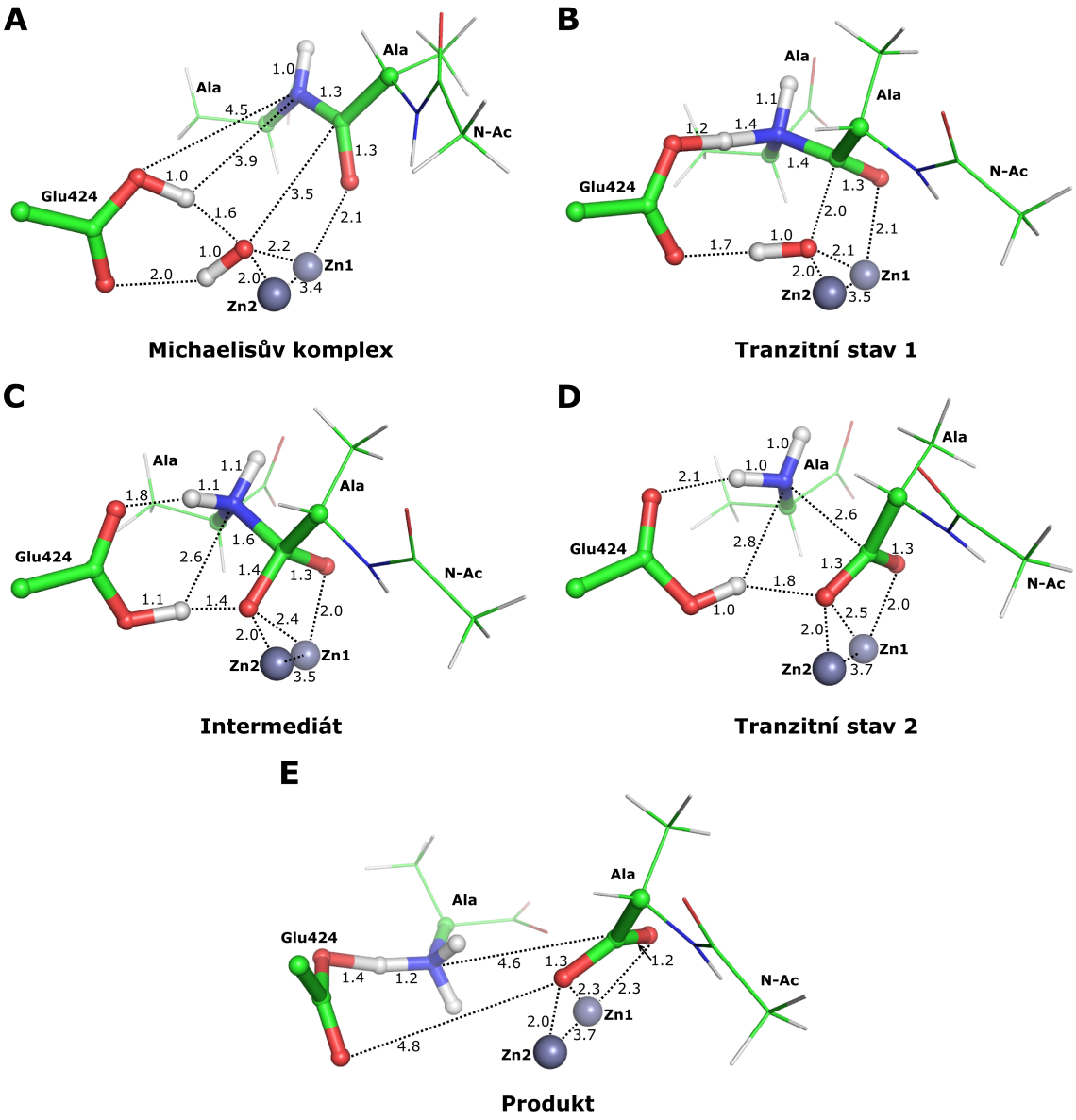
intermediátu. Ten je charakterizován kovalentní vazbou mezi O_{OH} a C_{pept}, přenesením protonu z $O_{\varepsilon 2}$ na N_{pept} a také protonu z O_{OH} na $O_{\varepsilon 1}$. Peptidová vazba se v této struktuře dále prodloužila na 1,56 Å. Tranzitní stav 2 vedoucí k disociaci peptidové vazby je charakterizován vibračním módem o poměrně malé frekvenci 66i cm^{-1}. Tento mód je spojen s natahováním peptidové vazby, jejíž délka je 2,61 Å a překlápěním karboxylové skupiny Glu424 tak, že se protonovaný kyslík karboxylové skupiny Glu424 $O_{\varepsilon 1}$ vzdaluje od O_{OH} a přibližuje N_{pept}. Při dalším vývoji systému směrem k produktu reakce (rozštěpení peptidové vazby) přechází proton z $O_{\varepsilon 1}$ na N_{pept} a stabilizuje tak volný elektronový pár na atomu dusíku.

Při hledání struktury tranzitního stavu 2 pouhým prodlužováním délky peptidové vazby se nepodařilo nalézt žádné energetické maximum, dokonce i při separaci 4 Å mezi N_{pept} a C_{pept} se systém při optimalizaci vracel do struktury intermediátu. Přibližná struktura tranzitního stavu 2 byla nalezena až metodou využívající lokální hessián (podrobnosti viz kapitola 3.3.3.). Tato struktura se vyznačovala odklonem protonovaného atomu $O_{\varepsilon 1}$ patřící Glu424 od atomu kuslíku O_{OH} (značení atomů viz obrázek 10 v kapitole 4.2.1.) a příklonu k peptidovému dusíku N_{pept}. Interakce mezi protonem na atomu $O_{\varepsilon 1}$ a atomem dusíku N_{pept} je pak zásadní pro stabilizaci produktu. Tato přibližná struktura byla dále optimalizována. Minimalizace nalezené struktury vedla ke struktuře intermediátu. Produkt reakce byl posléze získán zvětšením vzdálenosti $N_{pept}...C_{pept}$ o 0,2 Å, přeoptimalizováním postranního řetězce Glu 424 při zafixování poloh těchto dvou atomů a následné optimalizaci polohy i obou zmíněných atomů. Minimalizace struktur odvozených od tranzitního stavu manipulací $N_{pept}...C_{pept}$ vzdálenosti se vyznačovala velmi malými změnami energie během prvních několika desítek optimalizačních kroků (platí pro optimalizace směřující do intermediátu i do produktu), řádově $1 \cdot 10^{-5}$ Hartree, což naznačuje velmi plochý tvar hyperplochy potenciální energie pro tyto struktury.

Energie těchto struktur byla přepočtena pomocí DFT s použitím funkcionálů PBE, TPSS a B3LYP a různých bází atomových orbitalů, stejně tak jako pomocí HF, RIMP2 a RICC2. Vliv empirické disperze byl také zohledněn. Elektrostatické působení proteinového okolí bylo přibližně započteno pomocí modelu COSMO. Na základě vypočtených hessiánů byla stejným způsobem jako u QM/MM a ostatních klastrových modelů vypočtena také změna entalpie v průběhu reakce (hessián na úrovni PBE/SVP a energie na úrovni B3LYP/QZVP). Relativní energie tranzitních stavů, intermediátu a produktu vůči Michaelisovu komplexu shrnuje tabulka 12, entalpický profil reakce vypočtený metodou B3LYP/TZVP pak shrnuje tabulka 13.

Nejméně stabilní strukturou byl podle DFT výpočtů tranzitní stav 1, jehož B3LYP/QZVP energie byla o 23,1 kcal vyšší než energie Michaelisova komplexu. Bariéra na úrovni

TPSS/QZVP byla o 0,3 kcal.mol^{-1} nižší, zatímco PBE/QZVP bariéra o 1,2 kcal.mol^{-1}. Zohlednění disperze vedlo na B3LYP úrovni k relativní stabilizaci (oproti Michaelisovu komplexu) tranzitního stavu 1 (TS1) o 11,8 kcal.mol^{-1}, intermediátu (INT) o 12,7 kcal.mol^{-1}, zatímco tranzitního stavu 2 (TS2) o pouhých 1,9 kcal.mol^{-1}. Solvatace modelovaná pomocí COSMO vedla při zvolené dielektrické konstantě čtyři k relativní stabilizaci TS1 o 0,0 kcal.mol^{-1}, INT o 2,6 kcal.mol^{-1} a TS2 o 4 kcal.mol^{-1}. Referenční RICC2/TZVP výpočty ve vakuu pak vedly k reakční bariéře 16,8 kcal.mol^{-1} (tato bariéra odpovídá rozdílu mezi TS2 a INT). Rozdíl energie mezi nejméně stabilní strukturou a Michaelisovým komplexem je však pouhých 11,0 kcal.mol^{-1}, tzn. snížení o 10,8 kcal.mol^{-1} oproti metodě DFT(B3LYP)/TZVP.



Obrázek 16. Reakční mechanismus štěpení peptidové vazby GCPII založený na malém klastrovém modelu. Zinečnaté ionty jsou znázorněny šedými kuličkami. Postranní řetězec Glu424, hydroxidový aniont a štěpená peptidová vazba substrátu jsou znázorněny tyčinkovým a kuličkovým modelem, zbytek substrátu *N*-Ac-Ala-Ala je znázorněn pouze tenkými čárami. Vzdálenosti jsou uváděny v ångströmech.

Tabulka 12. Energetický profil štěpení peptidové vazby GCPII na základě malého klastrového modelu.

Relativní energie jsou udávány v kcal·mol^{-1}. MC značí Michaelisův komplex, TS1 a TS2 tranzitní stavy 1 a 2, INT intermediát a PROD produkt reakce.

Funkcionál	Báze	Okolí	MC	TS1	INT	TS2	PROD
B3LYP	SVP	vakuum	0.0	16.2	7.4	17.2	-7.8
B3LYP	TZVP	vakuum	0.0	21.8	14.3	17.1	-4.4
B3LYP	**QZVP**	**vakuum**	**0.0**	**23.1**	**16.3**	**17.5**	**-3.0**
B3LYP, disp.	TZVP	vakuum	0.0	10.0	1.5	15.2	-7.0
B3LYP	TZVP	COSMO, ε = 4	0.0	21.8	11.7	13.5	-3.1
B3LYP	TZVP	COSMO, ε = ∞	0.0	21.5	9.7	10.7	-2.1
TPSS	SVP	vakuum	0.0	16.1	5.6	15.3	-10.3
TPSS	TZVP	vakuum	0.0	21.4	12.2	15.3	-6.5
TPSS	**QZVP**	**vakuum**	**0.0**	**22.8**	**14.4**	**15.7**	**-5.0**
TPSS, disp.	TZVP	vakuum	0.0	10.1	0.0	13.4	-9.0
TPSS	TZVP	COSMO, ε = 4	0.0	21.4	9.6	11.6	-5.2
TPSS	TZVP	COSMO, ε = ∞	0.0	21.1	7.6	8.9	-4.2
PBE	SVP	vakuum	0.0	14.5	5.1	16.6	-9.0
PBE	TZVP	vakuum	0.0	20.5	12.4	16.6	-5.3
PBE	**QZVP**	**vakuum**	**0.0**	**21.9**	**14.5**	**17.0**	**-3.7**
PBE, disp.	TZVP	vakuum	0.0	12.1	3.2	15.2	-7.1
PBE	TZVP	COSMO, ε = 4	0.0	20.5	9.8	13.0	-3.9
PBE	TZVP	COSMO, ε = ∞	0.0	20.2	7.8	10.2	-3.0

Metoda	Báze	Okolí	MC	TS1	INT	TS2	PROD
Hartree-Fock	SVP	vakuum	0.0	34.3	19.3	11.2	-11.0
Hartree-Fock	TZVP	vakuum	0.0	38.8	24.7	11.1	-8.4
RIMP2	SVP	vakuum	0.0	11.6	-4.6	13.3	-13.5
RIMP2	TZVP	vakuum	0.0	11.6	-4.1	12.2	-11.5
RICC2	SVP	vakuum	0.0	7.9	-5.8	12.6	-14.3
RICC2	**TZVP**	**vakuum**	**0.0**	**7.4**	**-5.8**	**11.0**	**-12.3**

Tabulka 13. Entalpický profil štěpení peptidové vazby GCPII na základě malého klastrového modelu.

Relativní energie jsou udávány v kcal·mol^{-1}. MC značí Michaelisův komplex, TS1 a TS2 tranzitní stavy 1 a 2, INT intermediát a PROD produkt reakce. *ΔZPVE* značí příspěvek vibrační energie nulového bodu, *ΔH* pak součet entalpického příspěvku partiční funkce a *ΔZPVE*. Oprava k *ΔH* přibližně kompenzuje chybu vzniklou přítomností několika malých imaginárních vibračních módů.

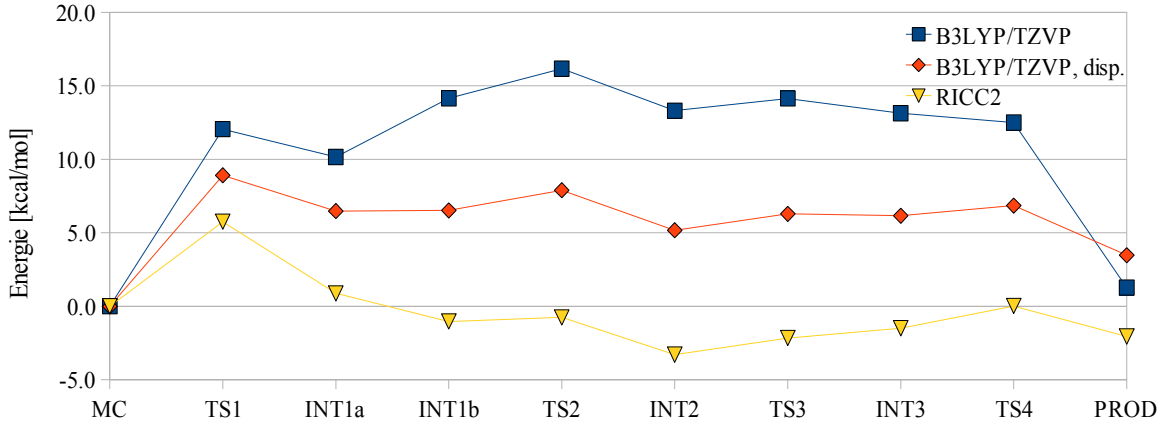
B3LYP/QZVP	MC	TS1	INT	TS2	PROD
energie	**0.0**	**23.1**	**16.3**	**17.5**	**-3.0**
ΔZPVE	0.0	-2.9	-0.5	-0.7	0.2
ΔH	0.0	-4.3	-1.7	-3.0	-0.1
oprava k ΔH	0.0	0.6	0.6	1.8	0.0
entalpie	**0.0**	**19.4**	**15.3**	**16.3**	**-3.0**

4.4.2. Aminopeptidasa z *Aeromonas proteolytica* (AAP)

Z předchozí kapitoly jsme získali srovnání metody DFT a metody RICC2 na klastrovém modelu aktivního místa GCPII. Abychom získali obecnější představu o vlivu použité metody na výslednou reakční bariéru, tak jsme se rozhodli provést stejné srovnání také na malých klastrových modelech dizinkové aminopeptidasy z *Aeromonas proteolytica* a monozinkové peptidasy thermolysinu (viz příští podkapitola).

Pro struktury malého klastrového modelu reakčního mechanismu štěpení peptidové vazby enzymem AAP převzaté z [Chen 2008] byly tedy vypočteny energie zcela stejnými metodami jako v případě malého modelu GCPII v předchozí podkapitole, tyto energie shrnuje tabulka 14. Hessiány a tudíž ani entalpie nebyly pro tyto struktury počítány.

Nejméně stabilní strukturou byl v B3LYP/QZVP výpočtech tranzitní stav 2, odpovídající reakční bariéře 17,2 kcal.mol^{-1}, přičemž vliv solvatace modelovaný pomocí COSMO tuto bariéru v případě dielektrické konstanty rovné čtyřem zvedl pouze o 0,9 kcal.mol^{-1}. Použití funkcionálů TPSS a PBE vedlo ke stejnému rychlost určujícímu tranzitnímu stavu, avšak o 5,0 kcal.mol^{-1} respektive 6,2 kcal.mol^{-1} nižší reakční bariéře. Započtení disperze vedlo k tomu, že rychlost určujícím krokem se stal průchod tranzitním stavem 1 při aktivační bariéře pouhých 8,9 kcal.mol^{-1} v případě B3LYP/TZVP výpočtu. Velmi podobné výsledky byly získány referenčním RICC2/TZVP výpočtem, který vedl k bariéře 5,8 kcal.mol^{-1}. Lze vypozorovat, že relativní DFT energie se započtením empirické disperze leží někde mezi samotnými DFT energiemi a RICC2 energiemi (viz graf 1) stejný trend lze vypozorovat i u GCPII.



Graf 1. Porovnání realtivních energií reakční koordináty klastrového modelu AAP

MC značí Michaelisův komplex; TS1, TS2, TS3 a TS4 jednotlivé tranzitní stavy; INT1a, INT1b, INT2 a INT3 jednotlivé intermediáty a PROD produkt reakce.

Tabulka 14. Relativní energie reakční cesty malého klastrového modelu reakčního mechanismu AAP.

Energie jsou udávány v kcal.mol^{-1}. MC značí Michaelisův komplex; TS1, TS2, TS3 a TS4 jednotlivé tranzitní stavy; INT1a, INT1b, INT2 a INT3 jednotlivé intermediáty a PROD produkt reakce. U výpočtů metodou DFT jsou použité funkcionály uvedeny v levém horním rohu tabulek, v případě „post-SCF" metod je příslušná metoda uvedena v řádku nad bázemi, disperze značí započtení empirické disperze.

B3LYP	SVP	TZVP	QZVP	disperze TZVP	COSMO, ε = 4 TZVP	COSMO, ε = ∞ TZVP
MC	0.0	0.0	0.0	0.0	0.0	0.0
TS1	6.3	12.1	12.4	8.9	14.1	15.4
INT1a	3.1	10.2	10.5	6.5	12.3	13.7
INT1b	6.2	14.1	14.8	6.5	15.5	16.4
TS2	7.8	16.2	17.2	7.9	18.1	19.4
INT2	6.1	13.3	14.2	5.2	14.0	14.6
TS3	6.9	14.1	15.2	6.3	14.9	15.4
INT3	6.4	13.1	13.9	6.2	13.4	13.6
TS4	6.7	12.5	13.3	6.8	12.8	12.9
PROD	-0.3	1.3	1.9	3.5	3.3	4.6

TPSS	SVP	TZVP	QZVP	disperze TZVP	COSMO, ε = 4 TZVP	COSMO, ε = ∞ TZVP
MC	0.0	0.0	0.0	0.0	0.0	0.0
TS1	4.2	9.4	9.8	6.4	11.3	12.5
INT1a	2.1	8.1	8.5	4.6	10.2	11.4
INT1b	2.9	10.2	10.8	2.9	11.4	12.2
TS2	3.4	11.2	12.2	3.4	13.0	14.2
INT2	1.6	8.3	9.2	0.6	8.9	9.3
TS3	2.0	8.9	9.8	1.4	9.4	9.8
INT3	1.8	8.2	8.9	1.5	8.4	8.5
TS4	1.9	7.6	8.3	2.2	7.8	7.9
PROD	-1.9	0.2	0.8	2.3	1.9	3.1

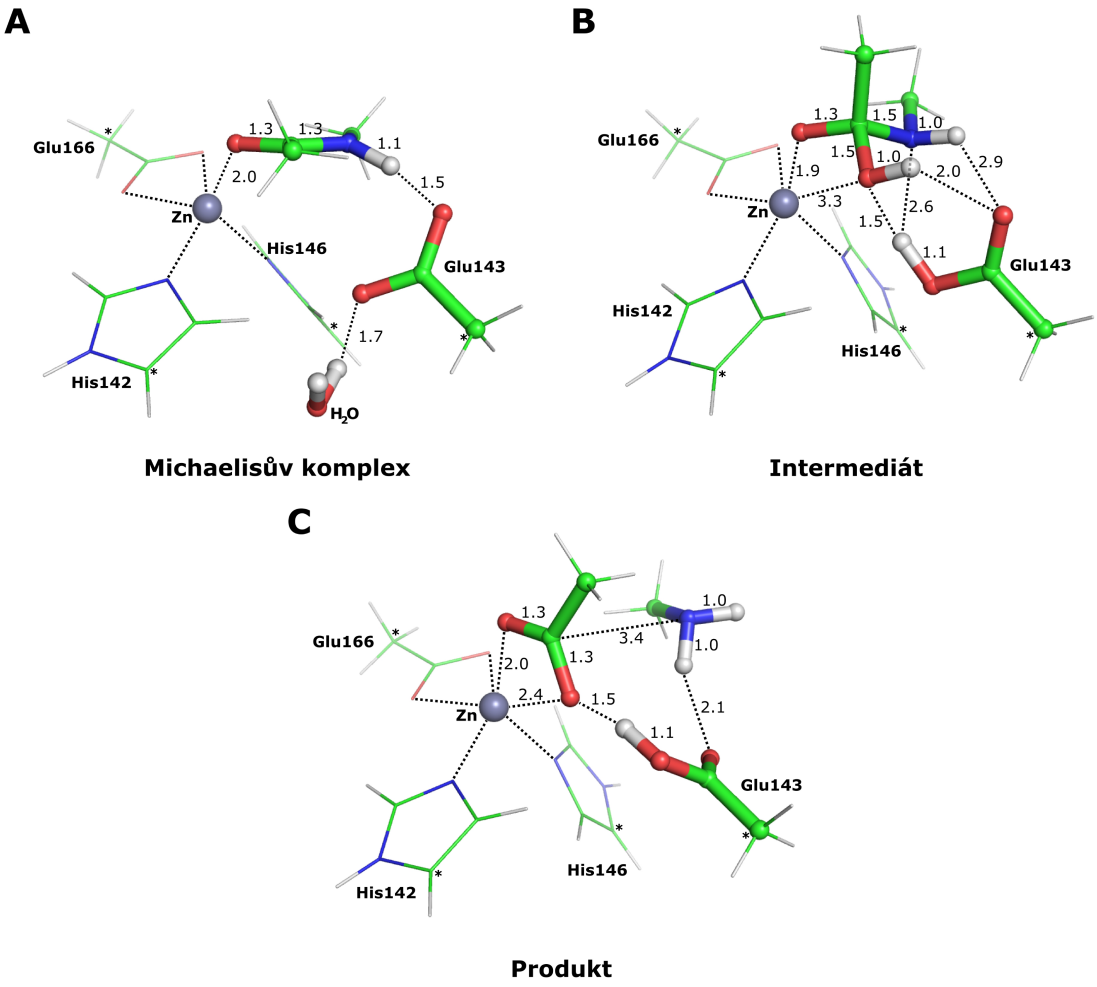
PBE	SVP	TZVP	QZVP	disperze TZVP	COSMO, ε = 4 TZVP	COSMO, ε = ∞ TZVP
MC	0.0	0.0	0.0	0.0	0.0	0.0
TS1	3.6	9.5	9.8	7.2	11.3	12.5
INT1a	0.8	7.7	8.0	5.0	9.6	10.9
INT1b	1.9	9.8	10.4	4.3	10.9	11.7
TS2	1.4	9.9	11.0	4.0	11.7	12.9
INT2	0.2	7.6	8.4	1.8	8.1	8.5
TS3	0.5	7.9	8.9	2.3	8.4	8.8
INT3	1.2	8.1	8.9	3.1	8.3	8.4
TS4	1.8	7.9	8.7	3.9	8.0	8.1
PROD	-0.4	1.6	2.4	3.2	3.3	4.5

	HF SVP	HF TZVP	RIMP2 SVP	RIMP2 TZVP	RICC2 SVP	RICC2 TZVP
MC	0.0	0.0	0.0	0.0	0.0	0.0
TS1	11.3	16.4	5.1	7.2	3.8	5.8
INT1a	5.5	11.3	-1.1	0.5	-0.5	0.9
INT1b	11.6	17.9	-3.3	-1.1	-3.2	-1.0
TS2	19.3	25.5	-2.1	-0.1	-2.7	-0.8
INT2	17.0	22.4	-2.7	-2.3	-3.7	-3.3
TS3	19.0	24.5	-1.4	-0.9	-2.5	-2.2
INT3	14.9	20.2	-1.5	-0.4	-2.4	-1.5
TS4	14.7	19.2	0.6	1.7	-1.0	0.0
PROD	-3.1	-3.1	-1.5	-1.2	-2.7	-2.1

4.4.3. Thermolysin

Pro malý klastrový model aktivního místa thermolysinu čítající 48 atomů byly nalezeny struktury Michaelisova komplexu, intermediátu a produktu štěpení modelového substrátu *N*-methyl-acetamidu (viz obrázek 17). Tento model byl vytvořen pouze za účelem porovnání metod DFT a RICC2, proto nebyly ani hledány tranzitní stavy.

Intermediát reakce vzniká po nukleofilní adici atomu kyslíku molekuly vodu na atom uhlíku peptidové vazby a současné extrakci protonu z této vody karboxylovou skupinou Glu143. V této struktuře jsou atomy uhlíku i dusíku štěpené peptidové vazby sp^3 hybridizovány. K rozštěpení této vazby vede rotace karboxylové skupiny Glu143. Následné přeskočení protonu z této skupiny na peptidový dusík a disociace vazby již probíhá po spádu gradientu. Tento reakční mechanismus dobře souhlasí s mechanismem popsaným v [Blumberger 2007].



Obrázek 17. Reakční mechanismus štěpení peptidové vazby thermolysinem.

Vzdálenosti jsou udávány v ångströmech, fixované atomy během optimalizací jsou označeny hvězdičkou. Štěpená peptidová vazba a Glu143 jsou znázorněny tyčinkovým a kuličkovým modelem, zinečnatý iont pak šedou kuličkou. Zbytek modelu aktivního místa je znázorněn pouze tenkými čárami.

Tabulka 15. Relativní energie modelu reakčního mechanismu štěpení peptidové vazby thermolysinu.

Energie intermediátu a produktu jsou vztažené k Michaelisovu komplexu a jsou udávány v kcal.mol^{-1}. QZVP* značí použití QZVPP báze pro atomy zinku a QZVP báze pro všechny ostatní atomy.

Funkcionál	Báze	Okolí	Intermediát	Produkt
B3LYP	SVP	vakuum	17.7	-3.7
B3LYP	TZVP	vakuum	20.5	-2.9
B3LYP	QZVP	vakuum	21.0	-1.9
B3LYP, disp.	TZVP	vakuum	20.1	2.2
B3LYP	TZVP	COSMO, $\varepsilon = 4$	20.3	-2.2
B3LYP	TZVP	COSMO, $\varepsilon = \infty$	19.9	-1.8
PBE	SVP	vakuum	17.8	-3.6
PBE	TZVP	vakuum	20.5	-2.8
PBE	QZVP	vakuum	21.1	-1.7
PBE, disp.	TZVP	vakuum	20.3	0.8
PBE	TZVP	COSMO, $\varepsilon = 4$	20.4	-2.1
PBE	TZVP	COSMO, $\varepsilon = \infty$	20.0	-1.7
TPSS	SVP	vakuum	18.0	-6.4
TPSS	TZVP	vakuum	20.6	-5.4
TPSS	QZVP	vakuum	21.0	-4.4
TPSS, disp.	TZVP	vakuum	20.3	-0.5
TPSS	TZVP	COSMO, $\varepsilon = 4$	20.5	-4.6
TPSS	TZVP	COSMO, $\varepsilon = \infty$	20.0	-4.2
Metoda	**Báze**	**Okolí**	**Intermediát**	**Produkt**
Hartree-Fock	SVP	vakuum	18.4	-9.7
Hartree-Fock	TZVP	vakuum	20.4	-10.4
Hartree-Fock	QZVP*	vakuum	20.7	-9.7
RIMP2	SVP	vakuum	17.2	-3.6
RIMP2	TZVP	vakuum	18.2	-2.1
RIMP2	QZVP*	vakuum	19.4	0.3
RICC2	SVP	vakuum	19.2	-3.4
RICC2	TZVP	vakuum	19.5	-2.0
RICC2	QZVP*	vakuum	20.7	0.4

Energie výsledných struktur byla vypočtena pomocí přesně stejných metod jako u modelů GCPII a AAP, pouze pro RICC2 výpočty byla použita i větší báze QZVP kvality. Relativní energie intermediátu a produktu vůči Michaelisovu komplexu shrnuje tabulka 15. Reakční bariéra aproximovaná energií intermediátu vychází na B3LYP/QZVP úrovni 21,0 kcal.mol^{-1}. Narozdíl od GCPII a AAP modelů vychází prakticky stejná bariéra i pro funkcionály TPSS a PBE a také pro RICC2 metodu (pouze o 0,3 kcal.mol^{-1} nižší). Vliv disperze či solvatace je také malý (do 1 kcal.mol^{-1}). Přeoptimalizování struktur metodou B3LYP/SVP a RIMP2/SVP nemělo žádný vliv na reakční bariéru vypočtenou danou metodou, přeoptimalizování metodou PBE/TZVP vedlo ke zvýšení reakční bariéry o 0,7 kcal.mol^{-1} vypočtené pomocí metody PBE/QZVP.

5. Diskuse

Pomocí metody PCR s reverzní transkripcí jsme v knihovnách cDNA dvou buněčných linií (LNCaP a MHH MED1), které jsou odvozené od lidských nádorových buněk a exprimují GCPII, hledali přítomnost dosud nepopsaných alternativně sestřižených variant GCPII. Studovaná oblast cDNA odpovídala přibližně prvním 8300 nukleotidům genu kódujícího GCPII (z 62000 nukleotidů celkem). V této oblasti bylo již popsány čtyři alternativně sestřižené varianty, proto jsme se na tento úsek zaměřili.

Podařilo se nám najít dohromady deset nových variant obsahujících šest doposud nepopsaných alternativních exonů. Jelikož se různé varianty GCPII liší v sekvenci mRNA poblíž iniciačního kodonu, tak by některé z nich mohly kódovat alternativní proteinové produkty. Protein kódovaný shodně variantami označovanými jako PSMD a PSMF [Schmittgen 2003, Cao 2007] by obsahoval namísto transmembránového helixu pravděpodobný signální peptid, a mohlo by se tedy jednat o sekretovanou formu GCPII. Přestože tyto varianty již byly v literatuře popsány (s drobnými odlišnostmi oproti námi nalezeným sekvencím, viz kapitola 4.1.), nebylo toto dosud nikým rozpoznáno. Přitom nahrazení membránového helixu signálním peptidem je u lidských genů poměrně časté [Modrek 2001, Stamm 2005]. Varianty nalezené v naší studii a označené jako Var12, Var13 a Var14 by pak obsahovaly různé aminokyselinové inserce (v délce 8 až 32 aminokyselin) mezi transmembránovou a proteasovou doménou (viz obrázek 7 a obrázek 8). Další studium těchto variant, například cestou rekombinantní exprese v lidských buněčných liniích, by mohlo vést k lepšímu pochopení biologické funkce GCPII, obzvláště přítomnost sekretované formy GCPII by mohla implikovat nové a dosud nepředpokládané funkce GCPII.

Nutno zdůraznit, že cílem rozhodně nebyla charakterizace buněčných linií, ale pouze a jen nalezení co nejvíce variant cDNA. Pro případnou charakterizaci různých buněčných linií a především pak lidských tkání a nádorů exprimujících GCPII je mnohem vhodnější kvantitativní RT PCR, které bude specificky amplifikovat jedinou variantu a umožní tak její přesnou kvantifikaci. Takovouto studii jsme ve spolupráci s firmou TipMolbio (Berlín, Německo) navrhli a v současné době testujeme její spolehlivost a specifitu. Studování množství různých variant GCPII v různých stadiích nádorů by mohlo mít přímé klinické využití při diagnostice, neboť bylo ukázáno, že množství varianty nazývané PSME koreluje s postupem karcinomu prostaty [Xu 2002, Cao 2007, Pajares 2007], a zároveň množství varianty PSMA (kódující GCPII divokého typu) je dobrým indikátorem množství cirkulujících mikrometastáz karcinomu prostaty v krvi, jejichž výskyt je pravděpodobně spojen s výrazně horší prognózou [Myiake 2007, Zhang 2008].

Na tyto experimentální výsledky navazovala naše snaha porozumět na teoretické úrovni co nejlépe fyzikálně-chemickým vlastnostem variatních (mutovaných) proteinů. Domníváme se, že spojení experimentu a teorie nám může poskytnout úplný obraz o studovaném systému. I když dnes ještě pokročilé metody teoretické chemie, reprezentované například technikami QM/MM, QM/MM-PBSA [Kaukonen 2008], či QM-MM-FEP [Hu 2008] (PBSA, Poisson-Boltzmann Surface Area; FEP, Free-energy perturbation) nejsou schopny poskytnout kvantitativně přesné charakteristiky systému, k tomuto cíli se pomalu přibližují. Teoretickou část této práce lze chápat jako malý příspěvek k tomuto úsilí.

V nedávné době byl popsán pomocí QM/MM modelu reakční mechanismus hydrolysy peptidové vazby dipeptidu *N*-Ac-Asp-Glu enzymem GCPII [Klusák 2009]. Kvantový systém obsahoval 217 atomů a byl popisován pomocí metody DFT(B3LYP)/TZVP. Jednalo se tedy o poměrně velký model, zahrnující kromě katalytického centra také postranní řetězce všech aminokyselin interagujících se substrátem. Rozhodli jsme se proto vyzkoušet, je-li tento model schopný správně popsat vliv záměn aminokyselin v aktivním centru. Pro námi zvolené záměny aminokyselin byly publikovány hodnoty rychlostních konstant k_{cat} rekombinantních proteinů, což nám umožňuje přímé srovnání s experimentem [Mlčochová 2007, Klusák 2009].

Vzhledem ke strukturní a energetické podobnosti obou tranzitních stavů a intermediátu jsme použili pro srovnání jednotlivých mutací s GCPII divokého typu relativní entalpii intermediátu vůči Michaelisovu komplexu. Rozdíly výsledných reakčních bariér mutantních forem oproti GCPII divokého typu velmi dobře odpovídaly rozdílům experimentálně stanovených rychlostních konstant k_{cat} (shrnutých v tabulce 2). Reakční bariéra se v případě záměny Tyr700Phe zvýšila o 2,4 kcal.mol^{-1} (z experimentálně pozorovaného patnáctinásobného snížení rychlostní konstanty k_{cat} vyplývá zvýšení bariéry o 1,7 kcal.mol^{-1}), v případě záměny Asn519Asp o 2,5 kcal.mol^{-1} (z experimentálně pozorovaného patnáctinásobného snížení rychlostní konstanty k_{cat} vyplývá zvýšení o 1,7 kcal.mol^{-1}) a v případě záměny Glu424Ala pak o 3,5 kcal.mol^{-1} (z experimentálně stanoveného k_{cat}<0,001 s^{-1} vyplývá zvýšení přinejmenším o 4,3 kcal.mol^{-1}).

V případě mutací Asn519Asp a Tyr700Phe se jednalo o záměny funkčních skupin interagujících s polárními skupinami substrátu mimo samotné reakční centrum. V případě mutace Glu424Ala se však jednalo o odstranění klíčové katalytické karboxylové skupiny aminokyseliny Glu424 účastnící se samotné hydrolysy peptidové vazby. Uvolněný prostor po této skupině je zaplněn dvěma molekulami vody, které svými interakcemi částečně nahrazují

funkci původního Glu424 (viz obrázek 11). Je zřejmé, že substituce klíčového aminokyselinového zbytku má za následek velké změny v katalytickém centru oproti GCPII divokého typu a nemusí proto platit náš předpoklad strukturní a energetické podobnosti obou tranzitních stavů a intermediátu odvozený z reakčního mechanismu GCPII divokého typu. Je důležité si také uvědomit, že na rozdíl od karboxylové skupiny Glu424 GCPII divokého typu, není orientace molekul vody (nahrazujících chybějící postranní řetězec Glu424) omezená žádnou kovalentní vazbou. Nelze tedy vyloučit existenci dalších lokálních geometrických minim a tudíž možný nezanedbatelný příspěvek konformační entropie k reakční bariéře, který nebyl při výpočtech zohledněn.

Přestože tento QM/MM model velmi dobře vystihuje rozdíly reakčních bariér u zde porovnávaných záměn funkčních skupin v kvantovém systému, vypočtená reakční bariéra (energie) GCPII divokého typu (29 kcal.mol^{-1}) je přibližně o 11 kcal.mol^{-1} vyšší než bariéra vyplývající z experimentálně stanovené rychlostní konstanty k_{cat} [Klusák 2009]. Hodnota přibližné entalpické bariéry je přitom pouze o 0,1 kcal.mol^{-1} nižší (entalpická korekce k energii byla dopočtena dodatečně pro již publikované struktury v [Klusák 2009] způsobem popsaným v kapitole 3.3.1).

Alternativní přístup k modelování enzymových reakcí spočívá v tom, že okolí centrálního systému, uvažovaného na kvantově mechanické úrovni, buď neuvažujeme vůbec nebo jej nahradíme dielektrickým solvatačním modelem (tzv. klastrový model). Tato metoda se ukázala jako velmi platná při popisu řady enzymových reakcí [Siegbahn 2009]. Použití modelů aktivních míst dizinkových metalohydrolas čítajících okolo 100 atomů vedlo k dobré shodě vypočtených reakčních bariér s experimentálními daty [Chen 2008, Liao 2009b]. V případě aminopeptidasy z *Aeromonas proteolytica* (AAP), která má stejné uspořádání dinukleárního zinkového centra včetně všech ligandů zinečnatých iontů a katalytického glutamátu, je možné porovnat reakční mechanismus získaný pomocí klastrového modelu [Chen 2008] a reakční mechanismus GCPII získaný pomocí QM/MM modelu [Klusák 2009]. Ukazuje se, že zatímco chemická podstata reakčního mechanismu je u obou modelů stejná, obě metody se liší v počtu tranzitních stavů a intermediátů.

Abychom mohli přímo porovnat oba přístupy k modelování enzymových reakcí, tak jsme vytvořili tři klastrové modely aktivního místa GCPII, lišící se počtem zahrnutých atomů (velikosti sytémů byly 78, 197 a 297 atomů). Výchozí strukturou pro tyto modely byla krystalová struktura GCPII s PDB kódem 3BXM [Klusák 2009]. Pro zahrnutí sterického vlivu

proteinového okolí, které zde není explicitně uvažováno byly krajní atomy modelu ponechány v krystalografických polohách. V prvním přiblížení jsme pro všechny tři modely nalezli strukturu intermediátu a vypočítali jejich relativní energii v porovnání s Michaelisovým komplexem. Tato energie (vypočtená na úrovni DFT(B3LYP)/TZVP) se zvětšujícím se počtem atomů klesala až na 17,2 kcal.mol^{-1} (20,9 kcal.mol^{-1} u nejmenšího modelu a 18,7 kcal.mol^{-1} u středního modelu). Uvážíme-li, že u QM/MM je výška celkové bariéry pouze o 2,9 kcal.mol^{-1} vyšší než energie intermediátu, můžeme předpokládat, že celková reakční bariéra u největšího klastrového modelu je okolo 20 kcal.mol^{-1}, což je v dobré shodě s experimentálními daty (18,1 kcal.mol^{-1}). Pro výsledné struktury jsme přibližně zohlednili proteinové okolí pomocí dielektrického kontinuálního solvatačního modelu COSMO. Při zvolené dielektrické konstantě ε=4 došlo ke snížení relativní energie intermediátu u nejmenšího modelu o 2,3 kcal.mol^{-1}, u středního o 4,0 kcal.mol^{-1} zatímco u největšího pouze o 0,8 kcal.mol^{-1}. Odlišné chování u nejmenšího modelu lze pravděpodobně připsat zmenšení substrátu z *N*-Ac-Asp-Glu na *N*-Ac-Ala-Ala, tzn. odstranění dvou záporně nabitých funkčních skupin. Malý vliv solvatačního modelu jsme tedy pozorovali teprve u největšího systému čítajícího téměř 300 atomů. U modelů aktivních míst jiných enzymů bylo přitom ukázáno, že vliv solvatačního modelu prakticky vymizí již při velikostech modelu necelých 200 atomů [Siegbahn 2008]. Odlišné chování modelů GCPII může být důsledkem přítomnosti tří záporně nabitých karboxylových skupin, jejichž náboj je dostatečně kompenzován teprve u největšího systému.

Pro nejmenší model, který velikostí odpovídá modelu aktivního místa AAP v [Chen 2008], byly dále nalezeny struktury obou tranzitních stavů a produktu reakce. Výchozí struktura, ze které byly odvozeny všechny ostatní, byla struktura intermediátu nejmenšího systému zmíněného v předchozím odstavci. Struktura Michaelisova komplexu se tedy liší od struktury Michaelisova komplexu nejmenšího systému v předchozím odstavci v poloze fixovaných krajních atomů (příprava modelů podrobně popsána v kapitole 3.3.2.) a proto se liší také relativní energie intermediátu od energie uvedené v předchozím odstavci (viz také tabulky 8 a 12). Proton nacházející se v Michaelisovu komplexu na karboxylové skupině Glu424 se ve struktuře intermediátu nachází již na dusíku štěpené vazby. Zároveň vodík nacházející se původně na hydroxidu se ve struktuře intermediátu nachází již na Glu424 (viz obrázek 14). Tímto se struktura intermediátu liší od struktury intermediátu v QM/MM modelu, kde jsou tyto protony vázány stejnými atomy jako v Michaelisově komplexu. K přeskoku protonu z Glu424 na dusík peptidové vazby dochází u klastrového modelu v tranzitním stavu 1, zatímco u QM/MM modelu v tranzitním stavu 2. Shodně u obou modelů je tento krok rychlost určujícím krokem,

což je ve shodě se zjištěním z klastrového modelu reakčního mechanismu AAP [Chen 2008], kde je tento fakt podložen experimentálními daty [Bienvenue 2002, Bzymek 2005]. Entalpická bariéra 19,4 kcal.mol^{-1}, odvozená od výpočtů na úrovni DFT(B3LYP)/QZVP (okolí modelu popsáno pomocí COSMO při ε=4), je v dobré shodě s 18,1 kcal.mol^{-1} odpovídající experimentálně stanovené rychlostní konstantě k_{cat}=1,15 s^{-1}.

Reakční bariéra získaná pomocí QM/MM modelu byla tedy nejen o 11 kcal.mol^{-1} vyšší než bariéra vyplývající z experimentálně stanovené rychlostní konstanty k_{cat}, ale také přibližně o 11 kcal.mol^{-1} vyšší než bariéra u klastrových modelů. Stejné přecenění reakční bariéry QM/MM metodou v porovnání s klastrovým modelem a experimentem bylo popsáno také u NiFe-hydrogenasy [Lill 2009], kde byla ale použit velmi malý kvantový systém v případě QM/MM metody (30 atomů). Je pro nás překvapením, že stejné chování pozorujeme i u kvantového systému čítajícího 217 atomů.

Srovnání metody DFT(B3LYP) pro klastrové modely metaloenzymů s experimentem často vede k dobré shodě [Siegbahn 2007, Chen 2008, Liao 2009b]. To ale může být dáno vzájemnou kompenzací chyb v důsledku použité metody a modelu, proto jsme se rozhodli porovnat na nejmenším modelu reakčního mechanismu GCPII metodu DFT s *ab initio* metodou RICC2. Stejné porovnání bylo provedeno pro klastrový model dizinkové peptidasy AAP převzatý z [Chen 2008]. Pro stejný účel jsme vytvořili malý klastrový model reakčního mechanismu monozinkové peptidasy thermolysinu. Vypočtená reakční bariéra hydrolysy peptidové vazby thermolysinem byla na úrovni DFT(B3LYP)/QZVP pouze o 0,3 kcal.mol^{-1} vyšší než bariéra vypočtená na úrovni RICC2/QZVP (viz tabulka 15). Naproti tomu u GCPII byla reakční bariéra pro štěpení peptidové vazby vypočtená metodou DFT(B3LYP)/TZVP o 10,8 kcal.mol^{-1} vyšší oproti hodnotě získané metodou RICC2/TZVP (viz tabulka 12), u AAP to bylo o 10,4 kcal.mol^{-1} více (viz tabulka 14).

Tyto nesoulady, které považujeme za významné, jsme dále studovali na modelových systémech hydrolysy peptidové vazby. Nejjednodušším modelem je zřejmě formamid (H_2NCHO) a jeho hydrolysa jednou či dvěma molekulami vody [Štrajbl 2000]. Dále byl vytvořen zcela minimalistický model reakčního centra se základními rysy zinkových metalopeptidas, systém $[Zn(NH_3)_2(HCOO)(OH_2)]^+$...(HCOO-)($H_2NCHO$). Tento systém by měl (alespoň *in silico*) reprezentovat stejný typ reakčního mechanismu jako u metalopeptidas. Obsahuje všechny základní stavební kameny reakčního centra: iont zinku s jeho první koordinační sférou, deprotonovanou karboxylovou skupinu vázanou ve druhé koordinační sféře

sloužící jako „proton shuttle“ a formamid jako substrát. Celková velikost je 26 atomů a lze na něm v řádu týdnů (při dnešních výpočetních možnostech) provést referenční CCSD(T) výpočet ve velké „triple zeta“ bázi.

Podle předběžných výsledků [Navrátil 2009] se pak ukazuje, že mnohé z metod kvantové chemie (CCSD, MP2, RICC2) aktivační bariéru při štěpení formamidu bez přítomnosti kovu zřetelně podceňují (chyba až 10 kcal.mol^{-1} u báze „triple zeta“ kvality), zatímco rozdíl energií mezi reaktanty a produkty je popisován vcelku konzistentně (rozdíly do 2 kcal.mol^{-1}). Považujeme-li u těchto systémů výpočet CCSD(T)/aug-cc-pVTZ jako referenční, můžeme bázi zredukovat na bázi def2-TZVP bez ztráty přesnosti. Právě CCSD(T)/def2-TZVP výpočty jsou pro nás v současnosti maximální možné pro minimalistický model kovem katalyzované hydrolysy. Metoda DFT ve spojení s nehybridními funkcionály použitými v této práci vedla také ke zřetelnému podcenění reakční bariéry, zatímco ve spojení s B3LYP k velmi dobré shodě s referenčními výpočty (chyba okolo 1 kcal.mol^{-1}). Na základě těchto dat považujeme funkcionál B3LYP za nejpřesnější z funkcionálů použitých v této práci, zatímco funkcionály GGA a „post-SCF“ metoda RICC2 dávají pravděpodobně jen spodní odhad reakční bariéry. Bohužel zatím nejsou dokončena srovnání pro hydrolysu modelového dipeptidu *N*-Ac-Ala-Ala a model štěpení formamidu za přítomnosti kovu.

Nakonec byly získané vědomosti a výpočetní postupy použity k predikci protonačního stavu zbytku Glu424 na počátku reakce. Pro největší klastrový model (čítající 297 atomů) byly namodelovány také dva intermediáty alternativního reakčního mechanismu, který se oproti doposud uvažovanému mechanismu liší protonačním stavem Glu424 v Michaelisově komplexu (viz obrázek 15). Na základě DFT(PBE)/TZVP výpočtů usuzujeme, že reakční bariéra pro tento mechanismus je přinejmenším o 14 kcal.mol^{-1} vyšší než u původního mechanismu. Proto považujeme za nepravděpodobné, že by se Glu424 nacházel v Michaelisově komplexu v deprotonovaném stavu.

6. Závěr

Pomocí experimentálních metod se nám podařilo v lidských nádorových buněčných liniích identifikovat celkem deset zcela nových alternativně sestřižených variant GCPII, které obsahují dohromady šest dosud neznámých alternativních exonů.

Medodou QM/MM jsme pak vypočetli relativní zvýšení reakční bariéry u vybraných mutací aminokyselin v aktivním místě v porovnání s GCPII divokého typu. Výsledné hodnoty pro mutace Glu424Ala, Asn519Asp a Tyr700Phe velmi dobře korespondují s experimentálně stanovenými rychlostními konstantami k_{cat}. Vycházeli jsme přitom z reakčního mechanismu hydrolysy peptidové vazby GCPII publikovaného v [Klusák 2009].

Reakční mechanismus GCPII jsme namodelovali také pomocí klastrového modelu aktivního místa. Popis absolutní hodnoty reakční bariéry pomocí klastrového modelu reakčního místa velmi dobře odpovídal experimentálně určené hodnotě k_{cat}=1,15 s^{-1}, zatímco použití QM/MM modelu vedlo k přecenění této bariéry o přibližně 10 kcal.mol^{-1}.

Pro všechny výpočty byla používána metoda DFT(B3LYP), která byla na závěr porovnána na minimalistických klastrových modelech enzymů GCPII, AAP a thermolysinu s metodou RICC2. S ohledem na existující porovnání metod DFT(B3LYP), RICC2 a CCSD(T) jsme tak získali přibližný odhad chyby metody DFT(B3LYP).

7. Literatura

Ahlrichs, R., Bar, M., Haser, M., Horn, H. and Kolmel, C. (1989) Electronic-Structure Calculations on Workstation Computers - the Program System Turbomole. *Chemical Physics Letters* **162**(3): 165-169.

Bander, N. H., Nanus, D. M., Milowsky, M. I., Kostakoglu, L., Vallabahajosula, S. and Goldsmith, S. J. (2003) Targeted systemic therapy of prostate cancer with a monoclonal antibody to prostate-specific membrane antigen. *Seminars in Oncology* **30**(5): 667-676.

Bařinka, C., Hlouchová, K., Rovenská, M., Majer, P., Dauter, M., Hin, N., Ko, Y. S., Tsukamoto, T., Slusher, B. S., Konvalinka, J.*, et al.* (2008) Structural basis of interactions between human glutamate carboxypeptidase II and its substrate analogs. *Journal of Molecular Biology* **376**(5): 1438-1450.

Bařinka, C., Rinnová, M., Šácha, P., Rojas, C., Majer, P., Slusher, B. S. and Konvalinka, J. (2002) Substrate specificity, inhibition and enzymological analysis of recombinant human glutamate carboxypeptidase II. *Journal of Neurochemistry* **80**(3): 477-487.

Bařinka, C., Rovenská, M., Mlčochová, P., Hlouchová, K., Plechanovová, A., Majer, P., Tsukamoto, T., Slusher, B. S., Konvalinka, J. and Lubkowski, J. (2007a) Structural insight into the pharmacophore pocket of human glutamate carboxypeptidase II. *Journal of Medicinal Chemistry* **50**(14): 3267-3273.

Bařinka, C., Šácha, P., Sklenář, J., Man, P., Bezouška, K., Slusher, B. S. and Konvalinka, J. (2004) Identification of the N-glycosylation sites on glutamate carboxypeptidase II necessary for proteolytic activity. *Protein Science* **13**(6): 1627-1635.

Bařinka, C., Starková, J., Konvalinka, J. and Lubkowski, J. (2007b) A high-resolution structure of ligand-free human glutamate carboxypeptidase II. *Acta Crystallografica Section F Structural Biology and Crystallization Communications* **63**(Pt 3): 150-153.

Bartlett, R. J. (1995) Coupled-Cluster Theory: An Overview of Recent Developments. in *Modern Electronic Structure Theory*. ed. Yarkony, D. R., World Scientific. **2:** 1047-1131.

Becke, A. D. (1988) Density-Functional Exchange-Energy Approximation with Correct Asymptotic-Behavior. *Physical Review A* **38**(6): 3098-3100.

Becke, A. D. (1993) Density-Functional Thermochemistry .3. The Role of Exact Exchange. *Journal of Chemical Physics* **98**(7): 5648-5652.

Bendtsen, J. D., Nielsen, H., von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology* **340**(4): 783-795.

Bienvenue, D. L., Mathew, R. S., Ringe, D. and Holz, R. C. (2002) The aminopeptidase from Aeromonas proteolytica can function as an esterase. *Journal of Biological Inorganic Chemistry* **7**(1-2): 129-135.

Blumberger, J., Lamoureux, G. and Klein, M. L. (2007) Peptide hydrolysis in thermolysin: Ab initio QM/MM investigation of the Glu143-assisted water addition mechanism. *Journal of Chemical Theory and Computation* **3**(5): 1837-1850.

Botchorishvili, G., Matikainen, M. P. and Lilja, H. (2009) Early prostate-specific antigen changes and the diagnosis and prognosis of prostate cancer. *Current Opinion in Urology* **19**(3): 221-226.

Bozulic, L. and Hemmings, B. A. (2009) PIKKing on PKB: regulation of PKB activity by phosphorylation. *Current Opinion in Cell Biology* **21**(2): 256-261.

Bzymek, K. P., Moulin, A., Swierczek, S. I., Ringe, D., Petsko, G. A., Bennett, B. and Holz, R. C. (2005) Kinetic, spectroscopic, and X-ray crystallographic characterization of the functional E151H aminopeptidase from Aeromonas proteolytica. *Biochemistry* **44**(36): 12030-12040.

Carter, R. E., Feldman, A. R. and Coyle, J. T. (1996) Prostate-specific membrane antigen is a hydrolase with substrate and pharmacologic characteristics of a neuropeptidase.

Proceedings of the National Academy of Sciences of the USA **93**(2): 749-753.
Casadesus, J. and Low, D. (2006) Epigenetic gene regulation in the bacterial world. *Microbiology and Molecular Biology Reviews* **70**(3): 830-856.
Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B. and Woods, R. J. (2005) The Amber biomolecular simulation programs. *Journal of Computational Chemistry* **26**(16): 1668-1688.
Copeland, R. A. (2000) Enzymes: A Practical Introduction to Structure, Mechanism and Data analysis., *Wiley-VCH*.
Cossi, M., Rega, N., Scalmani, G. and Barone, V. (2003) Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *Journal of Computational Chemistry* **24**(6): 669-681.
Cunha, A. C., Weigle, B., Kiessling, A., Bachmann, M. and Rieber, E. P. (2006) Tissue-specificity of prostate specific antigens: comparative analysis of transcript levels in prostate and non-prostatic tissues. *Cancer Letters* **236**(2): 229-238.
Davis, M. I., Bennett, M. J., Thomas, L. M. and Bjorkman, P. J. (2005) Crystal structure of prostate-specific membrane antigen, a tumor marker and peptidase. *Proceedings of the National Academy of Sciences of the USA* **102**(17): 5981-5986.
DeLano, W.L. (2002) The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, USA.
Donelly, J.G. (2001) Folic acid. *Critical Reviews in Clinical Laboratory Sciences* **38**: 183-223.
Fischer, W. H. and Spiess, J. (1987) Identification of a mammalian glutaminyl cyclase converting glutaminyl into pyroglutamyl peptides. *Proceedings of the National Academy of Sciences of the USA* **84**(11): 3628-3632.
Fracasso, G., Bellisola, G., Cingarlini, S., Castelletti, D., Prayer-Galetti, T., Pagano, F., Tridente, G. and Colombatti, M. (2002) Anti-tumor effects of toxins targeted to the prostate specific membrane antigen. *Prostate* **53**(1): 9-23.
Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J. A., Jr., Vreven, T., K. N. Kudin, K. N., Burant, J. C*., et al.* (2004). Gaussian03, Gaussian, Inc., Wallingford USA.
Ghosh, A. and Heston, W. D. (2003) Effect of carbohydrate moieties on the folate hydrolysis activity of the prostate specific membrane antigen. *Prostate* **57**(2): 140-151.
Greenblatt, H. M., Almog, O., Maras, B., Spungin-Bialik, A., Barra, D., Blumberg, S. and Shoham, G. (1997) Streptomyces griseus aminopeptidase: X-ray crystallographic structure at 1.75 A resolution. *Journal of Molecular Biology* **265**(5): 620-636.
Grimme, S. (2006) Semiempirical hybrid density functional with perturbative second-order correlation. *Journal of Chemical Physics* **124**(3): 034108.
Hattig, C. (2005) Optimization of auxiliary basis sets for RI-MP2 and RI-CC2 calculations: Core-valence and quintuple-zeta basis sets for H to Ar and QZVPP basis sets for Li to Kr. *Physical Chemistry Chemical Physics* **7**(1): 59-66.
Hattig, C. and Weigend, F. (2000) CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. *Journal of Chemical Physics* **113**(13): 5154-5161.
Haydon, P. G. (2001) GLIA: listening and talking to the synapse. *Nature Reviews Neuroscience* **2**(3): 185-193.
Hellweg, A., Hattig, C., Hofener, S. and Klopper, W. (2007) Optimized accurate auxiliary basis sets for RI-MP2 and RI-CC2 calculations for the atoms Rb to Rn. *Theoretical Chemistry Accounts* **117**(4): 587-597.
Heston, W. D. (1997) Characterization and glutamyl preferring carboxypeptidase function of prostate specific membrane antigen: a novel folate hydrolase. *Urology* **49**(3A Suppl): 104-112.

Hicke, L. (2001) Protein regulation by monoubiquitin. *Nature Reviews Molecular Cell Biology* **2**(3): 195-201.

Hlouchová, K., Bařinka, C., Klusák, V., Šácha, P., Mlčochová, P., Majer, P., Rulíšek, L. and Konvalinka, J. (2007) Biochemical characterization of human glutamate carboxypeptidase III. *Journal of Neurochemistry* **101**(3): 682-696.

Hlouchová, K., Bařinka, C., Konvalinka, J. and Lubkowski, J. (2009) Structural insight into the evolutionary and pharmacologic homology of glutamate carboxypeptidases II and III. *Febs Journal* **276**(16): 4448-4462.

Hohenberg, P. and Kohn, W. (1964) Inhomogeneous Electron Gas. *Physical Review B* **136**(3B): B864-B871.

Holm, L. and Park, J. (2000) DaliLite workbench for protein structure comparison. *Bioinformatics* **16**(6): 566-567

Holmes, E. H., Greene, T. G., Tino, W. T., Boynton, A. L., Aldape, H. C., Misrock, S. L. and Murphy, G. P. (1996) Analysis of glycosylation of prostate-specific membrane antigen derived from LNCaP cells, prostatic carcinoma tumors, and serum from prostate cancer patients. *Prostate Supplement* **7**: 25-29.

Horoszewicz, J. S., Kawinski, E. and Murphy, G. P. (1987) Monoclonal antibodies to a new antigenic marker in epithelial prostatic cells and serum of prostatic cancer patients. *Anticancer Research* **7**(5B): 927-935.

Howell, T., Hailey, D., *Use of In-111 Capromab Pentetide in detecting metastatic prostate cancer*. Alberta Heritage Foundation for Medical Research. 1999: Alberta.

Hu, H. and Yang, W. T. (2008) Free energies of chemical reactions in solution and in enzymes with ab initio quantum mechanics/molecular mechanics methods. *Annual Review of Physical Chemistry* **59**: 573-601.

Huang, K. F., Liu, Y. L., Cheng, W. J., Ko, T. P. and Wang, A. H. (2005) Crystal structures of human glutaminyl cyclase, an enzyme responsible for protein N-terminal pyroglutamate formation. *Proceedings of the National Academy of Sciences of the USA* **102**(37): 13117-13122.

Chandler, D. (1987) Introduction to Modern Statistical Mechanics, *Oxford University Press*.

Chang, S. S., O'Keefe, D. S., Bacich, D. J., Reuter, V. E., Heston, W. D. and Gaudin, P. B. (1999a) Prostate-specific membrane antigen is produced in tumor-associated neovasculature. *Clinical Cancer Research* **5**(10): 2674-2681.

Chang, S. S., Reuter, V. E., Heston, W. D., Bander, N. H., Grauer, L. S. and Gaudin, P. B. (1999b) Five different anti-prostate-specific membrane antigen (PSMA) antibodies confirm PSMA expression in tumor-associated neovasculature. *Cancer Research* **59**(13): 3192-3198.

Chen, S. L., Marino, T., Fang, W. H., Russo, N. and Himo, F. (2008) Peptide hydrolysis by the binuclear zinc enzyme aminopeptidase from Aeromonas proieolytica: A density functional theory study. *Journal of Physical Chemistry B* **112**(8): 2494-2500.

Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J. M., Moras, D. and Tarnus, C. (1994) Crystal structure of Aeromonas proteolytica aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. *Structure* **2**(4): 283-291.

Chu, T. M., Wang, M. C., Scott, W. W., Gibbons, R. P., Johnson, D. E., Schmidt, J. D., Loening, S. A., Prout, G. R. and Murphy, G. P. (1978) Immunochemical detection of serum prostatic acid phosphatase. Methodology and clinical evaluation. *Investigative Urology* **15**(4): 319-323.

Jackson, P. F., Cole, D. C., Slusher, B. S., Stetz, S. L., Ross, L. E., Donzanti, B. A. and Trainor, D. A. (1996) Design, synthesis, and biological activity of a potent inhibitor of the neuropeptidase N-acetylated alpha-linked acidic dipeptidase. *Journal of Medicinal Chemistry* **39**(2): 619-622.

Jakubowski, H. (1996) Proofreading in trans by an aminoacyl-tRNA synthetase: A model for single site editing by isoleucyl-tRNA synthetase. *Nucleic Acids Research* **24**(13): 2505-2510.

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T.*, et al.* (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Research* **36**: D480-D484.

Kaukonen, M., Soderhjelm, P., Heimdal, J. and Ryde, U. (2008) QM/MM-PBSA method to estimate free energies for reactions in proteins**.** *Journal of Physical Chemistry B* **112**(39): 12537-12548.

Kinoshita, Y., Kuratsukuri, K., Landas, S., Imaida, K., Rovito, P. M., Jr., Wang, C. Y. and Haas, G. P. (2006) Expression of prostate-specific membrane antigen in normal and malignant human tissues. *World Journal of Surgery* **30**(4): 628-636.

Klamt, A. and Schuurmann, G. (1993) Cosmo - a New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *Journal of the Chemical Society-Perkin Transactions 2*(5): 799-805.

Klusák, V., Bařinka, C., Plechanovová, A., Mlčochová, P., Konvalinka, J., Rulíšek, L. and Lubkowski, J. (2009) Reaction mechanism of glutamate carboxypeptidase II revealed by mutagenesis, X-ray crystallography, and computational methods. *Biochemistry* **48**(19): 4126-4138.

Kohn, W. and Sham, L. J. (1965) Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review* **140**(4A): 1133-1138.

Koch, W. and Holthausen, M. C. (2001) A Chemist's Guide to Density Functional Theory, *Wiley-VCH*.

Kožíšek, M., Bray, J., Řezácová, P., Šašková, K., Brynda, J., Pokorná, J., Mammano, F., Rulíšek, L. and Konvalinka, J. (2007) Molecular analysis of the HIV-1 resistance development: enzymatic activities, crystal structures, and thermodynamics of nelfinavir-resistant HIV protease mutants. *Journal of Molecular Biology* **374**(4): 1005-1016.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* **305**(3): 567-580.

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E. and Cech, T. R. (1982) Self-Splicing Rna - Auto-Excision and Auto-Cyclization of the Ribosomal-Rna Intervening Sequence of Tetrahymena. *Cell* **31**(1): 147-157.

Lawrence, C. M., Ray, S., Babyonyshev, M., Galluser, R., Borhani, D. W. and Harrison, S. C. (1999) Crystal structure of the ectodomain of human transferrin receptor. *Science* **286**(5440): 779-782.

Lee, C. T., Yang, W. T. and Parr, R. G. (1988) Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Physical Review B* **37**(2): 785-789.

Levine, I. N. (1991) Quantum Chemistry, *Prentice-Hall*.

Liao, R. Z. (2009a). Mechanistic Insights into Dinuclear Zinc Enzymes from Density Functional Studies. *Licentiate Thesis*, *Department of Theoretical Chemistry.* Royal Institute of Technology, Stockholm.

Liao, R. Z., Yu, J. G. and Himo, F. (2009b) Reaction Mechanism of the Dinuclear Zinc Enzyme N-Acyl-L-homoserine Lactone Hydrolase: A Quantum Chemical Study. *Inorganic Chemistry* **48**(4): 1442-1448.

Lill, S. O. N. and Siegbahn, P. E. M. (2009) An Autocatalytic Mechanism for NiFe-Hydrogenase: Reduction to Ni(I) Followed by Oxidative Addition. *Biochemistry* **48**(5): 1056-1066.

Liu, H., Moy, P., Kim, S., Xia, Y., Rajasekaran, A., Navarro, V., Knudsen, B. and Bander, N. H. (1997) Monoclonal antibodies to the extracellular domain of prostate-specific membrane

antigen also react with tumor vascular endothelium. *Cancer Res* **57**(17): 3629-3634.
Liu, H., Rajasekaran, A. K., Moy, P., Xia, Y., Kim, S., Navarro, V., Rahmati, R. and Bander, N. H. (1998) Constitutive and antibody-induced internalization of prostate-specific membrane antigen. *Cancer Research* **58**(18): 4055-4060.
Majer, P., Jackson, P. F., Delahanty, G., Grella, B. S., Ko, Y. S., Li, W., Liu, Q., Maclin, K. M., Polakova, J., Shaffer, K. A*., et al.* (2003) Synthesis and biological evaluation of thiol-based inhibitors of glutamate carboxypeptidase II: discovery of an orally active GCP II inhibitor. *J Med Chem* **46**(10): 1989-1996.
Mcclain, W. H. (1993) Rules That Govern Transfer-RNA Identity in Protein-Synthesis. *Journal of Molecular Biology* **234**(2): 257-280.
McQuarrie, D. (2000) Statistical Mechanics, *University Science books*.
Merz, K. M. (1992) Analysis of a Large Data-Base of Electrostatic Potential Derived Atomic Charges. *Journal of Computational Chemistry* **13**(6): 749-767.
Mesters, J. R., Bařinka, C., Li, W., Tsukamoto, T., Majer, P., Slusher, B. S., Konvalinka, J. and Hilgenfeld, R. (2006) Structure of glutamate carboxypeptidase II, a drug target in neuronal damage and prostate cancer. *Embo Journal* **25**(6): 1375-1384.
Miller, R.T. (2005) Immunohistochemical markers of prostate carcinoma, including prostate-specific membrane antigen. in *The Focus Immunochemistry*, ProPath.
Miyake, H., Hara, I., Kurahashi, T., Inoue, T. A., Eto, H. and Fujisawa, M. (2007) Quantitative detection of micrometastases in pelvic lymph nodes in patients with clinically localized prostate cancer by real-time reverse transcriptase-PCR. *Clinical Cancer Research* **13**(4): 1192-1197.
Mlčochová, P., Bařinka, C., Tykvart, J., Šácha, P. and Konvalinka, J. (2009) Prostate-Specific Membrane Antigen and Its Truncated Form PSM'**.** *Prostate* **69**(5): 471-479.
Mlčochová, P., Plechanovová, A., Bařinka, C., Mahadevan, D., Saldanha, J. W., Rulíšek, L. and Konvalinka, J. (2007) Mapping of the active site of glutamate carboxypeptidase II by site-directed mutagenesis. *Febs Journal* **274**(18): 4731-4741.
Modrek, B., Resch, A., Grasso, C. and Lee, C. (2001) Genome-wide detection of alternative splicing in expressed sequences of human genes**.** *Nucleic Acids Research* **29**(13): 2850-2859.
Morris, M. J., Divgi, C. R., Pandit-Taskar, N., Batraki, M., Warren, N., Nacca, A., Smith-Jones, P., Schwartz, L., Kelly, W. K., Slovin, S*., et al.* (2005) Pilot trial of unlabeled and indium-111-labeled anti-prostate-specific membrane antigen antibody J591 for castrate metastatic prostate cancer. *Clinical Cancer Research* **11**(20): 7454-7461.
Mount, S. M. (1982) A Catalog of Splice Junction Sequences**.** *Nucleic Acids Research* **10**(2): 459-472.
Murphy, G. P., Barren, R. J., Erickson, S. J., Bowes, V. A., Wolfert, R. L., Bartsch, G., Klocker, H., Pointner, J., Reissigl, A., McLeod, D. G*., et al.* (1996) Evaluation and comparison of two new prostate carcinoma markers. Free-prostate specific antigen and prostate specific membrane antigen. *Cancer* **78**(4): 809-818.
Navrátil, V. (2007) Lidská glutamátkarboxypeptidasa II a její paralogy. *Bakalářská práce, přírodovědecká Fakulta.* Univerzita Karlova v Praze
Navrátil, V., Klusák, V. and Rulíšek, L. (2009) *Rukopis v přípravě.*
Neale, J. H., Bzdega, T. and Wroblewska, B. (2000) N-Acetylaspartylglutamate: the most abundant peptide neurotransmitter in the mammalian central nervous system. *Journal of Neurochemistry* **75**(2): 443-452.
Nedergaard, M., Takano, T. and Hansen, A. J. (2002) Beyond the role of glutamate as a neurotransmitter. *Nature Reviews Neuroscience* **3**(9): 748-755.
Nelson, D. L. and Cox, M. M. (2004) Principles of Biochemistry, *W. H. Freeman*.
O'Keefe, D.S., Bacich, D., and Heston, W. (2001) Prostate specific membrane antigen.

in *Prostate cancer: biology, genetics and the new therapeutics*, ed. Chung, W.K., Issacs, W.B. and Simons, J.W. Humana Press. 307-326.
Pajares, M. J., Ezponda, T., Catena, R., Calvo, A., Pio, R. and Montuenga, L. M. (2007) Alternative splicing: an emerging topic in molecular and clinical oncology**.** *Lancet Oncology* **8**(4): 349-357.
Parr, R. G. and Yang, W. (1994) Density-Functional Theory of Atoms and Molecules, *Oxford University Press*.
Perdew, J. P., Burke, K. and Ernzerhof, M. (1996) Generalized gradient approximation made simple. *Physical Review Letters* **77**(18): 3865-3868.
Perner, S., Hofer, M. D., Kim, R., Shah, R. B., Li, H., Moller, P., Hautmann, R. E., Gschwend, J. E., Kuefer, R. and Rubin, M. A. (2007) Prostate-specific membrane antigen expression as a predictor of prostate cancer progression. *Human Pathology* **38**(5): 696-701.
Perozzo, R., Jelesarov, I., Bosshard, H. R., Folkers, G. and Scapozza, L. (2000) Compulsory order of substrate binding to herpes simplex virus type 1 thymidine kinase - A calorimetric study. *Journal of Biological Chemistry* **275**(21): 16139-16145.
Pinto, J. T., Suffoletto, B. P., Berzin, T. M., Qiao, C. H., Lin, S., Tong, W. P., May, F., Mukherjee, B. and Heston, W. D. (1996) Prostate-specific membrane antigen: a novel folate hydrolase in human prostatic carcinoma cells. *Clinical Cancer Research* **2**(9): 1445-1451.
Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J. and Barrett, A. J. (2008) MEROPS: the peptidase database. *Nucleic Acids Research* **36**(Database issue): D320-D325.
Robinson, M. B., Blakely, R. D., Couto, R. and Coyle, J. T. (1987) Hydrolysis of the brain dipeptide N-acetyl-L-aspartyl-L-glutamate. Identification and characterization of a novel N-acetylated alpha-linked acidic dipeptidase activity from rat brain. *Journal of Biological Chemistry* **262**(30): 14498-14506.
Ryde, U. (1996) The coordination of the catalytic zinc ion in alcohol dehydrogenase studied by combined quantum-chemical and molecular mechanics calculations. *Journal of Computer-Aided Molecular Design* **10**(2): 153-164.
Šácha, P., Zámečník, J., Bařinka, C., Hlouchová, K., Vícha, A., Mlčochová, P., Hilgert, I., Eckschlager, T. and Konvalinka, J. (2007) Expression of glutamate carboxypeptidase II in human brain. *Neuroscience* **144**(4): 1361-1372.
Sasaki, F. (1977) Effectiveness of Configuration Interaction Calculations for Large Molecules. *International Journal of Quantum Chemistry*: 125-130.
Sayers, E. W., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., DiCuccio, M., Edgar, R., Federhen, S.*, et al.* (2009) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research* **37**(Database issue): D5-D15.
Senn, H. M. and Thiel, W. (2007) QM/MM methods for biological systems. *Atomistic Approaches in Modern Biology: From Quantum Chemistry to Molecular Simulations* **268**: 173-290.
Shemarova, I. V. (2009) cAMP-dependent signal pathways in unicellular eukaryotes. *Critical Reviews in Microbiology* **35**(1): 23-42.
Schäfer, A., Horn, H. and Ahlrichs, R. (1992) Fully Optimized Contracted Gaussian-Basis Sets for Atoms Li to Kr. *Journal of Chemical Physics* **97**(4): 2571-2577.
Schäfer, A., Klamt, A., Sattel, D., Lohrenz, J. C. W. and Eckert, F. (2000) COSMO Implementation in TURBOMOLE: Extension of an efficient quantum chemical code towards liquid systems. *Physical Chemistry Chemical Physics* **2**(10): 2187-2193.
Schiffer, E. (2007) Biomarkers for prostate cancer. *World J Urol* **25**(6): 557-562.
Schilling, S., Hoffmann, T., Manhart, S., Hoffmann, M. and Demuth, H. U. (2004) Glutaminyl cyclases unfold glutamyl cyclase activity under mild acid conditions. *Febs Letters* **563**(1-3): 191-196.

Schmittgen, T. D., Teske, S., Vessella, R. L., True, L. D. and Zakrajsek, B. A. (2003) Expression of prostate specific membrane antigen and three alternatively spliced variants of PSMA in prostate cancer patients. *International Journal of Cancer* **107**(2): 323-329.

Schulke, N., Varlamova, O. A., Donovan, G. P., Ma, D., Gardner, J. P., Morrissey, D. M., Arrigale, R. R., Zhan, C., Chodera, A. J., Surowitz, K. G.*, et al.* (2003) The homodimer of prostate-specific membrane antigen is a functional target for cancer therapy. *Proceedings of the National Academy of Sciences of the USA* **100**(22): 12590-12595.

Schwabe, T. and Grimme, S. (2007) Double-hybrid density functionals with long-range dispersion corrections: higher accuracy and extended applicability. *Physical Chemistry Chemical Physics* **9**(26): 3397-3406.

Siegbahn, P. E. M. (2007) Hybrid density functional study of the oxidized states of NiFe-hydrogenase. *Comptes Rendus Chimie* **10**(8): 766-774.

Siegbahn, P. E. M. and Himo, F. (2009) Recent developments of the quantum chemical cluster approach for modeling enzyme reactions. *Journal of Biological Inorganic Chemistry* **14**(5): 643-651.

Silver, D. A., Pellicer, I., Fair, W. R., Heston, W. D. and Cordon-Cardo, C. (1997) Prostate-specific membrane antigen expression in normal and malignant human tissues. *Clinical Cancer Research* **3**(1): 81-85.

Slusher, B. S., Vornov, J. J., Thomas, A. G., Hurn, P. D., Harukuni, I., Bhardwaj, A., Traystman, R. J., Robinson, M. B., Britton, P., Lu, X. C.*, et al.* (1999) Selective inhibition of NAALADase, which converts NAAG to glutamate, reduces ischemic brain injury. *Nature Medicine* **5**(12): 1396-1402.

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y. S., Zhang, Z. Y., Toiber, D., Thanaraj, T. A. and Soreq, H. (2005) Function of alternative splicing. *Gene* **344**: 1-20.

Štrajbl, M., Florián, J. and Warshel, A. (2000) Ab initio/LD studies of chemical reactions in solution: Reference free-energy surfaces for acylation reactions occurring in serine and cysteine proteases. *International Journal of Quantum Chemistry* **77**(1): 44-53.

Su, S. L., Huang, I. P., Fair, W. R., Powell, C. T. and Heston, W. D. (1995) Alternatively spliced variants of prostate-specific membrane antigen RNA: ratio of expression as a potential measurement of progression. *Cancer Research* **55**(7): 1441-1443.

Szabo, A. and Ostlund, N. S. (1989) Modern Quantum Chemistry, *McGraw-Hill*.

Tao, J. M., Perdew, J. P., Staroverov, V. N. and Scuseria, G. E. (2003) Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids. *Physical Review Letters* **91**(14): 146401.

Truhlar, D. G., Garrett, B. C. and Klippenstein, S. J. (1996) Current status of transition-state theory. *Journal of Physical Chemistry* **100**(31): 12771-12800.

Volterra, A. and Meldolesi, J. (2005) Astrocytes, from brain glue to communication elements: the revolution continues. *Nature Reviews Neuroscience* **6**(8): 626-640.

Vosko, S. H., Wilk, L. and Nusair, M. (1980) Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin-Density Calculations - a Critical Analysis. *Canadian Journal of Physics* **58**(8): 1200-1211.

Vreven, T., Byun, K. S., Komaromi, I., Dapprich, S., Montgomery, J. A., Morokuma, K. and Frisch, M. J. (2006) Combining quantum mechanics methods with molecular mechanics methods in ONIOM. *Journal of Chemical Theory and Computation* **2**(3): 815-826.

Wang, J. M., Cieplak, P. and Kollman, P. A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry* **21**(12): 1049-1074.

Wang, M. C., Papsidero, L. D., Kuriyama, M., Valenzuela, L. A., Murphy, G. P. and Chu, T. M. (1981) Prostate antigen: a new potential marker for prostatic cancer. *Prostate* **2**(1): 89-96.

Weigend, F. (2006) Accurate Coulomb-fitting basis sets for H to Rn. *Physical Chemistry*

Chemical Physics **8**(9): 1057-1065.
Weigend, F. and Ahlrichs, R. (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics* **7**(18): 3297-3305.
Weigend, F., Furche, F. and Ahlrichs, R. (2003) Gaussian basis sets of quadruple zeta valence quality for atoms H-Kr. *Journal of Chemical Physics* **119**(24): 12753-12762.
Weigend, F., Haser, M., Patzelt, H. and Ahlrichs, R. (1998) RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chemical Physics Letters* **294**(1-3): 143-152.
Wheelan, S. J., Church, D. M. and Ostell, J. M. (2001) Spidey: a tool for mRNA-to-genomic allignments. *Genome Research* **11**(11):1952-1957
Wright, G. L., Jr., Grob, B. M., Haley, C., Grossman, K., Newhall, K., Petrylak, D., Troyer, J., Konchuba, A., Schellhammer, P. F. and Moriarty, R. (1996) Upregulation of prostate-specific membrane antigen after androgen-deprivation therapy. *Urology* **48**(2): 326-334.
Xu, Q., Modrek, B. and Lee, C. (2002) Genome-wide detection of tissue-specific alternative splicing in the human transcriptome**.** *Nucleic Acids Research* **30**(17): 3754-3766.
Yarkony, D. R., Ed. (1995) Modern Electronic Structure Theory, *World Scientific*.
Zhang, L., Wang, C. Y., Yang, R., Shi, J., Fu, R., Chen, L., Klocker, H. and Zhang, J. (2008) Real-time quantitative RT-PCR assay of prostate-specific antigen and prostate-specific membrane antigen in peripheral blood for detection of prostate cancer micrometastasis. *Urologic Oncology* **26**(6): 634-640.

8. Seznam publikovaných a nepublikovaných prací

Navrátil, V. (2007) Lidská glutamátkarboxypeptidasa II a její paralogy. *Bakalářská práce, přírodovědecká Fakulta.* Univerzita Karlova v Praze

Navrátil, V., Vícha, A., Šácha, P. and Konvalinka, J. (2009) Alternative splicing of the human FOLH1 gene near its 5' termini. *Febs Journal* **276**: 331-331.