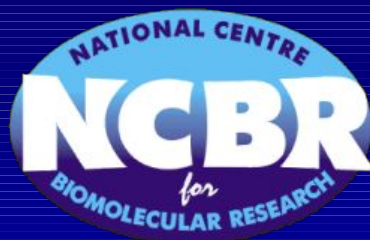


Racionální návrh proteinů

Martin Prokop

Národní centrum pro výzkum biomolekul
Přírodovědecká fakulta, Masarykova univerzita, Brno



Racionální návrh proteinů

Protein:

- známá struktura
- známá funkce

Modifikace proteinu:

- genetické techniky
- chemické techniky

Nový protein:

- modifikovaná struktura
- modifikovaná funkce

Jak modifikovat strukturu proteinu abychom dosáhli požadované změny jeho funkce?

Modifikované vlastnosti proteinu

Katalytické vlastnosti enzymů:

- zvýšení aktivity = zrychlení enzymové reakce
- ovlivnění specifity reakce
- změna pH optimum
- eliminace inhibičního centra
- eliminace residuí způsobujících nestabilitu

Strukturální vlastnosti proteinů:

- zlepšení termostability
- zlepšení stability v organických solventech
- modifikace specifity vázání ligandu

Vytvoření nových systémů:

- chimérické a multifunkční proteiny
- přidání značek pro purifikaci

Metody modifikace proteinů

- Místně-cílená mutageneze (site-directed)
- Chemické reakce na bočních řetězcích
- Totální syntéza proteinu
- Další metody

Místně-cílená mutagenese

- Záměna jednotlivých aminokyselin v DNA kódující protein
- Změnu provádíme pouze na místech klíčových pro danou funkci (reakční centrum enzymu, aktivní místo vázající ligand)

Výchozí informace:

- Sekvence DNA - sekvenování DNA - databáze sekvencí
- 3D struktura proteinu:
 - Krystalografie (rentgenová strukturní analýza)
 - NMR měření
 - Počítačové modelování (homologní modelování)
- Oblast proteinu klíčová pro modifikovanou funkci
 - Rtg. analýza s ligandem
 - Vyhledávání tunelů k reakčním centrům enzymů
 - Vyhledávání kavit na povrchu proteinu pro vázání ligandů

Experimentální metody návrhu proteinů

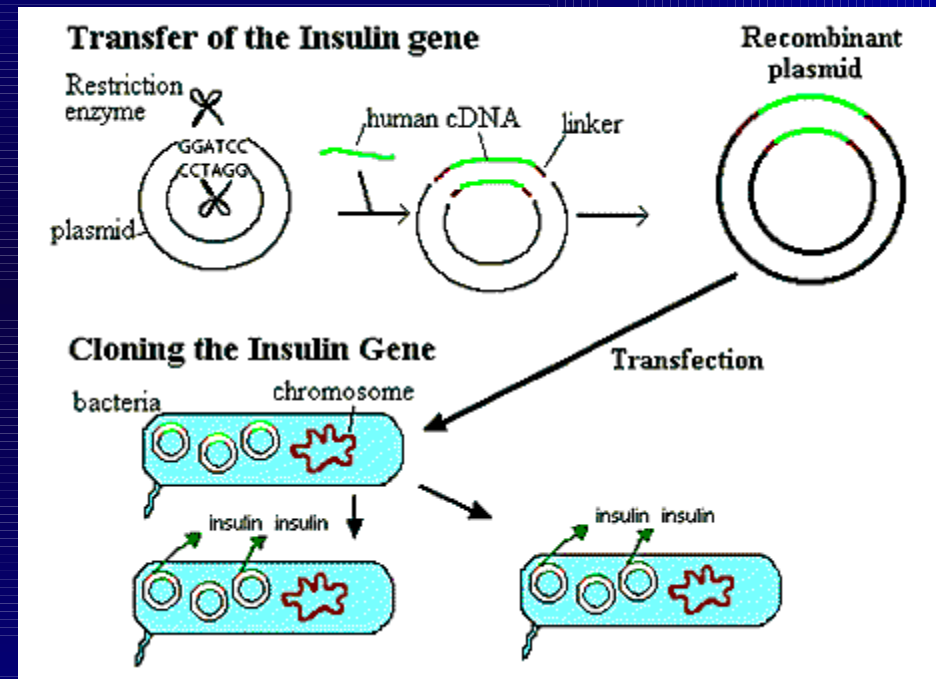
Experimentální příprava proteinu:

- syntéza oligonukleotidu
- příprava plasmidu
- vložení do buňky
- kultivace buněk
- izolace (purifikace proteinu)

Experimentální měření požadované vlastnosti:

- aktivita enzymu
- afinita receptor-ligand

Časová náročnost experimentálních metod => předvýběr vhodných mutantů počítačovými metodami



Počítačové metody návrhu proteinů

Modelování 3D struktury proteinu:

- homologní modelování

Počítačová simulace chování proteinu:

- modelování reakce enzymu => aktivita enzymu
- metody dockingu => vazebná energie ligand-protein

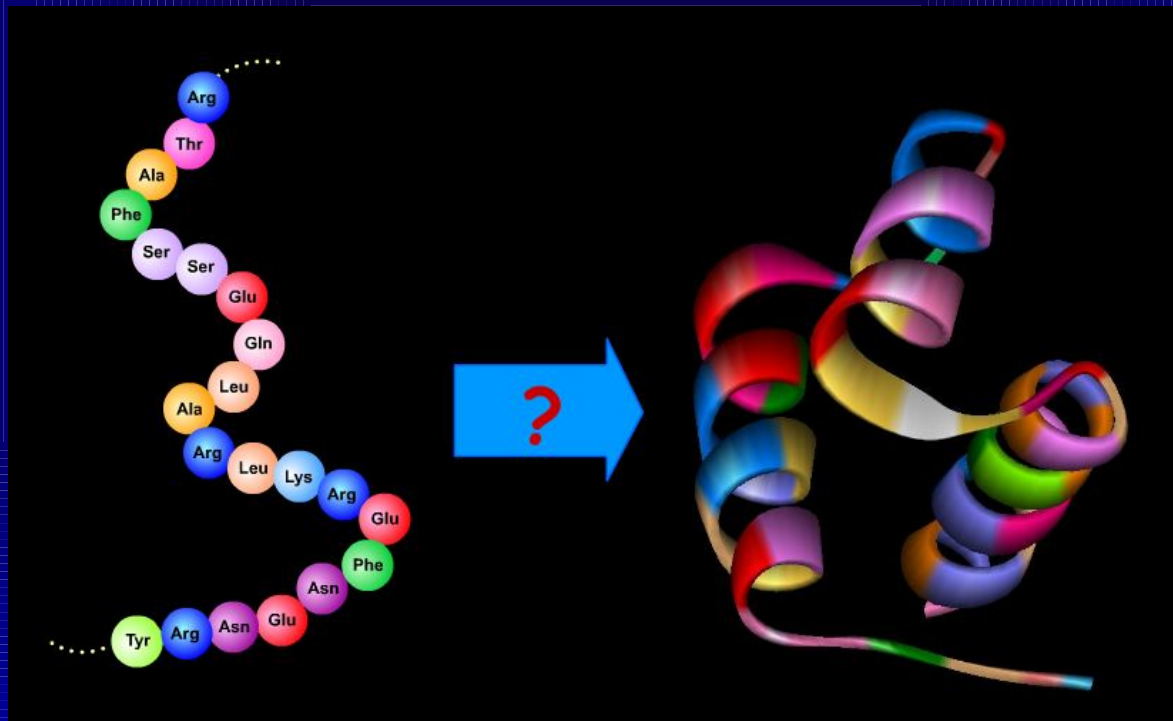
Homologní proteiny

- **evoluce organismů - mutace - změny v sekvenci aminokyselin proteinů**
- **3D struktura zůstává v průběhu evoluce konzervována více než sekvence aminokyselin (klíčová pro zachování funkce)**
- **proteiny vzniklé ze společného předka = homologní proteiny**
- **mají podobnou 3D strukturu a sekvenci**
- **větší variabilita aminokyselin na povrchu, hlavně ve smyčkách**
- **nejvíce jsou konzervovány úseky odpovědné za funkci proteinu (reakční centra, vazebná místa ligandů)**

Homologní modelování proteinů

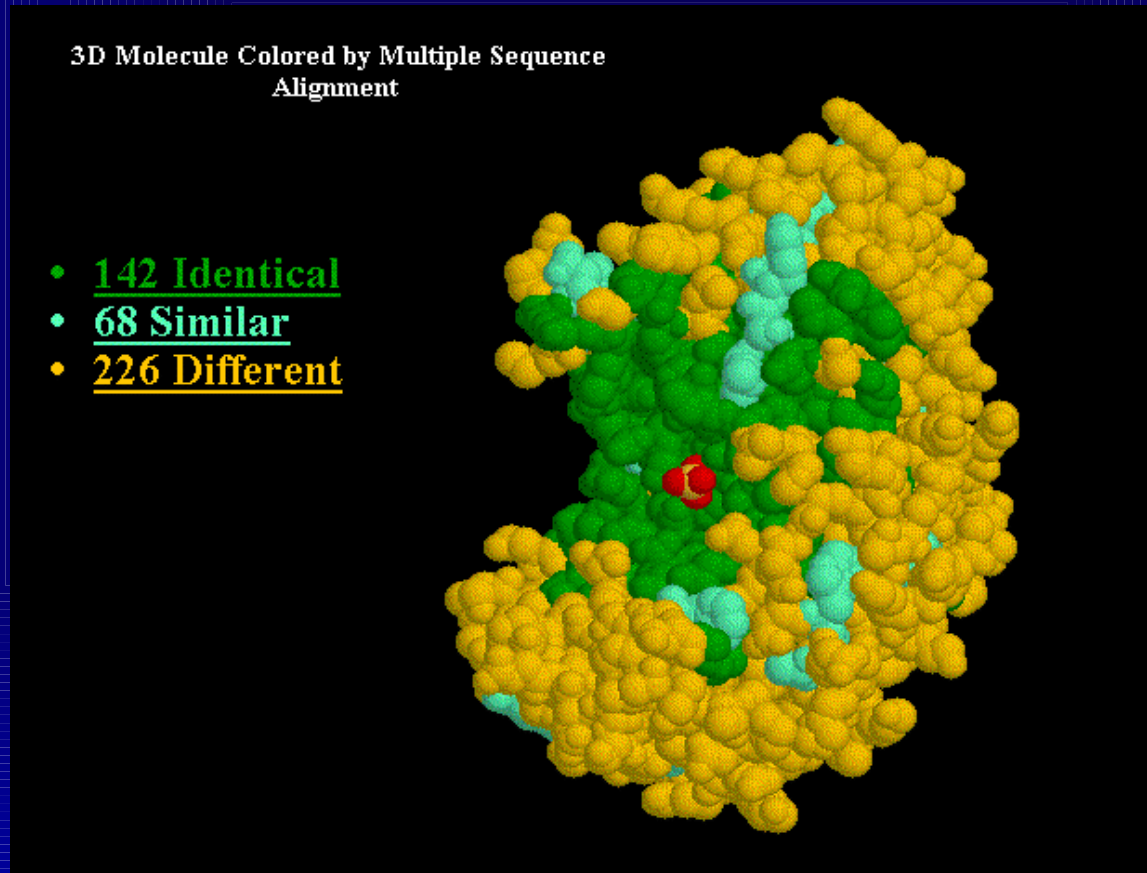
Slouží pro získání 3D struktury proteinu, u něhož známe pouze sekvenci aminokyselin (prot. se známou seq. > se známou 3D st.)

Nejdříve musíme najít vhodný homologní protein se známou (experimentálně zjištěnou) 3D strukturou



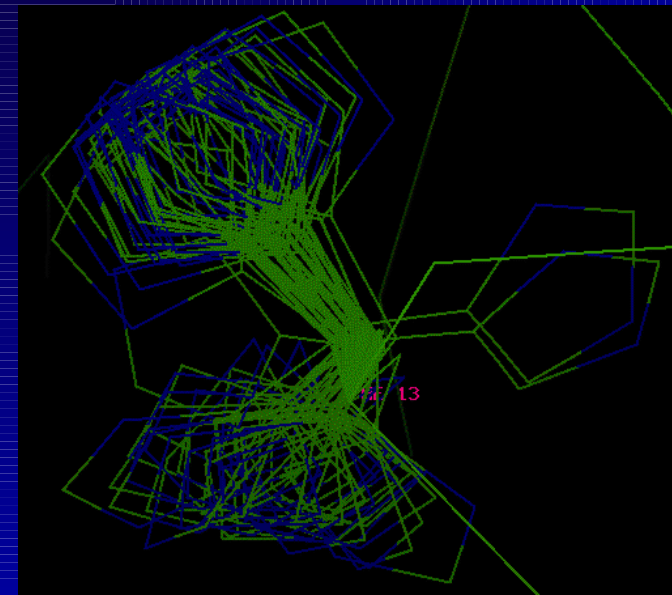
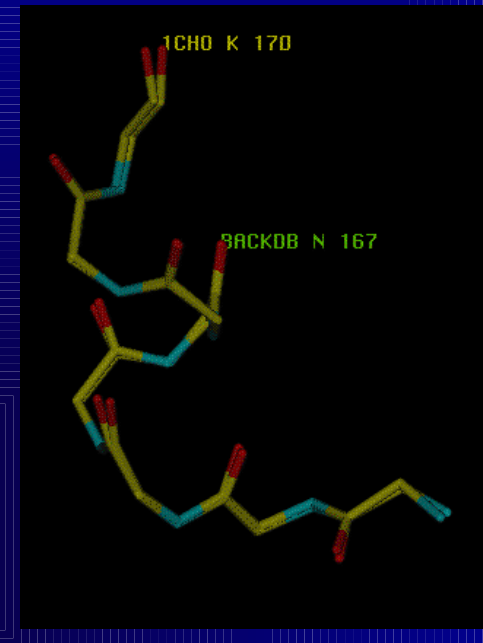
Postup při homologním modelování

1. Nalezení homologního proteinu (min. 30% homologie, optimálně 70%)
2. Identifikace strukturně konzervovaných regionů SCR a strukturně variabilních regionů SVR (srovnání sekvencí několika homolog. prot.)



Postup při homologním modelování

3. Přiložení sekvencí pro SCR
4. Konstrukce regionů SCR na základě koordinát šablonového proteinu (přikládají se na sebe atomy peptidické páteře)
5. Modelování SVR
6. Modelování postranních řetězců
 - konzervovaná residua - obvykle stejná pozice
 - ostatní residua - minimalizujeme prostorové kolize, vycházíme ze známých rotamerů



Srovnání sekvencí 6 dehalogenas

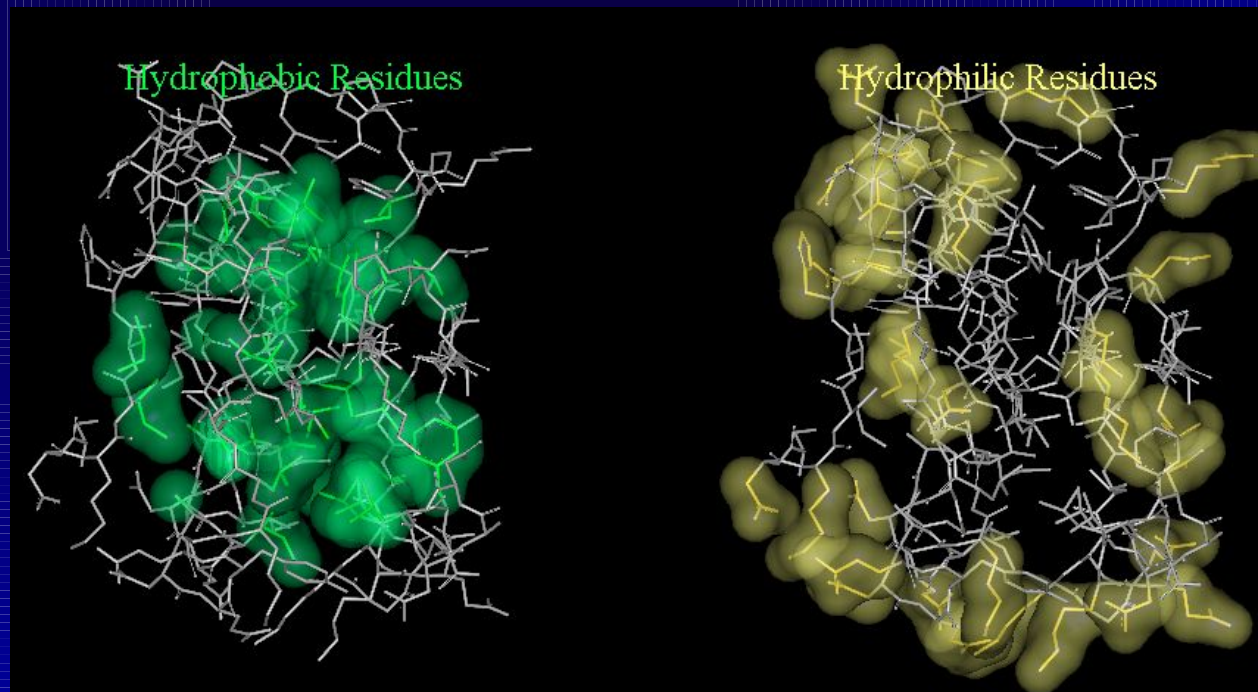
LinB	- - - - - M S L G A K - P F G E K K F I E I K - - - - - G R R M A Y I D E G - - T G D P I L F Q - H G N P T	40
Mtcy1	- - - - - M T A F G V E P Y G Q P K Y L - E I A - - - - - G K R M A Y I D E G K - - G D A I V F Q - H G N P T	41
DhaA	- - - - - M S E I G T G F P F D P - H Y V E V L - - - - - G E R M H Y V D V G P R D G T P V L F L - H G N P T	43
Mtcy3	- - - - - M S I D F T P D P Q L Y - P F E S R W F D S S R G - R I H Y V D E G T - - G P P I L L C - H G N P T	45
Mtcy2	M D - V L R T P D S R F E H L V G Y P F A P - H Y V D V T A G D T Q P L R M H Y V D E G P G D G P P I V L L - H G E P T	57
DhIA	M I N A I R T P D Q R F S N L D Q Y P F S P - N Y L D D L P G Y P G L - R A H Y L D E G N S D A E D V F L C L H G E P T	58
	$\beta 1$ $\beta 2$ $\beta 3$	
LinB	S S Y L W R N I M P H C A G L G - R L I A C D L I G M G D S D K L D P S - G P E R Y A Y A E H R D Y L D A L W E A L D L	98
Mtcy1	S S Y L W R N I M P H L E G L G - R L V A C D L I G M G A S D K L S P S - G P D R Y S Y G E Q R D F L F A L W D A L D L	99
DhaA	S S Y L W R N I I P H V A P S H - R C I A P D L I G M G K S D K - - P D - L - D - Y P F D D H V R Y L D A F I E A L G I	97
Mtcy3	W S F L Y R D I I V A L R D R F - R C V A P D Y L G F G L S E R - - P S - G F G - Y Q I D E H A R V I G E F V D H L G L	100
Mtcy2	W S Y L Y R T M I P P L S A A G H R V L A P D L I G F G R S D K - - P T R I E D - Y T Y L R H V E W V T S W F E N L D L	114
DhIA	W S Y L Y R K M I P V F A E S G A R V I A P D F F G F G K S D K - - P V D E E D - Y T F E F H R N F L L A L I E R L D L	115
	$\alpha 1$ Δ $\beta 4$ $\alpha 2$	
LinB	G D R V V L V V H D W G S A L G F D W A R R H R E R V O G I A Y M E A I A M - - P I E W A D F P E Q D - - - R D L F Q A	153
Mtcy1	G D H V V L V L H D W G S A L G F D W A N Q H R D R V O G I A F M E A I V T - - P M T W A D W P P A V - - - R G V F Q Q	154
DhaA	E E - V V L V I H D W G S A L G F H W A K R N P E R V K G I A C M E F I R P - I P - T W D E W P E F A - - - R E T F Q A	151
Mtcy3	- D R Y L S M G Q D W G G P I S M A V A V E R A D R V R G V V I G N - - - - - - - T W - F W P A D T L A M K - A F S R	149
Mtcy2	H D - Y T L F V Q D W G S L I G L R I A A E H G D R I A R L V V A N G - - - F L P A A Q G R T P L P F Y V W R - A F A R	169
DhIA	R N - I T L V V Q D W G G F L G L T L P M A D P S R F K R L I I M N A C L M T D P V T Q P A F S A F V T Q P A D G F T A	174
	$\beta 5$ $\alpha 3$ $\beta 6$ $\alpha 4$	
LinB	F R S Q A G E E L V L Q D - N V F V E Q V L P G L I L R P L S E A E M A A Y R E P F L A G G E A R R P T L S W P R O I P	212
Mtcy1	F R S P Q G E P M A L - E H N I F V E R V L P G A I L R Q L S D E E M N H Y R R P F V N G G E D R R P T L S W P R N L P	213
DhaA	F R T A D V G R E L I I D Q N A F I E G A L P K C V V R P L T E V E M D H Y R E P F L K P V D - R E P L W R F P N E L P	210
Mtcy3	V M S S P P V Q Y A I L R R N F F V E R L I P A G T E H R P S S A V M A H Y R A V Q P N A A A - R R G V A E M P K Q I -	207
Mtcy2	Y S P V L P A G R L V - - - N F G T V H R V P A G V - - R A G - - - - Y D A P F P D K T Y - Q A G A R A F P R L V P	217
DhIA	W K Y D L V T P S D L R L D Q F M K R - - - - W A P T - L T E A E A S A Y A A P F P D T S Y - Q A G Y R K F P K M V A	227
	$\alpha 5$ $\alpha 6$ Δ $\alpha 7$ $\alpha 8$	
LinB	I A G T P A D V V A I A R D Y A G W L S E S P I P K L F I N A E P G A L T T - G R M R D F C R T W P N Q - - - - T E I T	267
Mtcy1	I D G E P A E V V A L V N E Y R S W L E E T D M P K L F I N A E P G A I I T - G R I R D Y V R S W P N Q - - - - T E I T	268
DhaA	I A G E P A N I V A L V E A Y M N W L H Q S P V P K L L F W G T P G V L I P P A E A A R L A E S L P N C K - - - T V D I G	268
Mtcy3	L A A R P L L A R L A R E V P A T L G T - - - K P T L L I W G M K D V A F R - P K T I I P R L S A T F - - P D H V L V E	261
Mtcy2	T S P D D P A V P A N R A A W E A L G R W - D K P F L A I F G Y R D P I L G Q A D G P L I K H I P G A A G Q P H A R I K	276
DhIA	- Q R D Q A C I D I S T E A I S F W Q N D W N G Q T F M A I G M K D K L L G - P D V M Y P M K A L I N G C P E P L E I A	285
	Δ $\alpha 9$ $\beta 7$ $\alpha 10$ $\beta 8$	
LinB	V A - G A H F I Q E D S P D E I G A A I A A F V R R L R P A	
Mtcy1	V P - G V H F V Q E D S P E E I G A A I A Q F V R R L R S A A G V	
DhaA	- P - G L H Y L Q E D N P D L I G S E I A R W L P A L	
Mtcy3	L P N A K H F I Q E D A P D R I A A A I I E R F G	
Mtcy2	- - - A S H F I Q E D S G T E L A E R M L S W Q A T	
DhIA	D - - A G H F V Q E - F G E Q V A R E A L K H F A E T E	
	$\alpha 11$	

Postup při homologním modelování

7. Upřesnění struktury pomocí metod pro minimalizaci energie a molekulové dynamiky

8. Ověření kvality modelu

- zkoumáme distribuci hydrofobních residuí (ve vnitř. oblastech)
- detekujeme torsní úhly a atomové kontakty a srovnáváme s daty z databází



Využití homologním modelování pro návrh proteinů

- Jako templát slouží studovaný protein
- Zpravidla navrhujeme pouze několik bodových mutací => vysoká míra homologie => vysoká přesnost modelu
- Struktura podrobena QM minimalizaci resp. Simulaci MD => další zpřesnění struktury, hlavně postranních řetězců (netýká se dockingu)

Možné problémy:

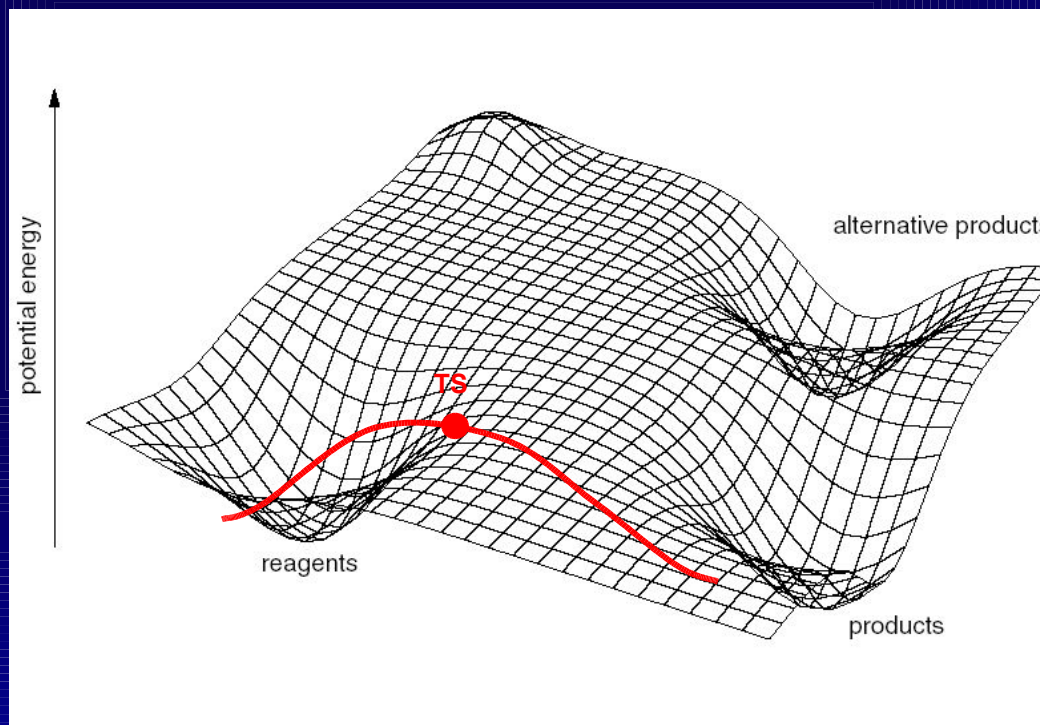
- některé mutace destabilizují protein - nedojde k jeho poskládání

Ověření vlastnosti modelovaného proteinu

- Docking - afinita nekovalentně vázaných ligandů
- Molekulová dynamika
- Modelování chemických reakcí

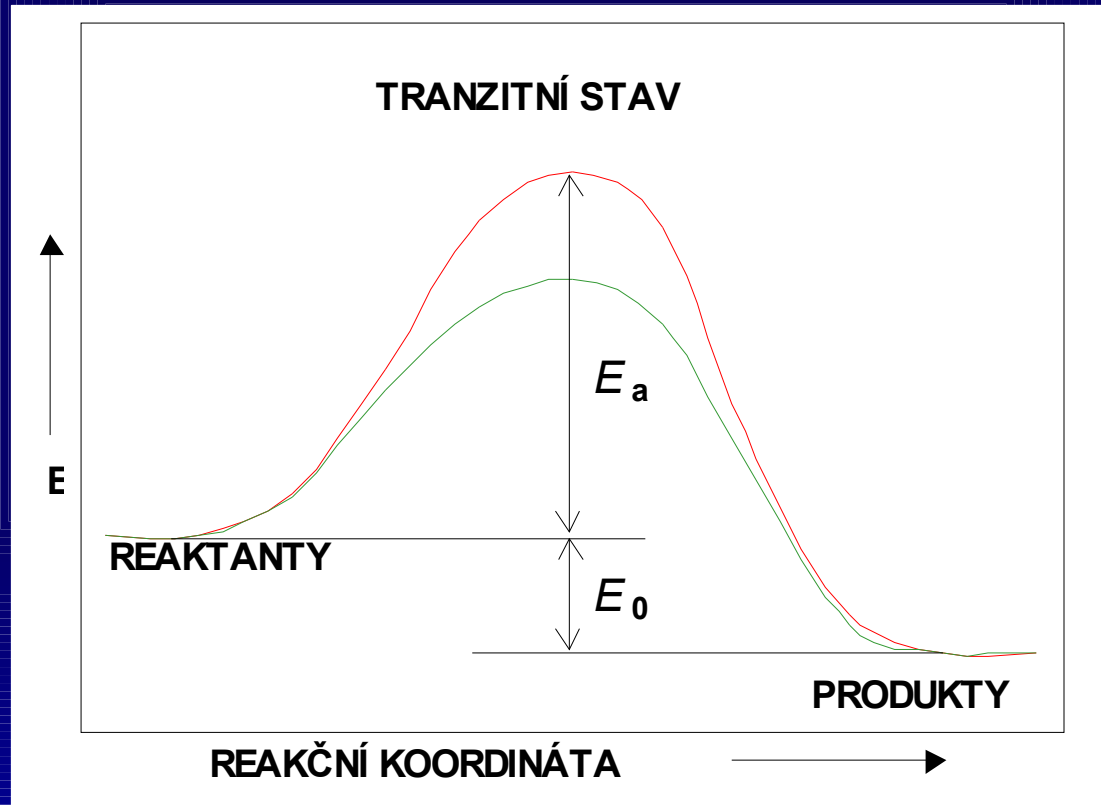
Modelování chemických reakcí

Chemické reakce - vznik a zánik kovalentních vazeb - pohyb po hyperploše potenciální energie



Chemická katalýza

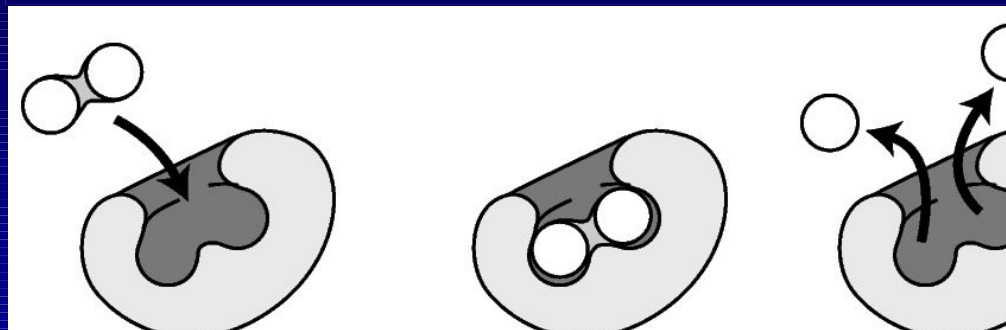
Snížení energie tranzitního stavu (aktivační energie) vlivem katalyzátoru



Enzymová katalýza

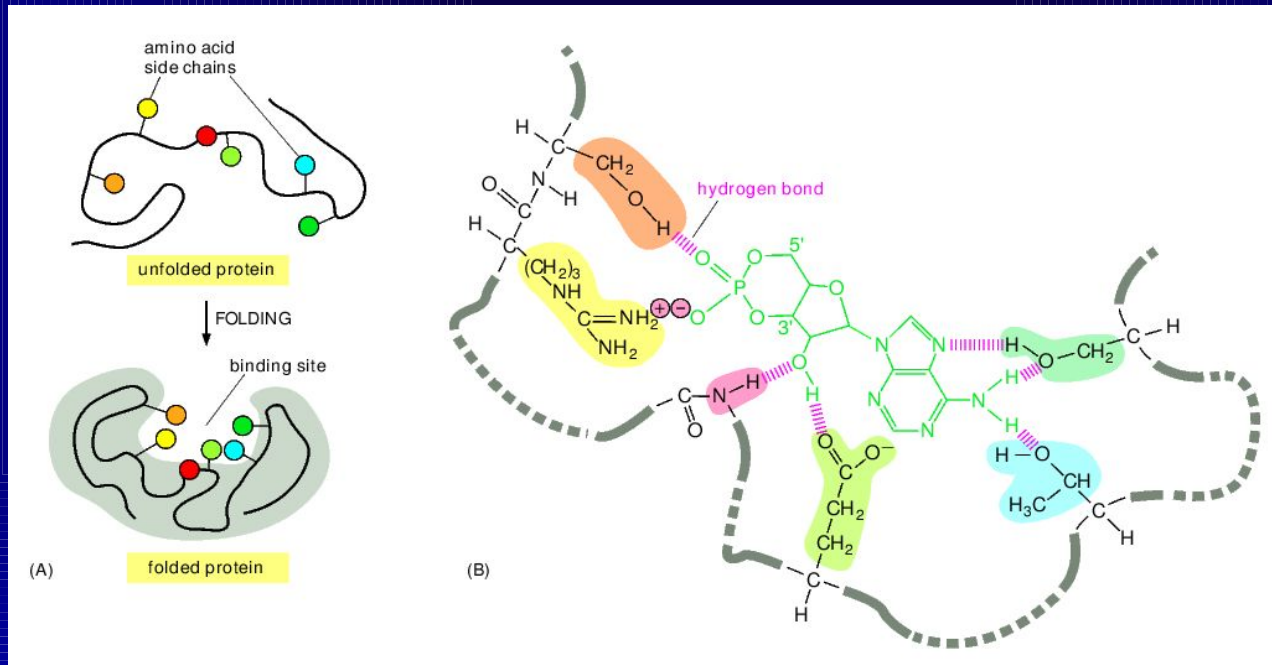
Enzymy:

- biologické katalyzátory
- globulární proteiny
- katalýza probíhá v aktivním centru
- snížení aktivační energie prostřednictvím nevazebných interakcí aktivního centra se substrátem
- roli hraje také dynamické chování enzymů



Nevazebné interakce v aktivních centrech enzymů

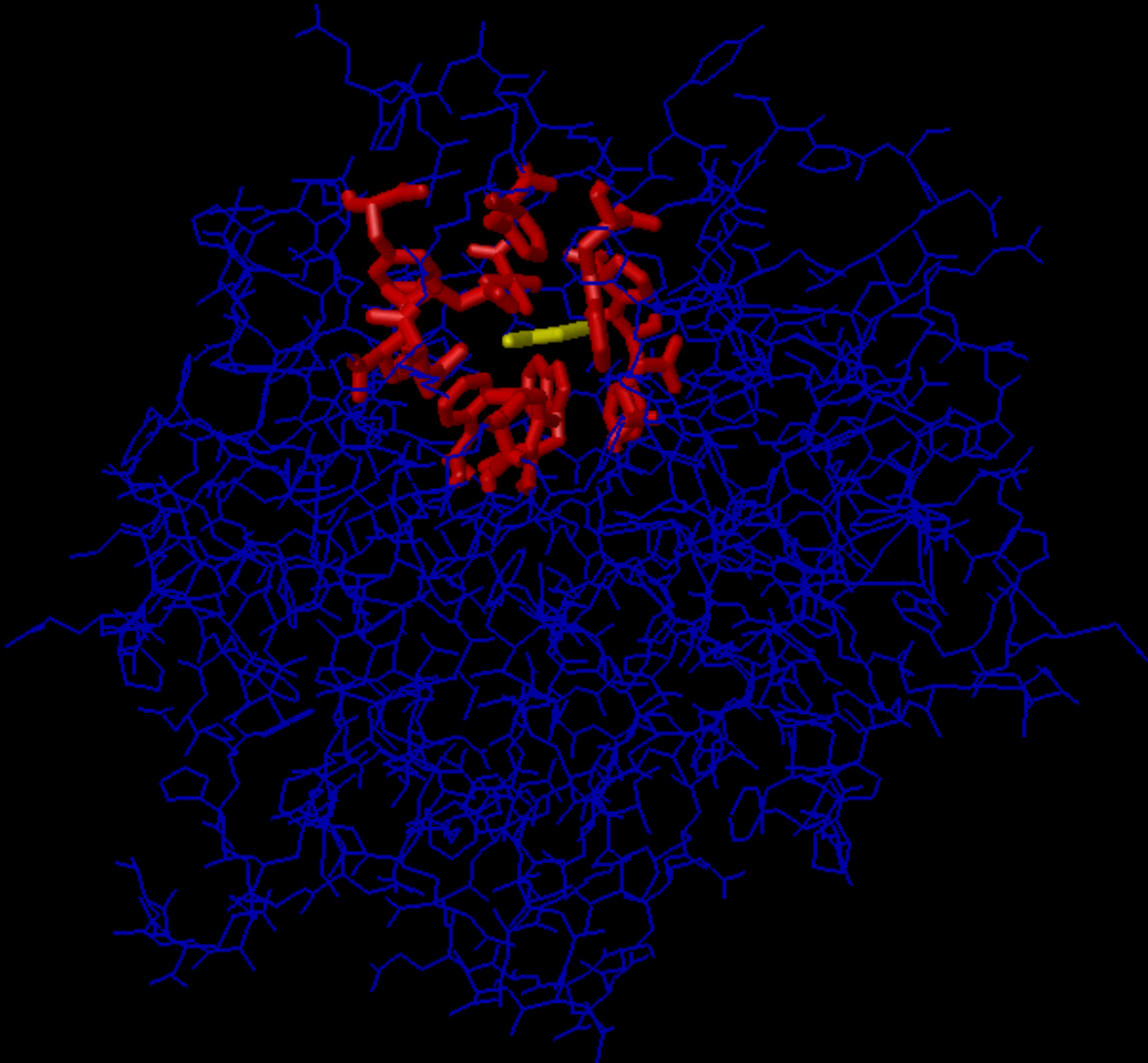
- vodíkové vazby
- elektrostatické interakce vč. dipolových
- van der Waalsovy interakce
- hydrofobní interakce
- stackingové interakce



Modelování enzymových reakcí

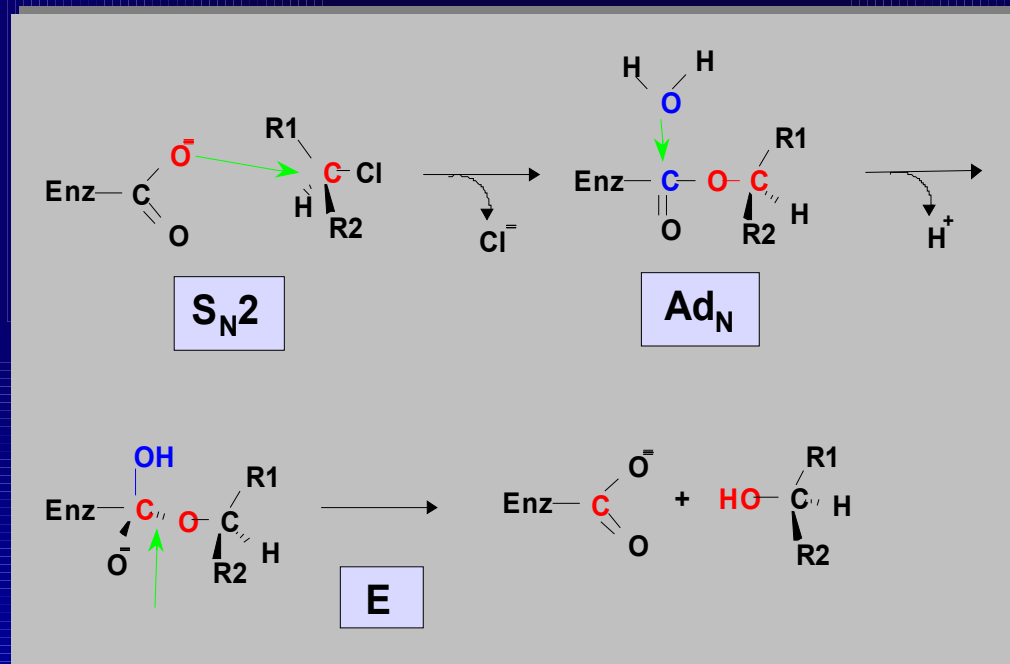
- vzniku nebo zánik kovalentních vazeb => nelze použít molekulovou mechaniku => kvantově-chemické metody
- velké molekuly - nelze použít *ab initio* metody => semiempirické kvantově-chemické metody (AM1, PM3)
- nelze zahrnout celou molekulu do výpočtu => použijeme aminokyseliny aktivního centra (kavita)
- napodobení situace v reálném proteinu - fixace atomů peptidické páteře

Aktivní centrum enzymu (kavita)

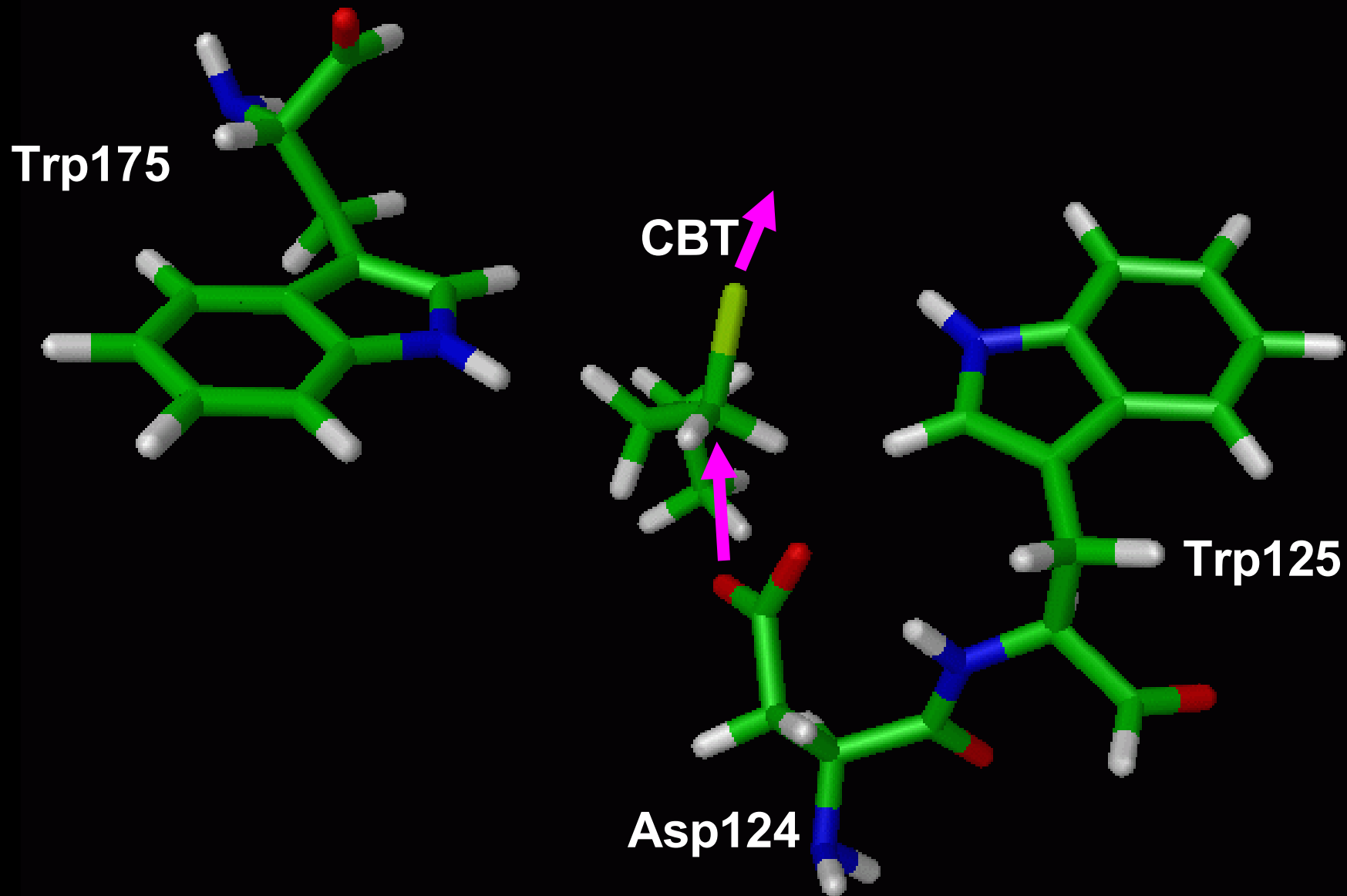


Haloalkan dehalogenasové enzymy

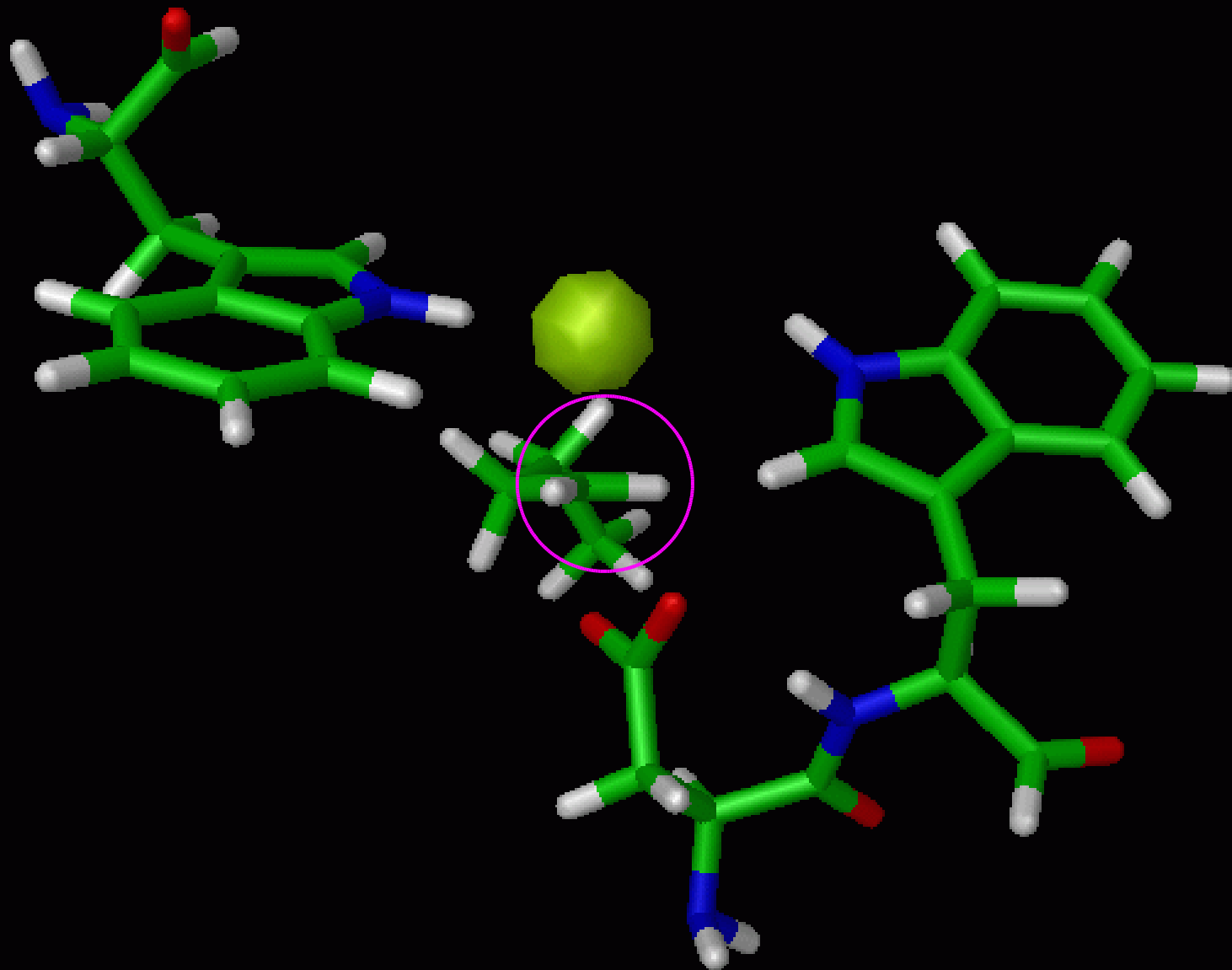
- Haloalkan dehalogenasy - odbourávají chlorované a bromované uhlovodíky
- Globulární protein ze skupiny α/β hydroláz - monomerní protein
- Tříkrokový mechanismus reakce
- Výpočet reakční cesty S_N2 kroku semiempirickými QM metodami



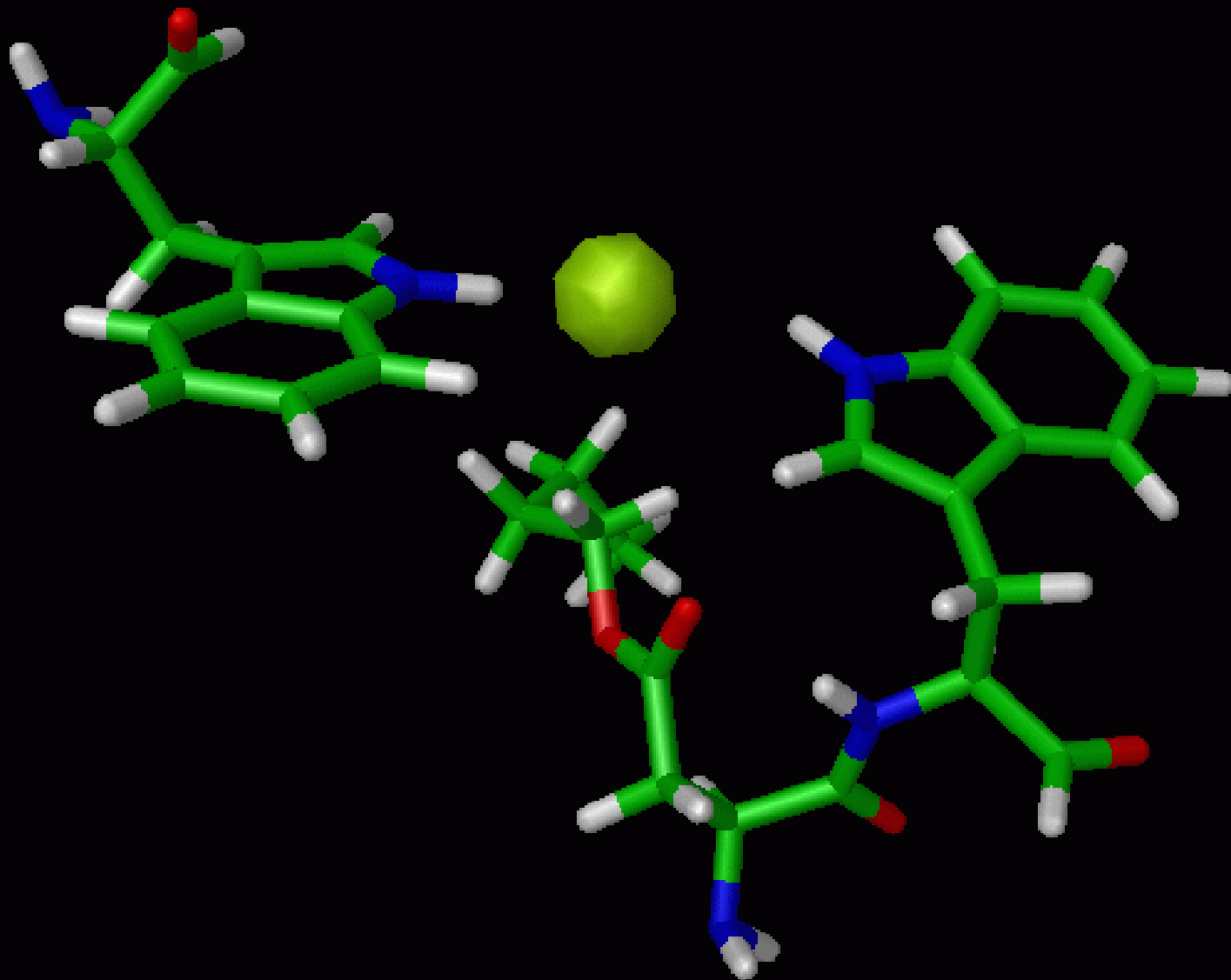
Reaktant



Tranzitní stav



Produkt



Výpočet reakční cesty

Program MOPAC - semiempirická QM metoda AM1

Program DRIVER - sledování reakční koordináty

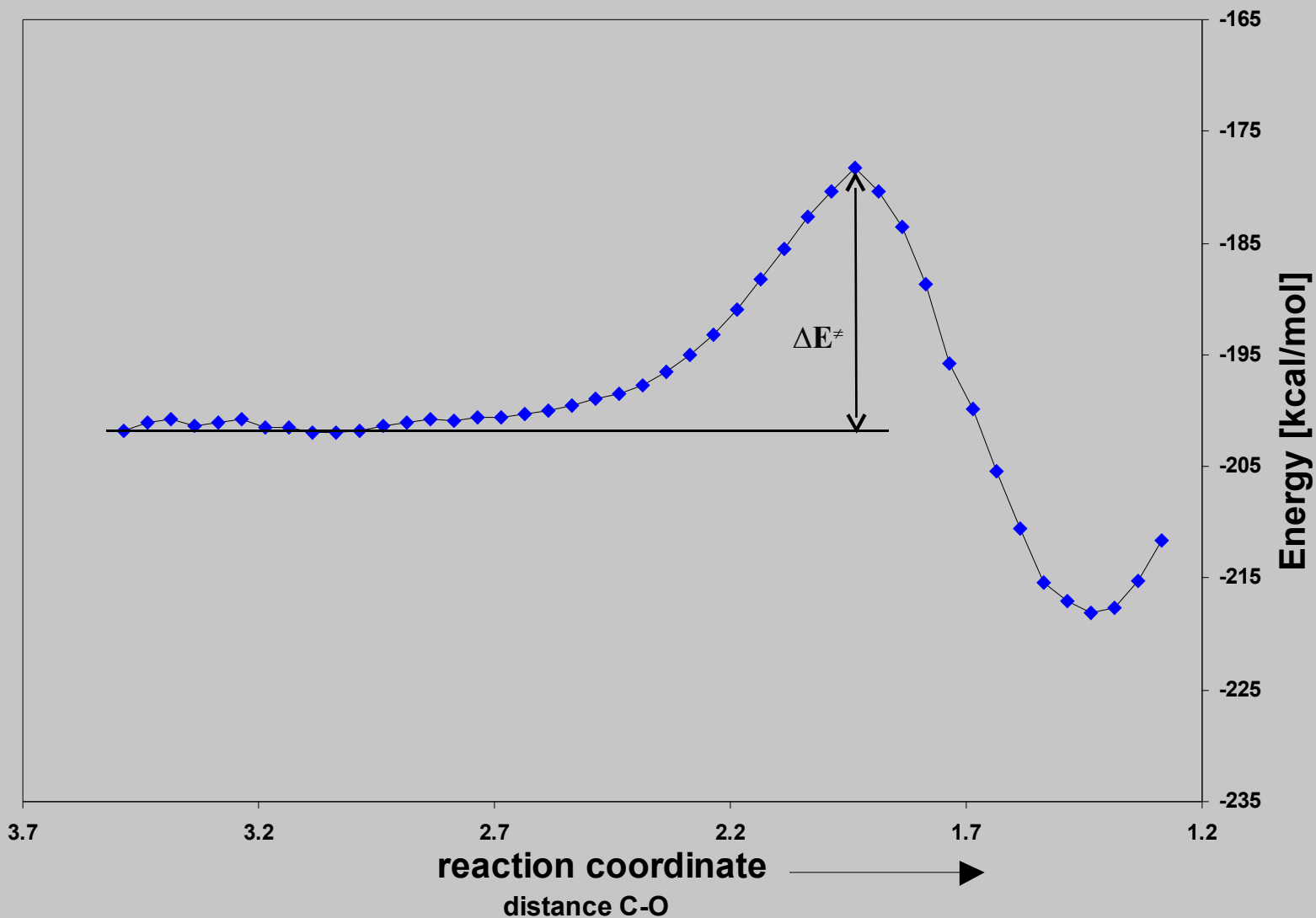
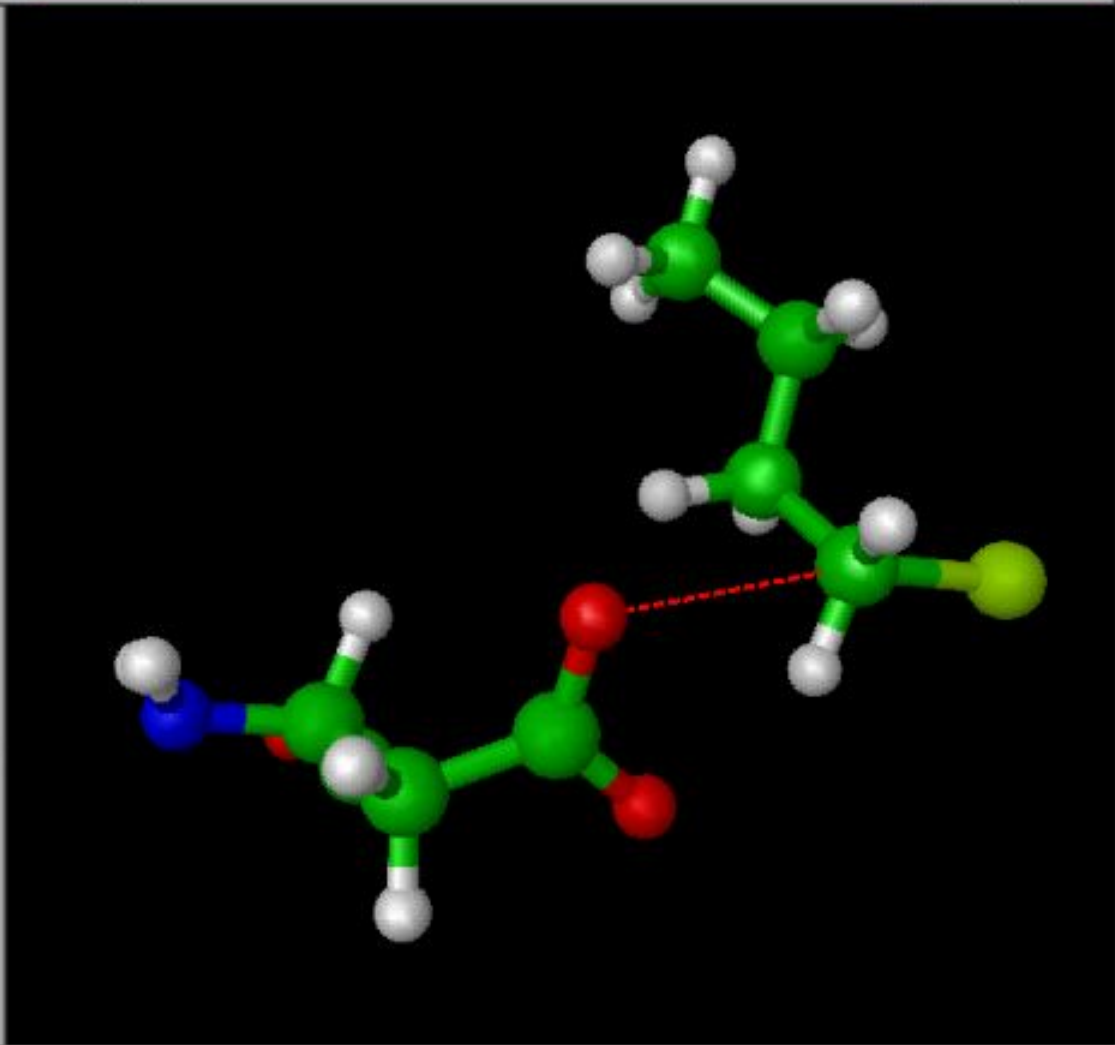
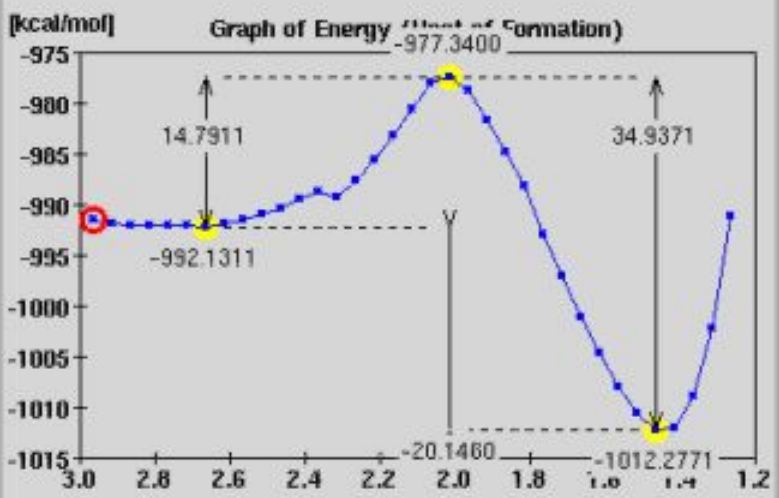
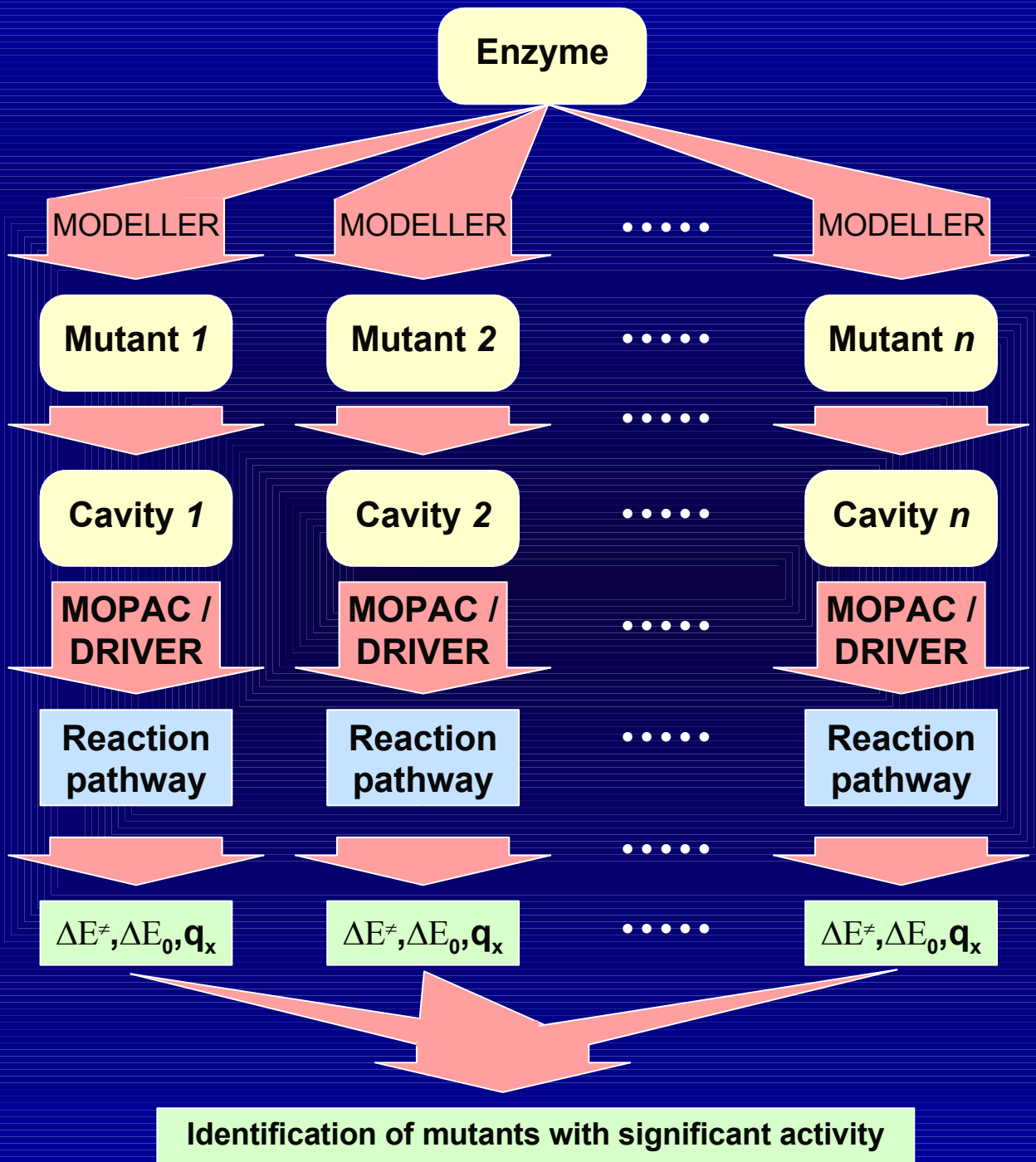




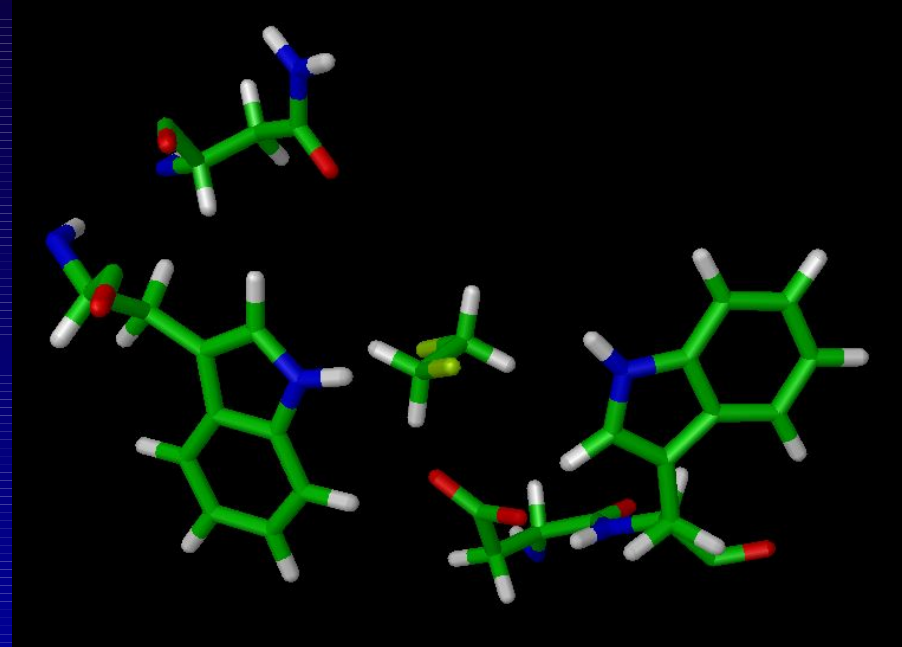
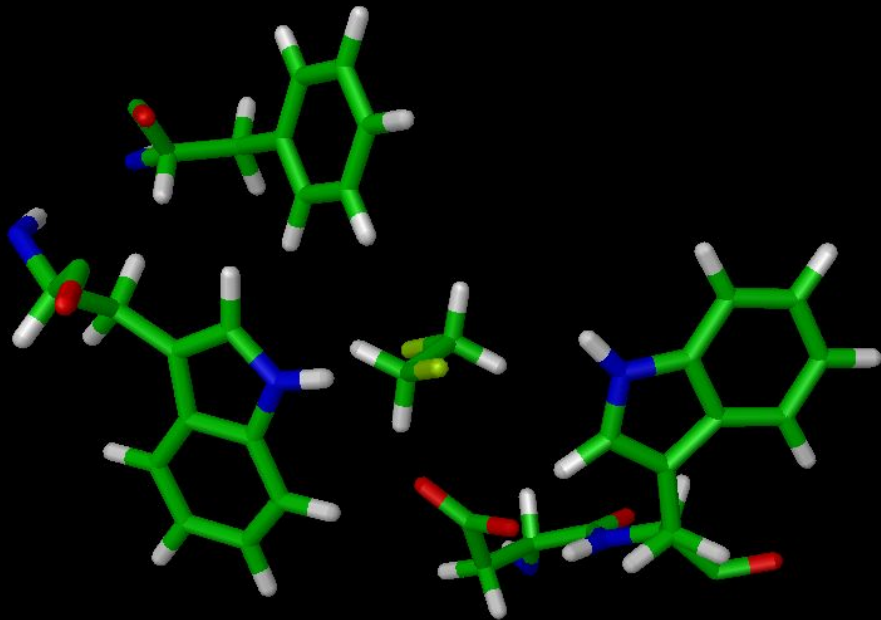
Image: 1 Distance: 2.9658 Energy: -991.4592 File: ne/marlinp/triton/release/linux/triton-3.0/triton-3.0/doc/tutorial/EXAMPLE4/DHLA_CBT/step00000



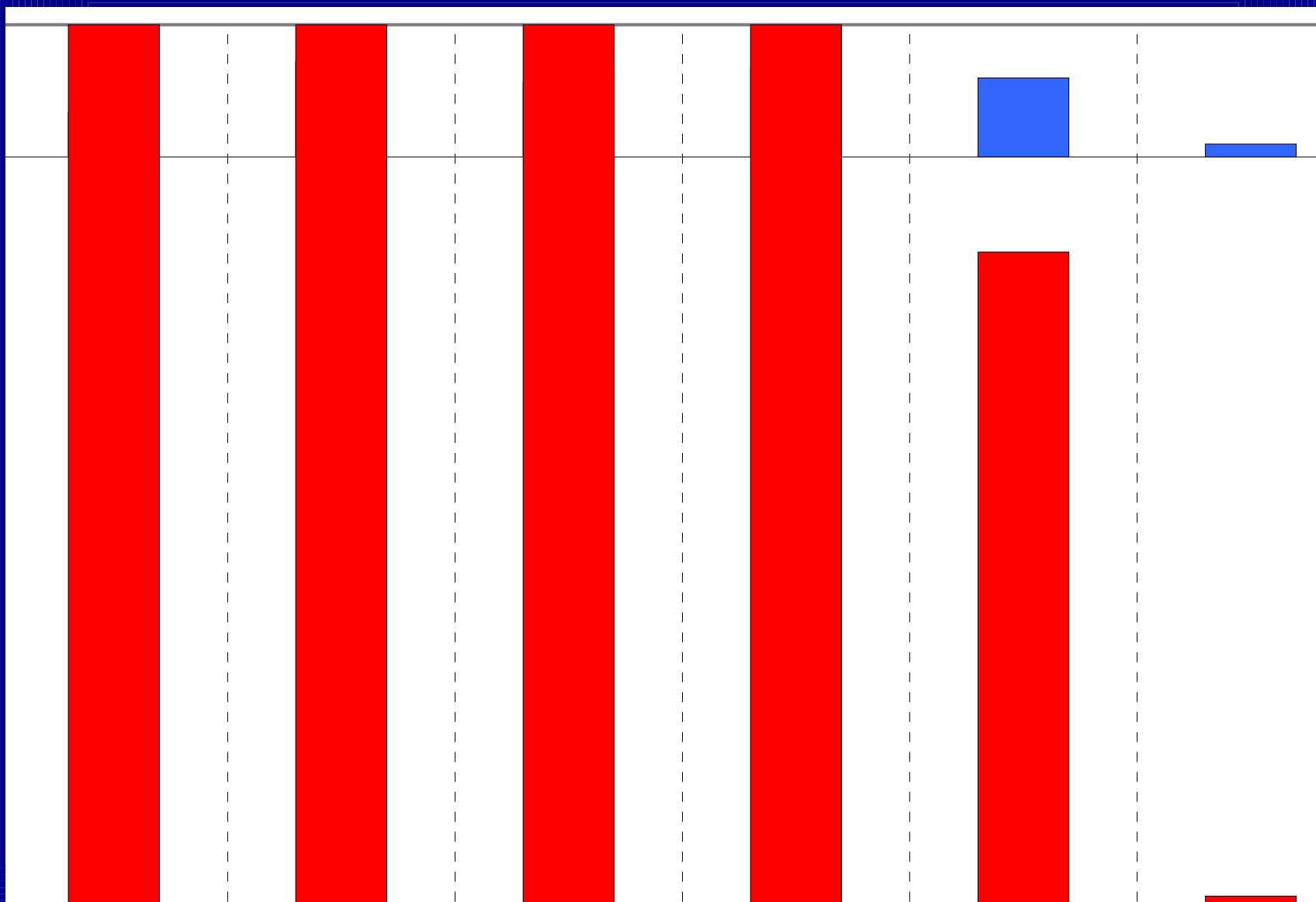


Mutagenese dehalogenasy DhIA

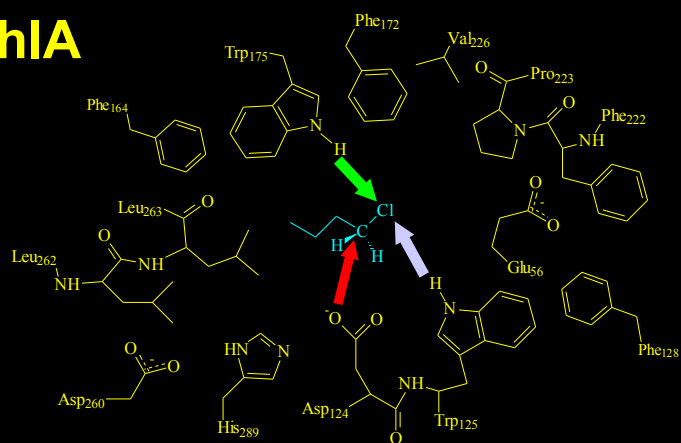
- Jednobodová mutace na pozici Phe172 - homologní modelování programem MODELLER
- Výpočet prvního S_N2 kroku reakční cesty programem MOPAC/DRIVER
- Aktivační energie - málo přesné a málo selektivní
- Sledování změny parciálního náboje na atomu substituujícího residua



Mutageneze dehalogenasy DhIA na pozici Phe172



DhaA

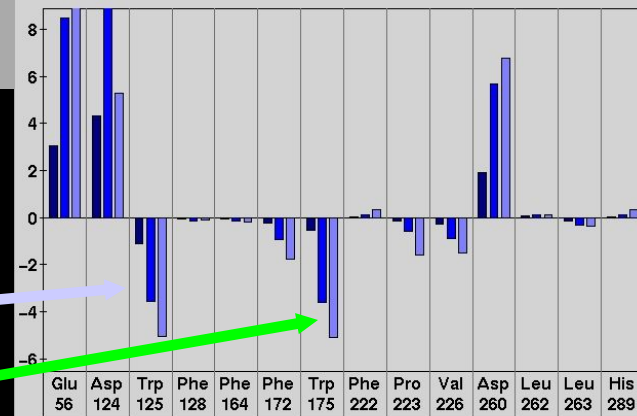


- Educ (E)
- Transition State (TS)
- Product (P)

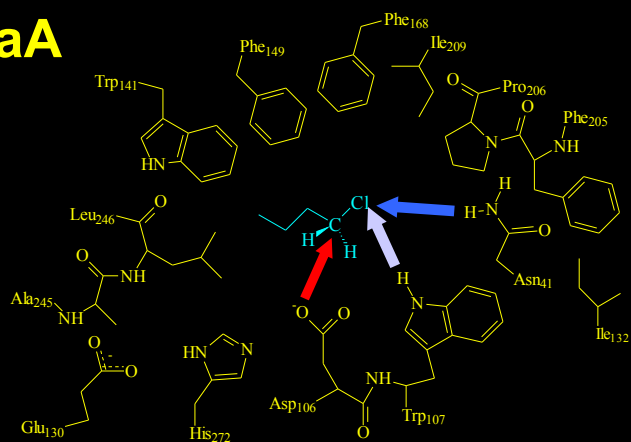
Trp125

Trp175

Interactions of Atom CL52 with Residues



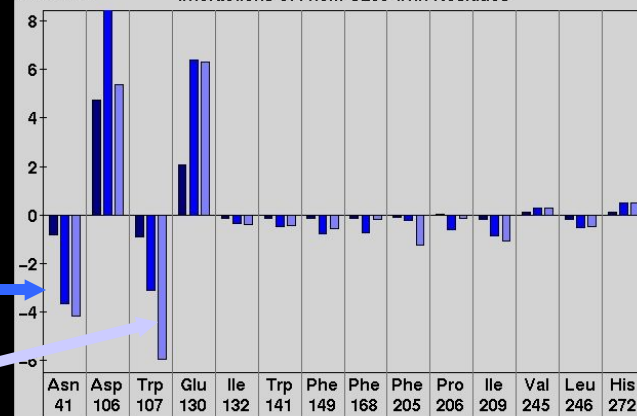
DhaA



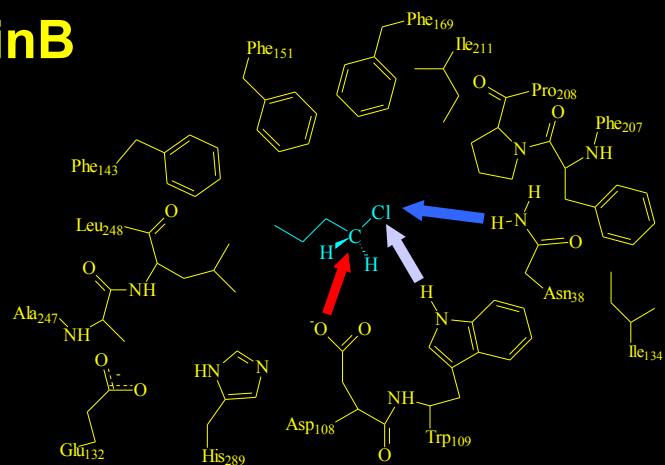
Asn41

Trp107

Interactions of Atom CL50 with Residues



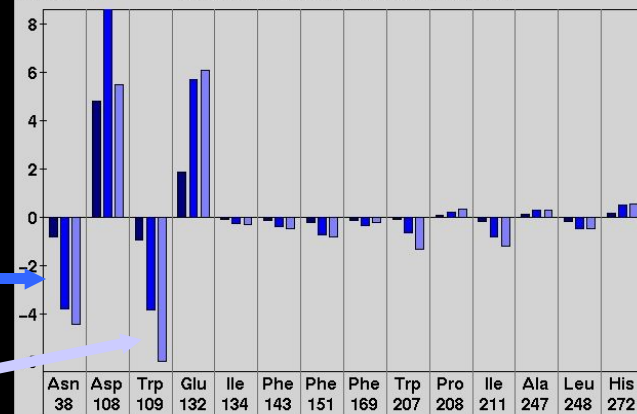
LinB



Asn38

Trp109

Interactions of Atom CL60 with Residues



Studie stabilizačního efektu residuí v hedalogenase LinB

- Jednobodové mutace Asn38, Trp109, Phe151, Phe169
- Modelování reakčního prvního S_N2 kroku - analýza elektrostatických interakcí
- Experimentálně změřené aktivity k_{cat}
- Potvrzen význam studovaných residuí pro katalýzu

	E_{elst} [kcal/mol]	k_{cat} [s^{-1}]
<i>residue</i>	<i>wt-mut</i>	
wt	-	1.3
N38D	-12.6	0.0
N38Q	-1.0	0.0
N38E	-11.7	0.0
N38F	-4.0	0.0
W109L	-3.8	0.0
F169L	0.2	0.4
F151L	-0.1	4.1
F151W	0.8	1.7
F151Y	-0.2	1.7

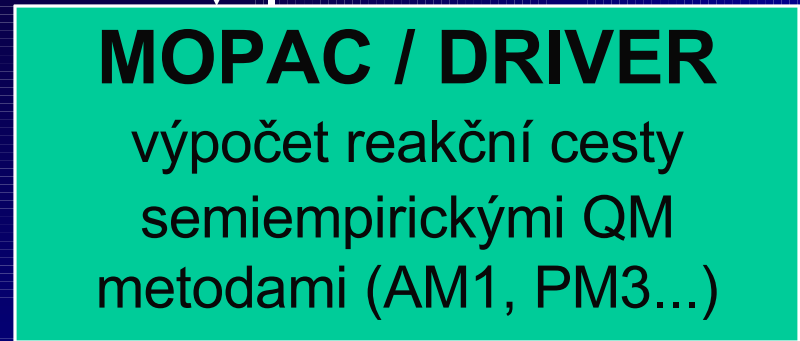
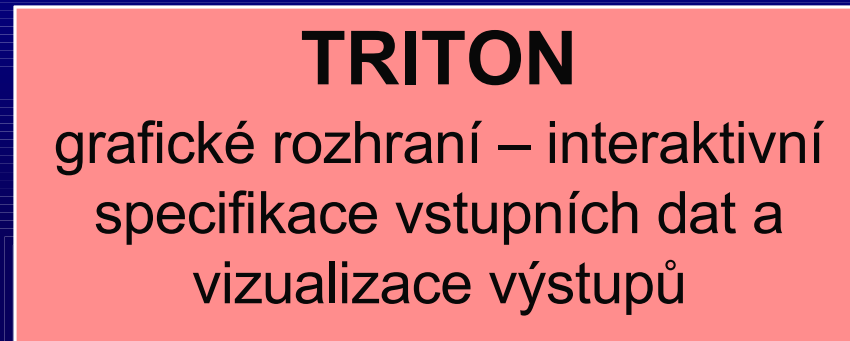
špatná stabilizace

žádná enzymatická aktivita

velmi malé změny
ve stabilizaci

enzymatická aktivita

Program TRITON



Spuštění programu:
module add triton
triton &

Program TRITON

<http://ncbr.chemi.muni.cz/triton/>

Cvičení 1

- Načtěte strukturu `/home/martinp/triton/structs/DhIA.pdb`
- Zobrazte seznam residuí
- Zobrazte pouze residua Asp124, Trp125, Trp175, Phe172, DCE. Ostatní atomy budou schované.
- Změňte model na *Tube*
- Obarvěte zobrazené atomy fialově
- Označte atomy DCE a obarvěte žlutě

Cvičení 2

- Vytvořte projekt *General* pojmenovaný *proj1*
- Vytvořte v něm projekt *Mutagenesis*
- Jako výchozí strukturu použijte `home/martinp/trito/structs/DhIA.pdb`
- Specifikujte bodovou mutaci na pozici Phe172
- Pro substituci použijte např. Tyr, Trp, Met, His, Cys
- Ostatní nastavení zůstávají implicitní
- Spust'ete výpočet

- Po skončení výpočtu zobrazte strukturu mutantu
- Načt'ete k ní strukturu výchozího enzymu DhIA
- Pro obě struktury zobrazte pouze residua Asp124, Trp125, Trp175, DCE, residuum 172
- Strukturu DhIA obarv'ete fialov'ě
- Zkoumejte pozici nového residua v pozici 172

Cvičení 3

- Vytvořte projekt *Reaction*
- Jako výchozí strukturu použijte `home/martinp/trito/structs/Dh1A.pdb`
- Počítejte enzymovou reakci
- Jako substrát zvolte DCE
- Residua cavity budou: Asp124
- Specifikujte reakční koordinátu mezi atomy O residua Asp124 (1985) a atomem C residua DCE (2)
- Step = -0.05 (stiskněte tlačítko *Default*)
- Specifikujte náboj -1
- Fixované atomy: peptidická páteř (backbone)
- Výpočet spust'ete manuaálně z terminálu (spustíte script `run.sh` v adresáři projektu)
- Po skončení výpočtu zobrazte graf energie a animaci
- Zkoumejte změny náboje v průběhu reakce