



UNIVERZITET U NOVOM SADU
PRIRODNO-MATEMATIČKI FAKULTET
DOKTOR BIOLOŠKIH NAUKA



PREDVIĐANJE TRODIMENZIONALNE STRUKTURE I KARAKTERIZACIJA AKTIVNOG MESTA ODABRANIH BETA- GALAKTOZIDAZA

- DOKTORSKA DISERTACIJA -

Mentor:
Doc. dr **Edward Petri**

Kandidat:
Vladimir Vukić, dipl. biol.-master

Novi Sad, 2015.

UNIVERZITET U NOVOM SADU
NAZIV FAKULTETA: PRIRODNO-MATEMATIČKI FAKULTET

KLJUČNA DOKUMENTACIJSKA INFORMACIJA

Redni broj: RBR	
Identifikacioni broj: IBR	
Tip dokumentacije: TD	Monografska dokumentacija
Tip zapisa: TZ	Tekstualni štampani materijal
Vrsta rada (dipl., mag., dokt.): VR	Doktorska disertacija
Ime i prezime autora: AU	Vladimir Vukić
Mentor (titula, ime, prezime, zvanje): MN	dr Edward Petri, docent
Naslov rada: NR	Predviđanje trodimenzionalne strukture i karakterizacija aktivnog mesta odabranih beta-galaktozidaza
Jezik publikacije: JP	Srpski, latinica
Jezik izvoda: JI	Srpski / engleski
Zemlja publikovanja: ZP	Republika Srbija
Uže geografsko područje: UGP	Vojvodina
Godina: GO	2015
Izdavač: IZ	Autorski reprint
Mesto i adresa: MA	Novi Sad, Trg dositeja Obradovića 2, 21000 Novi Sad

Fizički opis rada: FO	6 poglavlja / 138 stranica / 86 slika / 9 tabela / 242 reference
Naučna oblast: NO	Biološke nauke
Naučna disciplina: ND	Biohemija
Predmetna odrednica, ključne reči: PO	Beta-galaktozidaza, predviđanje strukture, molekularni doking, probiotici, tehnologija mleka
UDK	
Čuva se: ČU	U biblioteci Prirodno-matematičkog fakulteta u Novom Sadu, Trg dositeja Obradovića 2, 21000 Novi Sad
Važna napomena: VN	Nema
Izvod: IZ	<p>Proteini imaju osnovnu ulogu u životnim procesima, usled čega je ispitivanje raznovrsnosti njihove strukture jedna od ključnih potreba modernih bioloških istraživanja. Cilj ove doktorske disertacije je predviđanje trodimenzionalne strukture beta-galaktozidaza odabranih bakterijskih sojeva (<i>Streptococcus thermophilus</i>, <i>Lactobacillus acidophilus</i>, i <i>Bifidobacterium animalis</i> ssp. <i>lactis</i>) kao i karakterizacija njihovih aktivnih mesta i određivanje aktivnih rezidua. Takođe, cilj disertacije obuhvata i ispitivanje beta-galaktozidaze <i>Bacteroides thetaiotaomicron</i>-a, kao i ispitivanje aktivnosti beta-galaktozidaze <i>Kluyveromyces lactis</i>-a. Na osnovu sličnosti analizirane sekvence se mogu podeliti u dva klastera. Članovi klastera 1 pripadaju GH-2 subfamilijski, dok članovi klastera 2 pripadaju GH-42 subfamilijski. Na osnovu klaster analize i poravnanja sekvenci predložene su katalitički aktivne rezidue beta-galaktozidaza ispitivanih sojeva: GLU458/GLU546 kod <i>S. thermophilus</i>-a, GLU148/GLU307 kod <i>L. Acidophilus</i>-a i GLU164/GLU324 kod <i>B. animalis</i> ssp. <i>lactis</i>-a. Konstruisani su homologni modeli beta-galaktozidaza <i>S. thermophilus</i>, <i>L. acidophilus</i>, <i>B. animalis</i> ssp. <i>lactis</i>. Iz rezultata simulacije molekularnog dokinga svih beta-galaktozidaza uočljivo je da se supstrat veže za aktivno mesto koje se nalazi u regionu visoke elektronegativnosti. Simulacije dokinga su</p>

	<p>potvrdile aktivna mesta predložena poravnanjem sekvenci i otkrile rezidue koje učestvuju u vezivanju laktoze i galaktoze kao supstrata. Obzirom da je objavljena samo jedna kristalna struktura beta-galaktozidaze sa laktozom u aktivnom mestu, analiza molekularnog dokinga laktoze kao supstrata je sprovedena i sa beta-galaktozidazama koje imaju određenu trodimenzionalnu strukturu. Na osnovu slučnosti sekvenci može se zaključiti da je nukleofilna rezidua beta-galaktozidaze <i>B. thetaiotaomicron</i>-a GLU259. Analizom konzervativnih regiona, kao i doking simulacije, može se zaključiti da ulogu proton donora ima rezidua GLU182. Hemijski sastav mleka za ispitivanje aktivnosti beta-galaktozidaze <i>Kluyveromyces lactis</i>-a je u skladu sa Pravilnikom o kvalitetu proizvoda od mleka i starter kultura koji izdaje Ministarstvo poljoprivrede i zaštite životne sredine Republike Srbije. Celokupna laktoza se razgradila u permeatu na temperaturi od 40°C i pri koncentraciji enzima od 0,1% nakon 60 minuta od trenutka dodatka enzima. Maksimalna brzina postiže se toku prvih 10 minuta. Iz dobijenih rezultata može se zaključiti da beta-galaktozidaza <i>K. lactis</i>-a ima visoku efikasnost prilikom razgradnje laktoze što je čini pogodnom za komercijalnu upotrebu u industriji mleka. U ovoj disertaciji predstavljen je protokol za modelovanje komercijalno važnih beta-galaktozidaza i analiziranje njihovih interakcija sa laktozom kao supstratom. Dobijeni rezultati treba da pruže okvir za molekularna istraživanja vezivanja i razgradnje laktoze, kao i za buduće dizajniranje komercijalno interesantnih beta-galaktozidaza.</p>
Datum prihvatanja teme od strane Senata: DP	11.09.2014.
Datum odbrane: DO	

Članovi komisije:	
predsednik:	dr Spasenija Milanović, redovni profesor, Tehnološki fakultet Novi Sad, Univerzitet u Novom Sadu
član:	dr Danijela Kojić, docent, Prirodno-matematički fakultet, Univerzitet u Novom Sadu
član:	dr Anđelka Ćelić, docent, Prirodno-matematički fakultet, Univerzitet u Novom Sadu
član (mentor):	dr Edvard Petri, docent, Prirodno-matematički fakultet, Univerzitet u Novom Sadu

University of Novi Sad
Faculty of Sciences
Key word documentation

Accession number: ANO	
Identification number: INO	
Document type: DT	Monograph documentation
Type of record: TR	Textual printed material
Contents code: CC	PhD Thesis
Author: AU	Vladimir Vukić
Mentor: MN	Dr Edward Petri, Assistant Professor
Title: TI	Prediction of trodimensional structure and characterisation of active sites of selected beta- galaktosidases
Language of text: LT	Serbian

Language of abstract: LA	English / Serbian
Country of publication: CP	Republic of Serbia
Locality of publication: LP	Vojvodina
Publication year: PY	2015
Publisher: PU	Author reprint
Publication place: PP	Novi Sad, Trg dositeja Obradovića 2, 21000 Novi Sad

Physical description: PD	6 chapters / 139 pages / 86 figures / 9 tables / 242 references
Scientific field SF	Biology
Scientific discipline SD	Biochemistry
Subject, Key words SKW	Beta-galactosidase, structure prediction, molecular docking, probiotics, milk technology
UC	
Holding data: HD	Library of Faculty of Sciences, Trg dositeja Obradovića 2, 21000 Novi Sad
Note: N	None
Abstract: AB	Proteins have an essential role in all life processes. Therefore, investigation of their structure diversity is one of the key focuses of modern biological researches. The aim of this PhD thesis is prediction of three-dimensional structure of beta-galactosidases of selected bacterial strains (<i>Streptococcus thermophilus</i> , <i>Lactobacillus acidophilus</i> and <i>Bifidobacterium animalis</i> ssp. <i>lactis</i>), as well as characterization of their active sites and active residues. Furthermore, this thesis include investigation of

Bacteroides thetaiotaomicron and *Kluyveromyces lactis* beta-galactosidases. Based on sequence similarity analyzed sequences can be clustered in two clusters. Members of cluster 1 belong to GH-2 subfamily, while members of cluster 2 belong to GH-42 subfamily. Cluster analysis and sequence alignment revealed potential active residues: GLU458/GLU546 - *S. thermophilus*, GLU148/GLU307 - *L. Acidophilus* and GLU164/GLU324 - *B. animalis* ssp. *lactis*. Homologous models of beta-galactosidases from *S. thermophilus*, *L. acidophilus*, *B. animalis* ssp. *lactis* are constructed as a result of this PhD thesis. Molecular docking simulations of all beta-galactosidases revealed that substrate binds in region with high electronegativity. Furthermore, docking simulations confirmed active sites suggested by sequence alignment. As only one crystal structure with lactose in its active site has been published, docking simulations with lactose as a ligand were also performed with experimentally determined beta-galactosidases structures. Also, lactose and galactose binding residues are predicted. Based on sequence similarity GLU259 can be recognized as nucleophile of *B. thetaiotaomicron* beta-galactosidase. By analysis of conserved regions and docking simulation, GLU182 is predicted as a proton donor residue. Chemical composition of milk used for *Kluyveromyces lactis* beta-galactosidase research had appropriate quality according to directions of Ministry of agriculture and environmental protection of Republic of Serbia. Total lactose content was degraded in permeate after 60 min using enzyme concentration of 0,1% at 40°C. Maximal degradation was reached in the first 10 min. Based on the results, it can be concluded that *Kluyveromyces lactis* beta-galactosidase has high efficiency for lactose degradation, which makes it commercially interesting for milk industry. Protocol for modeling of commercially important beta-galactosidases and their interactions with lactose is developed as a result of this PhD thesis. Obtained results should

	provide the frame for molecular investigation of lactose binding and degradation, as well as for designing of commercially suitable beta-galactozidases.
Accepted on Senate on: AS	11.09.2014.
Defended: DE	
Thesis Defend Board:	
president:	dr Spasenija Milanović, redovni profesor, Tehnološki fakultet Novi Sad, Univerzitet u Novom Sadu
member:	dr Danijela Kojić, docent, Prirodno-matematički fakultet, Univerzitet u Novom Sadu
member:	dr Anđelka Čelić, docent, Prirodno-matematički fakultet, Univerzitet u Novom Sadu
member (menthor):	dr Edvard Petri, docent, Prirodno-matematički fakultet, Univerzitet u Novom Sadu

Sadržaj

1.	Uvod.....	1
2.	Opšti deo	3
2.1	Analiza sekvenci proteina.....	3
2.2	Energija u molekularnom modelovanju proteina	8
2.2.1	Dužina veze	9
2.2.2	Ugao veze	10
2.2.3.	Torzione veze.....	11
2.2.4.	Elektrostatičke interakcije.....	12
2.2.5.	Van der Valsove interakcije.....	12
2.2.6.	Minimizacija energije molekula	14
2.3.	Određivanje strukture proteina	15
2.4.	Homologo modelovanje.....	18
2.4.1.	Modelovanje bočnih lanaca.....	23
2.4.2.	Modelovanje petlji.....	24
2.4.3.	Modelovanje kompletne strukture proteina	25
2.4.4.	Značaj homologog modelovanja.....	27
2.5.	<i>De novo</i> predviđanje strukture proteina.....	29
2.6.	Eksperimentalno određivanje strukture proteina	30
2.7.	Predviđanje interakcija protein ligand (ligand doking)	37
2.7.1.	Funkcije za predviđanje afiniteta liganda za protein i jačinu veze	38
	<i>Funkcije afinitetna na osnovu energetske polja.....</i>	<i>39</i>
	<i>Empirijske funkcije afiniteta.....</i>	<i>40</i>
2.7.2.	Metode ligand dokinga	40
	<i>Doking rigidnih konformacija.....</i>	<i>40</i>
	<i>Algoritmi fleksibilnog dokinga.....</i>	<i>41</i>
	<i>Doking primenom simulacija</i>	<i>43</i>
2.7.3.	<i>De novo</i> sinteza liganda	46
2.7.4.	Validacija doking metoda	46
2.7.5.	Izazovi u ligand dokingu	47
2.8.	Tehnologija fermentisanih mlečnih proizvoda	47
2.8.1.	Starter kulture u fermentaciji mleka.....	47
2.8.2.	Razgradnja laktoze starter kulturom	49
2.9.	Beta-galaktozidaza.....	51

2.9.1. Klasifikacija beta-galaktozidaza	52
2.9.2. Ekspresija gena beta-galaktozidaze	52
2.9.3. Razgradnja laktoze.....	54
2.9.4. Struktura beta-galaktozidaze <i>E.coli</i>	59
2.9.5. Mehanizam razgradnje laktoze beta-galaktozidaze <i>E. coli</i>	63
2.9.6. Struktura beta-galaktozidaze <i>Kluyveromyces lactis</i> -a	68
2.9.7. Struktura beta-galaktozidaze <i>Bacillus circulans</i> -a	73
2.9.8. Mehanizam razgradnje laktoze beta-galaktozidaze <i>Bacillus circulans</i> -a.....	77
2.9.9. Beta-galaktozidaza <i>Streptococcus thermophilus</i> -a.....	78
2.9.10. Beta-galaktozidaza <i>Lactobacillus acidophilus</i> -a	79
2.9.11. Beta-galaktozidaza <i>Bifidobacterium animalis</i> ssp. <i>lactis</i>	81
2.9.12. Beta-galaktozidaza <i>Bacteroides thetaiotaomicron</i> -a.....	82
3. Materijal i metodi	83
3.1. Sekvence.....	83
3.2. Homologo modelovanje	83
3.3. Klaster analiza	84
3.4. Minimizacija energije.....	84
3.5. Simulacija dokinga.....	85
3.6. Permeat.....	85
3.7. Enzim i hidroliza laktoze.....	85
3.8. Fizičko-hemijske analize.....	86
3.9. Statistička analiza.....	87
4. Rezultati i diskusija	87
4.1. Klaster i filogenetska analiza beta-galaktozidaza koje pripadaju subfamilijama GH-2 i GH-42	87
4.2. Konstrukcija modela metodom modelovanja kompletne strukture proteina.....	91
4.3. Validacija konstruisanih modela.....	94
4.4. Analiza interakcije enzima sa supstratom (doking analiza).....	99
4.5. Beta-galaktozidaza <i>B. thetaiotaomicron</i> -a.....	112
4.6. Kinetika razgradnje laktoze beta-galaktozidazom <i>K. lactis</i> -a	116
5. Zaključci.....	119
6. Literatura	121

1. Uvod

Razvoj računarskih tehnika uveo je revoluciju u molekularnom modelovanju. Molekularno modelovanje predstavlja važan aspekt fundamentalnih istraživanja kako proteina tako i drugih molekula. Do početka 2015. godine određeno je 97946 trodimenzionalnih struktura proteina u bazama podataka (PDB – <http://www.rcsb.org>). U isto vreme, poznato je oko 170 miliona genskih sekvenci koje čine preko 15 miliona gena (NCBI – <http://www.ncbi.nlm.nih.gov>) i kodiraju preko 40 miliona proteina (UNIPROT). Ova razlika između broja sekvenci i određenih struktura nastavlja da raste uprkos sve uspešnijim metodama za određivanje trodimenzionalne strukture. Molekularno modelovanje predstavlja skup računarskih tehnika i pretežno se koristi da odgovori na neka od sledećih pitanja: 1. kako izgledaju molekuli koje ispituje u prostoru; 2. kako molekuli interaguju sa drugim molekulima; 3. kako izgleda okruženje molekula; 4. kako se molekuli i njihova okruženja menjaju tokom vremena. Određivanje trodimenzionalne strukture molekula je ključna za analize kao što su QSAR (Quantitative Structure–Activity Relationship), 3D analiza farmakofora, predviđanje enzimskih reakcija, itd. Sve veća ekspanzija farmaceutske industrije ima za cilj proizvodnju lekova čija se aktivnost u velikoj meri bazira na efikasnoj inhibiciji željenih enzima. Ispitivanje interakcija ligand-enzim svakako predstavlja ključni korak u cilju inhibicije ili povećanja efikasnosti reakcije. Favorizovanjem ili blokiranjem takve reakcije postiže se željena aktivnost farmaceutskih sredstava. S toga, kod lekova molekul inhibitor mora da ostvaruje jače veze sa enzimom od prirodnog supstrata. Sa druge strane, stimulacijom promena aminokiselina u aktivnom mestu mogu se predvideti konformacije koje će ostvarivati jače interakcije sa ligandom i time potencijalno rezultirati povećanjem efikasnosti reakcije. Molekularnim modelovanjem moguće je ispitivati i jedinjenja koja ne postoje u prirodi ili čak ni ne mogu postojati pod normalnim uslovima.

Analiza trodimenzionalne strukture proteina čija struktura nije eksperimentalno utvrđena, moguća je primenom molekularnog modelovanja i softverskih paketa za konstrukciju homologih modela. U poslednje vreme, povećanjem razvoja bioinformatičkih softvera, molekularno modelovanje trodimenzionalnih struktura proteina pruža sve pouzdanije rezultate koji se sa većom sigurnošću mogu koristiti za dalja ispitivanja. Kod homologog modelovanja, iako nije moguće predvideti trodimenzionalnu strukturu samo na osnovu sekvence, moguće je dobiti

precizne modele na osnovu homologije sekvenci proteina čija je struktura eksperimentalno utvrđena. Ovakvo modelovanje se bazira na činjenici da proteini sa sličnim sekvencama imaju slične trodimenzionalne strukture.

Laktoza kao dominantni šećer u mleku predstavlja primarni supstrat u tehnologiji fermentisanih mlečnih proizvoda. Fermentacijom laktoze dobija se niz fermentisanih mlečnih proizvoda kao što su: jogurt, kefir, kumis, fermentisani napitak od surutke, pavlaka, itd. Starter kulture su izolovani i sa željenim karakteristikama selekcionisani mikroorganizmi koji složenim biohemijskim procesima fermentišu supstrat i produkuju željeni proizvod. Starter kulture u tehnologiji fermentisanih mlečnih proizvoda primarno razgrađuju laktozu i dominantno produkuju mlečnu kiselinu (tipično za jogurt i fermentisane mlečne proizvode u tipu jogurta – homofermentacija) ili pored mlečne kiseline laktozu transformišu i u druge produkte: etanol, ugljen-dioksid (tipično za kefir i kumis – heterofermentacija).

Veliki napredak u savremenoj tehnologiji fermentisanih mlečnih proizvoda (prvenstveno zbog zdravstvenih aspekata finalnih proizvoda) predstavljao je razvoj nove vrste starter kultura koje sadrže probiotike, najčešće iz rodova *Lactobacillus* i *Bifidobacterium*. Probiotske starter kulture se (pored tradicionalnih starter kultura *Streptococcus thermophilus* i *Lactobacillus delbrueckii* ssp. *bulgaricus*) sve češće koriste u mlečnoj industriji. Osnovna odlika sojeva koji se koriste kao starter kulture je da poseduju enzim beta-galaktozidazu koji razlaže laktozu na glukozu i galaktozu i time otvara mogućnost fermentacije mleka i dalje razgradnje navedenih šećera. Beta-galaktozidaze spadaju među najprimenjenije enzime u industriji i tehnologiji mleka. Njihova glavna primena je u proizvodnji delaktoziranog mleka namenjenog osobama netolerantnim na laktozu. Beta-galaktozidaze sa transacetilacionom aktivnosti nalaze primenu u proizvodnji galakto-oligosaharida koji se koriste kao prebiotici koji stimulišu rast korisnih mikroorganizama u ljudskim crevima kao što su laktobacili i bifidobakterije. Razumevanje njihove strukture pruža bolji uvid u funkciju koju obavljaju, mehanizme razgradnje laktoze, kao i njihove interakcije sa laktozom. Trodimenzionalna struktura enzima može se koristiti za dalja ispitivanja i njegova poboljšanja.

Da bi se razumeli osnovni biološki procesi kao što su ćelijska deoba, rast, razvoj i ćelijska komunikacija, neophodno je poznavanje trodimenzionalne strukture aktivnih komponenata. Proteini imaju ključnu ulogu u ovim procesima i ispitivanje raznovrsnosti njihove strukture je jedna od ključnih potreba modernih bioloških istraživanja. Trodimenzionalne strukture beta-

galaktozidaza, sojeva *Streptococcus thermophilus*, *Lactobacillus acidophilus*, *Bifidobacterium animalis* ssp. *lactis* nisu ispitane u dosadašnjoj literaturi. Takođe, prema dostupnoj literaturi, strukture aktivnih mesta sa aktivnim reziduama navedenih i beta-galaktozidaze *Bacteroides thetaiotaomicron* do sada nisu utvrđene. Shodno sve pouzdanijim metodama molekularnog modelovanja javlja se mogućnost predviđanja njihovih struktura, određivanja aktivnih rezidua i ispitivanja kompleksa enzim – supstrat, što je od izuzetne važnosti za bolje razumevanje delovanja enzima.

Cilj doktorske disertacije je predviđanje trodimenzionalne strukture beta-galaktozidaza odabranih bakterijskih sojeva (*Streptococcus thermophilus*, *Lactobacillus acidophilus* i *Bifidobacterium animalis* ssp. *lactis*) kao i karakterizacija njihovih aktivnih mesta i određivanje aktivnih rezidua. Takođe, cilj disertacije obuhvata i ispitivanje beta-galaktozidaze *Bacteroides thetaiotaomicron*-a, kao i ispitivanje aktivnosti beta-galaktozidaze *Kluyveromyces lactis*-a.

2. Opšti deo

2.1 Analiza sekvenci proteina

Molekularno modelovanje i mnoge njegove tehnike usko su vezane sa analizom sekvenci proteina koji se žele modelovati. Za molekularno modelovanje veoma važnu ulogu imaju homologe sekvence. Genske i aminokiselinske sekvence se u evolutivnim istraživanjima smatraju homologim ukoliko imaju zajedničkog pretka. Njihova homologija ogleda se u sličnosti njihovih sekvenci. Homologe sekvence često (u zavisnosti od stepena sličnosti) imaju slične funkcije. Ispitivanje sličnosti među sekvencama se vrši iz dva osnovna razloga: 1. evolutivna istraživanja, 2. ispitivanje funkcije homologih proteina. Pri ispitivanju sličnosti među sekvencama (dve ili više) vrši se njihovo poravnanje.

Poravnanje sekvenci predstavlja šemu pisanja jedne sekvence iznad druge na taj način da residue na istim pozicijama imaju zajedničko evolutivno poreklo (Jermin i sar. 1997). Ukoliko se jedna rezidua nalazi na istoj poziciji kod obe sekvence smatra se da je ta pozicija konzervativna tokom evolucije. Ipak, moguće je da su slučajno evoluirale nezavisnim mutacijama, ali takva mogućnost se ne uzima u obzir tokom analize. Ukoliko se residue razlikuju na datoj poziciji,

postoje tri opcije: 1) da je prva promenjena tokom evolucije; 2) da je druga promenjena tokom evolucije; ili 3) da su obe promenjene tokom evolucije (slika 1).


Query	13	DPKIVSVNTVDAHSDHKYFESLEEF-SEGEMKLRQSLNGKWKIHYAQNTNQVLKDFYKTE	71
		+P + +N + AH + + EE ++ + +SLNG+W+ + V + + + +	
Sbjct	18	NPGVTLNRLAAHPPFASWRNSEEARTDRPSQQLRSLNGEWRFAWFPAPAEAVPESWLECD	77
Query	72	FDETDLNFINVPGHLELQGFSGSPQYVNVQYPWDGKEFLRPPQVPQESNAVASVYKHFTLN	131
		E D + VP + ++ G+ +P Y NV YP + PP VP E N Y F ++	
Sbjct	78	LPEADT--VVVP SNWQMHGYDAP IYTNVTYP----ITVNPPFVPT E-NPTGCYSLTFNVD	130
Query	132	DA-LKDKKVFISFQGVATSIFVWVNGNFVGYSEDSFTPSEFEISDYLVEGDNKLAVAVYR	190
		++ L++ + I F GV ++ +W NG +VGY +DS PSEF++S +L G+N+LAV V R	
Sbjct	131	ESWLQEGQTRIIFDGVNSAFHLWCNGRWVGYGQDSRLPSEFDLSAFLRAGENRLAVMVL R	190
Query	191	YSTASWLEDQDFWRLYGIFRDVYLYAIPKVHVQDLFVKGDYDYQTKAGQLDIDLKAVGDY	250
		+S S+LEDQD WR+ GIFRDV L P + D V ++ L+ +++ G+	
Sbjct	191	WSDGSYLEDDQDMWRMSGIFRDVSLHLKPTTQISDFHVATRFNDDFSRAVLEAEVQMC GEL	250

Slika 1. Primer poravnanja sekvenci urađen BLAST programom

Usled insercija i/ili delecija homologe sekvence mogu biti različitih dužina. Na slici 1 su delecije/insercije prikazane znakom “-“. Obzirom da sa stanovišta poravnanja sekvenci nije moguće utvrditi da li se radi o deleciji ili inserciji i da li se delecija u jednoj sekvenci može smatrati insercijom u drugoj, uveden je zajednički termin “gap” ili prevedeno “razmak”. Na slici 1 su lako uočljiva poklapanja i razlike u sekvencama. Ukoliko se sekvence poklapaju, uobičajena praksa je da se to prikaže oznakom rezidue između sekvenci ili crticom koja povezuje rezidue. Ukoliko se ne poklapaju ostavlja se praznina. Ukoliko se sekvence ne poklapaju, ali dele zajedničku osobinu (npr. obe su hidrofobne ili obe hidrofilne) mesto se obeležava plusom ili tačkama između rezidua.

Za analizu sličnosti dve sekvence, nije dovoljno utvrditi stepen sličnosti samo na osnovu broja promenjenih i očuvanih rezidua. Eksperimentalno je utvrđeno da su neke rezidue podložnije promenama i da češće mutiraju od drugih (Dayhoff i sar. 1968). Usled toga bilo je neophodno uraditi skaliranje mogućih mutacija. Na osnovu zapažanja na evolutivno bliskim proteinima Dayhoff i sar. konstruisali su seriju 20 x 20 matrica sa koeficijentima mutacija između rezidua. Za konstrukciju matrica korišćene su blisko srodne sekvence. Koeficijent mutacije predstavlja logaritam vrednosti odnosa verovatnoće da su dve rezidue na istoj poziciji homologe, u odnosu na slučaj da su slučajno na istoj poziciji. Ove matrice se nazivaju PAM (Point Accepted Mutation) matrice. PAM matrice koje se prave za sekvence koje imaju 1 mutaciju na 100 rezidua i nazivaju se PAM1. U PAM1 matrici su prokazane verovatnoće (log vrednosti) da jedna rezidua

BLAST (**B**asic **L**ocal **A**lignment **S**earch **T**ool) je jedan od najkorišćenijih programa za pretraživanje baza podataka (slika 4) (Altschul i sar. 1990; Altschul i sar. 1997). BLAST algoritam poredi primarne biološke sekvence tražeći kratke lokalne sličnosti. Metod se sastoji u traženju istih kratkih nizova u sekvencama koje se porede (Word metod). Kada ih pronade, program nastavlja da pretražuje njihovu okolinu u potrazi za daljim sličnostima koristeći BLOSUM matricu za vrednovanje sličnosti (slika 4). Navedeni metod ima prednost u odnosu na poređenje celokupnih sekvenci jer je značano brži i skraćuje vreme pri pretragama velikih baza podataka. Takođe, kao što je već napomenuto, za pretragu sekvenci koriste se skalirane vrednosti. BLOSUM62 matrica je podrazumevana za BLAST program. BLAST program je uključen u najveće i najsveobuhvatnije baze podataka za DNK i proteinske sekvence kao što su National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov>) i European Bioinformatics Institute (<http://www.ebi.ac.uk>).

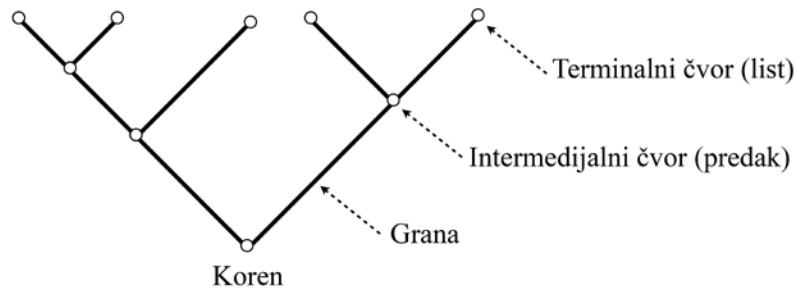
Ispitivana sekvenca	M	R	D	PYN	K	L	I	S
Baza podataka	M	H	E	PYN	D	V	P	W
	5	0	2	20	-1	1	-3	-3
								
	Ukupna vrednost = 24							

Slika 4. Pretraživanje BLAST programa

Kada se želi napraviti višestruko poravnanje sekvenci u bazi podataka, sa ciljem utvrđivanja konzervativnih regiona i sekvenci, najčešće se koristi model razvijen od strane Krogha i Baldija (Baldi i sar. 1994; Krogh i sar. 1994). Njihov model predstavlja probabilistički model koji identifikuje klastere sličnih sekvenci u bazama podataka. Model ima “stanja”, i svako stanje će se projektovati u simbol (npr. slovo aminokiselinske sekvence) u zavisnosti od distribucije vezane za dato stanje. Nakon projektovanja simbola, početno stanje prelazi u susedno stanje u zavisnosti od verovatnoće prelaska. Navedeni prelasci se nazivaju Markovljevi jer ne zavise od prethodnih stanja, već isključivo od sadašnjeg. Detaljno objašnjenje modela dato je u

radu (Durbin i et al. 1998). Navedene metode su višestruko i u potpunosti validovane (Brenner i sar. 1998).

Klaster analiza sekvenci podrazumeva ispitivanje njihove evolutivne udaljenosti na osnovu promena nastalih tokom vremena. Prilikom klaster analize konstruiše se filogenetsko stablo koje se sastoji od terminalnih čvorova (lišće), intermedijalnih čvorova (hipotetički preci), grana i korena (slika 5).



Slika 5. Filogenetsko stablo

Metode klaster analize se zasnivaju na principu maksimalne štedljivosti (maximal parsimony), što podrazumeva da se traga za što kraćim stablom, odnosno za stablom koje prikazuje udaljenost sekvenci sa što manje pretrpljenih mutacija. U tu svrhu konstruišu se matrice udaljenosti koje prikazuju udaljenost sekvenci na osnovu njihovih razlika. Metod prisajedinjenja suseda (neighbor joining) polazi od matrice udaljenosti i grupiše najbliže (susedne) sekvence sa zajedničkim pretkom (intermedijalni čvor) i računa njihovu udaljenost (dužinu grana) od zajedničkog pretka. Najbliže sekvence se računaju preko formule:

$$u_i = \frac{\sum D_{ij}}{(n-2)}$$

gde je u_i suma rastojanja određene sekvence od svih ostalih sekvenci – $\sum D_{ij}$, podeljena sa brojem sekvenci – n minus dva. Grupišu se one sekvence gde je vrednost D_{ij} od koje se oduzmu u_i i u_j najmanja. Dužina grana od zajedničkog pretka računa se prema:

$$v_i = 0,5 D_{ij} + 0,5 (u_i - u_j)$$

$$v_j = 0,5 D_{ij} + 0,5 (u_j - u_i)$$

Zatim se grupisane sekvence zamenjuju sa zajedničkim čvorom i ponovo računa matrica udaljenosti čvora od ostalih sekvenci. Vrednost razdaljine grupisanih sekvenci od neke sekvence računa se prema dole navedenoj formuli i proces se nastavlja dok se sve sekvence ne grupišu u drvo:

$$D_{ij,k} = \frac{(D_{ik} + D_{jk} - D_{ij})}{2}$$

2.2 Energija u molekularnom modelovanju proteina

Energetsko polje omogućuje brzo računanje potencijalne energije sistema u molekularnom modelovanju. Energetsko polje predstavlja matematičku jednačinu koja uzimajući u obzir dužinu intramoleulskih veza, njihovih uglova, torzionih uglova, rastojanje između atoma u koordinatnom sistemu i ostale relevantne parametre računa potencijalnu energiju sistema. Drugim rečima energetsko polje je površina koju zahvata optimizovana potencijalna energija molekula u okruženju u kojem se nalazi. Na energiju intermolekulske interakcije proteina i liganda najviše utiču elektrostatičke, van der Valsove i hidrofobne sile, kao i vodonične veze koje se grade u kompleksu. Elektrostatičke sile i vodonične veze određuju specifičnost liganda za receptor, dok su hidrofobne sile uglavnom odgovorne za jačinu veze. Ono što je važno napomenuti jeste da ne postoji univerzalno energetsko polje koje će tačno opisati svaki sistem, jer se za računanje svakog energetskog polja uzimaju empirijska predviđanja koja variraju u zavisnosti od uslova sredine. Dakle, ukoliko je jedna forma energetskog polja preciznija od druge za neki sistem, ona treba da bude korišćena.

Zajednički parametar koji se uzima u obzir za sve tipove energetskih polja jeste tip atoma. Neophodno je uzeti u obzir tip atoma za svaki atom u sistemu koji se ispituje. Tip atoma uključuje informacije o njegovom hibridizacionom stanju (sp^3 , sp^2 ili sp za ugljenikove atome) kao i uslove lokalnog okruženja. Na primer, AMBER (Assisted Model Building with Energy Refinement) energetsko polje ima definisano 40 tipova atoma. Ove informacije su važne za određivanje uglova u vezama koje gradi dati molekul.

Postoje dva osnovna pristupa računanju energija u molekularnom modelovanju proteina:

- Metode molekularne mehanike koje ne uključuju računanja kretanja elektrona, što ih čini brzim i u razumnoj meri preciznim. Prilikom računanja koristi se veliki broj

predviđanja čime se skraćuje postupak i vreme računanja, ali i preciznost same metode. Koriste se za velike molekule (nekoliko desetina, pa i stotina hiljada atoma) i najčešće za modelovanja i dokinge svih vrsta proteina.

- Metode kvantne mehanike koje uzimaju u obzir trodimenzionalni raspored elektrona oko atomskog jezgra, što ih čini izuzetno preciznim, ali sa druge strane veoma kopjuterski zahtevnim. Koriste se za proste sisteme i male molekule, najčešće do svega nekoliko (najčešće do 150) atoma. Takođe, računanja metodama kvantne mehanike se uglavnom rade na serverima super-računara koji obezbeđuju dovoljnu brzinu da se u razumnom vremenu izvrši analiza.

Obzirom da su predmet ove disertacije makromolekuli (proteini), njena sadržina su parametri i metode molekularne mehanike.

2.2.1 Dužina veze

Kriva potencijalne energije (Morzeov potencijal) tipične interatomske veze (slika 6) (McMurry i Fay 2003) ima oblik prikazan na slici 7 i njen uprošćen oblik predstavljen je kroz formulu Hukovog zakona:

$$V_B(l) = \frac{1}{2} k_l (l - l_0)^2$$

Gde je:

$V_B(l)$ = potencijalna energija

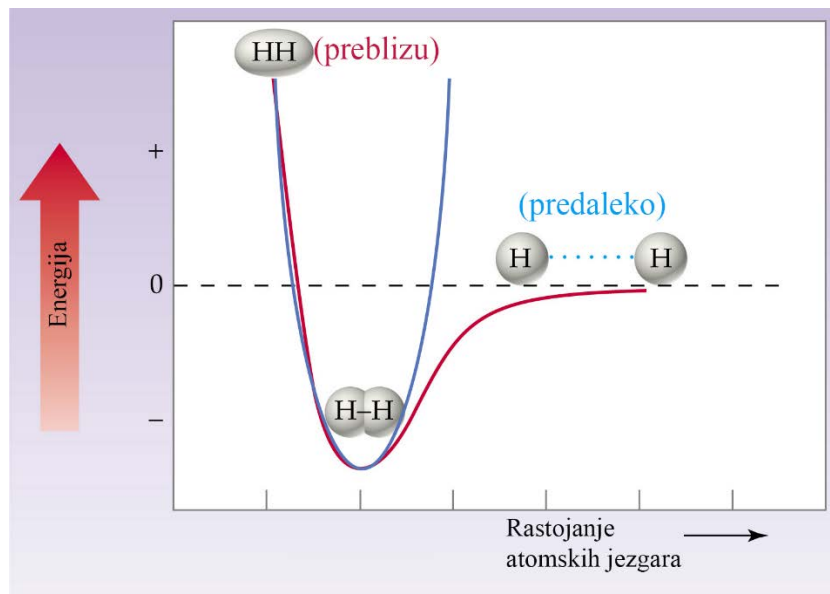
k_l = konstanta elastičnosti

l_0 = ravnotežna dužina veze

l = trenutna dužina veze



Slika 6. Dužina interatomske veze



Slika 7. Promena energije veze u zavisnosti od interatomskeg rastojanja; plava linija – harmonična oscilacija, crvena linija – Morzeov potencijal

U navedenoj jednačini istežanje veze se tretira kao klasični harmonijski oscilator. Na slici 6 je uočljivo da energija molekulske veze opada sa smanjenjem rastojanja između atomskih jezgara. Tom prilikom rastu odbijajuće sile, ali i privlačne, koje preovlađuju u ovoj fazi. Energija opada do ravnotežnog stanja kada je minimalna za datu interakciju. Sa daljim smanjenjem rastojanja atomskih jezgara energija raste, odbijajuće sile nastavljaju da rastu usled interakcije atoma (Paulijevo odbijanje), dok sa druge strane privlačne sile opadaju, što rezultira ukupnim odbijanjem atoma.

2.2.2 Ugao veze

Odstupanje ugla veze od njegove referentne vrednosti (slika 8) se takođe opisuje preko harmonijskog potencijala formulom:

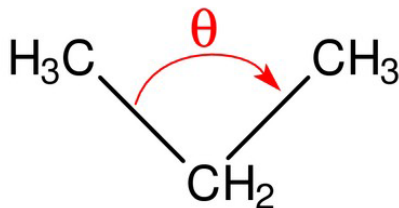
$$V_A(\theta) = \frac{1}{2} k_\theta (\theta - \theta_0)^2$$

$V_A(\theta)$ = potencijalna energija

k_θ = konstanta opruge za promenu ugla

θ_0 = ravnotežni ugao veze

θ = trenutni ugao veze



Slika 8. Ugao veze

Konstanta opruge i ravnotežni ugao veze zavise od tipa atoma koji grade vezu. Obzirom da je manje energije potrebno da se naruši ugao veze nego da se veza istegne ili sabije, vrednosti konstante je niža nego u prethodnoj formuli.

2.2.3. Torzione veze

Torzioni okreti predstavljaju potencijalne funkcije torzionih uglova koje modeluju prisustvo steričnih barijera između atoma koji su odvojeni kovalentnim vezama (slika 9) i računa se formulom:

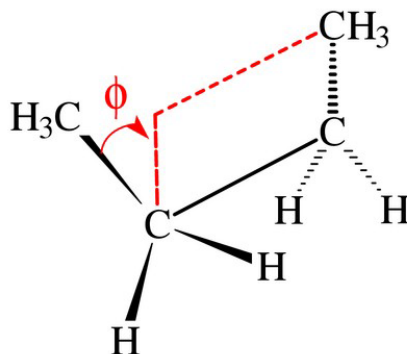
$$V_T(\Phi) = \frac{1}{2}k_\phi[1 + \cos(n\Phi - \delta)]$$

$V_T(\Phi)$ = potencijalna energija

k_ϕ = granična vrednost

n = periodičnost ($n = 3$ za etan, $n = 2$ za eten, itd.)

δ = fazni faktor – određuje minimalni ugao



Slika 9. Torzioni uglovi su definisani kao ugao Φ između ravni koje grade C atomi sa jednom i drugom metil grupom pojedinačno

2.2.4. Elektrostatičke interakcije

Interakcije između atoma koje nastaju usled njihovih trajnih momenta, u molekularnoj mehanici se opisuju smatrajući naelektrisane delove kao idealna tačkasta naelektrisanja. Molekul čije se naelektrisanje računa se postavi u trodimenzionalnu rešetku i zatim izračuna elektrostatički potencijal svake tačke u rešetci. Zatim se koristi Kulonov potencijal za idealna tačkasta naelektrisanja za procenu naelektrisanja samih atoma i između njih. Kulonov potencijal na sledeći način opisuje interakciju između dve naelektrisanje čestice:

$$V_E(i, j) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}}$$

$V_E(i, j)$ = potencijalna energija

q_{ij} = magniude naelektrisanja atoma

r_{ij} = rastojanje između atoma

ϵ_0 = električna provodljivost prostora

ϵ_r = relativna dielektrična konstanta medijuma

2.2.5. Van der Valsove interakcije

Van der Valsove interakcije između atoma proizilaze iz balansa između privlačnih i odbijajućih sila. Privlačne sile uglavnom nastalu usled fluktuacija u naelektrisanju unutar elektronskog oblaka. Sa druge strane, odbijajuće sile nastaju usled prevelike blizine atoma, gde dolazi do interakcije atoma (Paulijevo odbijanje). Najpoznatija i najčešće korišćena funkcija van der Valsovog potencijala je Lenard-Džonsova 12-6 funkcija, koja opisuje interakciju između dva atoma na sledeći način:

$$V_{vdV}(r) = 4\epsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$

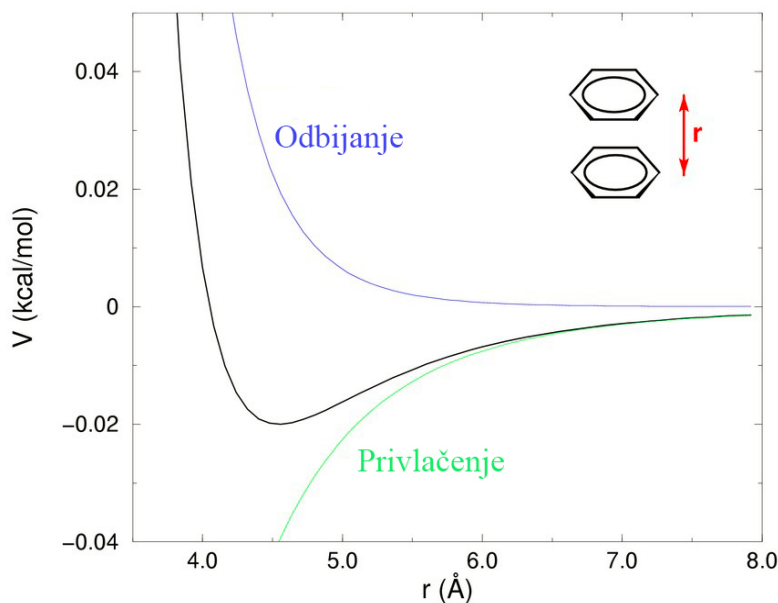
$V_{vdV}(r)$ = potencijalna energija

σ_{ij} = rastojanje pri kojem dolazi do sudara atoma (energija je jednaka nuli)

r_{ij} = rastojanje između atoma

ϵ = dubina jame potencijalne energije

Problem prilikom računanja nevezujućih energija (npr. van der Valsove i elektrostatičke) predstavlja činjenica da se energija interakcije približava nuli u obliku logaritamske funkcije sa povećanjem rastojanja (slika 10). Iz ovoga sledi da je empirijski veoma zahtevno izračunati rastojanje gde prestaje interakcija, posebno ukoliko se računa za veliki broj atoma u molekulu. Ovaj problem se rešava postavljanjem granice iza koje uzima da je vrednost interakcije jednaka nuli (cutoff vrednost). Najčešće se kao vrednost uzima rastojanje od 8 do 10Å.



Slika 10. Promena nevezujućih energija sa promenom rastojanja

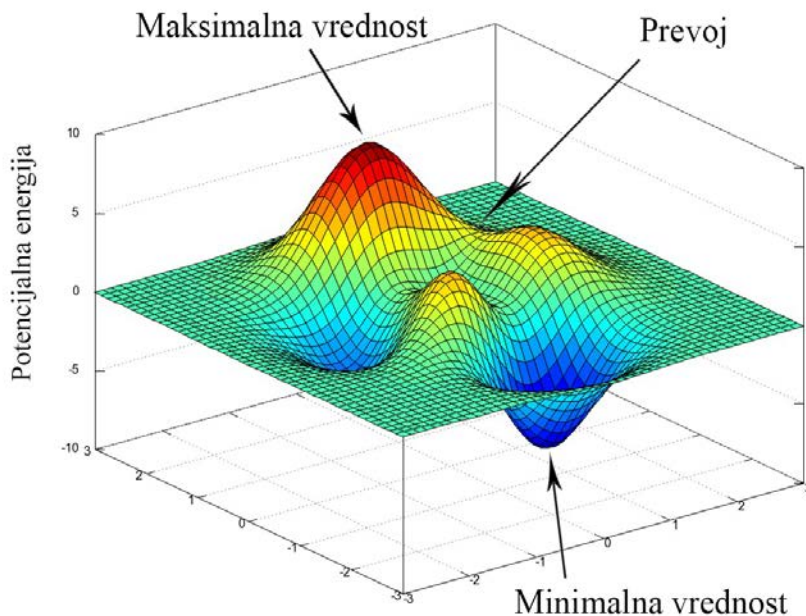
Postoje i drugi parametri koji se koriste za definisanje energija u molekularnom modelovanju, ali oni prevazilaze okvire ove doktorske disertacije.

Na kraju se računa ukupna potencijalna energija sabiranjem svih parametara uzetih u obzir:

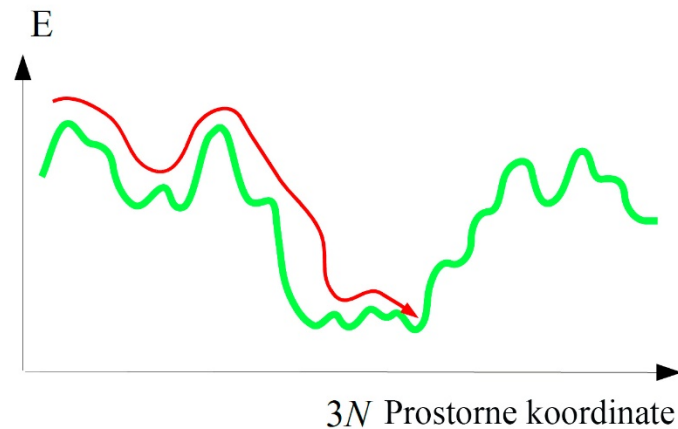
$$E_{potenc.} = \sum_{veza} \frac{1}{2} k_l (l - l_0)^2 + \sum_{uglova} \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{torz.uglovi} \frac{1}{2} k_\phi [1 + \cos(n\Phi - \delta)] \\ + \sum_{i,j} \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}} + \sum_{ij} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$

2.2.6. Minimizacija energije molekula

Pri minimalnoj energiji biomolekuli se nalaze u svom prirodnom stanju (najstabilnija konformacija). Zato je prilikom molekularnog modelovanja neophodno odrediti minimalnu energiju molekula. Na slici 11 je prikazan plot potencijalne energije u zavisnosti od nekoliko hipotetičkih parametara. Minimalne vrednosti energije (lokalni minimum) nalaze se na onim mestima gde će se pomeranjem u bilo kom pravcu postići veća energija od početne. Metode minimizacije energije sadrže algoritme koji prihvataju nove konformacije ukoliko im je energija manja od prethodne (metoda najoštrijeg smanjenja energije). Međutim, usled broja i složenosti parametara koji utiču na izgled plota, u mnogim slučajevima neophodne su modifikacije navedenih algoritama koje preskaču lokalne minimume u potrazi za globalnim minimumom – na primer metode molekularne dinamike (slika 12). To se uglavnom postiže ili pokretanjem više minimizacija istovremeno ili nasumičnim preskakanjem konformacija u potrazi za najoptimalnijom (Monte Karlo metode).



Slika 11. Zavisnost potencijalne energije od nekoliko hipotetičkih parametara



Slika 12. Potraga za globalnim minimumom

2.3. Određivanje strukture proteina

Do početka 2014. godine poznato je oko 90000 trodimenzionalnih struktura proteina u bazama podataka (PDB). U isto vreme, poznato je oko 170 miliona genskih sekvenci koje čine preko 15 miliona gena (NCBI) i kodiraju preko 40 miliona proteina (UNIPROT). Ova razlika između broja sekvenci i određenih struktura nastavlja da raste uprkos sve uspešnijim metodama za određivanje trodimenzionalne strukture. Usled toga, javlja se potreba za razvijanje metoda za predviđanje trodimenzionalne strukture proteina.

Teoretski, trodimenzionalna struktura proteina mogla bi se predvideti pomoću osnovnih zakona fizike samo iz sekvence proteina (Anfinsen 1973). Međutim, problem predstavljaju razlike u energijama nativnog i nesavijenog proteina (oko 1 kcal/mol), kao i neodređenost konstanti neophodnih za proračun. Danas postoje softveri koji mogu predvideti strukturu na osnovu sekvence, međutim njihova pouzdanost je ograničena do određenog nivoa strukture (to su *ab initio* metode, dok je od softvera najpoznatiji Robetta server – <http://robetta.bakerlab.org>). Određivanje preciznog položaja bočnih lanaca, rotamera i stepena protonacije, nije pouzdano ili moguće navedenim metodama.

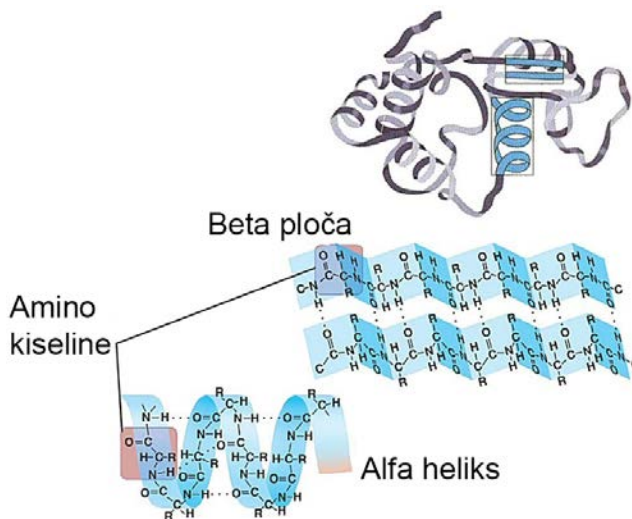
Iako nismo u mogućnosti predvideti detaljnu strukturu proteina iz sekvence, postoje metode homologog modelovanja koje pružaju mogućnost predviđanja strukture na osnovu homologih sekvenci proteina čija je struktura eksperimentalno utvrđena. Ovakvo modelovanje se

zasniva na činjenici da proteini sa sličnom sekvencom imaju slične trodimenzionalne strukture (Liu i Rost 2002).

Prvi koraci u predviđanju trodimenzionalne strukture proteina napravljeni su u predviđanju njihove sekundarne strukture. Razvijene metode i dobijeni rezultati pokazali su se kao vrlo pouzdani i korisni u mnogim biološkim istraživanjima. Razvijene metode su dalje poslužile za razviće drugih metoda za predviđanje proteinskih struktura (organizaciju domena, kontakt sa rastvaračem, itd.). Sekundarna struktura proteina predstavlja lokalne konformacije najčešće u vidu heliksa i/ili ploče. Sekundarna struktura nastaje usled delovanja vodoničnih veza između atoma polipeptidnog kostura. Takođe, ulogu u stabilizaciji strukture ima i rastvarač, kao i ostali atomi u polipeptidu. Stvaranje sekundarne strukture predstavlja prvi korak u krajnjoj konformaciji polipeptida, te s toga ima važnu ulogu za njegovu krajnju strukturu i funkciju.

Zanimljivo je da je prvi pokušaj predviđanja sekundarne strukture proteina bio pre pojave trodimenzionalnih eksperimentalnih podataka (Pauling i Corey 1951). Metod koji su razvili Pauling i saradnici bazirao se na interakcijama vodoničnih veza unutar proteinskog kostura. U Kasnije su razvijene metode zasnovane na neuronskim mrežama kao što su PSI-PRED, JPRED ili Predict Protein, koje su dostupne na veb serverima (<http://bioinf.cs.ucl.ac.uk/psipred>, <http://www.compbio.dundee.ac.uk/www-jpred>, <http://www.predictprotein.org>). Takođe, razvijene su i metode koje prepoznaju sekundarne strukture, a baziraju na elektrostatičkim interakcijama (DSSP metoda), ϕ - ψ torzionim uglovima (STRIDE metoda), $C\alpha$ koordinatama (DEFINE metoda) ili geometrijskoj analizi proteinskog zakrivljenja (P-Curve metod) (Kabsch i Sander 1983; Richards i Kundrot 1988; Sklenar i sar. 1989; Frishman i Argos 1995).

Sekundarna struktura proteina ima osam različitih formi: tri tipa heliksa (α heliks - čine je četiri rezidue, π heliks - čine je pet rezidua i 3_{10} heliks - čine je tri rezidue), β ploča, β most, okret, traka i neodređene strukture koje ne pripadaju ni jednoj navedenoj grupi, Međutim, obzirom da α heliks i β ploča forme čine u proseku 50% strukture proteina, metode koje automatski određuju sekundarnu strukturu proteina prilagođene su da predvide tri forme: α heliks, β ploču i ostalo (slika 13). Pouzdanost svake razvijene metode se proverava i izražava kroz SOV vrednost (Segment Overlap score). SOV predstavlja prosečno preklapanje predviđene i eksperimentalno utvrđene strukture. Savremene metode postižu pouzdanost od oko 76% SOV.



Slika 13. Sekundarna struktura proteina

Značaj predviđanja sekundarne strukture ogleda se u pretrazi strukturalnih sličnosti između proteina čije sekvence nisu homologe ili nemaju dovoljnu sličnost da bi se uzele u obzir kao homologe (Rost 1995; Fischer i Eisenberg 1996; Russell i sar. 1996). Naime, utvrđeno je da neki proteini mogu imati i do 70% različitu AK sekvencu, dok im je struktura slična i funkcija ista (Zvelebil i sar. 1987; Benner i Gerloff 1991; Abagyan i Batalov 1997; Park i sar. 1998). Dakle, struktura proteina konzervativnija je od aminokiselinske sekvence, što ukazuje da funkcija prvenstveno zavisi od strukture. Usled toga predviđanje sekundarne strukture može biti upotrebljeno za pretraživanje proteina sa istom ili sličnom funkcijom. Razvijene metode daju dovoljno precizne rezultate da se mogu koristiti dalje u predviđanju trodimenzionalne strukture proteina, kao i njihove funkcije.

Razvijeni su mnogi softveri za predviđanje sekundarne strukture proteina. Prvi internet servis za predviđanje sekundarne strukture proteina koji je postao dostupan online je PredictProtein (Sen i sar. 2005; Schneider 1989). Postao je dostupan 1992. godine i sadrži PHD/PROF automatske metode predviđanja. Postoje i drugi javno dostupni servisi kao što su: PROFsec, PHDsec, PHDpsi, PSIPRED, itd. (Rost i sar. 1994; Jones 1999; Przybylski i Rost 2002; Rost 2005).

Usled trodimenzionalne konformacije proteina, neke rezidue se nalaze na površini, izložene su kontaktu sa rastvaračem i dostupne su za interakcije sa drugim molekulima. S toga je njihova uloga u proteinu funkcionalna. Sa druge strane, rezidue koje nisu izložene rastvaraču

verovatnije imaju ulogu u njegovim intramolekulskim vezama i stabilnosti proteina. Određivanjem rezidua koje su dostupne rastvaraču i onih koje to nisu, mogu se odrediti funkcionalno i strukturalno važne rezidue, što u značajnoj meri može olakšati određivanje funkcije celog proteina.

Iako predviđanje sekundarne strukture predstavlja važan korak, ipak nije dovoljan za predviđanje trodimenzionalne strukture proteina, neophodno je utvrditi i udaljenost između sekvenci, što se postiže konstruisanjem 2D mapa udaljenosti. Za njihovu konstrukciju potrebne su mape interakcije rezidua. Najveći problem u konstrukciji mapa interakcije predstavljaju poteškoće u određivanju razdaljine gde prestaje interakcija između rezidua. Drugim rečima, neophodno je postaviti graničnu razdaljinu gde se smatra da rezidue više nemaju kontakta. Ta granica se naziva razdaljina razgraničenja (cut-off distance – detaljnije u poglavlju 2.2.5). Ne postoji sveopšte prihvaćena granica razgraničenja, ali je najčešće prihvaćena razdaljina 8-10Å.

Veliki proteini se često (preko tri četvrtine) sastoje od subjedinica koje imaju kompaktnu strukturu i uglavnom se savijaju nezavisno jedna od druge (Liu i Rost 2004). Veoma često, određeni domeni su povezani sa određenim biološkim funkcijama. Poznavanje organizacije domena u proteinu nepoznate trodimenzionalne strukture, značajno može pomoći i u eksperimentalnim i u računarskim postupcima za određivanje njegove strukture.

2.4. Homologo modelovanje

Još daleke 1960. godine nagovešteno je da na trodimenzionalnu strukturu proteina presudnu ulogu ima aminokiselinska sekvenca proteina. Sproveden je jednostavan eksperiment u kojem je denaturišuća supstanca dodata u proteinski rastvor, što je uslovalo narušavanje strukture i funkcije proteina. Nakon uklanjanja supstance proteini su povratili svoju funkciju, na osnovu čega je zaključeno da su ponovo zauzeli aktivnu konformaciju u kojoj su bili pre tretmana (Anfinsen i sar. 1961; Anfinsen 1973). Naravno, kasnije je utvrđeno da to nije uvek potpuno vraćanje strukture proteina i da su mnogim proteinima potrebni još neki faktori da bi povratili strukturu, ali navedeni eksperiment je otvorio mogućnost ispitivanja mehanizama povezanosti trodimenzionalne strukture i sekvence proteina. Međutim, mehanizmi koji dovode do savijanja proteina nisu do danas u potpunosti poznati da bi mogli samostalno poslužiti kao osnova za

predviđanje trodimenzionalne strukture. S toga se upotrebljavaju metode koje koriste eksperimentalno utvrđene trodimenzionalne strukture čije sekvence se poravnavaju sa ispitivanim sekvencama prilikom modelovanja željenog proteina.

Homologi proteini nastaju evolucijom i imaju zajedničkog pretka. Razlike tokom evolucije nastaju putem mutacija, insercija ili delecija. Ipak, navedene promene često su vrlo male, a time je i dugo evolutivno vreme potrebno da se značajnije promeni struktura i funkcija proteina (u zavisnosti od značaja uloge proteina koju ima za organizam koji ga sintetiše). Usled toga, kroz filogenetsko stablo se javlja niz homologih sekvenci koje imaju sličnu strukturu i mogu imati sličnu funkciju.

Čim je ustanovljeno da proteini mogu imati sličnu strukturu počele su se razvijati kompjuterske metode za predviđanje strukture proteina na osnovu sličnosti sa eksperimentalno utvrđenim strukturama (Browne i sar. 1969). Većna metoda za pravljenje homologih modela sastoji se od sledećih koraka:

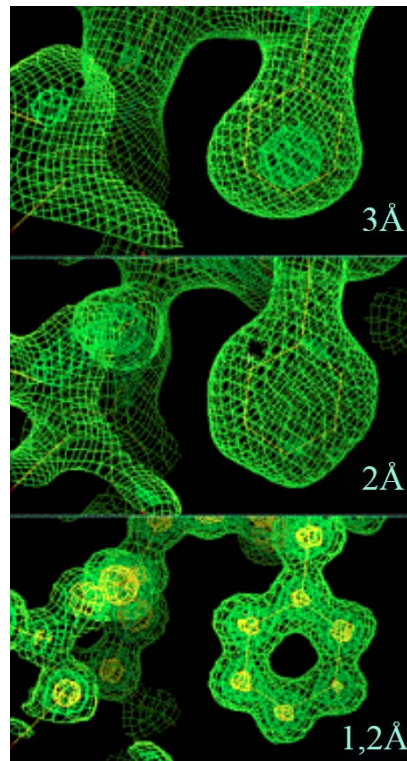
1. Identifikacija srodne strukture koja će poslužiti kao šablon za pravljenje modela i njeno poravnanje sa ispitivanom sekvencom
2. Na osnovu poravnanja sekvenci uzimaju se koordinate podudarajućih regiona kostura strukture kao i podudarajućih sekundarnih struktura
3. Regioni koji se ne podudaraju sa šablon sekvencom (usled delecija/insercija), što uglavnom čine petlje, konstruišu se metodama *de novo* predviđanja strukture
4. Konstruišu se bočni lanci aminokiselina na osnovu podudaranja sa šablon sekvencom i konstruisanim petljama
5. Male popravke modela do dobijanja krajnje strukture
6. Validacija modela

Značajan napredak predstavlja metod razvijen od strane Blundell i sar. (1987), koji su predložili da se kao šablon sekvenca koristi prosečna sekvenca nekoliko odabranih šablon sekvenci (Sutcliffe i sar. 1987; Sutcliffe i sar. 1987). Kasnije su razvijene i kompleksnije metode koje se baziraju na rekonstrukciji struktura primenom minimizacije energije i molekularne dinamike (Sali i Blundell 1993; Li i sar. 1997; Sanchez i Sali 1997; Sahasrabudhe i sar. 1998). Vremenom su navedene metode usavršavane i razvijen je veliki broj softverskih paketa kako za

pojedinačne korake tako i za automatsko pravljenje kompletnih modela. Razvoj softvera omogućen je i razvojem hardverskih komponenti kao što su vešejezgarni procesori i superračunari, koje su neophodne za pokretanje visokozahtevnih aplikacija i izvršavanje proračuna.

Homologi modeli se koriste u slučajevima kada ekperimentalni podaci o trodimenzionalnoj strukturi željenog proteina nisu dostupni. Oni mogu dati i uvid u promene strukture nastale tačkastim mutacijama, bilo da su nastale prirodnim procesima ili laboratorijskim manipulacijama. Takođe, homologi modeli se mogu koristiti za predviđanje efekata potencijalnih mutacija na strukturu, a time i funkciju proteina i za davanje smernica u kom pravcu manipulirati proteinom sa ciljem dobijanja ili poboljšanja željene funkcije. Homologi modeli imaju veoma važnu ulogu i u dizajniranju lekova. Da bi model bio pogodan za takva istraživanja potrebno je da sličnost između sekenci bude veća od 30%(Chothia i Lesk 1986). Primer uspešne primene homologih modela u dizajniranju lekova je otkrivanje inhibitora za HIV proteazu (Weber i sar. 1989; Weber 1990).

Veliki značaj za razvoj homologog modelovanja ima razvoj i napredak u sekvencioniranju genoma živih organizama. Što je veći broj dostupnih sekvenci to je veća verovatnoća da će protein koji modelujemo imati poznatu sekvencu, a time i mogućnost da se napravi homologi model njegove trodimenzionalne strukture. Takođe, veoma važnu ulogu ima razvoj metoda i napredak u strukturalnoj genomic i eksperimentalnom određivanju trodimenzionalne strukture proteina. Povećanjem rezolucije kristalografije rendgenskim zracima povećava se preciznost dobijenih struktura, a time i homologih modela napravljenih na osnovu tih struktura (slika 14) (Lowe 2013). Sve veći broj dostupnih eksperimentalno određenih struktura proteina povećava verovatnoću za pronalaženje odgovarajuće šablon sekvence koja će poslužiti za modelovanje, što predstavlja ključni momenat u homologom modelovanju. Ukoliko se sličnost ispod 30%, u dobijenom modelu se mogu pojaviti značajna odstupanja od realnosti (Chothia i Lesk 1986). Ukoliko je sličnost oko 20%, odstupanja u kosturu modela i šablon sekvence su velika i daju nepozdane rezultate. Kao što je već napomenuto, sličnost sekvenci preko 30% daje pouzdane rezultate za primenu za ispitivanje interakcija proteina i liganda (ligand doking).



Slika 14. Primer primene različitih rezolucija na istom uzorku

Nakon što odaberemo protein čiji trodimenzionalni model želimo da napravimo, prvi korak u konstrukciji modela je pretraživanje genomskih i proteomičkih baza podataka da ustanovimo da li je njegova aminokiselinska sekvenca poznata. Pretraživanje se vrši pomoću specijalizovanih programa. Najčešće korišćeni i do sada najprecizniji program je PSI-BLAST (Altschul i sar. 1997). Program pretražuje najsveobuhvatnije baze podataka kao što su NCBI baza podataka proteinskih sekvenci ili UniProt baza podataka (Bairoch i sar. 2005; Benson i sar. 2005).

Nakon odabira željene sekvence vrši se pretraga baze trodimenzionalnih struktura proteina – PDB database u potrazi za homologim sekvencama čija je trodimenzionalna struktura određena. Ukoliko je više struktura homologo željenoj sekvenci vrši se odabir najpogodnije (ili više njih) za konstrukciju modela. Odabir se vrši na osnovu procenta sličnosti sekvenci i e vrednosti (koja predstavlja verovatnoću da sekvence nisu homologe). Takođe, pregledanjem poravnanja upoređenih sekvenci oučavaju se delovi željene sekvence koji su dobro očuvani u odnosu na pretraživane. Može se javiti slučaj da je željena sekvenca u različitim regionima slična sa različitim sekvencama. U tom slučaju pravi se hibridna struktura (Sutcliffe i sar. 1987). Ukoliko je struktura jednog proteina najpogodnija za pravljenje modela, vrši se pretraga da li

postoji više eksperimentalno određenih struktura datog proteina. U tom slučaju najpogodnija se bira na osnovu najveće rezolucije ili najpogodnijeg liganda (u zavisnosti od cilja istraživanja). Na kraju, uzimaju se u obzir i drugi eksperimentalni podaci koji mogu biti relevantni za željenu ili šablon sekvencu pri konstrukciji modela kao što su vezivna mesta sa ligandima ili supstratima, inhibitori, metal ili drugi joni uključeni u strukturu, itd.

Tokom vremena razvijene su brojne metode za konstrukciju modela. Razlikuju se kako u složenosti tako i u preciznosti rezultata koje daju.

1. Jednostavan model – zadržava se kostur i konzervirane bočne lance šablon sekvence u čije koordinate se ubacuje željena sekvenca koristeći poravnanje sekvenci kao vodič i grade preostali bočni lanci koristeći odgovarajući program za tu namenu, na primer SCWRL (Bower i sar. 1997; Dunbrack 1999; Canutescu i Dunbrack 2003);
2. Stepnasti model – zadržava srž kostura od šablon sekvence, bez delova sa insercijama i delecijama koje gradi iznova. Zatim gradi bočne lance srži, predviđa strukture regiona sa delecijama i insercijam primenom metoda za predviđanje petlji i gradi bočne lance predviđenih delova kostura. Na kraju se ceo molekul optimizuje primenom minimizacije energije, Monte Karlo metoda ili metodom molekularne dinamike;
3. Model slagalice (Jigsaw model) – zadržava delove srži kostura koji su zajednički zankoliko strukturalno poravnatih šablon sekvenci, koristeći regione petlji od različitih šablon sekvenci, zadržavajući one petlje gde ne postoje rupe u poravnatoj sekvenci. Ukoliko ne postoje petlje bez rupa, one se moraju *de novo* modelovati;
4. Globalni model – gradi ceo protein iz prostorno fiksiranih delova preuzetih od poznatih struktura i poravnanja sekvenci, na primer Modeller ili Rosetta (Sali i Blundell 1993; Sanchez i Sali 1997).

Nije uvek složeniji metod i bolji metod. Ukoliko metod dozvoljava pomeranje sekundarne strukture, nastaju velike promene strukture u poređenju sa početnom i povećava se verovatnoća da model ima greške. Takođe, postoje istraživanja koja pokazuju da prekomerna upotreba minimizacije energije i molekularne dinamike mogu udaljiti model od njegove prave strukture (Koehl i Levitt 1999; Fan i Mark 2004).

2.4.1. Modelovanje bočnih lanaca

Obzirom da struktura bočnih lanaca ima ključnu ulogu u građenju aktivnih i vezivnih mesta enzima i proteina njihovo modelovanje predstavlja veoma važan korak u homologom modelovanju proteina. Objavljeno je nekoliko metoda koje rešavaju problem modelovanja bočnih lanaca (Levitt i sar. 1997; Liang i Grishin 2002; Peterson i sar. 2004).

Iako su metode za predviđanje konformacione strukture bočnih lanaca dobro opisane u literaturi, mali broj programa je javno dostupan (tabela 1). Skoro svi programi pretpostavljaju da je kostur proteina čiji se bočni lanci modeluju fiksiran. Kostur protein može biti kako homologi model, tako i eksperimentalna struktura određena rendgenskim zracima. Samo mali broj metoda koristi koordinate neidentičnih bočnih lanaca iz šablon sekvence.

Tabela 1. Javno dostupni programi za predviđanje konformacije bočnih lanaca proteina

Program	Dostupnost	Internet adresa
SMD	download	http://condor.urbb.jussieu.fr/Smd.php
Confmat, Decorate	web	http://lorentz.immstr.pasteur.fr/website/projects
CARA/GeneMine	download	http://www.bioinformatics.ucla.edu/genemine
RAMP	download	http://www.ram.org/computing/ramp
SCAP	download	http://honiglab.cpmc.columbia.edu/programs/sidechain
SCWRL	download	http://dunbrack.fccc.edu/scwrl
Maxsprout/Torso	web	http://www.ebi.ac.uk/maxsprout
SCATD	download	http://www.bioinformatics.uwaterloo.ca/%7Ej3xu
PLOP	download	http://francisco.compbio.ucsf.edu/~jacobson/plop_manual

Skoro sve metode predviđanja bočnih lanaca zasnivaju se na konceptu rotamera bočnih lanaca koji predstavljaju njihove niskoenergetske konformacije (Dunbrack 2002). Analizom strukture organskih molekula i polipeptida, utvrđeno je da bi proteinski bočni lanci trebalo da imaju ograničen broj konformacija usled prostornih i torzionih sila interakcija unutar svakog lanca i između lanca i kostura proteina (Ramachandran i sar. 1963; Sasisekharan i Ponnuswamy 1970; 1971). Povećanjem broja eksperimentalno određenih kristalnih struktura proteina, pojavila se mogućnost pravljenja biblioteka rotamera, čija se preciznost i poizdanost povećava sa razvojem tehnika kristalografije (povećanje rezolucije utvrđenih struktura) (Bhat i sar. 1979; Benedetti i sar. 1983; Dunbrack i Karplus 1994; Dunbrack i Cohen 1997; Lovell i sar. 2000). Novije biblioteke imaju preko 850 struktura rezolucije od 1,7Å.

2.4.2. Modelovanje petlji

Prilikom stepenaste konstrukcije homologih modela, moraju se konstruisati delovi kostura koji se razlikuju između željene i šablon sekvence (na osnovu poravnanja sekvenci). Prilikom konstrukcije petlji u homologom modelu neophodno je rešiti nekoliko pitanja: 1) koliki deo šablon sekvence se uklanja pre modelovanja petlji, 2) da li sve bočne lance zadržanog jezgra modelovati pre modelovanja petlji i 3) da li spojene petlje modelovati zajedno ili pojedinačno. Koliki deo šablon sekvence će se ukloniti zavisi od rezultata poravnanja sekvenci. Ukoliko postoje insercije i delecije, uobičajeno rešenje je da se taj deo briše i ponovo modeluje. Na primer, algoritmi za poravnanje sekvenci daju poravnanja koja kao rezultat imaju najkraću sekvencu. Stoga će poravnanje sledećih sekvenci izgledati kao na slici 15a, a ne kao na slici 15b:

a) AGVEPMENYKLS	b) AGVEPMEN - - - YKLS
SG - - - L DDFKLT	SGL - - - - - LDDFKLT

Slika 15. Poravnanje sekvenci

Međutim, poravnanje sa slike 15b je verovatnije, što će reći da je 5 rezidua iz prve sekvence zamenjeno sa 3 rezidue u drugoj sekvenci (Alexandrov i Luethy 1998). Usled ovih poteškoća u određivanju realnog stanja, deo između konzervativnih rezidua se ponovo modeluje.

I pored činjenice da se petlje u proteinskoj strukturi često ne mogu odrediti ni eksperimentalnim putem kristalografijom rendgenskim zracima, predviđanje njihove strukture najčešće se bazira na podacima eksperimentalno određenih struktura. Petlje se dele po veličini na kratke (1 – 4 rezidue), srednje (5 – 8 rezidua) i dugačke (9 i više rezidua). Sibanda i Thornton (1991) klasifikovali su beta okret konformacije u homologom modelovanju. One predstavljaju kratke petlje koje povezuju dve anti paralelne beta ploče. Pojavljuju se u malom broju konformacija u zavisnosti od sekvence same petlje, posebno od prisustva glicina i prolina na određenim mestima. Njihove strukture se razvrstavaju na osnovu položaja aminokiselina u Ramačandranovom plotu (detaljnije o Ramačandranovom plotu u poglavlju 2.4.4). Najčešće se razvrstavaju kao α_R , β_P , β_E , γ_R , α_L i γ_L . Programi kao što su BTPRED i BHAIRPRED predviđaju položaje specifičnih tipova beta okreta na osnovu proteinske sekvence i sekundarne strukture proteina (Kumar i sar. 2005; Shepherd i sar. 1999). Petlje srednje dužine se takođe

klasifikuju po konformaciji rezidua u kosturu (α R, β P, β E, γ R, α L i γ L) (Donate i sar. 1996). Dugačke petlje mogu povezivati susedne strukture, ali i one koje su prostorno udaljene. Dugačke petlje se klasifikuju na trakaste (veoma dugačke), Ω petlje, koje podsećaju na grčko slovo omega i ζ petlje, koje su nepolarne i imaju cik-cak strukturu.

Najčešće metode koje se koriste u modelovanju petlji koriste postojeće podatke od drugih proteinskih struktura (Jones i Thirup 1986; Fechteler i sar. 1995; Bates i sar. 1997; Heuser i sar. 2004; Fernandez-Fuentes i sar. 2005). Program meri položaj i orijentaciju delova kostura neposredno pored dela koji treba da se modeluje i zatim pretražuje PDB bazu podataka u potrazi za segmentima iste dužine koji se protežu u istom rasponu i imaju istu orijentaciju (Jones i Thirup 1986). Pored metoda koje se zasnivaju na bazama podataka postoje i metode koje predviđaju strukturu petlji bez početne strukture (ab initio). Ukoliko je početna konformacija slučajno odabrana, program kreće sa N ili C kraja lanca, tako da drugi kraj nije vezan za protein. Zatim se petlja povezuje sa drugim krajem lanca (zatvara) primenom metoda minimizacije energije ili metodama koje se zasnivaju na jednačinama iz kinematike koje omogućavaju određivanje parametara koji obezbeđuju krajnju poziciju petlje (Canutescu i Dunbrack 2003; Boomsma i Hamelryck 2005). Takođe, postoje i metode koje prave Ramačandranove plotove nasumičnih konformacija petlje koja želi da se modeluje i uzimaju one delove koji se uklapaju u prostor koji imaju i tako zatvaraju petlju. MODELLER koristi metodu koja prvo zatvori petlju nasumičnom konformacijom i onda je relaksira primenom minimizacije energije (Fiser i sar. 2000).

Prilikom razvoja metoda koje se baziraju na slučajno odabranim početnim konformacijama važan aspekt predstavlja funkcija koja računa slobodnu energiju na osnovu koje se biraju zadovoljavajuće konformacije koje povezuju proteinske krajeve i grade petlju. Funkcije slobodne energije koje računaju potencijalnu energiju molekularnom mehanikom i određivanjem površine dostupne rastvaraču (primenom MM/PBSASA metode – Molecular Mechanics energetsko polje / Poisson-Boltzmann Solvent Accessible Surface Area), imaju mogućnost da identifikuju lažne pozitive koji imaju strukturu vrlo sličnu nativnoj (Fogolari i Tosatto 2005).

2.4.3. Modelovanje kompletne strukture proteina

Programi koji automatski prave modele počinjući sa proteinskom sekvencom i završavajući sa kompletnim modelom imaju mogućnost greške i njihove rezultate treba naknadno

pažljivo i temeljno proveriti. Takođe, većina takvih programa ima korake u kojima dopušta manualnu intervenciju (na primer odabir šablon sekvence ili metod poravnanja sekvenci). Takve intervencije mogu imati važnu ulogu u krajnjem rezultatu. Postoji nekoliko programa koji prave kompletne modele iz početne sekvence. Najpoznatiji i najprimenjeniji su MODELLER i MolIDE (Sali i Blundell 1993; Sanchez i Sali 1997; Canutescu i Dunbrack 2005). Takođe, postoji i nekoliko on-line programa koji pružaju usluge homologog modelovanja. Najpoznatiji i najčešće korišćeni je SWISS-MODEL (Peitsch 1996; Guex i Peitsch 1997). Lista dostupnih programa data je u tabeli 2.

Tabela 2. Javno dostupni programi za kompletno homologo modelovanje proteina

Program	Dostupnost	Internet adresa
3d-JIGSAW	web	http://www.bmm.icnet.uk/servers/3djigsaw
CPHmodels	web	http://www.cbs.dtu.dk/services/CPHmodels
EsyPred	web	http://www.fundp.ac.be/urbm/bioinfo/esypred
FAMS	E-mail server	http://www.pharm.kitasato-u.ac.jp/fams
Geno3D	web	http://geno3d-pbil.ibcp.fr
MODELLER	download	http://salilab.org/modeller
ModWeb	web	http://salilab.org/modweb
Modzinger	web	http://peyo.ulb.ac.be/mz/index
Nest	download	http://honiglab.cpmc.columbia.edu/programs/nest
Parmodel	web	http://laboheme.df.ibilce.unesp.br/cluster/parmodel_mpi
Robetta	web	http://robeta.bakerlab.org
SDSC	web	http://cl.sdsc.edu/hm
SWISS-MODEL	Web/download	http://swissmodel.expasy.org//SWISS-MODEL

MODELLER koristi šablon strukture za određivanje granične vrednosti (udaljenosti između atoma, torzione uglove, itd.) koji se kombinuju sa statističkim vrednostima distribucije atoma preuzetih od više homologih struktura kombinovanih u pdb fajl. Kao rezultat dobija se uslovna funkcija verovatnoće distribucije za stepene slobode datog proteina. Na primer, funkcija verovatnoće za torziju neke rezidue u kosturu pravi se na osnovu kombinovanih podataka iz poznate strukture i informacija o Ramačandranovoj distribuciji date rezidue. Granične vrednosti se kombinuju sa CHARMM (Chemistry at HARvard Macromolecular Mechanics) energetskim poljem i pravi se funkcija koja se optimizuje gradijentalnom minimizacijom i molekularnom dinamikom (Brooks i sar. 2009).

SWISS-MODEL je pravljen da bude internet program za kompletno modelovanje od sekvence do modela koji se e-mailom šalje poručiocu (Peitsch 1996; Guex i Peitsch 1997). Danas

podtoji i instalaciona verzija pod nazivom Deep View. SWISS-MODEL prati standardni protokol identifikacije homologih struktura, poravnanja sekvenci, pravljenja kostura, i modelovanja bočnih lanaca i petlji. Za pretragu baza podataka koristi BLAST program i prihvata homologu strukturu za šablon ukoliko je sličnost najmanje 30%. Postoji mogućnost ručnog odabira šablon sekvence i poravnanja za pravljenje homologog modela. Ukoliko postoji više pogodnih struktura, program ih kombinuje tako da šablon struktura bude što sličnija sekvenci koju želimo da modelujemo. Srž kostura se pravi na osnovu poravnanja sekvenci. Uzima se srednja vrednost struktura na osnovu lokalnih sličnosti sa sekvencom koja se modeluje. Petlje se prave na osnovu pretrage PDB baza podataka koristeći $C\alpha$ atome na oba kraja kao početne pozicije. Obzirom da se ova metoda koristi samo za $C\alpha$ atome, kostur se kompletira pretragom pentapeptidnih segmenata u PDB bazi koji se uklapaju u $C\alpha$ nit petlje. Zatim se konstruišu bočni lanci za one rezidue koje se ne poklapaju sa šablon sekvencom koristeći najčešći rotamer date rezidue, uz uslov da ne postoji prepolkapanja struktura. Na kraju se pomoću GROMOS programa vrši minimizacija energije (van Gunsteren i sar. 1995).

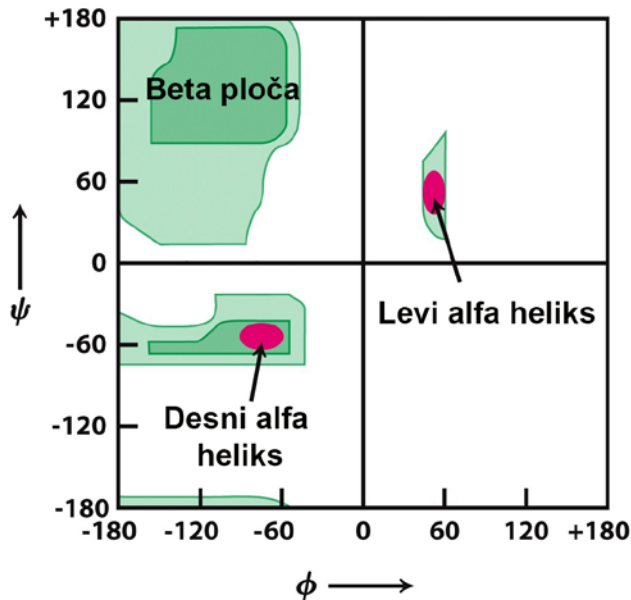
2.4.4. Značaj homologog modelovanja

Kao što je već istaknuto, značaj homologog modelovanja se prvenstveno ogleda u predviđanju strukture, funkcije i evolucije proteina. U nedostatku eksperimentalnih podataka homologni modeli u mnogim situacijama mogu zadovoljiti kriterijume potrebne za dobijanje pouzdanih informacija o ispitivanom proteinu. Prilikom primene homolognih modela mora se voditi računa o pouzdanosti modela koji se koristi. Što je veća evolutivna udaljenost šablon i ispitivane sekvence, raste verovatnoća da je pouzdanost modela manja.

Obzirom na postojanje velikog broja programa koji prave homologe modele, postavlja se pitanje pouzdanosti pojedinih programa i koji daje najbolje rezultate. Izvršena je studija koja je poredila nekoliko programa koji prave homologe modele na osnovu šablon sekvenci i poravnanja sa ispitivanom sekvencom (Wallner i Elofsson 2005). Ispitivanje je uključilo sledeće programe: MODELLER, SegMod/ENCAD, SWISS-MODEL, 3D-JIGSAW, Nest, Builder i SCWRL. Rezultati su pokazali da MODELLER, Nest i SegMod/ENCAD daju pouzdanije rezultate od ostalih ispitivanih programa.

Sa razvojem metoda za modelovanje strukture proteina, razvijale su se i metode koje proveravaju pouzdanost dobijenih modela (validacija modela) u odsustvu eksperimentalnih

podataka. One se zasnivaju na proveri stereoheimijskih parametara koji moraju zadovoljiti određene kriterijume. Najčešće korišćeni programi za ovu namenu su ProCheck i WhatCheck (Vriend 1990). ProCheck proverava stereoheimijski kvalitet proteina analizirajući geometriju rezidua jednu po jednu i zatim geometriju cele strukture. WhatCheck vrši obimnu proveru mnogih stereoheimijskih parametara rezidua koje grade model. Ovi i još neki programi uključeni su u veb server SAVES koji proverava pouzdanost modela dobijenih kako eksperimentalno tako i primenom homologog modelovanja (<http://services.mbi.ucla.edu/SAVES>). MolProbity je takođe veb server koji ima široku zastupljenost (Davis i sar. 2004). On traga za statistički malo verovatnim strukturama u poređenju sa setom veoma pouzdanih struktura visoke rezolucije. Ispituju se loše konformacije rotamera, rezidue koje su preblizu, loši bočni lanci, vodonične veze i ostali mogući nedostaci u ispitivanoj strukturi. MolProbity je takođe server koji ispituje kako eksperimentalne tako i homologe modele. Jedan od prvih metoda za proveru strukture proteina je provera geometrije kostura koji mora da se uklapa u prihvatljive regione Ramačandranovog plotu. Ramačandranov plot je prikaz torzionih uglova koje grade veze N-C α (ugao koji se naziva Phi, ϕ) i C α -C (ugao koji se naziva Psi, ψ) u aminokiselinskim reziduama polipeptidnog kostura (Ramachandran i sar. 1963). Ramačandran je utvrdio da pojedine strukture u kosturu proteina zahvataju uglove u određenom rasponu, kao što je prikazano na slici 16. Ukoliko bi odstupile od datog raspona (ukoliko je rastojanje između nevezanih atoma manje od rastojanja pri van der Valsovim interakcijama) došlo bi do sudara između atoma u strukturi što nije moguće. S toga, delovi strukture koji su izvan preporučenih delova najverovatnije ne odgovaraju realnom stanju. Ramačandranov plot meri rastojanja između atoma i prikazuje njihov raspored u odnosu na dozvoljene vrednosti. Kao što se vidi na slici samo tri regiona na plotu su fizički dozvoljena i svi tipovi sekundarne strukture grade uglovi koj ipripadaju ovim regionima. Sa druge strane, neophodno je naglasiti da geometrijski korektne strukture ne znači da su i ispravne i da odgovaraju realnom stanju, već je to potvrda da su fizički moguće.



Slika 16. Ramačandranov plot

2.5. *De novo* predviđanje strukture proteina

Dok se kod homologog modelovanja koristi struktura homologa ispitivanoj za konstrukciju modela, *de novo* predviđanje strukture gradi model na osnovu same ispitivane sekvence i poznatih pravila o savijanju proteina. Usled toga, ove metode su nepouzdanije od homologog modelovanja, ali ukoliko ne postoji homologa struktura odgovarajuće sličnosti (sličnost je manja od 30%) mogu biti upotrebljene za predviđanje željene strukture sa dozom rezerve u pozdanost rezultata i ograničenom primenom.

Najpoznatiji i najkorišćeniji program za *de novo* modelovanje strukture proteina je Rosetta program (Simons i sar. 1997; Bonneau i sar. 2001; Rohl i sar. 2004). Rosetta je prvenstveno razvijena kao program koji vrši samo *de novo* predviđanje strukture, ali su kasnije dodate i funkcije dokinga, određivanja strukture na osnovu eksperimentalnih podataka, protein-protein interakcija i protein-DNK interakcija.

Rosetta vrši predviđanje strukture proteina u dve faze: 1. pravi model male rezolucije gde pomoću statističkih metoda i uklapanja malih delova gradi opšti izgled proteina; 2. pravi detaljan model proteina korišćenjem detaljnijih statističkih metoda za određivanje položaja rotamera i

poboljšanja strukture kostura proteina. Navedene statističke metode se zasnivaju na Monte Karlo principu.

U prvoj fazi program koristi podatke dostupne u PDB bazi podataka za odabir mogućih konformacija lokalnih delova sekvence. Pravi se biblioteka lokalnih delova sekvence preuzetih iz PDB baze podataka na osnovu sličnosti sa sekvencom proteina koji se modeluje. Ovaj metod daje zadovoljavajuće rezultate za sekvence koje nemaju homologe strukture sa dovoljnom evolutivnom bliskosti u poznatim bazama podataka. Rosetta zatim sklapa lokalne strukture u globalnu strukturu koristeći statističke metode koje uzimaju u obzir hidrofobno pakovanje, uparivanje lanca, kompaktnost i verovatnoću uparivanja rezidua. U drugoj fazi bočni lanci se modeluju do nivoa atoma i rotameri se uzimaju u obzir. Statističke metode uključene u ovu fazu uzimaju u obzir rastvarač, vodonične veze i ostale relevantne parametre za fizičke interakcije između atoma.

Sa povećanjem veličine proteina (dužine sekvence) produžuje se vreme potrebno za konstrukciju *de novo* modela. Takođe, sa povećanjem složenosti proteina povećava se i mogućnost greške u konstrukciji modela. Usled toga konstrukcija *de novo* modela je ograničena na mali broj rezidua, uglavnom do 150. Ipak, postoji mogućnost da se ovo ograničenje donekle premosti zasebnim modelovanjem posebnih domena. Pored ograničenja u dužini lanca, jedan od najvećih problema prilikom *de novo* konstrukcije modela predstavljaju stepeni slobode položaja bočnih lanaca proteina. Uspešno rešavanje položaja bočnih lanaca je ključni korak za ovaj tip modelovanja.

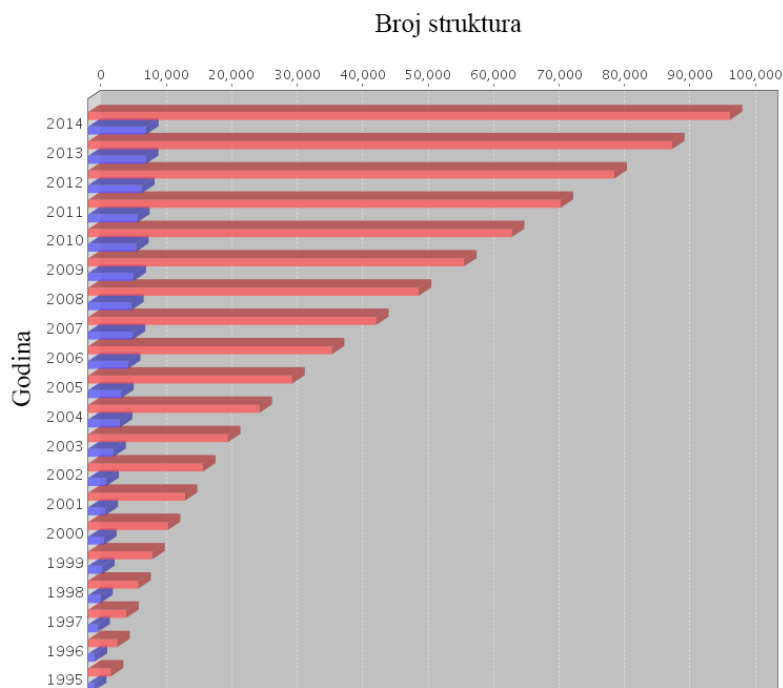
2.6. Eksperimentalno određivanje strukture proteina

Jedan od velikih nedostataka eksperimentalnog određivanja strukture proteina je visoka cena eksperimenata (40000€ do 80000€ po strukturi). Takođe, problem predstavljaju izolacija čistog proteina u dovoljnoj koncentraciji kao i dobijanje kristala, koji su vremenski vrlo zahtevni. S toga je prvenstveni cilj takvih eksperimenata sa fokusa na određivanju strukture proteina koji želimo da ispitamo, pomeren na pokrivanje što šireg dijapazona struktura koje će omogućiti modelovanje što većeg broja trodimenzionalnih struktura proteina (Brenner 2001; Vitkup i sar. 2001). Naravno, strukture nekih proteina se i dalje određuju sa specifičnim ciljem kao što je povezanost sa bolestima i traganje za potencijalnim lekovima. Sve strukture koje se

ekperimentalno odrede zajedno sa referencama koje ih prate, deponuju se u PDB bazu podataka (tabela 3). Porast broja proteinskih struktura deponovanih u PDB baza podataka prikazan je na slici 17 (Berman i sar. 2000).

Tabela 3. Strukture u PDB na dan 27.01.2015.

Experimentalni method	Proteini	Nukleinske kiseline	Protein/NKkompleks	Ostalo	Ukupno
Rendgenskim zracima	88186	1598	4343	4	94131
Nuklearna magnetna rezonanca	9468	1104	222	7	10801
Elektronska mikroskopija	522	29	165	0	716
Hibridna metoda	68	3	2	1	74
Ostalo	161	4	6	13	184
Ukupno	98405	2738	4738	25	105906



Slika 17. Unete strukture proteina u PDB do početka 2015. godine (www.rcsb.org)

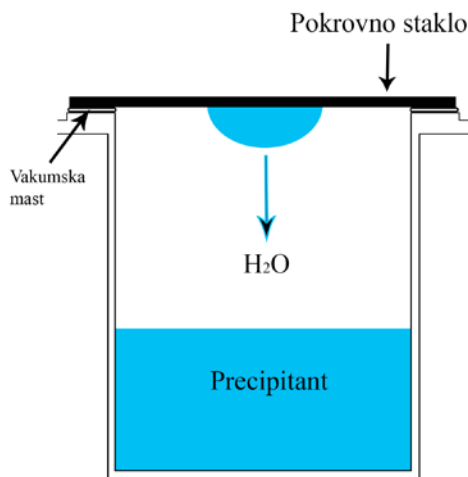
U eksperimentalnom određivanju trodimenzionalne strukture koristi se više metoda: kristalografija rendgenskim zracima, nuklearna magnetna rezonanca (NMR), krio-elektronska mikroskopija (krio-EM) i neutron difrakcija. Međutim, kristalografija rendgenskim zracima se

pokazala kao optimalna uzimajući u obzir metodu i rezultate koje pruža, na prvom mestu visoku rezoluciju (primenjena u skoro 90% ispitivanja struktura u PDB bazi podataka). Pored kristalografije koriste se i nuklearna magnetna rezonanca (čija je prednost ispitivanje native strukture molekula u rastvoru i mogućnost ispitivanja njegove dinamike, jer se kao rezultat dobija 15-20 preklopljenih slika slične strukture) i u ređim slučajevima krio-EM mikroskopija (primenjuje se za velike molekule, organele, vlakna ili cele ćelije), dok se ostale metode izuzetno retko primenjuju.

Prilikom svakog posmatranja pod uveličanjem potrebno je imati sočivo koje će uveličati objekat koji posmatramo i time pojačati signal koji želimo da detektujemo. Nažalost, ne postoji sočivo za rendgenske zrake koji imaju veoma malu talasnu dužinu – između 0.01 i 10 nanometara. S toga je detekcija pojedinačnog molekula njihovom promenom nemoguća, jer je signal suviše slab. Da bi se prevazišao ovaj problem koristi se više molekula koji moraju biti orijentisani na isti način, čime se signal pojačava. Tako orijentisani molekuli predstavljaju kristal. Kristalografija rendgenskim zracima zasnovana je na difrakciji rendgenskih zraka i sadrži četiri glavna koraka: 1. Rekombinantna ekspresija i prečišćavanje proteina, 2. kristalizacija, 3. prikupljanje podataka i 4. konstrukcija modela.

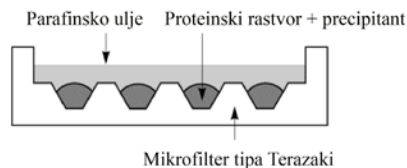
1. Rekombinantna ekspresija i prečišćavanje proteina – prvi korak podrazumeva odabir, kloniranje gena (ukoliko je potrebno) i izolaciju proteina koji se želi ispitati;
2. Kristalizacija – predstavlja kritični korak jer je potrebno optimizovati metodu kristalizacije koja je specifična za svaki protein. Vreme potrebno za ovu fazu može varirati u intervalu od 1 do 12 meseci. Prilikom kristalizacije proteina javljaju se brojne poteškoće. Jedna od njih je konformaciona fleksibilnost proteina usled čega se svega 25% molekula nalazi u istoj konformaciji. Obzirom da proteini imaju nepravilan oblik, njihovo stavljanje u ravnomeran raspored u kristalnoj rešetci (koji je neophodan radi dobrog signala kao što je gore napomenuto) predstavlja veliki izazov. Veze koje održavaju proteinsku trodimenzionalnu strukturu su prilično slabe (vodonične, hidrofobne i elektrostatičke). Proteinski kristali takođe sadrže i rastvarač u procentu oko 50%. Usled toga, njihovi kristali su meki, lako lomljivi i osetljivi na najmanje promene u okolini, što značajno otežava fazu kristalizacije i njihovu manipulaciju. Najčešći postupak kristalizacije je primenom tehnike viseće kapljice

(slika 18). Prilikom procesa kristalizacije, počinje se sa proteinskim rastvorom visoke koncentracije (2-50mg/ml). U čašu se sipa precipitirajući agens, dok se na stakleni pokrov stavi kap proteinskog rastvora ($v/v = 1/1$) i poklopi. Obzirom da je koncentracija agensa u kapljici manja u odnosu na rastvor u čaši (usled dodatka precipitirajućeg agensa), rastvarač isparava, raste koncentracija proteina i u uslovima specifičnim za dati protein počinju da se pojavljuju kristali. Ovi uslovi predstavljaju najveći izazov u procesu kristalografije i potrebno je mnogo pokušaja da bi se oni postigli. U tu svrhu se koriste roboti koji pod mikroskopom fotografišu uspešne kristalizacije (veličina kristala 10-100 μm) koje se vrše u mnogo ponavljanja sa različitim uslovima i na taj način biraju optimalne uslove kristalizacije. Variranje temperatura može u značajnoj meri olakšati kristalizaciju. Postoje i druge varijacije tehnike kristalizacije kao što su tehnika sedeće ili sendvič kapljice, tehnika mikro gomila ili kristalizacija dijalizom (slika 19) (Drenth 2007).

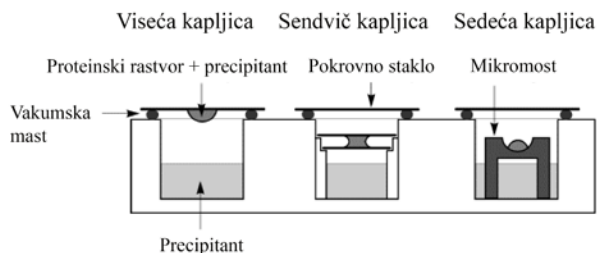


Slika 18. Tehnika viseće kapljice

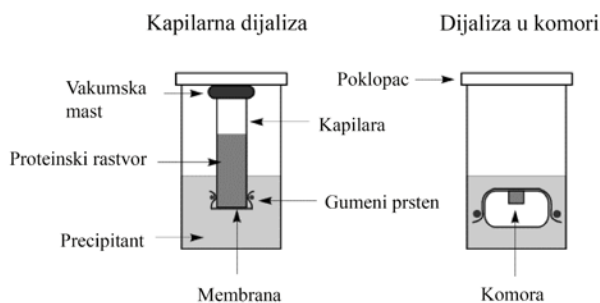
a) Tehnika mikrogomila



b) Difuziono isparljivujuće tehnike



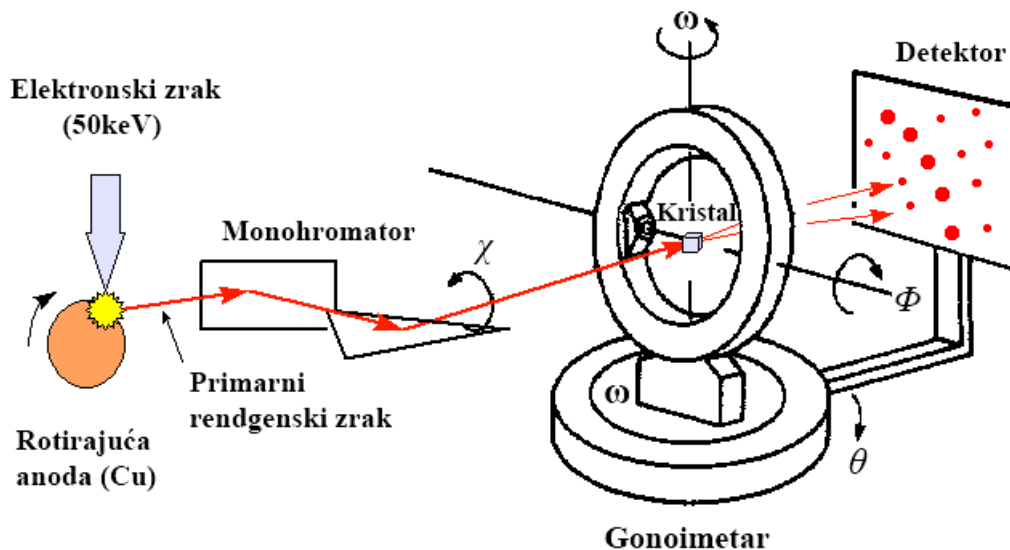
c) Tehnika dijalizom



Slika 19. Različite tehnike kristalizacije proteina

3. Prikupljanje podataka – za izvođenje eksperimenta potreban je izvor rendgenskih zraka (slika 20) (Drenth 2007). Najčešće je to rotirajuća anoda (Cu) koja proizvodi rendgenske zrake velike energije (talasne dužine $1,5\text{\AA}$, što otprilike odgovara dužini C–C veze – $1,2\text{--}1,5\text{\AA}$). Fokusirajuća ogledala ili monohromator se koriste za izdvajanje zraka željene talasne dužine. Zrak dalje prolazi kroz uzorak koji je postavljen na glavu igle gonoimetra koji omogućava precizno pozicioniranje uzorka. Difraktovani zraci se beleže na detektoru (višeslojni detektori ili CCD) i skladište za dalju analizu. Primena sinhrotronskog zračenja značajno unapređuje i olakšava potupak. U svetu postoji nekoliko sinhrotrona, kao na primer u Trstu ili Grenoblu. Kristalne strukture se postavljaju u sinhrotron u specijalno dizajnirane cilindrične, aluminijumske kasete u koje staje 96 kristala. Uzorci se automatski fiksiraju i

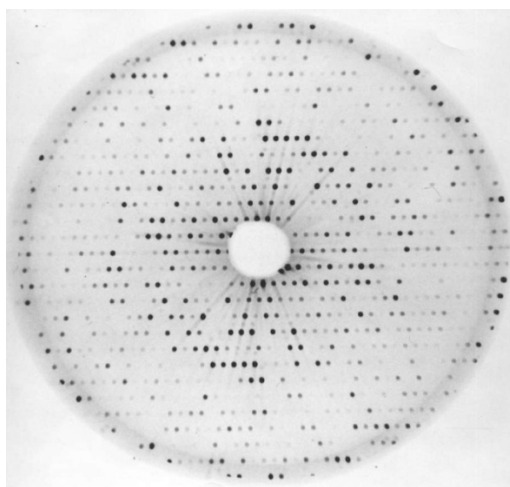
propuštaju se visokoenergetski rendgenski zraci. Detektor beleži podatke i skladišti za dalju analizu. Ceo proces je automatizovan i traje od 30 sekundi do nekoliko sati.



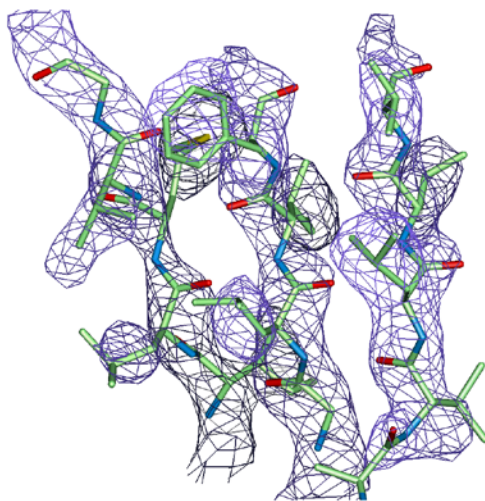
Slika 20. Uopštena šema kristalografije rendgenskim zracima

4. Konstrukcija modela – podaci se prikupljaju u vidu difrakcionih tačaka nastalih difrakcijom rendgenskih zraka prilikom prolaska kroz ravnomerno raspoređen uzorak u kristalnoj rešetci (slika 21) (Drenth 2007). Pozicije i razmak između tačaka su definisane veličinom i tipom kristalne rešetke. Međutim, intenzitet tačaka je funkcija njihovih amplituda koje zavise od tipa atoma, dok se njihove faze baziraju na položaju atoma. Nedostatak faza na dobijenoj slici se u kristalografiji naziva fazni problem. Ovaj problem se uglavnom rešava metodom primene teških atoma ili u sinhrotronima uspostavljanjem početnog seta faza sa različitim talasnim dužinama koje imaju neravnomernu raspodelu (Hendrickson 1991). Prelamanje rendgenskog zraka kroz oblak elektrona specifičnog atoma omogućuje dobijanje početnog seta faza potrebnih da se odredi položaj atoma u kristalnoj rešetci. Nedostatak ovog metoda je što se podaci prikupljaju za svaku talasnu dužinu što produžuje vreme eksperimenta. Takođe, moguće je koristiti i model šablon strukture koji ima određene položaje atoma od koga se uzima početni set faza (molecular replacement). Nakon utvrđivanja faza pravi se mapa gustine elektrona koja dalje služi za konstrukciju modela (slika 22)

(Leach 2001). Nedostatak kristalografije rendgenskim zracima može predstavljati činjenica da protein u kristalnoj rešetki nema uvek identičnu strukturu kao u fiziološkim uslovima. Nakon konstrukcije modela sledi njegova optimizacija postavljanjem kostura i bočnih lanaca u optimalni položaj i finalnom minimizacijom molekula. Krajnji modeli se konstruišu i deponuju u baze podataka bez vodonikovih atoma koji se usled male veličine ne vide rendgenskom kristalografijom ukoliko je rezolucija veća od 1Å. Naknadnim modelovanjem moguće ih je dodati željenom molekulu.



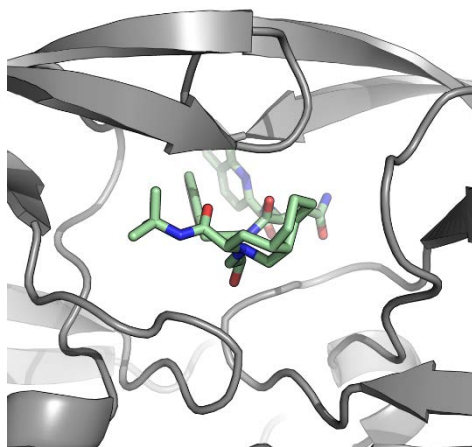
Slika 21. Slika proteina dobijena kristalografijom rendgenskim zracima.



Slika 22. Gustina elektrona i model proteina konstruisana osnovu gustine elektrona

2.7. Predviđanje interakcija protein ligand (ligand doking)

Da bi se odredila i razumela ili čak predvidela funkcija nekog proteina važno je razumeti ili predvideti njegove interakcije sa drugim molekulima (ligandima, inhibitorima, supstratima, itd.). Prilikom ispitivanja interakcija određuje se mogućnost pravljenja kompleksa, njegova stabilnost i geometrija. Potrebno je odrediti gde se veže ligand i koju orijentaciju zauzima. Interakcije se dešavaju između svih tipova molekula, ali ova doktorska disertacija će obraditi prvenstveno interakcije proteina (enzima) sa malim molekulima koje su od posebnog interesa kako za industrijsku enzimologiju tako i za farmaceutsku industriju u dizajniranju lekova (slika 23). Ispitivanje ovih interakcija daje važan uvid u funkciju proteina (inhibicija ili aktivacija funkcije proteina). Pretraga potencijalnih liganada virtuelnim pretraživanjem baza podataka putem dokinga, pored primene u ispitivanju ciljanog molekula, služi i za sužavanje izbora jedinjenja za detaljnija eksperimentalna ispitivanja, čime se štedi vreme i novac.



Slika 23. Protein-ligand kompleks HIV proteaze sa XK263 inhibitorom

Računarske metode za predviđanje interakcija protein-ligand sastoje se iz dva glavna dela: algoritam koji predviđa geometriju potencijalnog kompleksa i funkcija koja predviđa afinitet liganda za protein i jačinu veze koja utiče na stabilnost kompleksa.

Predviđanje geometrije kompleksa podrazumeva uzimanje u obzir stepene slobode. Najznačajniji su orijentacija molekula koji interaguju i konformacija liganda (koji može

interagovati u različitim konformacijama sa proteinom). Usled mogućnosti različitih konformacija liganda, sama geometrija molekula nije dovoljna da bi se odredio niskoenergetski kompleks sa proteinom koji predstavlja aktivnu konformaciju. Takođe i konformacija proteina može varirati, molekuli vode mogu biti pozicionirani u aktivnom mestu i naelektrisanje molekula se može menjati u zavisnosti od sredine u kojoj se nalaze. I najmanje promene u navedenim promenljivima mogu uticati na ceo kompleks i interakciju molekula kroz vodonične ili elektrostatičke veze.

Iako se radi o malim molekulima, njihova veličina ima uticaja na određivanje interakcije sa proteinom, jer ne reaguje uvek ceo molekul liganda. Usled toga u nekim slučajevima postoji mogućnost ispitivanja interakcije samo dela liganda. U tim slučajevima, prostor koji taj deo zauzima je veoma mali i njegova konformacija ne utiče značajno na vezu sa proteinom. Tada se primenjuju algoritmi koji uzimaju u obzir samo orijentaciju molekula (algoritmi rigidnih konformacija). Postoje i algoritmi koji dozvoljavaju fleksibilnost liganda, kao i oni koji konstruišu ligand iz dostupnih baza podataka. U virtuelnom svetu nismo ograničeni na postojeća jedinjenja, tako da se primenom dokinga mogu ispitati i konstruisati hipotetički molekuli koji bi davali željene rezultate.

2.7.1. Funkcije za predviđanje afiniteta liganda za protein i jačinu veze

Funkcija ima zadatak da predvidi različite orijentacije liganda u aktivnom mestu merenjem određenih parametara i utvrdi koja je najverovatnije prisutna u realnom svetu.

Na osnovu zakona o dejstvu masa može se izračunati vezujuća slobodna energija iz koncentracija i hemijske aktivnosti. Iz vrednosti vezujuće slobodne energije možemo izračunati da li će se ligand vezati ili ostati u rastvoru. Odnos između vezujuće slobodne energije $\Delta G_{vezujuća}$ i pojedinačnih koncentracija molekula dat je jednačinom:

$$\Delta G_{vezivanja} = \Delta H_{vezivanja} - T\Delta S_{vezivanja} = -RT \ln \frac{[R-L]}{[R][L]}$$

$[R-L]$ – protein-ligand kompleks

P – protein

L – ligand

R – gasna konstanta

T – absolutna temperatura

Slobodna vezujuća energija sardži entalpijski i entropijski doprinos. Vezujuća entropija proizilazi iz promene stepena uređenosti formacija kompleksa usled promene u translacionim, rotacionim i konformacionim stepenima slobode. Vezujuća entalpija sumira promene u intra- i intermolekulskim silama. U njih spadaju elektrostatičke sile, kao i indukcioni i disperzioni efekat. Sa smanjenjem razdaljine između atoma raste njihovo odbijanje usled sudaranja elektronskih omotača. Vodonične veze (koje pripadaju elektrostatičkim) imaju važnu ulogu u vezivanju molekula (Bursulaya i sar. 2003).

Funkcije afinitetna na osnovu energetskih polja

Funkcije energetskih polja generalno imaju sledeću formu:

$$E_{ukupna} = \sum_{veza} \frac{1}{2} k_l (l - l_0)^2 + \sum_{uglova} \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{torz.uglovi} \frac{1}{2} k_\phi [1 + \cos(n\Phi - \delta)] \\ + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$

Kao što se vidi u jednačini, ukupna energija se računa na osnovu unutrašnjih energija istezanja i vezivanja (nekovalentnih). Energije dužine veza (l_{eq}) u uglova veze (θ_{eq}) računaju se preko harmoničnih potencijalnih funkcija. Rotacije uglova se uključuju primenom periodične funkcije, dok se intramolekulske sile između atoma i i j računaju Lenard-Džonsonovom, kao i Kulonovom jednačinom. Računanje energetskog polja uključuje mnogobrojne parametre karakteristične za određeni sistem. Za biomolekule najčešće se koristi AMBER, zatim CHARMM i GROMOS (Marsden i sar. 2004; Oostenbrink i sar. 2004; Crippen 2013). Nedostatak ovoj jednačini je svakako doprinos entropije, kao i efekta rastvora. Zato se energetska polja uglavnom koriste u molekularnoj dinamici ili Mote Karlo metodu. Tom prilikom se slobodna energija i entropija uključuju primenom velikog broja stanja sistema (kroz Bolcmanovu jednačinu) (Beveridge i DiCapua 1989).

Empirijske funkcije afiniteta

Empirijske funkcije se baziraju na korelaciji geometrijskih parametara kompleksa protein ligand sa izmerenim slobodnim energijama vezivanja. Bohm (1992) je prvi predstavio ovaj metod. U *de novo* programu LUDI, primenio je sledeću formulu na kompleks protein ligand:

$$\begin{aligned} \Delta G_{vezivanja} = & \Delta G_0 \\ & + \Delta G_{hb} \sum_{H-veze} f(\Delta R, \Delta \alpha) \\ & + \Delta G_{jon} \sum_{jonske\ inter} f(\Delta R, \Delta \alpha) + \Delta G_{lipo} |A_{lipo}| + \Delta G_{rot} NROT \end{aligned}$$

Vezujuća slobodna energija se izračunava kao suma vodoničnih veza, jonskih interakcija, lipofilnih dodirnih površina i broja rotirajućih veza u ligandu. Pojedinačni parametri se dobijaju iz multiple regresione analize 45 kompleksa receptor ligand iz PDB baze podataka.

2.7.2. Metode ligand dokinga

Metode dokinga su usko povezane sa alatima softverskog paketa koji se koristi i u značajnoj meri zavise od njih.

Doking rigidnih konformacija

Algoritmi koji koriste rigidne konformacije su prvi korišćeni za pretragu liganada koji se uklapaju u ciljni protein. Metode rigidne konformacije podrazumevaju da molekuli koji interaguju ne menjaju svoju konformaciju i baziraju se na poklapanju lokalnih karakteristika dva molekula u prostoru, sprečavaju njihovo prepoklapanje i maksimizuju kompatibilnost oblika molekula (Lazar i sar. 2004). Jednostavnost ovih metoda čini ih brzim usled čega su bile prihvaćene u vreme kada računari nisu imali mogućnost obrade većeg broja podataka. Jedan od najstarijih programa koji je i dalje i najširoj upotrebi je program DOCK. Algoritmi traže najveći broj podudaranja trodimenzionalnih konformacija, u smislu da se ligand što bolje postavi u ciljanu proteinsku strukturu. Neki od algoritama se baziraju na traganju za kompatibilnostima u rastojanjima. Prateći površinu proteina algoritmi traže pogodna mesta za aktivno mesto i grade nekoliko sfera unutar njega. Date sfere predstavljaju prostor koji može biti popunjen ligandom

(koji je predstavljen ili sferama ili direktno atomima). U DOCK programu algoritmi traže četiri kompatibilna mesta koja služe za početno postavljanje liganda u protein. Zatim se pozicija dalje optimizuje i izračunava afinitet vezivanja. Knegtel i sar. (1997) su dali model po kojem se doking vrši primenom nekoliko proteina istovremeno. Korišćenjem srednje vrednosti moguće je predstaviti sve strukture istovremeno i time dobiti neku vrstu fleksibilnosti proteina. Za razliku od DOCK-a neki programi koriste hemijske interakcije (vodonične veze) za doking liganda.

Postoje metode koje ugrađuju rigidni ligand deo po deo u aktivno mesto (Friesner i sar. 2004). Takođe primenjuje se i metoda (LUDI model) koja pravi klastere različitih poza oko atoma koji interaguju i traga za trouglima koji se preklapaju tako da grade vezu proteina i liganda (Bohm 1992). Metode rigidnog dokinga su se pokazale neupotrebljive u nekim slučajevima, pa se sa razvojem računara sve više primenjuju metode fleksibilnog dokinga.

Algoritmi fleksibilnog dokinga

Najveći nedostatak rigidnog dokinga je što ne uzima u obzir fleksibilnost liganda. Mali molekuli (čak i oni sa svega dve rotirajuće veze) često imaju više niskoenergetskih konformacija koje treba uzeti u obzir prilikom dokinga.

Prvi metod koji se pojavljuje logični je nastavak rigidnog dokinga pri kojem se koristi više rigidnih struktura jednog liganda i na taj način dobija uvid u doking različitih konformacija (Miller i sar. 1994). Lorber i Shoichet (1998) su razvili nešto drugačiji metod u kojem su koristili oko 300 konformacija po molekulu koje imaju jedan rigidni deo (npr. aromatični prsten) zajednički za sve konformacije. Preklapanjem rigidnog dela dobija se prikaz fleksibilnog liganda. Za doking su koristili DOCK algoritam. Postoje i drugi programi koji koriste preklapanja kao što su često korišćeni GLIDE, FRED i EUDOC (Pang i sar. 2001; Merlitz i sar. 2003; Friesner i sar. 2004).

Drugi tip fleksibilnog dokinga je primenom fragmentacije liganda. Ligand se deli na nekoliko fragmenata koji se zatim rigidnim dokingom (ili sa malim brojem konformacija) pozicioniraju u aktivno mesto. Razlikuju se dva tipa fleksibilnog dokinga fragmentacijom: u prvom se doking fragmenata vrši nezavisno jedan od drugoga, dok se u drugom tipu jedan fragment postavi na predviđeno mesto i na njega nadovezuju ostali fragmenti. Prilikom nezavisnog dokinga fragmenata algoritmi traže atome koji se preklapaju i povezuju ih tako da grade početni molekul (Sandak i sar. 1994). Algoritmi fragmentišu molekul na način da se dobiju

fragmeni srednje i male veličine. Ukoliko su svi fragmenti mali, javiće se poteškoće u njihovom nezavisnom pozicioniranju na odgovarajuće mesto. Drugi tip algoritama počinje pozicioniranjem takozvanog baznog fragmenta u aktivno mesto. Odabira se mali broj potencijalnih baznih fragmenata, i algoritam računa najpogodniji na osnovu broja interagujućih grupa (što više to bolje) i niskoenergetskih konformacija (što manje to bolje). Zatim se dodaju preostali fragmenti i među njima odabiraju oni koji se mogu povezati sa baznim fragmentom. Dodavanje se nastavlja dok se ne rekonstruiše ceo početni molekul. Krajnje pozicije fragmenata se proveravaju i računa interakcija sa proteinom primenom metoda baziranih na energetskom polju. FlexX program koristi MIMUMBA model za određivanje konformacija liganda (Krammer i sar. 2005). Za svaku rotirajuću vezu uzima se set niskoenergetskih torzionih uglova preuzetih iz baza podataka, dok se za računanje afiniteta liganda za protein vrši primenom Bomovom empirijskom funkcijom afiniteta (Bohm 1992). Za dodavanje vode i metalnih jona u aktivno mesto FlexX koristi "koncept čestica" koji je razvijen za tu namenu (Rarey i sar. 1999). Za razliku od FlexX programa koji ligand fragmentiše na male delove (između svake rotirajuće veze), Surfex program pravi mali broj velikih fragmenata (Jain 2003). Tokom rekonstrukcije liganda, dodaju se oni fragmenti čiji se atomi poklapaju sa baznim fragmentom i koji istovremeno interaguju sa proteinom. Usled mogućih varijacija uglova veza, uzimaju se u obzir i visokoenergetske konformacije i daje veći stepen slobode za određivanje interakcija. Iako su fragmentišuće metode jednostavne i efikasne, mora se voditi računa i o vremenu potrebnom za takvu analizu. Metode sa baznim fragmentom su se pokazale kao optimalne uzimajući u obzir vreme trajanja i preciznost rezultata koje daju.

Od sredine poslednje dekade dvadesetog veka u primeni su i genetički algoritmi za fleksibilni doking liganada u aktivno mesto proteina (Gehlhaar i sar. 1995; Jones i sar. 1997). Genetički algoritam je optimizaciona šema koja imitira proces evolucije (Goldberg 1989). U prostoru koji pretražujemo, konfiguracije predstavljaju individue i takozvana funkcija fitnesa određuje koje će konformacije preživeti i dati potomstvo. Prvo je potrebno opisati sve stepene slobode problema koji ispituujemo i predstaviti ih linearnim zapisom, što predstavlja najteži korak u procesu modelovanja. Zatim se definiše funkcija fitnesa koja pomenuti linearni zapis interpretira na način da se može primeniti funkcija afiniteta. Ovaj korak predstavlja analognu radnju prevođenju genotipa u fenotip i zahteva najviše vremena u opisanom procesu. Genetički algoritam koji se najčešće koristi u molekularnom dokingu je implementiran u program GOLD i

razvijen je od strane Jones i sar. (1997). Program radi u dva odvojena procesa. Prvi zapisuje konformaciju liganda i odabranih bočnih lanaca proteina koji učestvuju u vezivanju liganda definišući torzione uglove svih rotirajućih veza. Drugi proces sakuplja podatke o prepoklapanjima atoma koji učestvuju u građenju vodoničnih veza u proteinu i ligandu. Za procenu fitnesa konstruiše se trodimenzionalna struktura iz linearnog opisa stepena slobode. Na osnovu prepoklapanja sakupljenog u drugom procesu, vodonične veze se poklapaju i postavljaju na mesta predviđenim za njih u aktivnom mestu. Na kraju funkcija afiniteta računa jačinu vodoničnih veza i unutrašnju energiju molekula. Kao funkcija fitnesa koristi se Van der Vals-ova energija nastala interakcijom proteina i liganda.

Geometrija rastojanja je još jedna dobro poznata tehnika koja se koristi za određivanje strukture primenom Nuklearne Magnetne Rezonance (NMR) (Havel i sar. 1983; Havel i sar. 1983). Tehnika ne koristi Euklidov koordinatni sistem za opisivanje molekula, već primenjuje matrice rastojanja koje sadrže sve udaljenosti između atoma sa određenim odstupanjima. Računajući rastojanja za svaki par atoma detaljno se može opisati set željenih konformacija. Mana ovog postupka je što se usled ograničavajućeg faktora intervala rastojanja atomskog para, mali broj matrica može prevesti u trodimenzionalnu strukturu.

Većina navedenih programa u svojim novijim verzijama imaju dostupnu i mogućnost primene fleksibilnih H atoma u receptoru. To zahteva intenzivnije računanje, ali daje preciznije rezultate. Sledeći stepen složenosti je fleksibilnost receptora. Metode koje dopuštaju fleksibilnost receptora uglavnom primenjuju simulacije za doking liganada.

Doking primenom simulacija

Za razliku od do sada prikazanih kombinatoričkih algoritama optimizacije, postoje i metode koje koriste tehnike simulacije za rešavanje postavljenog problema dokinga. Umesto nabiranja računskih niskoenergetskih stanja, ove metode počinju računanje od startne pozicije i lokalno se pomeraju ka konfiguracijama u niskoenergetskom stanju.

Najjednostavniji oblik simulacije je simulacija takozvanim anilingom prilikom kojeg se kompleks optimizuje bez fizičke interpretacije simulacije (Kirkpatrick i sar. 1983). Proces kreće od početne konfiguracije A koja ima određenu energiju $E(A)$. Nasumičnim lokalnim promenama konformacije molekula dobijamo novu konformaciju B koja ima energiju $E(B)$. Ukoliko je $E(B) \leq E(A)$ (Metropolis kriterijum) nova konformacija se prihvata i od nje proces kreće ponovo.

Program AUTODOCK koje je jedan od najprimenjenijih za doking malih molekula koristi ovakav tip algoritma (Goodsell i Olson 1990; Goodsell i sar. 1996). AUTODOCK dozvoljava odabir bočnih lanaca amino kiselina kojima se dopušta fleksibilnost prilikom računanja interakcija sa ligandom.

Molekularna dinamika (MD) je prvi metod kojim se primenom računara predvidela interakcija proteina i liganda i koji se i danas primenjuje za doking liganada (Platzer i sar. 1972). Metoda koristi energetska polja za izračunavanje energija svakog atoma u sistemu koji se simulira. Zatim se računaju njihove brzine i ubrzanja i atomi pomeraju u skladu sa izračunatim vrednostima i proteklom vremenom. Najveći nedostatak MD predstavlja izuzetno dugačko vreme potrebno za analizu, posebno velikih molekula. Da bi se skratilo vreme potrebno za simulaciju razvijene su metode koje preskaču lokalne minimume povećanjem temperature za određene stepene slobode (Di Nola i sar. 1994). Sličan efekat se dobija skupljanjem i širenjem liganda u aktivnom mestu proteina.

Simulacije molekularne dinamike se baziraju na drugom Njutnovom zakonu prema formuli:

$$F_i = m_i \cdot a = m_i \cdot \frac{dx_i}{dt} = m_i \cdot \frac{d^2x}{dt^2}$$

Zakon opisuje kretanje čestice koja se silom F_i kreće duž koordinate x i ima masu m_i . Navedena formula se koristi za računanje određenog broja atoma i molekula pod uticajem energetske polja (opisano u poglavlju 2.2) koje opisuje interakcije unutar sistema sa potencijalnom energijom $V(x)$, gde x odgovara koordinatama svih atoma u sistemu. Odnos između potencijalne energije i drugog Njutnovog zakona je:

$$F(\vec{x}_i) = -\nabla_i V(\vec{x})$$

gde je $F(x)$ sila koja deluje na česticu usled potencijalne energije $V(x)$. Kombinovanjem jednačina dobija se jednačina koja govori da se potencijalna energija sistema menja sa promenom položaja atoma tokom vremena:

$$\frac{dV(\vec{x})}{dx_i} = -m_i \cdot \frac{d^2x_i}{dt^2}$$

Obzirom da je ubrzanje $a = -dV/(m \cdot dx)$ možemo uzračunati promenu u sistemu tokom vremena poznavajući potencijalnu energiju, početne koordinate i početne brzine atoma. Stoga, ovaj sistem predstavlja deterministički sistem, što znači da možemo predvideti njegovo stanje u bilo kom prošlom ili budućem vremenu. Usled zavisnosti potencijalne energije od više kompleksnih faktora koje možemo sve uzeti u obzir, njeno računanje predstavlja numeričko predviđanje vrednosti uz izvesni stepen poverljivosti. Armen i sar. (2009) su ispitali uticaj različitih stepena slobode fleksibilnosti receptora (fleksibilnost petlji, segmenata kostura proteina ili kompletnog proteina) primenom MD. Rezultat njihovog istraživanja bio je da se fleksibilnošću bočnih lanaca i delova kostura povećava preciznost dokinga, dok fleksibilnost celog proteina vodi ka smanjenju pouzdanosti rezultata.

Za razliku od MD gde se lokalna pomeranja atoma vrše na osnovu sila, u Monte Karlo simulaciji (MC) lokalna pomeranja atoma se dešavaju po principu slučajnosti. Za MC simulaciju stepeni slobode predstavljaju jedan od najvažnijih parametara i cilj je da se razviju metode koje izbegavaju prihvatanje visokoenergetskih stanja. Metode se zasnivaju na promeni dužina i uglova veza, kao i torzionih uglova. Obzirom da se energije molekula različito menjaju promenom navedenih promenljivih, za svaku promenjivu pravi se posebna energetska skala. Primenom internih koordinata i definisanjem promena za svaku skalu energija promenljivih, dolazi se do niskoenergetskih stanja molekula i željene konformacije.

Pored navedenih, postoje i hibridne metode dokinga simulacijom, čiji je cilj da prevaziđu nedostatke pojedinačnih metoda.

Postoje tri osnovna tipa simulacija (ne samo vezanih za doking, već uopšteno za molekularno modelovanje) u zavisnosti od nivoa na kojem računaju parametre:

- Metode molekularne mehanike koje ne uključuju računanja kretanja elektrona (npr. program NAMD – Molecular Dynamics Simulator)
- Metode kvantne mehanike koje uzimaju u obzir trodimenzionalni raspored elektrona oko atomskog jezgra (ne koriste se za simulacije celih proteina, već samo delova; npr. program Jaguar koji računa parcijalnu pK vrednost u molekulu)

- Poluempirijske metode koriste metode kvantne mehanike, ali sa velikim brojem predviđanja kao što se dešava u metodama molekularne mehanike (npr. program MOPAC – Molecular Orbital PACkage). Naelektrisanja molekula se uglavnom računaju poluempirijskim metodama. Takođe, postoje i poluempirijske metode koje kvantnom mehanikom računaju samo deo koji interaguje, dok ostale delove sistema računaju po metodama molekularne mehanike.

2.7.3. De novo sinteza liganda

Primenom savremenih tehnika i programskih paketa, moguće je konstruisati ligand na osnovu parametara aktivnog mesta. Razlika u odnosu na doking pretraživanjem baza je u tome što su u ovom sučaju baze podataka sačinjene od velikih do malih delova, pa čak i pojedinačnih atoma. Svaki deo u bazi ima jedno ili nekoliko vezivnih mesta koja su uglavnom obeležena nedostatkom atoma na mestu vezivanja. Povezivanjem fragmenata koja mogu graditi vezu sa proteinom konstruišu se molekuli tako da odgovaraju sferi aktivnog mesta i predstavljaju potencijalne ligande. Na ovaj način se mogu predvideti i ligandi koji ne postoje u prirodi, što daje veliki značaj ovom tipu dokinga. Algoritmi koji se koriste za *de novo* konstrukciju liganda usko su povezani sa algoritmina dokinga i razlikuju se samo u sitnim detaljima.

2.7.4. Validacija doking metoda

Svaka nova razvijena metoda, algoritam ili program za doking mora proći validaciju primenom eksperimentalnih podataka. Najčešće se koriste podaci kristalografije rendgenskim zracima i odabere se test uzorak. Nakon računanja dokinga test uzorka, rezultati se porede sa eksperimentalnim i potvrđuje ili opovrgava preciznost postupka. Da bi se doking algoritam smatrao pouzdanim, mora da rekonstruiše strukturu dobijenu kristlografijom rendgenskim zracima. Najčešće se odabere nekoliko desetina uzoraka za test i većina doking programa ima uspešnost kod oko 70% struktura, sa odstupanjem od 2Å RMSD.

U nekoliko radova dat je pregled najčešće korišćenih programa za doking (Yuriev i sar. 2011; Sousa i sar. 2013). U zavisnosti od predmeta istraživanja (molekuli koji interaguju i metode koje se koriste) variraju i rezultati ispitivanja. Neki programi su se pokazali kao dobri za određena istraživanja, dok za neka druga ne daju toliko precizne rezultate. Ipak, kao najpouzdaniji najčešće se navode Autodock (kao najzastupljeniji) i SurflexDock.

2.7.5. Izazovi u ligand dokingu

Problem tautomera, stereoizomera i protonacije liganda predstavlja veliki izazov u rešavanju dokinga. U zavisnosti od razlika u navedenim stanjima, ligand se može vezati ili ne vezati za receptor (Davis i sar. 2009). Kao rešenje autori su predložili uzimanje u obzir svih tautomera i stanja protonacije, ali se kao problem pojavljuje vreme potrebno za doking svih oblika liganda. Iako postoji problem sa dužinom trajanja analiza, tautomere, stereoizomere i stanje protonacije bi trebalo uzimati u obzir dok se ne razviju metode koje će biti u stanju da reše navedeni problem (ten Brink i Exner 2009). Razvijeni su softveri koji generišu tautomere željenog jedinjenja koji se mogu koristiti za doking (Kochev i sar. 2013). Mnogi softveri koji se primenjuju za doking dozvoljavaju ručno podešavanje stanja protonacije.

U prirodi se interakcija liganda i receptora dešava u rastvoru. Usled toga neophodno je uzeti u obzir i uticaj rastvarača na doking molekula. Mnoga ispitivanja su pokazala da dodavanje rastvarača značajno utiče na rezultat dokinga (Cincilla i sar. 2009; Fong i sar. 2009; Lang i sar. 2009). Usled toga, u mnoge programe (kao što je npr. AutoDock) dodat je uticaj rastvarača u doking proceduru. Takođe, u obzir se uzimaju i razlike između polarnih i nepolarnih heteroatoma u zavisnosti od hibridizacije i veza koje grade. Metode MD standardno uzimaju u obzir uticaj rastvarača na vezivanje liganda.

2.8. Tehnologija fermentisanih mlečnih proizvoda

2.8.1. Starter kulture u fermentaciji mleka

Starter kulture u tehnologiji fermentisanih mlečnih proizvoda su kombinacije različitih vrsta i sojeva mikroorganizama, uglavnom bakterija mlečne kiseline. Prvi preduslov koji treba da ispuni potencijalna starter kultura jeste da sadrži mikroorganizme koji razgrađuju laktozu (poseduju enzim beta-galaktozidazu), što predstavlja osnovu za proizvodnju fermentisanih mlečnih napitaka. Starter kulture su prirodni izolati selekcionisani na osnovu svojih fizioloških karakteristika tj. sposobnosti da stvaraju određene količine poželjnih komponenti (laktata, acetata, diacetila, acetaldehida, etanola, ugljendioksida). Takođe, selekcija se vrši na osnovu rezistencije na bakteriofage, sposobnosti da rastu u međusobnim kombinacijama, a u poslednjih nekoliko godina naglasak je dat i njihovim probiotskim svojstvima (Čurić 1997).

Bakterije mlečne kiseline koje se koriste kao starter kulture generalno se dele na:

- Mezofilne – koriste se za proizvodnju maslaca, kisele pavlake, svežih sireva, sirnih namaza
- Termofilne – koriste se za proizvodnju fermentisanih napitaka tipa jogurta.

Veliki napredak u tehnologiji fermentisanih mlečnih proizvoda postignut je razvojem nove vrste proizvoda za specijalne namene, inkorporiranjem probiotskih mikroorganizama u starter kulture ili u već fermentisano mleko (Milanović 1997).

Probiotici su vijabilni mikroorganizmi izolovani iz intestinalnog trakta čoveka, koji povoljno utiču na rad creva i celokupno ljudsko zdravlje preko metaboličkih puteva varenja hrane, sinteze enzima, smanjenja nivoa holesterola, smanjenja rizika od pojave raka, kao i stimulacije imunog sistema (McDonough i sar. 1987; Fuller 1989; Holzapfel i sar. 1998; Sanders 1998; Klaenhammer i Kullen 1999; Holzapfel i sar. 2001). Mlečni proizvodi sa probiotskim svojstvima su najčešće proizvedeni korišćenjem kombinacije nekoliko sojeva kod kojih su dominantno zastupljeni probiotici, kao što su *Bifidobacterium* spp., *Lactobacillus acidophilus*, itd. Neka istraživanja ukazuju na terapijski efekat oralno konzumiranih bifidobakterija, čije pozitivno dejstvo se ogleda u antitumorskoj aktivnosti, imunološkoj stimulaciji regulisanju balansa crevne flore, kao i smanjenju nivoa holesterola u krvi (Lee i sar. 1993; Reddy i Rivenson 1993; S.A. i Bezkorovainy 1993; Hamilton-Miller 1999). Takođe, razvoj stabilne crevne flore je neophodan za razvoj crevnog trakta kod novorođenčadi (Stappenbeck i sar. 2002; Rakoff-Nahoum i sar. 2004). Ubrzo nakon rođenja ljudski gastrointestinalni trakt nastanjuju mikroorganizmi koji do treće godine života uspostavljaju mikrofloru koja ostaje i kod odraslog čoveka (Morgan i sar. 2013). U crevnom traktu ima oko 10^{14} mikroorganizama (preko 500 vrsta), dok ljudsko telo ima 10 puta više bakterijskih nego eukariotskih ćelija. Funkcionalna genomika je pokazala da ljudski geni koji kodiraju metaboličke funkcije predstavljaju samo 20 posto ukupnih gena, dok ostale funkcije kodiraju geni mikroorganizama (Candela i sar. 2010). Usled navedenih činjenica neophodno je vršiti ispitivanja funkcionalne genomike i metaboličkih puteva bakterija u ljudskom organizmu.

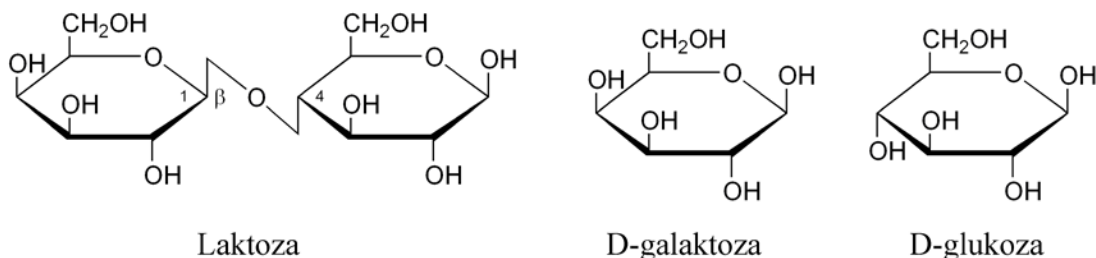
Aktivnost probiotika zavisi od pH sredine, sastava supstrata i temperature. U industriji prerade mleka probiotici imaju prednost, s obzirom da zadržavaju visoku vijabilnost ćelija u procesu fermentacije kao i tokom skladištenja gotovih proizvoda (Milanović 1997).

Novorazvijeni fermentisani mlečni napici pokazuju brojna profilaktična i/ili terapijska svojstva, zahvaljujući multifunkcionalnim karakteristikama intestinalnih bakterija (antimikrobna, biohemijska, imunološka, fiziološka) (Tamime i Robinson 1999; 2007; Tamime 2008).

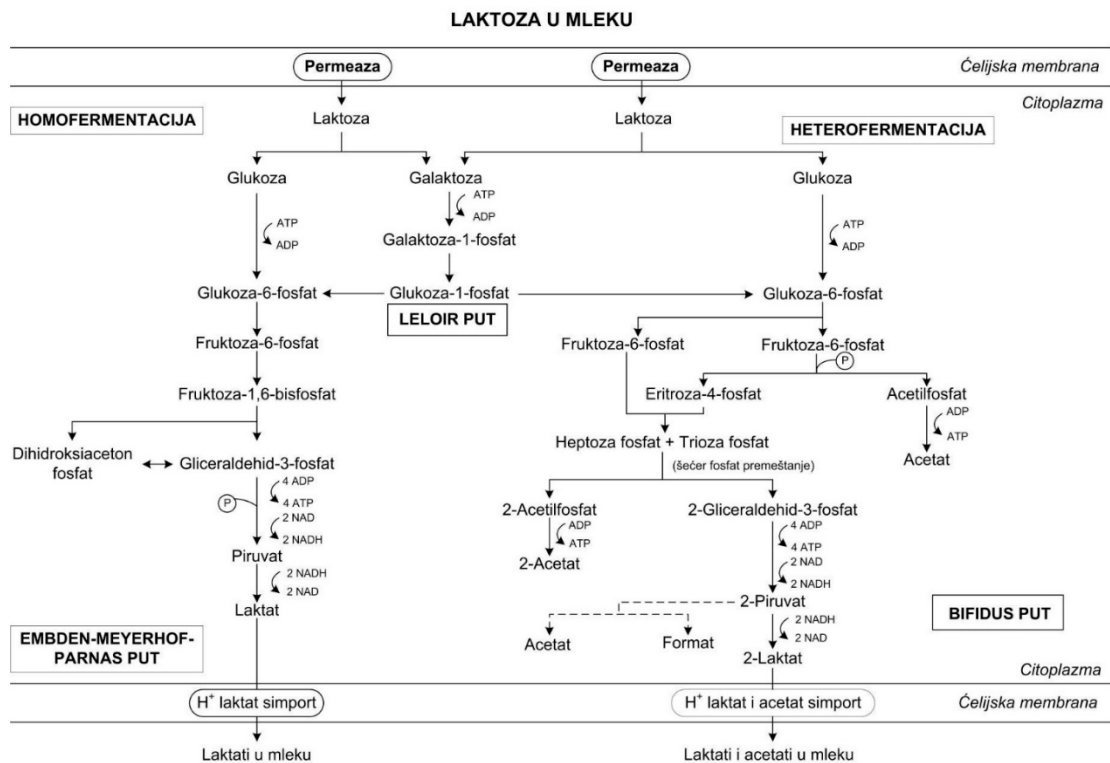
2.8.2. Razgradnja laktoze starter kulturom

Laktoza je disaharid sastavljen od jednog molekula glukoze i jednog molekula galaktoze (slika 24). Prva i najvažnija reakcija u proizvodnji fermentisanih mlečnih proizvoda je razgradnja molekula laktoze. Laktoza se razgrađuje intracelularno od strane mikroorganizama prisutnih u starter kulturi. Postoje dva puta ulaska laktoza u ćeliju koja se javljaju kod različitih mikroorganizama:

1. putem laktoza fosfotransferaznog sistema pri kojem se vrši fosforilacija pomoću fosfoenol piruvata i fosfotransferazni sistem. Navedeni složeni sistem translocira laktozu iz spoljašnjeg u unutrašnji deo citoplazmine membrane i u ćeliji se formira laktoza-fosfat. Laktoza-6-fosfat se hidrolizuje pomoću β -fosfogalaktozidaze u monosaharide ili mnogo češći slučaj;
2. Putem proton gradijent permeaznog sistema koji prenosi laktozu kroz bakterijsku ćelijsku membranu. Za ovu distribuciju od značaja je permeazni sistem pri kojem laktozu nakon ulaska u ćeliju razgrađuje enzim beta-galaktozidaza. Beta-galaktozidaza razgrađuje laktozu na jedan molekul glukoze i jedan molekul galaktoze. Dalji put razgradnje ovih heksoza može biti putem homofermentacije ili heterofermentacije i prikazan je na slici 25 (Tamime i Robinson 2007).



Slika 24. Molekuli laktoze, D-galaktoze i D-glukoze



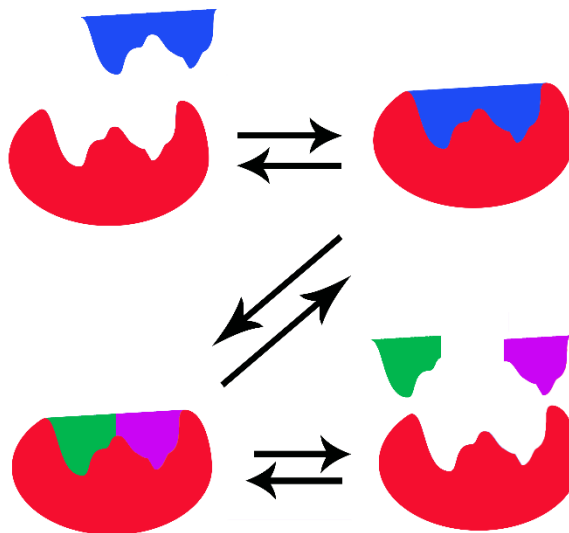
Slika 25. Homofermentacija i heterofermentacija laktoze u jogurtu primenom jogurtne i probiotske starter kulture nakon translokacije permeazom

Homolitička fermentacija mleka korišćenjem *S. thermophilus* i *L. acidophilus* odvija se Embden-Meyerhof-Parnas-ovim putem (EMP put) katabolizma glukoze. Citoplazmatični proteini mikroorganizama transportuju laktozu bez hemijske modifikacije putem permeaze. Nefosforilovani disaharid se hidrolizuje putem β -galaktozidaze na nefosforilovanu glukozu i galaktozu. Glukoza se kataboličkim putem transformiše do piruvata, a galaktoza izlazi iz ćelije. Kada se glukoza potpuno istroši od strane 3 navedena mikroorganizma, galaktoza se razlaže Leloir putem.

Probiotske bakterije *Bifidobacterium* spp. fermentišu laktozu i glukozu heterofermentacijom. Katabolizmom glukoze produkuje se CO_2 usled odvijanja procesa dekarboksilacije. Laktoza se prvo transportuje u ćeliju pomoću permeaze i hidrolizuje se na glukozu i galaktozu. Navedene heksoze se katabolizuju preko fruktoza-1,6-bifosfata, koji se razgrađuje pomoću fruktoza-bifosfat aldolaze. Krajnji produkti fermentacije su laktati i acetati (fermentacijom 2 mola glukoze nastaju 3 mola acetata i 2 mola laktata).

2.9. Beta-galaktozidaza

Laktoza se u živim organizmima razgrađuje enzimskim putem. Enzimi predstavljaju katalizatore hemijskih reakcija koji skraćuju njihovo vreme sniženjem energije aktivacije. Enzimi imaju veoma visoku specifičnost za supstrat i deluju po principu ključa i brave (slika 26) (Koolman i Röhm 2005). Enzim obezbeđuje optimalnu orijentaciju molekula koji reaguju u njegovom aktivnom mestu, čime se povećava verovatnoća interakcije u odnosu na slobodne molekule u rastvoru.



Slika 26. Princip ključa i brave prilikom enzimске reakcije – supstrat se specifično veže za enzim, nakon čega dolazi do njegove razgradnje i otpuštanja produkata reakcije u medijum

Enzim beta-galaktozidaza (EC 3.2.1.23) katališe hidrolizu $\beta(1-3)$ i $\beta(1-4)$ glikozidne veze i transglikozilaciju beta-D-glikopiranozida (kao što je galaktoza). Njena primena je široka u industriji mleka jer pored vodeće uloge u fermentaciji mleka olakšava njegovo varenje razgradnjom laktoze koja izaziva probleme kod ljudi netolerantnih na laktozu. Takođe, beta-galaktozidaza se može koristiti i za rešavanje problema kao što je kristalizacija laktoze (Juajun i sar. 2011). Produkti razgradnje laktoze, glukoza i galaktoza, mogu se primenjivati kao zaslađivači ili obogaćivači nutritivno siromašnih proizvoda (Gänzle i sar. 2008). U poslednje vreme akcenat se daje na iskorišćenost produkata industrijskih procesa, čime se povećava

produktivnost procesa, njegova isplativost, kao i ekološka efikasnost. Prilikom procesa proizvodnje sira nastaju velike količine surutke koje predstavljaju veliki ekološki problem. Procenjuje se da se u Svetu godišnje proizvede oko 160 miliona tona surutke (Guimarães i sar. 2010). Surutka je bogata laktozom i primena beta-galaktozidaze za njenu razgradnju čini je primenjivom u daljim procesima prehrambene industrije. Osobina beta-galaktozidaze da vrši transglikozilaciju upotrebljava se i prilikom proizvodnje galaktooligosaharida (GOS). GOS imaju široku primenu u prehrambenoj industriji zahvaljujući njihovim prebiotičkim osobinama (Pivarnik i sar. 1995; Chen i sar. 2008).

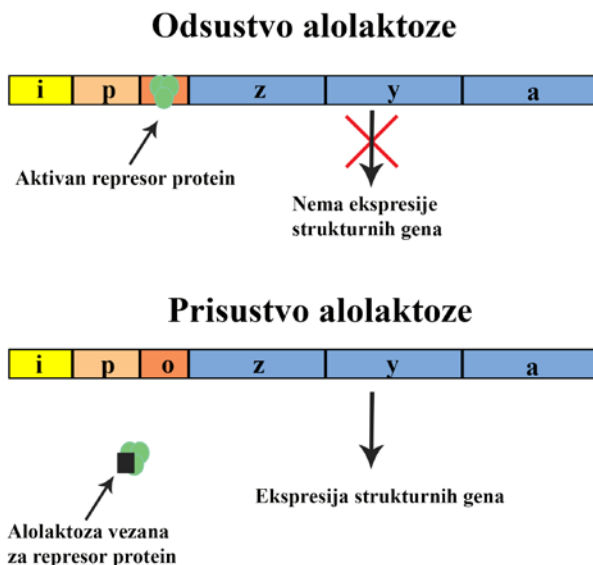
2.9.1. Klasifikacija beta-galaktozidaza

Sve poznate beta-galaktozidaze pripadaju GH-A superfamiliji glikozidhidrolaza. Dalje se dele u GH-2, GH-35, GH-42 i GH-59 subfamilije (Cantarel i sar. 2009). Ova klasifikacija je napravljena prema sličnosti sekvenci koje kodiraju enzim. Beta-galaktozidaze eukariota su klasifikovane u subfamiliju 35 sa izuzecima beta-galaktozidaza *Klyveromices lactis*, *K. marxianus* (99% identične) i *Aspergillus nidulans* koje pripadaju subfamiliji GH-2 zajedno sa prokariotskim beta-galaktozidazama. Beta-galaktozidaza *Streptococcus thermophilus*-a pripada subfamiliji GH-2, dok beta-galaktozidaze *Lactobacillus acidophilus*-a i *Bifidobacterium animalis* ssp. *lactis*-a pripadaju GH-42 subfamiliji.

2.9.2. Ekspresija gena beta-galaktozidaze

Ekspresija beta-galaktozidaze je najbolje istražena kod *E. coli* (Müller-Hill 1996). Ustanovljeno je da prisustvo laktoze ili nekih drugih galaktozida (npr. orto-nitrofenil- β -D-galaktozid), indukuje ekspresiju beta-galaktozidaze koja hidrolizuje laktozu na glukozu i galaktozu. Međutim, ustanovljeno je da stimulatori ekspresije nisu samo supstrati, niti su supstrati i njihovi analozi uvek stimulatori. Studije primenom radioaktivnog obeležavanja su pokazale da se beta-galaktozidaza sintetiše *de novo* samo u prisustvu stimulatora ekspresije. Detaljnija ispitivanja pokazala su da je produkcija beta-galaktozidaze kontrolisana negativnom regulacijom, što je rezultat prisustva operona u njenoj ekspresiji (Pardee i sar. 1959; Jacob i Monod 1961). Operon se sastoji iz dva dela: strukturni geni koji kodiraju enzime (LacZ – beta-galaktozidazu, LacY – beta-galaktozid permeazu i LacA – beta-galaktozid transacetilazu), i regulatorni geni (operator gen i regulator gen). Operator i regulator su regulatorni geni operona

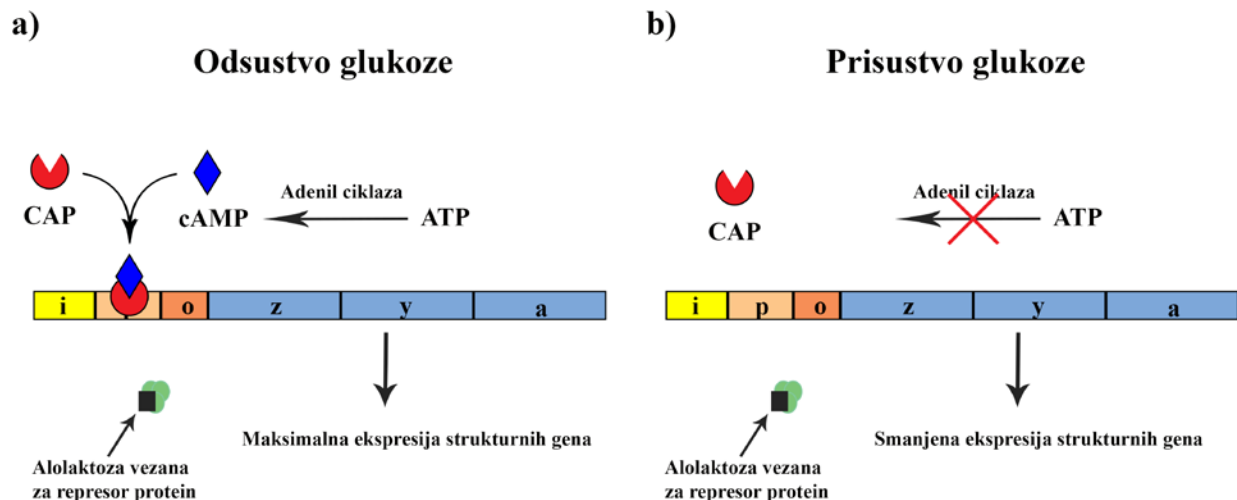
koji kontrolišu ekspresiju strukturalnih gena, s tim što regulatorni gen takođe kodira protein. Mehanizam regulacije Lac operona prikazan je na slici 27.



Slika 27. Ekspresija Lac operona; i – regulatorni gen (ekspresuje represor protein), p –promotor, o – operator, z, y, a – strukturalni geni

Represor protein se konstantno ekspresuje i vezuje za operator gen. Time se sprečava vezivanje RNK polimeraze i ekspresija strukturalnih gena. U prisustvu alolaktoze, koja nastaje delovanjem beta-galaktozidaze na laktozu, represor menja svoju konformaciju i gubi sposobnost vezivanja za operator. Dakle, beta-galaktozidaza proizvodi sopstveni stimulator ekspresije. Ovaj naizgled paradoks je rešen otkrivanjem mogućnosti da se represor odvoji od operatora, čak i u odsustvu stimulatora ekspresije (Berg i sar. 1981; Wong i sar. 1997). Dakle, u ćeliji je uvek prisutna izvesna količina permeaze i beta-galaktozidaze potrebnih za transport i razgradnju laktoze (po oko 10 molekula proteina). Međutim, nije dovoljno samo da laktoza bude prisutna da bi došlo do maksimalne ekspresije gena, već je potrebno i da glukoza ne bude prisutna u ćeliji. Niame, pored opisane negativne, postoji i pozitivna kontrola Lac operona. Da bi došlo do njegove maksimalne ekspresije potrebno je da se za promotor veže CAP protein (Catabolite Activator Protein), kome je za tu vezu neophodan cAMP molekul (cyclic Adenosine

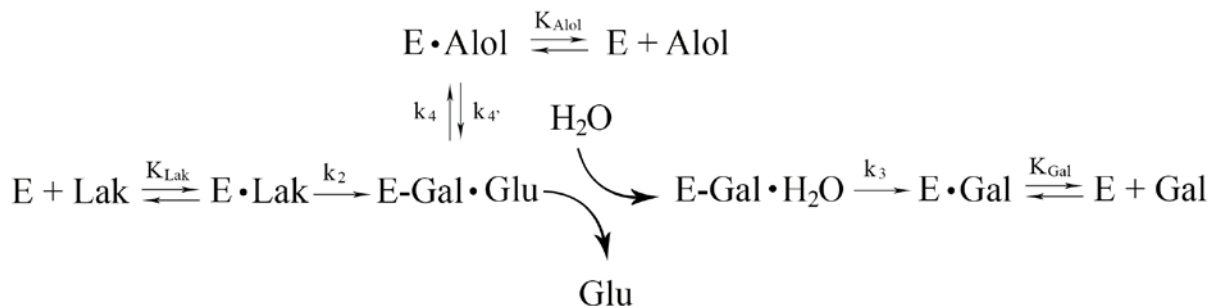
MonoPhosphate). Prilikom prisustva glukoze u ćeliji koncentracija cAMP-a je niska, čime se onemogućava vezivanje CAP proteina za promotor, a time i ne postiže maksimalna ekspresija strukturalnih gena operona. Prilikom odsustva glukoze, cAMP se sintetizira kao signalni molekul, CAP se veže i u prisustvu laktoze maksimalno ekspresuju proteini potrebni za njenu razgradnju (slika 28).



Slika 28. Ekspresija Lac operona u prisustvu i odsustvu glukoze; *i* – regulatorni gen (ekspresuje represor protein), *p* – promotor, *o* – operator, *z*, *y*, *a* – strukturalni geni

2.9.3. Razgradnja laktoze

Iako nije jedina reakcija koju katalizira, razgradnja laktoze predstavlja dominantnu i sa tehnološke strane najinteresantniju osobinu beta-galaktozidaze. Kao što se vidi na slici 29, laktoza se hidrolizuje na glukozu i galaktozu. Zatim postoje dva puta, glukozu disosuje iz kompleksa (označeno kao ireverzibilna reakcija jer se pokazalo da se u alolaktozu ne mogu ugraditi slobodni molekuli glukoze, već samo neposredno pre toga nastali razgradnjom laktoze) ili nastaje alolaktoza koja se razgrađuje na glukozu i galaktozu (Jobe i Bourgeois 1972).



Slika 29. Razgradnja laktoze; E + X označava odvojene molekule u rastvoru, E•X označava nekovalentne veze, dok E-X označava kovalentno vezivanje molekula

Kinetika razgradnje supstrata enzimskim putem prati se kroz nekoliko parametara. Michaelis-Menten model enzimске razgradnje opisuje njenu kinetiku kroz definisanje parametra K_m , koji predstavlja koncentraciju supstrata pri kojoj enzim postiže polovinu svoje maksimalne brzine (Murray i sar. 2006). Michaelis-Menten jednačina matematičkim putem predstavlja odnos između početne brzine reakcije i koncentracije supstrata:

$$v_i = \frac{V_{max}[S]}{K_m + [S]}$$

v_i – početna brzina reakcije

V_{max} – maksimalna brzina reakcije

K_m – Michaelis-Mentenova konstanta

[S] – koncentracija supstrata

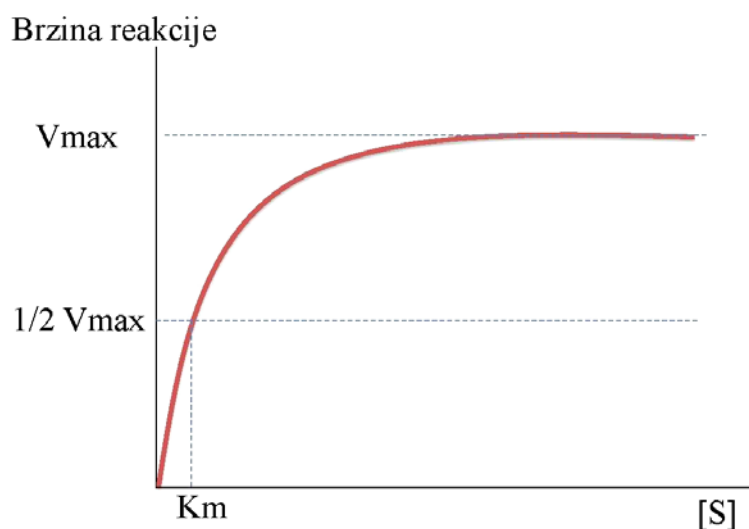
Parametar K_m je karakterističan za svaku reakciju i predstavlja koncentraciju supstrata pri kojoj enzim postiže polovinu maksimalne brzine razgradnje ($v = V_{max} / 2$) (slika 30). Maksimalna brzina reakcije se postiže kada je sav enzim vezan u enzim-supstrat kompleksu, odnosno kada dođe do zasićenja enzima. Što je vrednost K_m manja to je afinitet enzima za supstrat veći, i obrnuto, što je vrednost veća, afinitet je manji. Ipak, iz ovoga ne sledi da će reakcija sa manjim K_m biti efikasnija od reakcije sa većim K_m , jer reakcija sa većim K_m može imati i veću V_{max} . Usled toga, sledeći parameter koji se koristi za karakterizaciju enzimске

reakcije je k_{cat} , koji predstavlja maksimalan broj molekula supstrata koje jedno aktivno mesto enzima prevodi u produkt u jedinici vremena:

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

V_{max} – maksimalna brzina reakcije

$[E]$ – koncentracija enzima

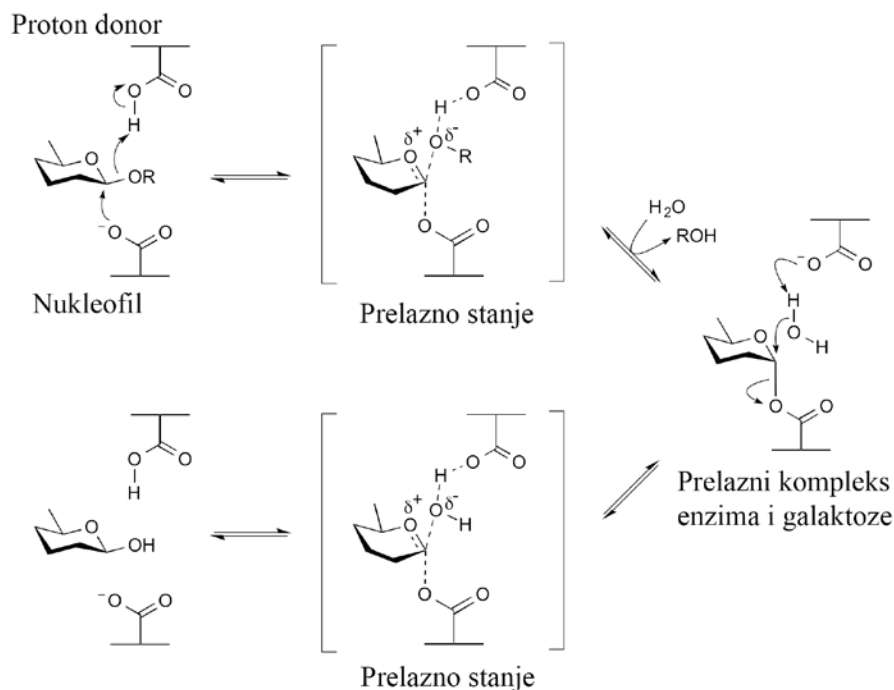


Slika 30. Michaelis-Menten model enzimske razgradnje

[S] – koncentracija supstrata

Prilikom katalize supstrata ključni korak predstavlja njegova interakcija sa enzimom. Sve do sada ispitane beta-galaktozidaze imaju (α/β)8 barrel (takozvani TIM barel) katalitički domen, sa dve aktivne glutaminske kiseline (Davies i Henrissat 1995). Beta-galatozidaze pripadaju neinvertujućim glikozidazama jer razgrađuju laktozu na takav način da nakon razgradnje ne dolazi do promene anomerne konfiguracije. Proces razgradnje laktoze se sastoji iz dva koraka (slika 31). Reakcija se dešava uz prisustvo jedne kiselo/bazne (proton donor) i jedne nukleofilne amino kiseline bočnog lanaca enzima. U beta-galaktozidazama to su dva molekula glutaminske kiseline udaljena približno 5,5Å (McCarter i Stephen Withers 1994; Wang i sar. 1994). Prvi korak

je glikozilacija u kojoj jedna rezidua ima ulogu nukleofila napadajući anomerni centar, pri čemu se oslobađa glukoza i gradi prelazni kompleks enzima i galaktoze. U isto vreme druga rezidua protonuje slobodni kiseonik koji nastaje prilikom raskidanja veze. Drugi korak se naziva deglikozilacija. Tom prilikom kompleks enzima i galaktozidaze se hidrolizuje vodom posredstvom proton donor rezidue koja deprotonizuje vodu i raskida kompleks oslobađajući galaktozu.



Slika 31. Neinvertujući put razgradnje laktoze

Brzina neenzimske razgradnje laktoze i najčešćih supstituenata orto-nitrofenil- β -D-galaktopiranozida (ONPG) i para-nitrofenil- β -D-galaktopiranozida (PNPG) na sobnoj temperaturi iznosi $3,6 \times 10^{-11}$ i $2,2 \times 10^{-9}$, $3,2 \times 10^{-9} \times \text{s}^{-1}$. Međutim, prilikom enzimske razgradnje beta-galaktozidazom iz *E. coli* (u optimalnim uslovima) vrednosti iznose 60, 1200 i $90 \times \text{s}^{-1}$, što predstavlja povećanje brzine razgradnje od 6×10^{11} , 5×10^{11} i 3×10^{10} puta. Biohemijske osobine beta-galaktozidaza iz različitih sojeva su detaljno ispitane u literaturi. Deo tih istraživanja dat je u tabeli 4.

Tabela 4. Biohemijske osobine beta-galaktozidaza

Mikroorganizam	Temp (°C)	pH	Km (mM)	Vmax (U/mg)	kcat (s ⁻¹)	Potrebni metalni joni	Referenca
<i>S.thermophilus</i> LMD9	25-40	6,5-7,5	1,1	464	234	Mn ²⁺ , Mg ²⁺	(Rhim i sar. 2010)
<i>Lactobacillus bulgaricus</i>	35-50	5,0-5,5	0,98	72,5	45,7	Mn ²⁺ , Co ²⁺	(Rhim i sar. 2009)
<i>Pseudoalteromonas haloplanktis</i>	50	-	2,4	-	33	-	(Hoyoux i sar. 2001)
<i>Lactobacillus acidophilus</i>	55	6,5-8,0	4,0	28,8	50,4	Mg ²⁺	(Nguyen i sar. 2007)
<i>A.psychrolactophilus</i> F2	10	8,0	42,1	-	-	-	(Nakagawa i sar. 2006)
<i>Bifidobacterium infantis</i>	60	7,5	2,6	262	-	Na ⁺ , K ⁺	(Hung i Lee 2002)
<i>Lactobacillus reuteri</i>	50	7,0	-	-	-	-	(Nguyen i sar. 2007)
<i>Thermus</i> sp. A4	70	6,5	19	-	12,5	Zn ²⁺ , Mn ²⁺ , Co ²⁺	(Ohtsu i sar. 1998)
<i>Bacillus licheniformis</i>	45	5,7	-	-	-	-	(Tran i sar. 1998)
<i>Klyveromyces lactis</i>	40	7,0	20	570	-	Mn ²⁺	(Kim i sar. 2003)
<i>Thermus aquaticus</i> YT1	80	5,5	6,6	-	5,7	Ca ²⁺	(Berger i sar. 1997)
<i>Flavobacterium</i> sp.	42	7,5	0,65	62	-	-	(Sorensen i sar. 2006)
<i>E. coli</i>	35	6,6-7,2	5,5	-	60	Na ⁺ , Mg ²⁺	(Wallenfels i Weil 1972)
<i>Thermotoga maritima</i>	80	6,5	1,6	63,3	-	Mn ²⁺	(Kim i sar. 2004)
<i>Pseudoalteromonas</i> sp. 22b	40	6,0-8,0	3.3	-	157	Na ⁺ , K ⁺ , Mg ²⁺ , Mn ²⁺	(Cieslinski i sar. 2005)
<i>Bacillus circulans</i>	-	-	41,7	-	-	-	(Nakanishi i sar. 1983)

Međutim, sve navedene osobine beta-galaktozidaza treba uzeti sa velikom dozom opreza, jer na njihovu aktivnost utiču brojni fizički faktori kao što su pH vrednost sredine, temperatura i pritisak (Kuby i Lardy 1953; Tenu i sar. 1971; Degraeve i Lemay 1997; Rhimi i sar. 2010). Usled toga se u literaturi i bazama podataka mogu za isti soj naći značakno različiti rezultati. Pored navedenih faktora važnu ulogu ima i prisustvo metalnih jona u aktivnom mestu koji učestvuju u pozicioniranju proton donora prilikom reakcije katalize (Jacobson i sar. 1994; Maksimainen i sar.

2012). Kod *E. coli* to su Mg^{2+} i Na^+ joni, dok je kod *Bacillus circulans* to Zn^+ jon. Iako su metalni joni potrebni za maksimalnu aktivnost enzima, nisu neophodni za njegovo delovanje.

2.9.4. Struktura beta-galaktozidaze *E.coli*

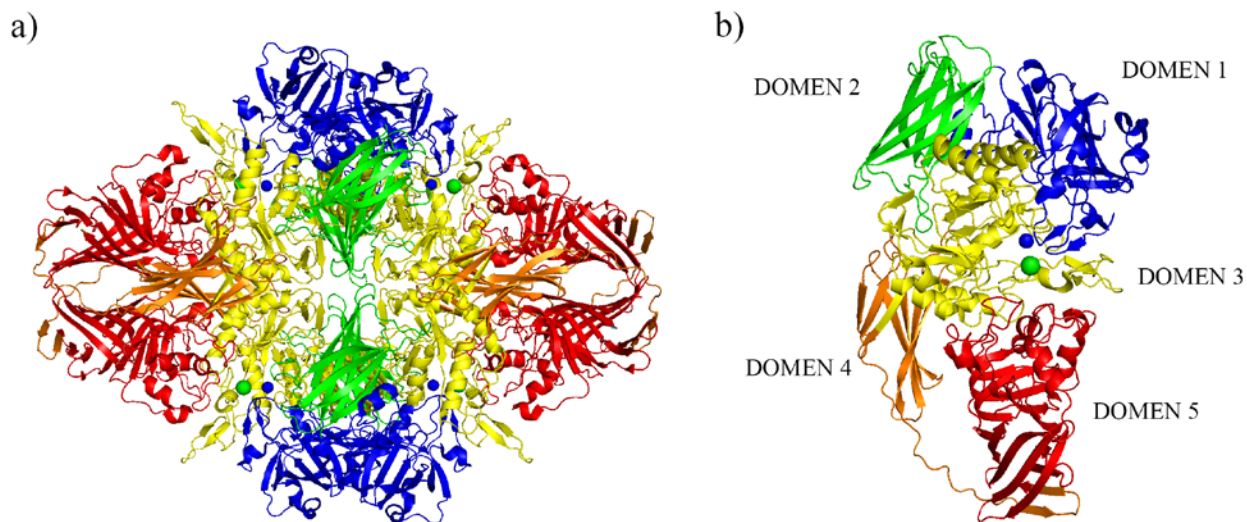
Eksperimentalno je određena kristalna struktura nekoliko beta-galaktozidaza čiji su rezultati deponovani u PDB bazu podataka (tabela 5) (www.rcsb.org).

Tabela 5. Beta-galaktozidaze sa određenom kristalnom strukturom

Organizam	GH-subfamilija	PDB kod	Rezolucija (Å)	Referenca
<i>Sulfolobulus solfataricus</i>	GH-1	1UWQ	2,02	(Aguilar i sar. 1997)
<i>Escherichia coli</i>	GH-2	1DP0	1,70	(Juers i sar. 2000)
<i>Arthrobacter</i> sp.	GH-2	1YQ2	1,90	(Skalova i sar. 2005)
<i>Klyveromices lactis</i>	GH-2	3OBA	2,75	(Pereira-Rodriguez i sar. 2012)
<i>Aspergillus oryzae</i>	GH-35	4IUG	2,60	(Maksimainen i sar. 2013)
<i>Penicillium</i> sp.	GH-35	1TG7	1,90	(Rojas i sar. 2004)
<i>Bacteroides thetaiotaomicron</i>	GH-35	3DEC	2,80	Neobjavljen
<i>Trichoderma reesei</i>	GH-35	3OG2	1,20	(Maksimainen i sar. 2011)
<i>Homo sapiens</i>	GH-35	3THC	1,80	(Ohto i sar. 2012)
<i>Thermus thermophilus</i>	GH-42	1KWG	1,60	(Hidaka i sar. 2002)
<i>Bacillus circulans</i> ssp. <i>alkalophilus</i>	GH-42	3TTS	2,40	(Maksimainen i sar. 2012)

Aktivne rezidue beta-galaktozidaze iz *E. coli* su eksperimentalno određene. Iako se prvobitno smatralo da rezidua GLU461 deluje kao nukleofil, ispitivanja primenom dinitrofenil-2-F-galaktozida je utvrđeno da se ipak GLU537 nukleofil, dok se GLU461 proton donor rezidua (Gebler i sar. 1992). Iako je kristalna struktura beta-galaktozidaze *E. coli* prvobitno utvrđena 1994 godine (Jacobson i sar. 1994), rezolucija modela (2,4Å) nije zadovoljavala potrebe istraživača, pa je istraživanje ponovljeno što je rezultiralo strukturom u rezoluciji od 1,7Å (Juers i sar. 2000). Beta-galaktozidaza *E. coli* se sastoji od četiri identična monomera, dok se svaki sastoji od 1023 amino kiseline i podeljen je u pet strukturalnih domena: jelly-roll barel (Domen 1 sa oko 170

rezidua), dva fibronektin 3 slična domena (Domeni 2 i 4, oko 110 rezidua), jedan veliki beta sendvič domen (Domen 5, oko 300 rezidua) i centralni TIM barel domen koji ima funkcionalnu ulogu u razgradnji supstrata (Domen 3, oko 300 rezidua) (slika 32) (Juers i sar. 2000). Prvih 50 rezidua polipeptidnog lanca ne pripadaju ni jednom navedenom domenu, već predstavlja izduženu nezavisnu konformaciju. Ovaj region interaguje sa prva tri domena svog i prva tri domena susednog monomera, što ukazuje na njegovu ulogu u stabilizaciji strukture kako monomera, tako i celog enzima.

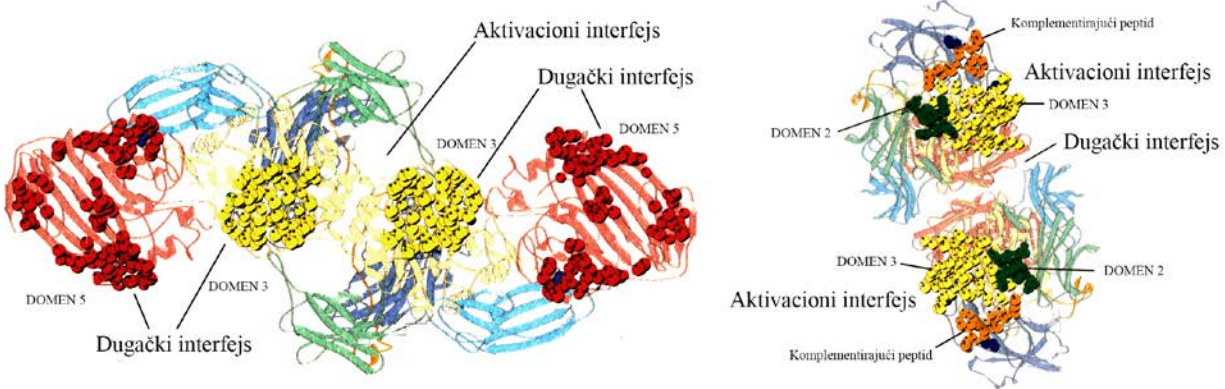


Slika 32. Beta-galaktozidaza *E. coli*, a) struktura tetramera, domen 1 – svetlo plavo, domen 2 – zeleno, domen 3 – žuto, domen 4 – tamno plavo, domen 5 – crveno, metalni katjoni su prikazani sferično; joni u aktivnom mestu su Na⁺ - zeleni krug i Mg²⁺ - plavi krug; b) struktura monomera, prikazani su domen 1, 2, 3, 4 i 5

Kao i kod svih enzima koji sadrže TIM barel domen, aktivno mesto beta-galaktozidaze *E. coli* se nalazi na C-terminalnom kraju centralnog jezgra ovog domena. Aktivno mesto je smešteno u jami koja duboko zalazi u centralni deo TIM barela. Međutim, pored ovog domena, u aktivnom mestu se nalaze i delovi petlje prvog, drugog i petog domena datog monomera.

Tetramer beta-galaktozidaze *E. coli* ima elipsoidnu strukturu sa dimenzijama: 175 x 135 x 90Å. Kroz tetramer se pruža kontinuirani sistem kanala (5-20Å) koji omogućuje dostupnost aktivnog mesta rastvaraču koji kroz njega prolazi. Postoje dva osnovna interfejsa subjedinice: “dugački” interfejs koji zauzima oko 4000Å² i “aktivacioni” interfejs koji zauzima oko 4600Å².

Postoji i treći interfejs, koji je znatno manji sa 230\AA^2 (slika 33) (Juers i sar. 2000). Dugački interfejs grade domeni 3, 4 i 5 i ima dva različita regiona. Prvi region se nalazi u regionu kontakta dva domena 3 iz dva monomera i čini oko 40% interfejsa. U njemu se nalaze dve ARG561 rezidue (od oba monomera po jedna). Drugi region gde dolazi do kontakta uključuje domene 4 i 5. Aktivacioni interfejs uključuje uglavnom domene 2 i 3, kao i komplementirajući peptid. Oko polovina aktivacionog interfejsa je formirana od dve ekvivalentne interakcije između domena 3 i petlje koja uključuje rezidue 272-288 koje pripadaju domenu 2 susedne subjedinice. Drugu polovinu grade interakcije domena 3 i komplementarne petlje.



Slika 33. Dva interfejsa beta-galaktozidaze *E. coli*

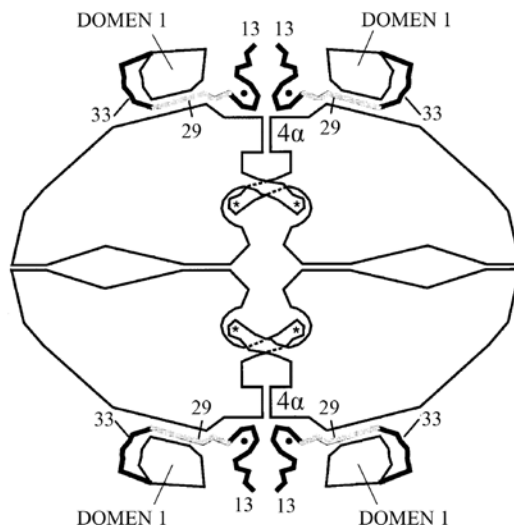
Ispitivani model beta-galaktoridaze *E. coli* (1DP0) imao je oko 4900 atoma rastvarača uključujući vodu, dimetil-sulfoksid (DMSO), Na^+ i Mg^{2+} jone. Oko 80% molekula vode nalazi se na rastojanju manjem od $3,5\text{\AA}$ od polarnih atoma proteina. U modelu je prisutno 112 DMSO molekula koji su posledica korišćenja velike koncentracije za kristalizaciju (30%).

Model sadrži sedam cis-peptidnih veza u svakom monomeru koje se nalaze na PRO87, PRO112, ASP164, HIS-391, PRO422, TRP-568 i PRO902 amino kiselinama. Sve navedene rezidue se nalaze na okretima ili drugim prelomima sekundarne strukture. HIS-391 i TRP-568 se nalaze u aktivnom mestu, dok se PRO112 i PRO422 nalaze u aktivnom interfejsu. Navedeni rezultati pokazuju da su cisX-PRO veze retke, dok su cis X-HIS i cis X-TRP veze veoma retke u ispitivanoj beta-galaktozidazi.

Svaki domen ima dobro izdefinisano hidrofobno jezgro. Međutim, unutar nekih domena se nalaze soni mostovi. To je posebno izraženo u domenu 3 (TIM barel) i domenu 5 (beta-supersendvič). Domen 3 ima mrežu sonih mostova kroz spoljašnje jezgro (ARG 356-ARG 375-

ARG611). Takođe, kao sastavni deo domena 3, u aktivnom mestu je prisutna i mreža koju grade četiri rezidue HIS 391, ASP412, ARG388 i GLU537. Sona mreža domena 5 uključuje šest rezidua (ARG-786, ASP-792, ARG-881, GLU-934, ASP-987 i HIS-990) čiji su bočni lanci u potpunosti uronjeni u unutrašnjost domena. Hidrofobnost domena se smanjuje od sredine ka periferiji zajedno sa smanjenjem broja vodoničnih veza. Isti trend nastavlja se i ka periferiji monomera. Ovakav raspored hidrofobnosti i vodoničnih veza ukazuje na veću stabilnost središnjih u odnosu na periferne delove domena i monomera.

Pokazano je da je beta-galaktozidaza u velikoj meri tolerantna na promene sekvenci pri čemu zadržava svoju funkciju. Beta-galaktozidaza ostaje aktivna nakon hibridizacije sa drugim polipeptidima do dužine od 26 rezidua (Ullmann 1992). Nakon delecija delova lanca 23-31 i 11-41 amino kiselina enzim gubi aktivnost, koji se može dopuniti određenim peptidom (komplementarni peptid, alfa donor) nakon čega se ponovo uspostavlja njegova funkcija (Ullmann i sar. 1967; Ullmann 1992). Deo N kraja lanca od prvih 50 rezidua ima nepravilnu strukturu i prostire se po površini proteina (slika 34). Međutim, njegove rezidue 13 i 15 učestvuju u građenju aktivacionog interfejsa što ih čini funkcionalno izuzetno važnim. Takođe, rezidue od 29 do 33 prolaze kroz tunel koji je formiran spajanjem 3 domena. Magnezijumovi joni (koji su na slici 30 prikazani kao mali puni krugovi) povezuju komplementarni peptid sa ostatkom proteina. Pozicija aktivnih mesta (obeležena zvezdicom) ukazuju na važnost dimerne strukture u tetrameru. Delovi 4 α heliksa aktivacionog interfejsa (prostire se vertikalno po sredini slike) imaju važnu ulogu u spajanju dimera i formiranju aktivacionog interfejsa. 4 α heliks stabilizuje rezidue 22-31. Rezultati istraživanja Fowler i Zabin (1983) pokazala su da rezidue 27-31 imaju ključnu ulogu u formiranju tetramera, što je u saglasnosti sa prikazanom skicom. Kada se formira aktivacioni interfejs, četiri ekvivalentne petlje koje grade rezidue od 272 do 288 se pružaju duž interfejsa kompletirajući aktivna mesta. Za funkcionisanje aktivnih mesta neophodna je tetramerna struktura. Disocijacijom tetramera uklanja se GLU281 petlja iz aktivnog mesta što dovodi do inaktivacije enzima.



Slika 34. Skica ključnih delova beta-galaktozidaze *E. coli* (vidi objašnjenje u tekstu)

Kristalografijom je utvrđeno da postoji 5 vezivnih mesta za Na^+ . Obzirom da se jedno mesto nalazi blizu aktivnog mesta (vezuje se za ASP202 i ASN605), izostanak natrijuma dovodi do smanjenja aktivnosti enzima, ali ne i do inaktivacije. Takođe, i magnezijumovi joni su otkriveni u aktivnom mestu (Douglas H Juers, Beatrice Rob, Megan L Dugdale, Nastaron Rahimzadeh, Clarence Giang, Michelle Lee, Brian W Matthews, and Reuben E Huber 2009.). Utvrđeno je da Mg^{2+} joni imaju strukturalnu i funkcionalnu ulogu u enzimu. U aktivnom mestu se nalaze dva Mg^{2+} jona od kojih se jedan veže za GLU461, GLU416, HIS418 i tri molekula vode koji obezbeđuju interakcije 3-OH i 4-OH grupe. Njegova konstanta disocijacije je mala ($K_d=10^{-7}\text{M}$). On je pozicioniran najbliže mestu vezivanja supstrata i ima najveći uticaj na aktivnost enzima. Moguća je njegova zamena sa Mn^{2+} jonom bez značajnijeg gubitka enzimske aktivnosti. Drugi Mg^{2+} jon interaguje sa GLU797 i stabilizuje GLY794 – PRO803 petlju u aktivnom mestu. Njegova konstanta disocijacije je $K_d=10^{-4}\text{M}$ i ima manji uticaj na aktivnost enzima.

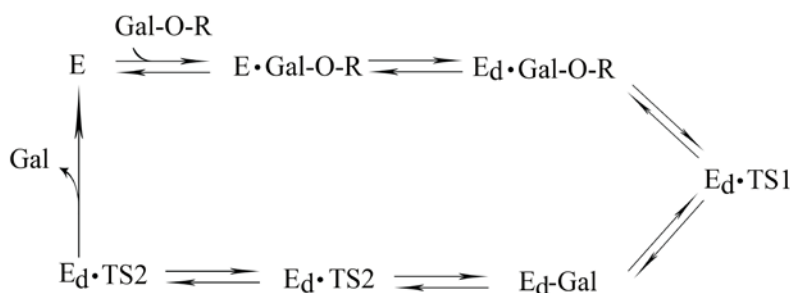
2.9.5. Mehanizam razgradnje laktoze beta-galaktozidaze *E. coli*

Da bi došlo do razgradnje supstrata, on prvo mora dospeti u aktivno mesto. U beta-galaktozidazi *E. coli* supstrat difunduje do aktivnog mesta (Fersht 1985). Kada se nađe u aktivnom mestu, supstrat se razgrađuje brzinom od 15 ms. Da bi se postigao zadovoljavajući model razgradnje, potrebno je da što više molekula enzima u kristalnoj strukturi bude sinhronizovan u procesu razgradnje supstrata. Da bi se to postiglo, često se koriste substituenti

supstrata koji će se ugraditi u aktivno mesto, ali ne dolazi do reakcije, i supstrat ostaje zaustavljen u aktivnom mestu. Time se obezbeđuje da veliki broj molekula imaju supstrat vezan za aktivno mesto, čime se poboljšava signal za kristalografiju. Takođe, da bi se odredilo mesto vezivanja liganda uglavnom se primenjuje tehnika sa dva enzima, gde je u jednom vezan ligand dok je drugi u nativnom stanju.

Utvrđeno je da se aktivnom mestu nalazi petlja sastavljena od rezidua GLY794 do PRO803 koja ima dve različite konformacije u zavisnosti od faze vezivanja i vrste supstrata (Juers i sar. 2000). Pokazano je da petlja ima otvorenu konformaciju prilikom vezivanja supstrata, dok je zatvorena u nekim fazama vezivanja određenih analoga supstrata. Nakon ulaska laktoze u aktivno mesto petlja ima otvorenu konformaciju.

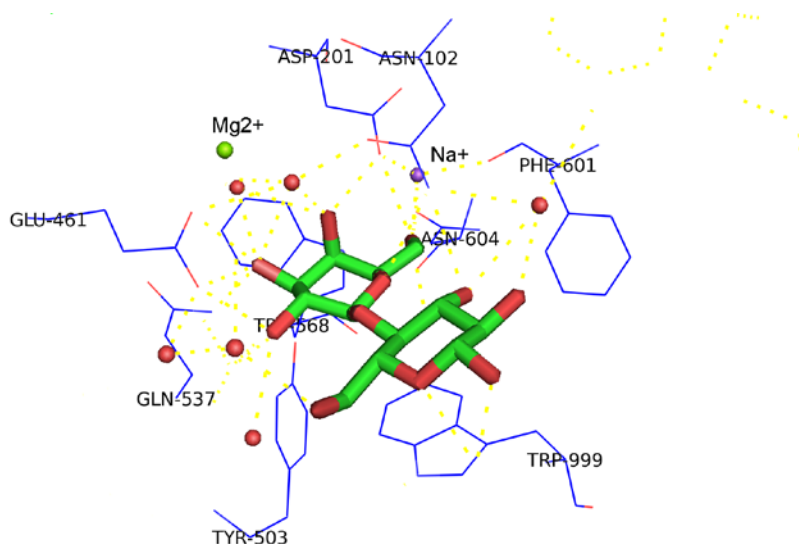
Šema razgradnje galaktozida beta-galaktozidazom prikazana je na slici 35. Galaktozid (Gal-OR) se vezuje za enzim (E) na takozvanoj plitkoj poziciji u aktivnom mestu (E•Gal-OR). Zatim, usled konformacionih promena prelazi na duboku poziciju aktivnog mesta (Ed•Gal-OR). Prva reakcija je transgalaktozilacija koja počinje dok je supstrat u dubokom mestu i nastavlja se tokom tranzicionog stanja (Ed•TS1) koje vodi do nastanka intermedijeta (Ed-Gal), prilikom čega dolazi do otpuštanja molekula glukoze. Intermedijet se otpušta u vodeni rastvor preko drugog tranzicionog stanja (Ed•TS2), pri čemu nastaje kompleks produkt-enzim, koji se dalje odvaja na produkt i enzim. U slučaju sinteze alolaktoze, molekul glukoze se ne otpušta, već prima molekul galaktoze i oslobađa se disaharid alolaktoza.



Slika 35: reakciona šema beta-galaktozidaze (objašnjenje u tekstu)

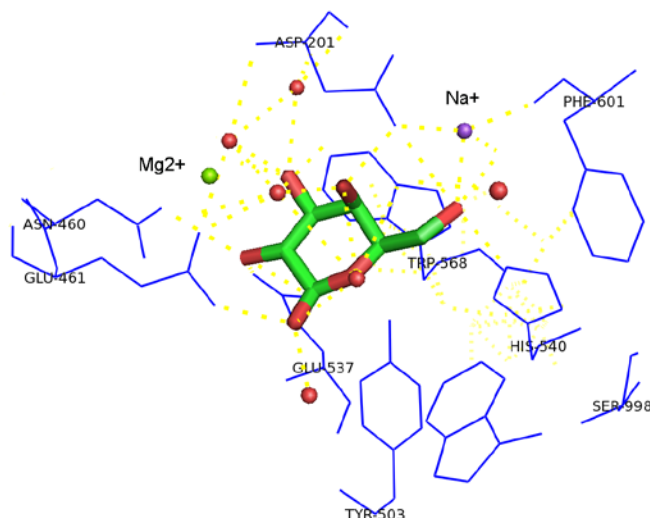
Vezivanje laktoze za beta-galaktozidazu je ispitano upotrebom inaktivne varijante enzima (E537Q) koji nema sposobnost razgradnje (Juers i sar. 2012). Laktoza u počenim fazama vezivanja naleže na reziduu TRP999, dok njen galaktozilhidroksil gradi veze sa proteinom i

rastvaračem (slika 36). 6-OH grupa se vezuje za Na⁺joni HIS540. 3-OH i 4-OH grupe se vezuju za protein (ASP201) i Mg²⁺ jon, dok 2-OH ima jedan polarni kontakt sa GLU461. Glukozni deo molekula koji se nalazi izvan aktivnog mesta ne gradi polarne veze sa enzimom. C6 atom gradi hidrofobnu vezu sa PHE601. U opisanoj početnoj fazi vezivanja C1 atom se nalazi na udaljenosti od oko 6Å od nukleofilnog GLU537 koji gradi vodoničnu vezu sa C3 atomom na udaljenosti od 3,1Å. O4 takođe udaljen oko 6Å od proton donora GLU461, koji gradi vodoničnu vezu sa C2 na udaljenosti od oko 2,6Å.



Slika 36. Prikaz vezivanja laktoze u aktivno mesto beta-galaktozidaze *E.coli*

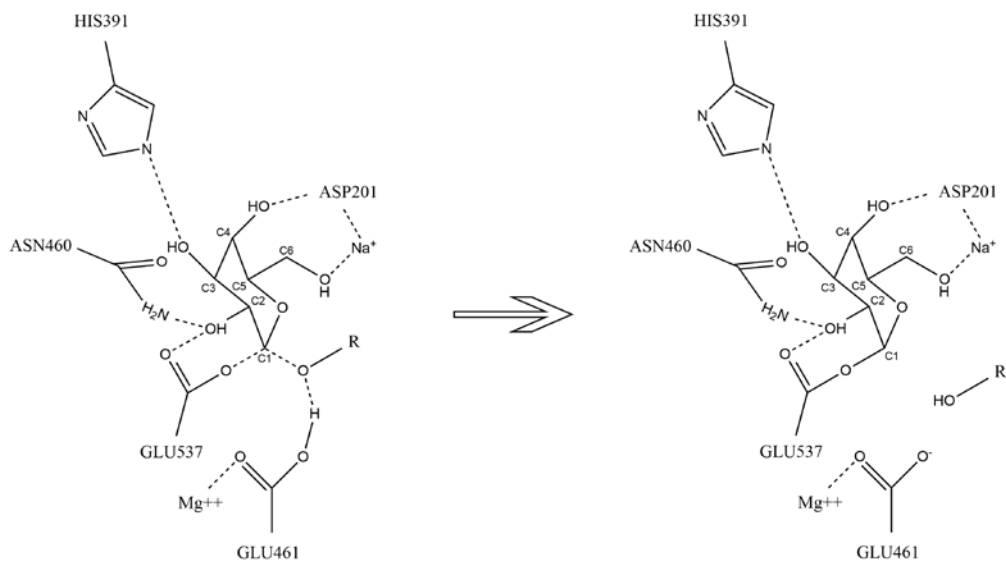
Iz gore opisanog početnog položaja laktoze u aktivnom mestu beta-galaktozidaze sledi da ona ne zauzima potrebnu konformaciju za katalitičko delovanje aktivnih rezidua GLU461 i GLU537. Laktoza treba da se pomeri oko 3Å dublje u aktivno mesto da bi došlo do katalize. Vezivanje u toj poziciji je ispitano pomoću vezivanja galaktoze za slobodan enzim (slika 37) (Juers i sar. 2001; Juers i sar. 2012). Galaktoza je pogodna za ovo ispitivanje jer pored toga što je produkt razgradnje laktoze predstavlja i supstrat u reakcijama sinteze GOS-a (Huber i Hurlburt 1986). Galaktoza se vezuje za enzim na način kao i laktoza, što potvrđuje isti položaj petlje i rezidua koje utiču na njen položaj (ARG559 i PHE601).



Slika 37. Prikaz vezivanja galaktoze u aktivno mesto beta-galaktozidaze *E.coli*

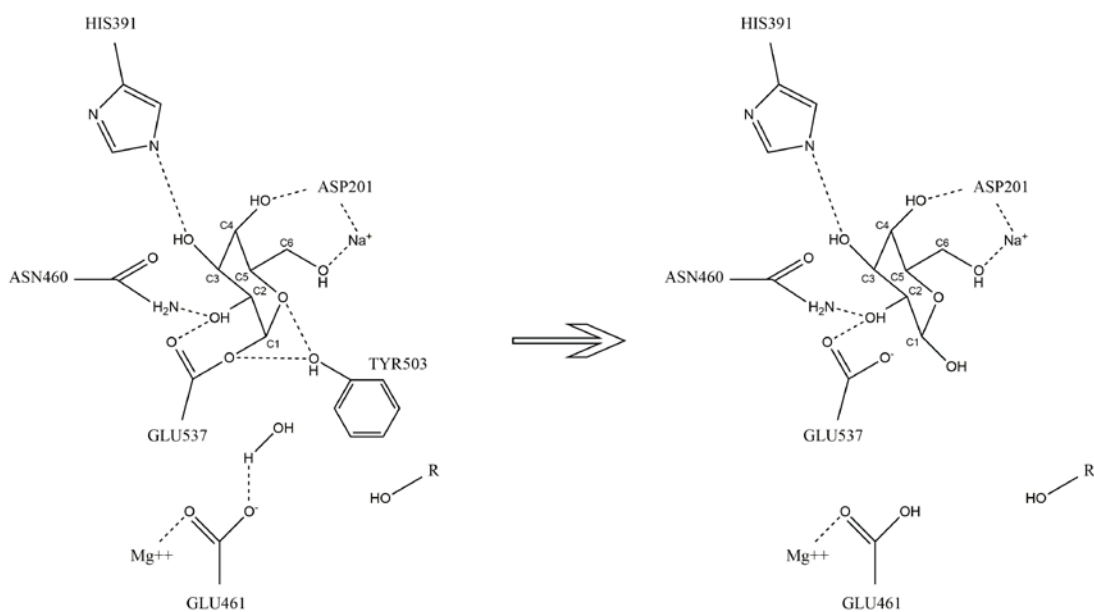
Da bi došlo do pomeranja laktoze dublje u aktivno mesto dolazi do rotacije oko ose koja povezuje 4-OH i 6-OH (Juers i sar. 2012). Tom prilikom održavaju se interakcije 4-OH sa ASP 201 i Mg^{2+} , dok 6-OH interaguje sa Na^+ jonom, HIS540 i ASN604. Sa druge strane 3-OH dislocira molekul vode pri čemu on interaguje sa HIS391 i dva druga molekula vode - jednim koji je povezan sa Mg^{2+} jonom u aktivnom mestu i drugim koji je povezan sa HIS357. 2-OH postaje dovoljno blizu ASN460, GLU461 i GLU537 ($<3.2 \text{ \AA}$) usled čega dolazi do interakcije. Zatim 1-OH zauzima poziciju koju je imao 2-OH gradeći vodoničnu vezu sa GLU461, što omogućuje proton donorsko delovanje GLU461 na 1-OH grupu.

Faza galaktozilacije počinje kada GLU461 donira proton na glikozidnu vezu u dubokom mestu. Veza se raskida između C1 atoma galaktoznog dela molekula i kiseonika koji gradi beta 1→4 glikozidnu vezu (slika 38). Kiseonik postaje pozitivno naelektrisan, privlači elektrone i pozitivno elektriše C1 atom. Uspostavljanjem ravnoteže dolazi do parcijalne dvostruke veze između O5 i C1 atoma. GLU537 donira elektron čime uspostavlja vezu sa C1 atomom i gradi enzim-supstrat intermedijer, dok se glukozni deo oslobađa. Tranziciono stanje je stabilisano interakcijama između 2-OH i GLU537 i ASN460.



Slika 38. Razgradnja supstrata beta-galaktozidazom – prvi deo reakcije

Drugi deo reakcije prenosi galaktozilnu grupu od nukleofilnog GLU537 na receptorski molekul (slika 39). Receptorski molekul može biti voda ili glukoza u slučaju sinteze alolaktoze. GLU461 napada vezu GLU537 i galaktoze odvajajući proton čime nadoknađuje svoj izgubljen u prvom delu reakcije. U slučaju vode kao receptora, nakon raskidanja veze, preostala OH grupa se vezuje za galaktozil i gradi galaktozu koja se oslobađa i napušta enzim.

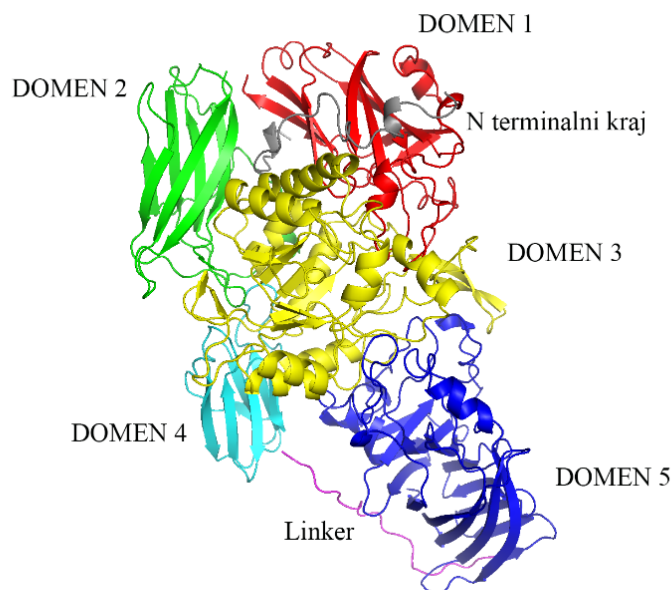


Slika 39. Razgradnja supstrata beta-galaktozidazom – drugi deo reakcije

2.9.6. Struktura beta-galaktozidaze *Kluyveromyces lactis*-a

Beta-galaktozidaza iz *Kluyveromyces lactis*-a je jedan od najprimenjenijih enzima u industriji mlečnih proizvoda. Obzirom da se procenjuje da oko 70% svetske populacije ima neki oblik netolerancije na laktozu, našla je široku primenu u proizvodnji delaktoziranog mleka (Husain 2010). Iako potiče iz eukariota ne pripada GH-35 subfamiliji, već zajedno sa beta-galaktozidazama *E.coli* i mnogim drugim prokariota pripada GH-2 subfamiliji glikozidhidrolaza (tabela 5). Njena trodimenzionalna struktura je određena kako u nativnoj formi, tako i u kompleksu sa galaktozom u rezoluciji 2,8Å (Pereira-Rodriguez i sar. 2012). Struktura enzima sa laktozom u aktivnom mestu nije određena usled velike aktivnosti enzima u uslovima kristalografije, što je rezultiralo razgradnjom laktoze i strukturom sa galaktozom u aktivnom mestu. Kao i beta-galaktozidaza *E. coli* (sa kojom deli stepen homologije od 48%, RMSD = 6,0Å), beta-galaktozidaza *Kluyveromyces lactis*-a se sastoji iz četiri identične monomerne subjedinice (A, B, C i D) od po 1024 amino kiseline sa molekulskom masom 119kDa.

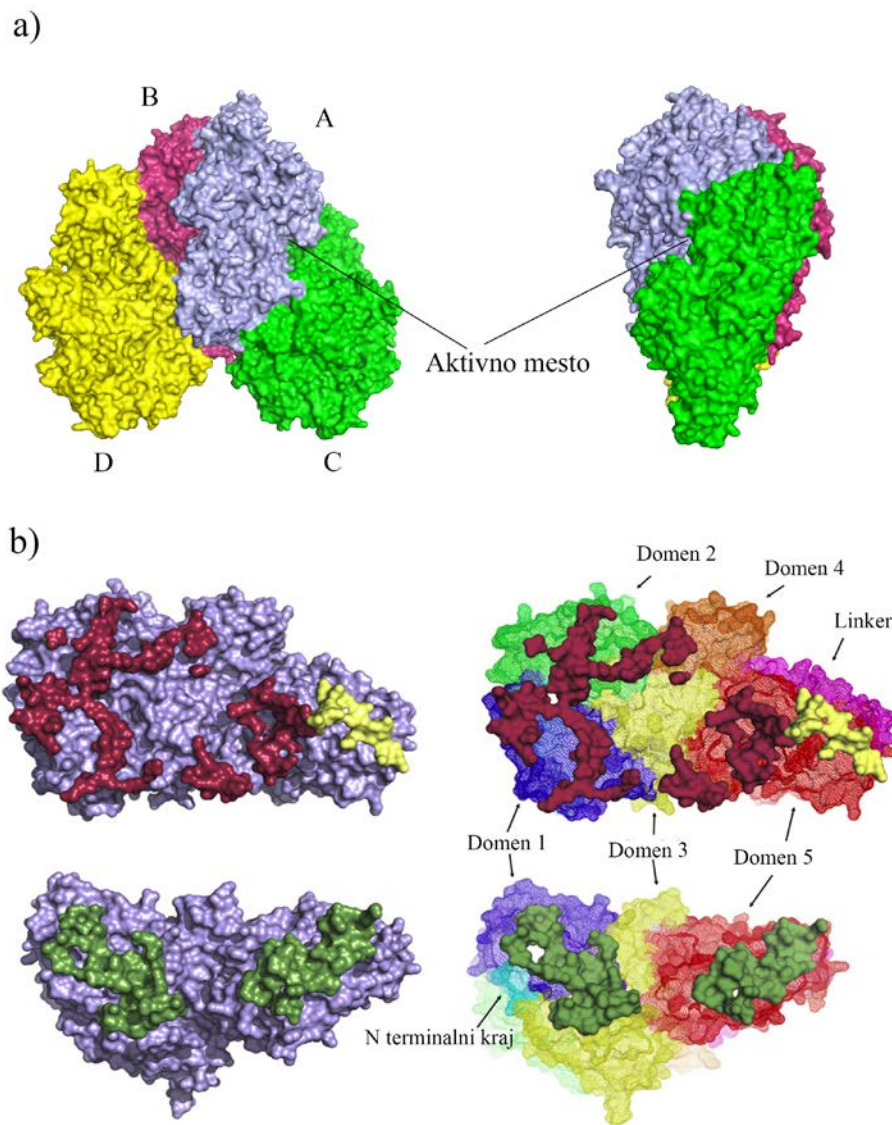
Monomer beta-galaktozidaze *Kluyveromyces lactis*-a se sastoji iz pet domena od kojih jedan (domen 3) ima katalitičku funkciju (slika 40). Domen 1 sadrži rezidue 32-204, ima formu rolata i ulogu u vezivanju supstrata. Domen 2 (rezidue 205-332) i domen 4 (rezidue 643-720) grade beta sendvič domene slične imunoglobulinu. Kao što je već pomenuto kod beta-galaktozidaze *E. coli*, domen 3 (rezidue 333-642) je katalitički TIM barel domen u kojem se nalazi aktivno mesto. Domen 5 (741-1024) predstavlja mali lanac beta-galaktozidaze. Dva regiona se ne mogu svrstati u postojeće domene: N terminalni region (rezidue 2-31) i deo koji povezuje domene 4 i 5 (rezidue 721-740).



Slika 40. Struktura monomera beta-galaktozidaze *Kluyveromyces lactis*-a

Ispitivanja aktivnosti beta-galaktozidaze *Kluyveromyces lactis*-a pokazala su da postoje dve forme aktivnog enzima. Nakon elektroforetske karakterizacije zaključeno je da se enzim može nalaziti u formi dimera i tetramera (Becerra M i sar. 1998). Kristalografska i ispitivanja energija disocijacija potvrdila su ove rezultate takozvanog dimera dimera. Naime, pokazano je da je energija disocijacija tetramera u dva dimera ($\Delta G^{\text{int}} = 6\text{kcal/mol}$) značajno manja od energije disocijacije samog dimera ($\Delta G^{\text{int}} = 20\text{kcal/mol}$).

Aktivna mesta monomera su u interfejsu dimera okrenuta jedno prema drugom (Pereira-Rodriguez i sar. 2012). Dimeri nastaju interakcijama A i C, kao i B i D subjedinice (slika 41a). Površine interakcija između njih su identične i zahvataju površinu od 2521\AA^2 (slika 41b). Interakcije A i B subjedinice su ključne za formiranje tertramera (2438\AA^2), iako postoje i interakcije između subjedinica A i D, kao i B i C koje učestvuju u stabilizaciji tertramerne strukture (350\AA^2). Interakcije između interfejsa su uglavnom nepolarne. Interakcije A-C i B-D se događaju preko domena 1, 3 i 5, dok interakcije A-B monomera uključuju domene 1, 2, 4 i 5, kao i petlju 8 katalitičkog domena 3.

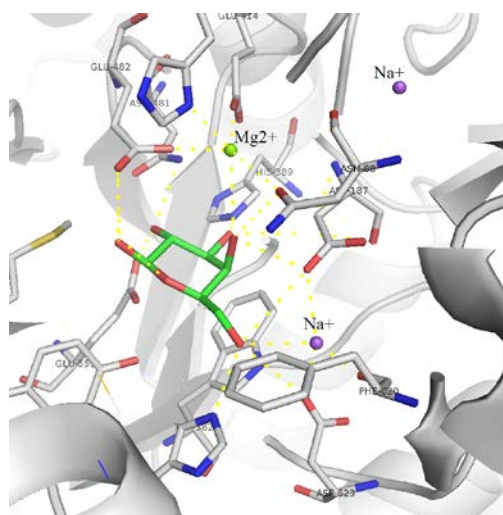


Slika 41. Struktura tetramera beta-galaktozidaze *Kluyveromyces lactis*-a; a) domeni, b) površine interakcija između domena

Aktivno mesto beta-galaktozidaze *K. lactis*-a nalazi se u TIM barel domenu. Zauzima centralnu poziciju u svakom monomeru i okruženo je reziduama iz domena 1, 3 i 5. Aktivne rezidue su identifikovane na osnovu homologije sekvenci sa do sada poznatim beta-galaktozidazama GH-2 subfamilije (GLU482 – proton donor i GLU551 – nukleofil). Stvaranjem dimera aktivna mesta postaju još zaštićenija reziduama jer su okrenuta jedno ka drugom. I pored fizičke bliskosti nisu zabeležene interakcije među njima koje bi uticale na aktivnost enzima.

Ovakvom organizacijom aktivno mesto je pristupačno samo kroz uski kanal promera od oko 10Å.

Ispitivanjem strukture beta-galaktozidaze *K. lactis*-a kristalografijom rendgenskim zracima u nativnom enzimu nisu zabeleženi metalni joni, već samo molekuli vode (Pereira-Rodriguez i sar. 2012). Sa druge strane u enzimu u kompleksu sa galaktozom nalaze se jedan Mg^{++} i 2 Na^{+} jona. Magnezijumov i jedan natrijumov jon su nađeni u blizini galaktoze vezane za enzim (slika 42), što je takođe homologo sa beta-galaktozidazom *E. coli* (Juers i sar. 2009). Pored pomenutih metalnih jona u aktivnom mestu, pronađeni su još po jedan Na^{+} i Mn^{++} jon za koje se smatra da imaju strukturnu ulogu jer se nalaze u delovima koji grade interfejsne interakcije dimera. Obzirom da su metalni joni u aktivnom mestu konzervativni u odnosu na beta-galaktozidazu *E. coli*, pretpostavlja se da im je i funkcija ista (Lo i sar. 2010).

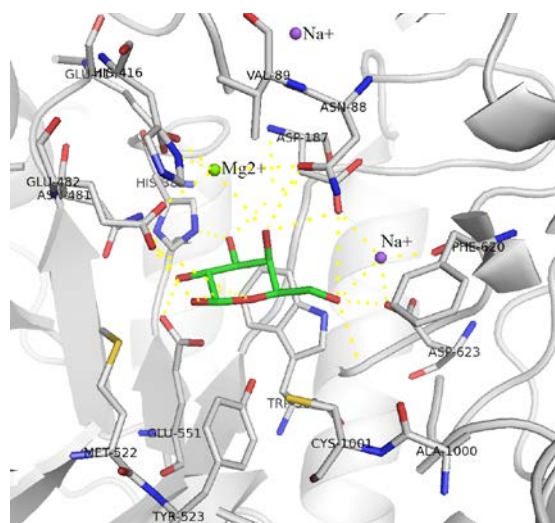


Slika 42. Aktivno mesto beta-galaktozidaze *Kluyveromyces lactis*-a sa molekulom galaktoze i metalnim Na^{+} i Mg^{2+} jonima

Iako postoji visok stepen homologije sa beta-galaktozidazom iz *E. coli*, poređenjem struktura utvrđeno je odstupanje od 3,2Å, što se prepisuje odstupanjima u orijentaciji domena. Komplementarni peptid koji je otkriven kod beta-galaktozidaze *E. coli* nije pronađen kod beta-galaktozidaze *K. lactis*-a, odnosno za taj region peptida nije pronađena funkcija. U domenu 2 beta-galaktozidaze *K. lactis*-a postoji insercija na poziciji 246-274 u odnosu na beta-galaktozidazu *E. coli*, koja gradi veći deo dodirne površine subjedinica A i B, odnosno učestvuje u stabilizaciji tertramera. Takođe, pomenuta interakcija se događa i između C i D subjedinica. U

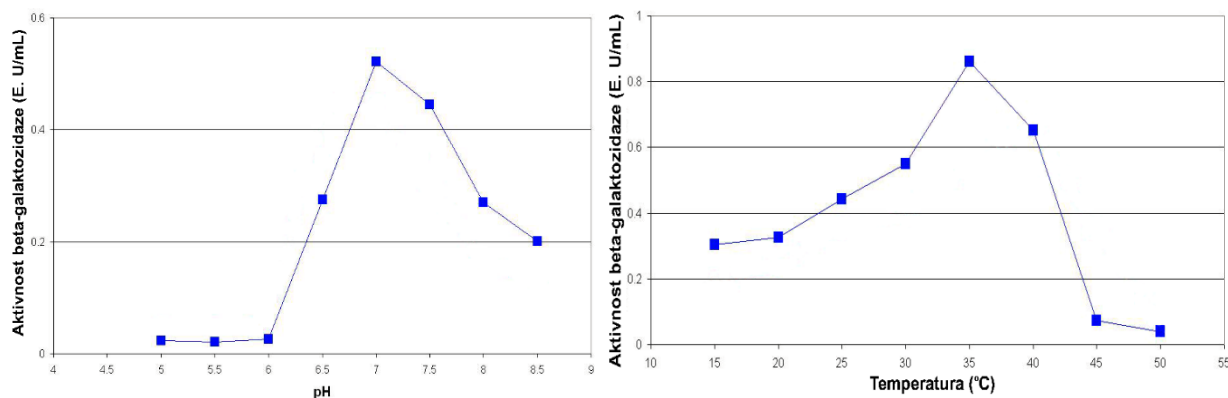
katalitičkom TIM barel domenu takođe postoji insercija dela peptida (rezidue 420-443). Ona se prostire preko ulaza u aktivno mesto skrivajući ga dublje u džep u kojem se nalazi. Tokom stvaranja dimera ona gradi kanal putem kojeg je aktivno mesto obe subjedinice dostupno rastvaraču. Domen 4 je značajno manji nego kod beta-galaktozidaze *E. coli* u kojem je većina petlji i beta ploča smanjena. Takođe, dugački lanac koji povezuje domene 4 i 5 je pomeren više ka domenu 5. Interesantan je podatak da male promene tipa insercija i delecija mogu imati značajan uticaj na biohemijske karakteristike enzima koji globalno imaju slično savijanje peptida. Insercije i delecije dovode do malih promena u nekim petljama koje uslovljavaju kompletno drugačiju oligomerizaciju peptida (Skalova i sar. 2005).

Pored hidrolize laktoze (beta-galaktozidaza *E. coli*), beta-galaktozidaza *K. lactis*-a je sposobna da u većoj meri sintetiše i 6 galaktobiozu, kao i trisaharid 6 galaktozil laktozu (Martínez-Villaluenga i sar. 2008). Globalno gledano struktura aktivnog mestabeta-galaktozidaze *K. lactis*-ane razlikuje se značajno od beta-galaktozidaze *E. coli*. Ipak, postoje neke promene koje mogu uticati na vezivanje supstrata i aktivnost enzima. Aktivne residue zauzimaju gotovo identične pozicije u oba enzima što ukazuje na istovetan mehanizam reakcije razgradnje supstrata. Međutim, aktivno mesto galaktozidaze *K. lactis*-a pored domena 3 grade i neke residue domena 1 (ASN88, VAL89, ASP187) i domena 5 (ALA1000, CYS1001) (slika 43). Posebnu važnost ima uloga CYS1001, na čijem mestu se kod *E. coli* nalazi TRP999. Ispitivanjem mutacija kod *E. coli* pokazano je da ova promena pozitivno utiče na aktivnost enzima.



Slika 43. Aktivno mesto beta-galaktozidaze *K. lactis*-a sa ključnim reziduama zanjegovu aktivnost

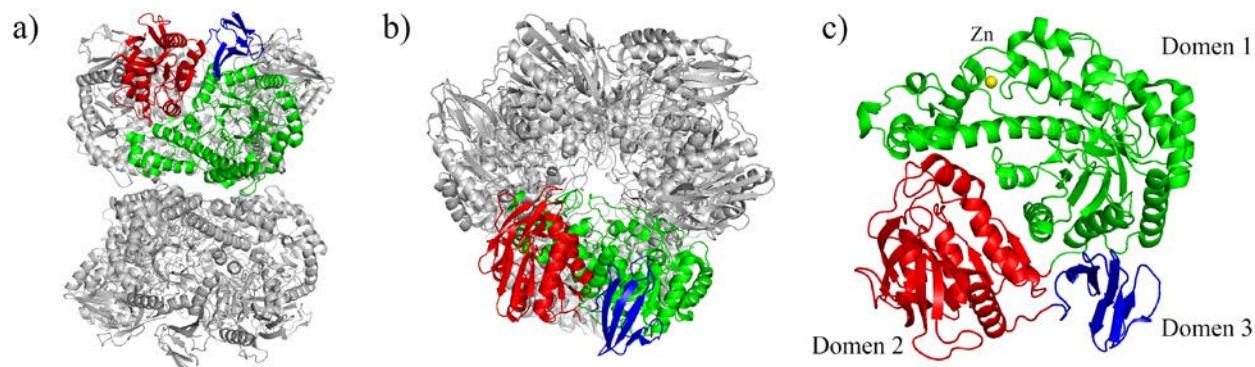
Ispitivanja aktivnosti beta-galaktozidaze *K. lactis*-a pokazala su da je K_m vrednost 20mM, dok je V_{max} iznosila 570 U/mg (Rodriguez i sar. 2006). Kao optimalna pH vrednost utvrđena je 7,0 do 7,5. Sa smanjenjem pH opada i aktivnost enzima usled promena u naelektrisanim reziduama. Optimalna temperatura je 30-35°C, iako je enzim pokazao stabilnost na njenu promenu od 25°C do 41°C (slika 44).



Slika 44. Aktivnost enzima beta-galaktozidaze *K. lactis*-a u zavisnosti od pH vrednosti i temperature

2.9.7. Struktura beta-galaktozidaze *Bacillus circulans*-a

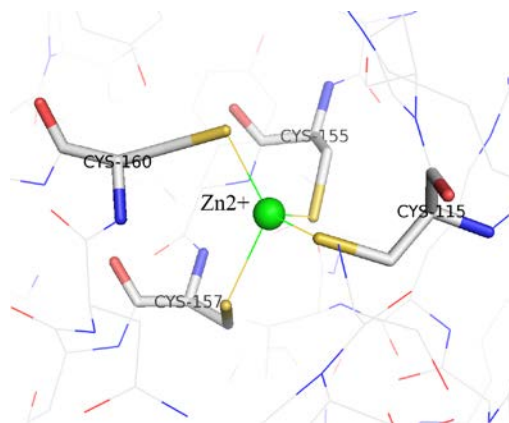
Struktura beta-galaktozidaze iz *Bacillus circulans*-a (predstavnik iz GH-42 podfamilije) je eksperimentalno utvrđena primenom kristalografije rendgenskim zracima (Maksimainen i sar. 2012). Određena je struktura nativnog enzima, kao i u kompleksu sa galaktozom u rezoluciji od 2,4Å. Za razliku od beta-galaktozidaze *E. coli* (GH-2 subfamilija) koja je tetramerne strukture, beta-galaktozidaza *Bacillus circulans*-a (GH-42) se sastoji od šest monomera koji su izgrađeni od po 675 amino kiselina (slika 45 a, b). Monomeri su simetrično raspoređeni, tako da gornji i donji deo predstavljaju dva trimera. Pojedinačni monomeri su imali odstupanje strukture od 0,18Å (RMSD), dok je poređenjem celih heksamera zabeleženo odstupanje od 0,33Å, što ukazuje na promene relativnog položaja monomera u kvarternoj strukturi.



Slika 45. Struktura beta-galaktozidaze *Bacillus circulans*-a; a) i b) strukture heksamera iz različitih uglova, c) struktura monomera

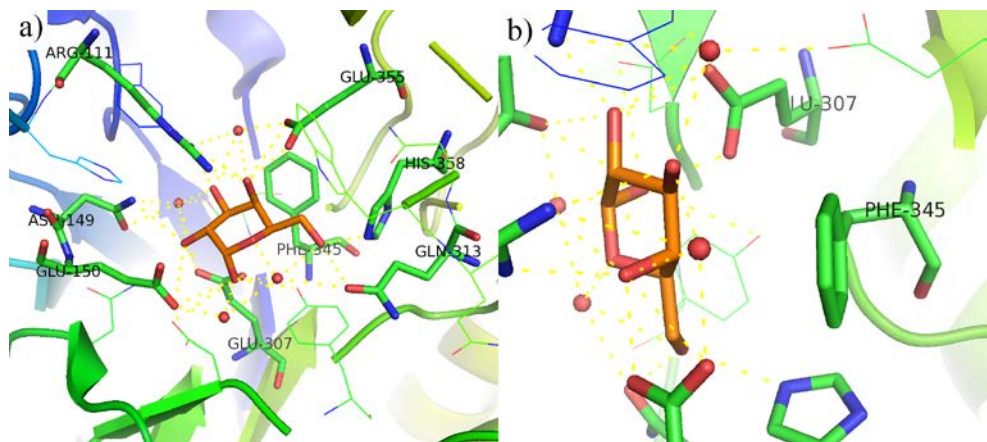
Monomer beta-galaktozidaze iz *Bacillus circulans*-a se sastoji iz 3 domena (slika 45c). Domen 1 je obeležen zelenom bojom i predstavlja katalitički domen sa TIM barelom. Domen 1 je najkonzervativniji među homologim sekvencama uz zabeleženu RMSD od svega 1,03Å sa beta-galaktozidazom iz *Thermus thermophilus*-a. Petlje koje povezuju C terminalne krajeve alfa heliksa i N terminalne krajeve beta ploča su kratke čineći N terminalne krajeve TIM barela relativno glatkim. Petlje koje povezuju C terminalne krajeve beta ploča i N terminalne krajeve alfa heliksa su duže i često sadrže elemente sekundarne strukture. Petlja koju grade rezidue THR80–ASN117 sadrži vezivno mesto za Zn^{2+} jone. Takođe, u njihovom vezivanju učestvuje i petlja VAL147–PHE214 koja takođe ima važnu ulogu u formiranju kvarterne strukture jer učestvuje u interakcijama između monomera. Domen 2 (GLU396–PRO609) ima strukturu sastavljenu od centralno postavljenih sedam uspravnih beta ploča i pet alfa heliksa. Njegova funkcija je strukturna jer formira interfejs sa drugim monomerom stabilišući na taj način enzim. Treći domen (LEU610–ALA675) je najmanji i ima antiparalelnu beta-sendvič strukturu od dve beta ploče koje grade strukturu sličnu rolatu. Pretpostavlja se daje i njegova uloga strukturna.

U aktivnom mestu monomera se nalazi vezivno mesto za metalni jon Zn^{2+} . Njega grade četiri cisteina CYS, CYS155, CYS157 i CYS160 (slika 46). Pretpostavlja se da cink nema funkcionalnu ulogu u aktivnom mestu, već da mu je uloga strukturna (Maksimainen i sar. 2012).



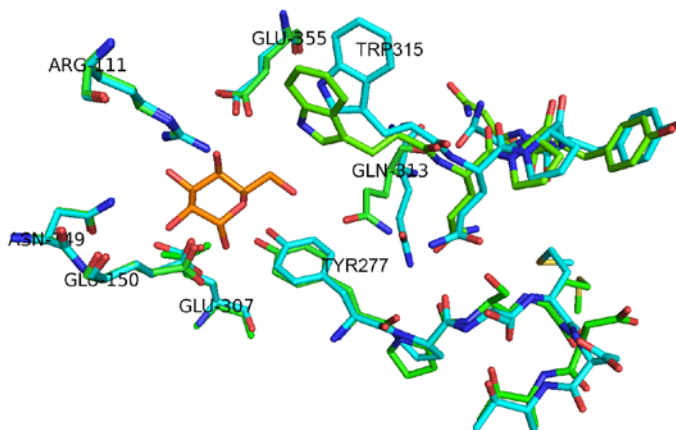
Slika 46. Vezivno mesto za Zn^{2+} jon u aktivnom mestu beta-galaktozidaze *B. circulans*

Ispitivanja kristalografijom rendgenskim zracima (analizom gustine elektrona) i primenom galaktoze kao liganda utvrđeno je da se u svakom katalitičkom domenu nalazi po jedan molekul galaktoze (slika 47a) (Maksimainen i sar. 2012). Međutim, svi molekuli galaktoze su se nalazili u alfa anomernoj konfiguraciji. U vezivanju galaktoze učestvuje sedam rezidua (ARG111, ASN149, GLU150, GLU307, GLN313, GLU355 i HIS358). Takođe, PHE345 se nalazi u blizini C4 atoma galaktoze na kojem je hidroksilna grupa u poprečnom (aksijalnom) položaju, što objašnjava vezivanje galaktoze, dok se glukoza nije favorizovana (slika 47b). Mreža vodoničnih veza uključuje interakcije i sa okolnim molekulima vode. Na primer TRP187 gradi vodonične veze sa alfa-D-galaktosom preko molekula vode. Na osnovu sličnosti sa već eksperimentalno utvrđenim osobinama beta-galaktozidaza od drugih organizama utvrđeno je da su GLU150 i GLU307 aktivne residue (Hidaka i sar. 2002). Hidaka i sar. (2002) su odredili aktivne residue (GLU141 and GLU312) superpozicioniranjem rezidua aktivnog mesta sa reziduama iz beta-galaktozidaze *E. coli*, čije osobine su diskutovane u prethodnim poglavljima. GLU150 je proton donor, dok je GLU307 nukleofilna rezidua. ASP275 je identifikovana kao važna rezidua za katalitičke osobine obe aktivne residue jer se nalazi u njihovoj neposrednoj blizini i sa njima gradi kratke vodonične veze (2,5Å). Iako ne postoje eksperimentalni podaci kao ni doking studije, autori Maksimainen i sar. (2012) pretpostavljaju da je TYR277 najverovatnije vezivno mesto za laktozu.



Slika 47. Molekul alfa-D-galaktoze u aktivnom mestu beta-galaktozidaze *B. circulans*: a) galaktoza u aktivnom mestu sa reziduama koje interaguju sa njom; b) pozicija PHE345 u blizini C4 atoma galaktoze

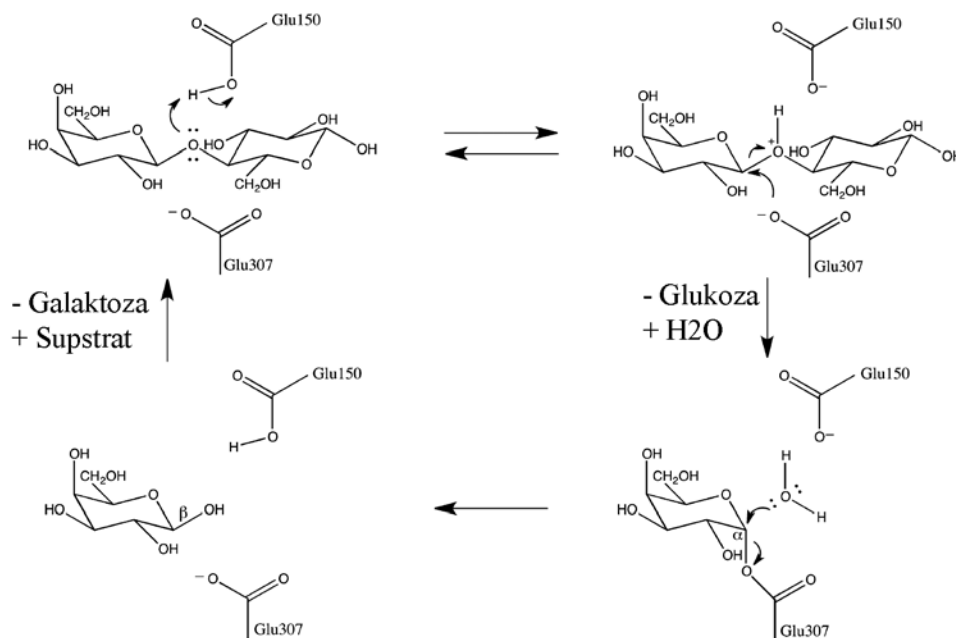
Poređenjem struktura nativnog enzima i enzima u kompleksu sa alfa-D-galaktozom uočene su strukturalne promene kod dve petlje (slika 48). Prva promena je u petlji koja povezuje C terminalni kraj beta ploče i N terminalni kraj alfa heliksa u regionu TYR277-THR282 (petlja 6). Tokom konformacione promene, njena pozicija se menja za 1-3Å. Petlja 6 interaguje sa petljom 7 u kojoj se takođe događa konformaciona promena u regionu GLN313-GLN316. Tokom promene TRP315 se pomera bliže ka alfa-D-galaktozi. Navedene konformacione promene se događaju na istim lokacijama kao kod beta-galaktozidaze iz *Trichoderma reesei* (Maksimainen i sar. 2011).



Slika 48. Konformaciona promena u aktivnom mestu nativnog (svetlo plavaboja) i enzima beta-galaktozidaze *B. circulans* u kompleksu sa alfa-D-galaktozom (zelena boja).

2.9.8. Mehanizam razgradnje laktoze beta-galaktozidaze *Bacillus circulans*-a

Činjenica da je galaktoza u aktivnom mestu zabeležena u alfa anomernoj konformaciji deluje interesantno na prvi pogled. Reakcija hidrolize započinje (kao i kod beta-galaktozidaze *E. coli*) protonacijom glikozidnog kiseonika koju inicira proton donor GLU150 rezidua (slika 49). Zatim dolazi do nukleofilne reakcije GLU307 usled čega dolazi do raskidanja galaktozidne veze i formira se kovalentna veza između galaktoze i GLU307 (nastaje intermedijer), dok glukoza napušta aktivno mesto. Nakon reakcije sa vodom koja nadoknađuje H atom GLU150 i OH grupu galaktozi, beta-D-galaktoza se otpušta u rastvor gde se procesom mutarotacije prevodi u alfa-D-galaktozu, koja se ponovo veže za aktivno mesto, usled čega je u takvoj konformaciji zabeležena u aktivnom mestu enzima.



Slika 49. Razgradnja laktoze beta-galaktozidazom iz *Bacillus circulans*-a

Pored razgradnje laktoze beta-galaktozidaza *Bacillus circulans*-a je sposobna i da sintetiše galaktooligosaharide, što je posebno značajno za sintezu prebiotika (Li i sar. 2009). Međutim, utvrđeno je da je nivo sinteze prilično slab i ukoliko se koriste velike koncentracije supstrata. Beta-galaktozidaza *Bacillus circulans*-a sintetiše GOS putem reverzne hidrolize i

proces se može termodinamički kontrolisati (Maksimainen i sar. 2012). Proces sinteze podrazumeva promenu u reakciji hidrolize laktoze. Nakon nastanka intermedijera njegova kovalentna veza se raskida molekulom receptora koji može biti glukoza ili GOS (kod hidrolize je taj molekul voda) i galaktoza se ugrađuje u dati molekul gradeći ponovo laktozu ili nadograđujući GOS molekul. Međutim, važno je napomenuti da je hidroliza laktoze egzotermna reakcija, usled čega je nastanak monosaharida favorizovan hemijskim ekvilibrijumom.

Obzirom da oblik beta-galaktozidaza može značajno varirati, posebno između pripadnika različitih subfamilija enzima, neke beta-galaktozidaze mogu biti pogodnije od drugih za sintezu GOS-a. Takođe, uslovima se može značajno menjati efikasnost određenog enzima i kao najbolji metod pokazao se uklanjanje produkta iz reakcione smeše (Torres i sar. 2010). Nakanishi i sar. (1983) su ispitali beta-galaktozu iz *Bacillus circulans*-a i utvrdili da je $K_m = 41,7\text{mM}$ prilikom razgradnje laktoze.

2.9.9. Beta-galaktozidaza *Streptococcus thermophilus*-a

Obzirom da pripada subfamiliji GH-2 i da ima visok stepen homologije sa beta-galaktozidazom *E. coli*, beta-galaktozidaza *S. thermophilus*-a je takođe tetramerni protein sastavljen od 4 identične subjedinice. Svaka subjedinica se sastoji od 1026 amino kiselina i ima molekularnu masu 117,04 kDa (Rhim i sar. 2010).

Prečišćen protein beta-galaktozidaze *S. thermophilus*-a je pokazao specifičnu aktivnost od $V_{\max} = 464$ i 126 U/mg sa OPNG i laktozom kao supstratima. Enzim je imao izvanrednu aktivnost razgradnje laktoze sa brzinom od $k_{\text{cat}} = 234\text{ s}^{-1}$, dok je K_m vrednost iznosila $1,1\text{ mM}$. Ispitivanja optimalne temperature za aktivnost enzima pokazala su široku temperaturnu valencu (24°C do 40°C) u okviru koje je enzim pokazao maksimalnu aktivnost. Takođe, enzim je pokazao visoke vrednosti aktivnosti od čak 80% na temperaturi od 5°C , dok su visoke temperature izraženije inhibirale njegovu aktivnost (75% na 65°C). Široka temperaturna valenca se poklapa sa nekim beta-galaktozidazama, dok se od nekih razlikuje (Cieslinski i sar. 2005; Nakagawa i sar. 2006). Ova osobina čini beta-galaktozidazu *S. thermophilus*-a veoma interesantnom za primenu u prehrambenoj industriji, gde se u različitim tehnološkim procesima primenjuju različite temperature. Analiza uticaja pH na aktivnost enzima pokazala je da je enzim najaktivniji na $\text{pH} = 6,5 - 7,5$, što enzim čini primenjivim u tehnologiji mleka, imajući u vidu da je pH mleka oko 6,7. Aktivnost beta-galaktozidaze iz *S. thermophilus*-a u značajnoj meri zavisi

od prisustva Mn^{2+} i Mg^{2+} jona. U odsustvu navedenih jona enzim gubi aktivnost za 48%, dok se njihovim ponovnim uvođenjem vraća puna aktivnost enzima (Rhim i sar. 2010). Joni Ni^{2+} i Cu^{2+} pokazali su inhibitornu aktivnost prema beta-galaktozi *S. thermophilus*-a (Rao i Dutta 1981; Rhimi i sar. 2010).

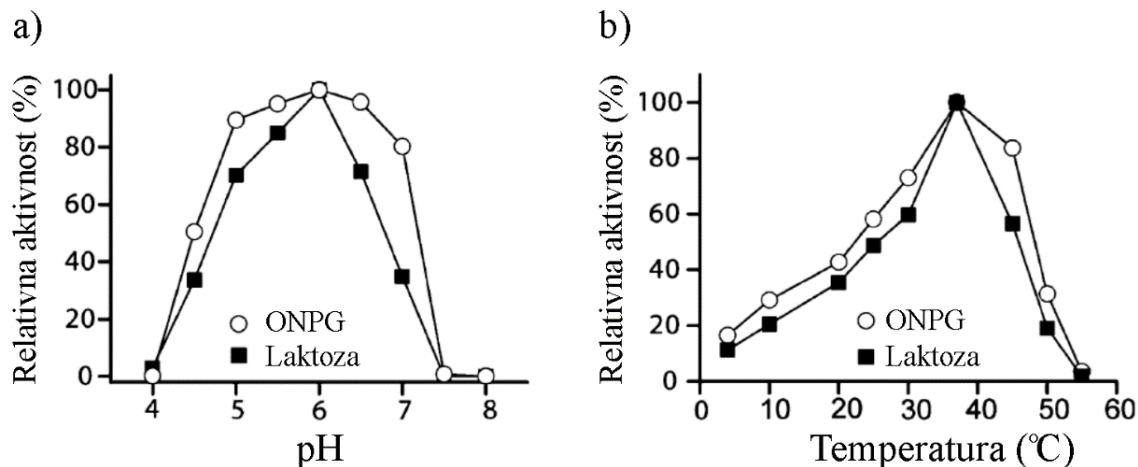
2.9.10. Beta-galaktozidaza *Lactobacillus acidophilus*-a

Lactobacillus acidophilus je probiotska bakterija koja se može naći u ljudskom tankom crevu (intestinum), gde nastanjuje kako lumen tako i zidove creva (Itsaranuwat i sar. 2003). *L. acidophilus* u crevnom traktu između ostalog razgrađuje laktozu čime poboljšava njeno varenje, posebno kod ljudi sa niskom aktivnošću endogene beta-galaktozidaze.

Kompletnim sekvencioniranjem genoma ustanovljeno je da postoje dva različita gena koja kodiraju beta-galaktozidaze kod *L. acidophilus*-a – LacZ i LacML (Altermann i sar. 2005; Nguyen i sar. 2007; Pan i sar. 2010). Njihovi enzimi se razlikuju kako po sekvenci i strukturi, tako i po aktivnosti. Beta-galaktozidaze *L. acidophilus*-a pripadaju subfamilijama GH-42 (LacZ) i GH2 (LacML). Ispitivanja su pokazala da se oba gena pojačano ekspresuju u prisustvu laktoze, što je ukazalo da su oba aktivna kao beta-galaktozidaze.

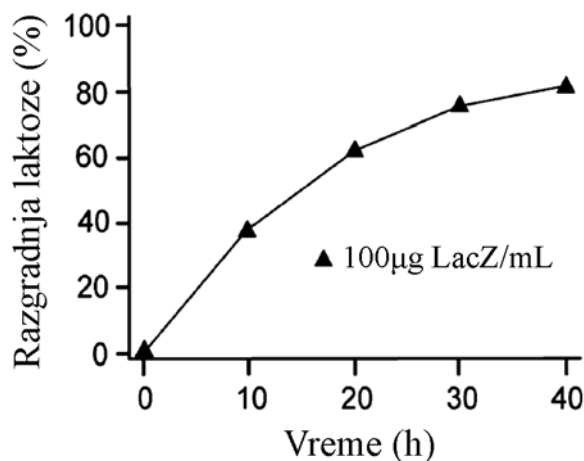
Beta-galaktozidazu *L. acidophilus*-a (kodirana LacZ genom) koja pripada subfamiliji GH-42 čini sekvenca od 667 amino kiselina koje grade protein sa molekulskom masom od 76,5 kDa (Barrangou i sar. 2006).

Optimalna pH vrednost LacZ beta-galaktozidaze je oko 6 za oba supstrata (laktoza i ONPG). Međutim, promene pH vrednosti su značajnije uticale na hidrolizu laktoze nego ONPG-a (Slika 50a). Na slici 50b može se videti da je optimalna temperatura za oba supstrata oko 37°C, što je očekivano obzirom da je *L. acidophilus* probiotska bakterija koja živi u ljudskim crevima čija je temperatura približno 37°C. Na temperaturi od 4°C enzim zadržava 11,2% aktivnosti pri razgradnji laktoze i 16,5% aktivnosti pri razgradnji ONPG-a.



Slika 50. Uticaj pH i temperature na aktivnost LacZ beta-galaktozidaze *L. acidophilus-a*

Ispitivanja stabilnosti enzima pokazala su da enzim na optimalnoj temperaturi gubi 50% aktivnosti nakon jednog sata. Na 45°C aktivnost pada na 50% nakon 10 min, dok se prekida nakon 40 minuta. Na temperaturama od 10°C do 25°C enzim je pokazao veliku stabilnost kada zadržava 95-98% aktivnosti. Termostabilnost enzima je kao kod vrsta prilagođenih hladnim uslovima kao što su *Carnobacterium piscicola* BA (BgaB) i *Planococcus* sp. L4 (BgaP) (Coomb's i Brenchley 1999; Hu i sar. 2007). Kinetička ispitivanja LacZ beta-galaktozidaze pokazala su da je vrednost $K_m = 3,84$ mM za ONPG, dok je za laktozu $K_m = 88,9$ mM. Aktivnost enzima je ispitana i na razgradnji laktoze u mleku (10°C 4,32U enzima, 1mL mleka) (slika 51). U roku od 30 min hidrolizovano je 73% laktoze.



Slika 51. Razgradnja laktoze u mleku LacZ beta-galaktozidazom *L. acidophilus-a*

Strukturu beta-galaktozidaze LacML *L. acidophilus*-a koja pripada subfamiji GH-2 čini heterodimer sa molekulskom masom od oko 105 kDa (Nguyen i sar. 2007). SDS-PAGE elektroforezom je utvrđeno da velika subjedinića ima oko 72 kDa, dok je mala oko 35 kDa. Poređenjem sekvenci različitih sojeva *L. acidophilus*-a – NCFM, JCM1229 i R22, pokazalo je da je velika subjedinića (LacL gen) 100% identična i sastoji se od 628 rezidua, dok je mala subjedinića (LacM gen) pokazala sličnost od 99% (što je posledica varijabilnosti rezidua 302) i sastoji se od 316 rezidua.

Merenjem aktivnosti LacML beta-galaktozidaze *L. acidophilus*-a (povećavanjem koncentracije laktoze) utvrđeno je da je za razgradnju laktoze $V_{\max} = 28,8 \text{ U/mg}$, dok je $K_m = 4,04 \text{ mM}$, što je značajno veći afinitet u odnosu na LacZ enzim ($K_m = 88,9 \text{ mM}$). Enzim je pokazao brzinu razgradnje laktoze od $50,4 \text{ s}^{-1}$, dok je ta vrednost za ONPG vrednost 632 s^{-1} . Katalitička efikasnost za laktozu je iznosila $12,5 \text{ mM}^{-1}\text{s}^{-1}$, dok je za ONPG bila $865 \text{ mM}^{-1}\text{s}^{-1}$, što ukazuje da LacML beta-galaktozidaza *L. acidophilus*-a preferira ONPG. Optimalna pH vrednost se kretala od 6,5 do 8 za oba supstrata. Najveća stabilnost aktivnosti enzima zabeležena je na pH 6,5, gde je enzim zadržavao 80% svoje aktivnosti tokom inkubacije od 48 sati na temperaturi od 37°C . Kao optimalna za kratkoročnu aktivnost (10 min) pokazala se temperatura od 55°C za oba supstrata. Optimalne vrednosti temperature LacML enzima su više u poređenju sa LacZ enzimom (37°C). Za optimalnu aktivnost enzima potrebni su i Mg^{2+} joni, u čijem odsustvu enzim brzo gubi aktivnost. Mn^{2+} , Cu^{2+} i Zn^{2+} joni su u koncentracijama od 10 mM potpuno inhibirali aktivnost enzima. Najveća stabilnost enzima zabeležena je na temperaturi od 4°C . LacML enzim je u potpunosti povratio aktivnost nakon skladištenja od 30 dana i više od 90% aktivnosti nakon skladištenja od 45 dana na 4°C . Pored razgradnje laktoze LacML beta-galaktozidaza *L. acidophilus*-a sposobna je i da sintetiše GOS (Nguyen i sar. 2007).

Za ovu doktorsku disertaciju od značaja je beta-galaktozidaza kodirana LacZ genom, obzirom da su produkti LacZ gena ispitani i kod drugih vrsta.

2.9.11. Beta-galaktozidaza *Bifidobacterium animalis* ssp. *lactis*

B. animalis ssp. *lactis* je probiotska bakterija koja je originalno izolovana iz fermentisanog mleka (Meile i sar. 1997). Usled optimalnih tehnoloških karakteristika, dokumentovanih probiotskih osobina, pozitivnog uticaja na imuni sistem, kao i pozitivnog uticaja

na digestivni trakt, našla je široku primenu u medicini i idustriji (Guyonnet i sar. 2009; Rizzardini i sar. 2012; Loquasto i sar. 2013).

Genom *B. animalis* ssp. *lactis* kodira dve beta-galaktozidaze iz subfamilije GH-42 (kodirane genima BIF_00799 i BIF_01474) i jednu iz subfamilije GH-2 (kodiranu genom BIF_01526) (Barrangou i sar. 2009). Jedna od beta-galaktozidaza iz subfamilije GH-42 se ekspresuje u prisustvu GOS-a (Andersen i sar. 2011).

2.9.12. Beta-galaktozidaza *Bacteroides thetaiotaomicron*-a

Bacteroides thetaiotaomicron je gram negativna anaerobna, simbiotski komensalna bakterija koja nastanjuje ljudski crevni trakt. Ona čini 6% od ukupnih bakterija i 12% od bakterija roda *Bacteroides* u ljudskim crevima. *Bacteroides thetaiotaomicron* vrši važnu funkcionalnu ulogu u crevima kroz metaboličke puteve koji nisu kodirani ljudskim genomom (Sonnenburg i sar. 2005; Gill i sar. 2006). Poseban značaj imaju putevi koji metabolišu ugljene hidrate, kao na primer laktozu (Henrissat i Bairoch 1996). Sa druge strane, *Bacteroides thetaiotaomicron* je fakultativni patogen sa vrlo ozbiljnim posledicama. Kao patogen se najčešće javlja u postoperativnim infekcijama trbušne duplje (Wexler 2007). Genom *Bacteroides thetaiotaomicron*-a se sastoji od 6,26 Mb i kodira 4779 gena (Xu i sar. 2003). Njeni geni kodiraju 226 glikozid hidrolaza i 15 polisaharid liaza. Najispitivaniji proteini su glikozidaze, obzirom na njihov veliki značaj za metabolizam glukooligosaharida (Wang i sar. 2012; Guillotin i sar. 2014). Veliki značaj *Bacteroides thetaiotaomicron*-a ogleda se u doprinosu postnatalnog razvoja crevne flore putem horizontalnog transfera gena (Wang i sar. 2000). U pogledu tehnologije mleka njena najznačajnija karakteristika je razgradnja laktoze beta-galaktozidazom, čime se poboljšava varenje u ljudskim crevima. Pripada subfamiliji GH-35, i o njenoj strukturi, funkciji i mehanizmu delovanja nema dovoljno podataka u naučnoj literaturi, iako je njena struktura eksperimentalno određena (PDB ID: 3DEC) i sekvenca dostupna u bazama podataka (NCBI ID: AAO78285.1) (Cantarel i sar. 2009).

Kristalna struktura beta-galaktozidaze *B. thetaiotaomicron*-a je utvrdila da se radi o homodimeru sastavljenom iz dve identične subjedinice od po 1024 amino kiseline. Svaki monomer se sastoji od 3 domena. Obzirom da pripada subfamiliji GH-35 poseduje TIM barel katalitički domen sa dve aktivne glutaminske kiseline, jednom kao proton donorom i drugom kao nukleofilom. Pripada neinvertujućim glikozidazama i ne dolazi do promene anomerne

konfiguracije oslobođene galaktoze. Domeni 1 i 3 su domeni sa izgledom rolata. U enzimu je prisutan i K⁺jon vezan za enzim kao ligand preko rezidua 926, 968, 969 i 971. Iako ne postoje rezultati vezani za aktivnost beta-galaktozidaze *B. thetaiotaomicron*-a, na osnovu sličnosti sa ostalim članovima GH-35 kojima su eksperimentalno određene aktivne rezidue mogu se napraviti određena predviđanja.

3. Materijal i metodi

3.1. Sekvence

Koordinate trodimenzionalnog modela strukture beta-Dlaktoze (CID 6134) i D-galaktoze (CID: 6036) preuzete su iz PubChem baze podataka (<https://pubchem.ncbi.nlm.nih.gov>). Za istraživanje su odabrane beta-galaktozidaze sledećih sojeva mikroorganizama: *E. coli* (Protein Data Bank (PDB) ID: 1DP0:A), *Thermus thermophilus* (PDB ID: 1KWG, GenBank: CAB07810.1), *Arthobacter* sp. (PDB ID: 1YQ2, GenBank: CAD29775.1), *Kluyveromyces lactis* (PDB ID: 3OBA:A), *Bacillus circulans* ssp. *alkalophilus* (PDB ID: 3TTS:A), *Streptococcus thermophilus* (GenBank: ACA96932.1), *Lactobacillus acidophilus* (GenBank: ACC38288.1), *Bifidobacterium animalis* ssp. *lactis* BB12 (NCBI Reference Sequence: WP_004268781) *B. thetaiotaomicron* (UniProt: Q8AB22. PDBID: 3DEC), *Aspergillus oryzae* Q2UCU3, *Penicillium* sp. (UniProt: Q700S9), *Trichoderma reesei*, (UniProt: Q70SY0), *Homo sapiens* (UniProt: P16278). Koordinate svih trodimenzionalnih struktura beta galaktozidaza preuzete su iz PDB baze podataka. Obzirom da je istraživanje u ovoj disertaciji fokusirano na početne faze vezivanja laktoze za beta-galaktozidazu, odabrane su kristalne strukture enzima bez supstrata ili produkta u aktivnom mestu.

3.2. Homologo modelovanje

Homologi modeli su primenom programa Modeller v. 9.11 (Eswar i sar. 2007) konstruisani za beta-galaktozidaze čija trodimenzionalna struktura nije do sada ekperimentalno određena: *S. thermophilus*, *L. acidophilus* i *B. animalis* ssp. *lactis*. Aminokiselinske sekvence za navedene beta-galaktozidaze su preuzete iz National Center for Biotechnology Information

(NCBI) baze podataka. Izvršeno je pretraživanje PDB baze podataka BlastP programom u potrazi za kristalnim strukturama čije aminokiselinske sekvence imaju dovoljan procenat sličnosti (>30%) da bi mogle služiti kao šablon sekvence za pravljenje homologih modela. Nakon odabira najpogodnijih sekvenci, identifikovane željene i šablon sekvence korišćene su za konstrukciju homologih modela u Modeller programu. Najbolji modeli su odabrani na osnovu najmanje vrednosti optimizovane energije molekula (DOPE vrednost) (Shen i Sali 2006). Konačne trodimenzionalne strukture verifikovane konstrukcijom Ramačandranovih plotova i primenom web servera SAVES (Structural Analysis and Verification Server) (<http://services.mbi.ucla.edu/SAVES/>).

3.3. Klaster analiza

Klaster analiza je urađena metodom prisajedinjenja suseda (neighbor joining) primenom programa UGENE 1.11.3 (Okonechnikov i sar. 2012). Kao filter korišćena je sličnost sekvenci od 30%. Poravnanje sekvenci je urađeno za sve klastere primenom programa MUSCULE (Edgar 2004). Visokokonzervativni regioni i potencijalna aktivna mesta identifikovana su poravnanjem sekvenci.

3.4. Minimizacija energije

Energije preuzetih trodimenzionalnih struktura proteina i liganada su minimizovane primenom MMFF94 energetskeg polja, Powell metoda, distance dependent dielectric constant i convergence gradient metoda sa kriterijumom konvergencije od 0,005 kcal/mol. Atomska naelektrisanja su dodata primenom MMFF94 metoda. Modelovanim strukturama proteina je minimizovana energija primenom AMBER7 energetskeg polja i AMBER atomskog naelektrisanja. Prilikom minimizacije energije modelovanih proteina primenjeni su sledeći parametri: Powell metod, distance dependent dielectric constant i convergence gradient metod sa kriterijumom konvergencije od 0,005 kcal/mol. Sve proteinske strukture su pripremljene za molekularnu doking simulaciju dodavanjem vodonikovih atoma, podešavanjem protonacije, popravljanjem krajeva lanaca i dodavanjem parcijalnih naelektrisanja primenom MMFF94 metoda.

3.5. Simulacija dokinga

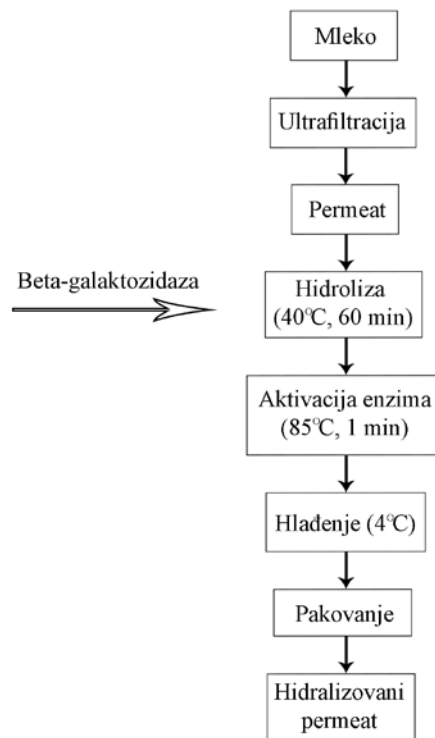
Simulacije dokinga su izvođene primenom Surfex-Dock programa sa omogućenom fleksibilnosti vodonikovih atoma (Jain 2003). Za doking su korišćeni monomeri proteina obzirom da isključivo rezidue pojedinačnih monomera interaguju direktno sa supstratom. Takođe, iako beta-galaktozidaze iz obe subfamijile poseduju metalne jone u aktivnim mestima, oni ne utiču na vezivanje supstrata, usled čega nisu uključeni u simulaciju dokinga. Za određivanje aktivnog mesta primenjen je Residues mode sa odabirom rezidua na osnovu dostupne literature. Za doking su korišćene četiri dodatne početne konformacije po molekulu, sa maksimalnih 20 konformacija po fragmentu, maksimalnih 20 poza po ligandu i minimalnim RMSD od 0,5Å između krajnjih poza liganda. Prilikom dokinga dozvoljena je fragmentacija molekula.

3.6. Permeat

Permeat je dobijen tokom procesa ultrafiltracije mleka kada se laktoza, minerali i vitamini odvajaju od proteina koji se koriste za proizvodnju feta sira. Permeat je u proizvodnji feta sira nusproizvod i kao takav predstavlja potencijalnu sirovinu u daljim tehnološkim postupcima. Permeat ne dobijen ultrafiltracijom mleka sa 3,7% mlečne masti (postrojenje mlekare Šabac, Srbija). Prilikom proizvodnje korišćena je ultrafiltraciona jedinica (DDS Danska) sapolisulfonskom membranom (hollow fibre modul) kapaciteta od 5000L mleka/h.

3.7. Enzim i hidroliza laktoze

Za hidrolizu laktoze korišćena je beta-galaktozidaza dobijena iz kvasca *Kluyveromyces lactis* – Maxilact® LG5000 (DSM Food Specialties, Holandija). Pripremljen enzim je dodat u permeat na temperaturi od 40°C u koncentraciji od 0,1%, 0,3% i 0,5%. Uzorci za merenje stepena razgradnje laktoze uzimani su tokom 60 minuta u intervalima od 10 minuta. Na slici 52 prikazan je tehnološki proces hidrolize laktoze u permeatu mleka.



Slika 52. Tehnološki proces hidrolize laktoze u permeatu mleka

3.8. Fizičko-hemijske analize

Standardnim analitičkim metodama ispitani su sledeći fizičko-hemijski parametri mleka, permeata i hidrolizovanog permeata (Carić i sar. 2000):

- pH vrednost (pH-metar Consort C830. Belgija)
- mlečna mast (po Gerberu)
- suva materija (sušenjem u sušnici)
- pepeo (spaljivanjem)
- ukupni proteini (metod po Kjeldahl-u)

Sadržaj šećera je analiziran primenom HPLC metode (Liquid Chromatograph Agilent Technologies 1200 Series), sa ugljenohidratnom kolonom (Zorbax Carbohydrate Column 4,6 x 250mm, 5 μ m) i ELSD (Evaporative Light Scattering Detector) detektorom (Agilent Technologies). Od svakog uzorka mleka 5g je rastvoreno sa 10 mL destilovane vode u

volumetrijskom sudu od 25mL. Rastvori su inkubirani u vodenom kupatilu na 50°C u periodu od 15 min. Nakon hlađenja dodato je 0,5 mL Karez I, 0,5 mL Karez II rastvora i 1 mL 100mM NaOH. Sodovi su dopunjeni sa destilovanom vodom, izmućkani i uzorci su filtrirani kroz filter papir br 381. Po 10 µL uroka je injektovano u HPLC sistem primenom autosampler-a. Protok je podešen na 1,000 mL/min, na sobnoj temperaturi u toku 15 min. Kao mobilna faza korišćen je acetonitril/voda u odnosu 70/30 v/v. ELSD parametri su bili: temperatura 40±1°C, pritisak azota: 4,5±0,1 bara.

3.9. Statistička analiza

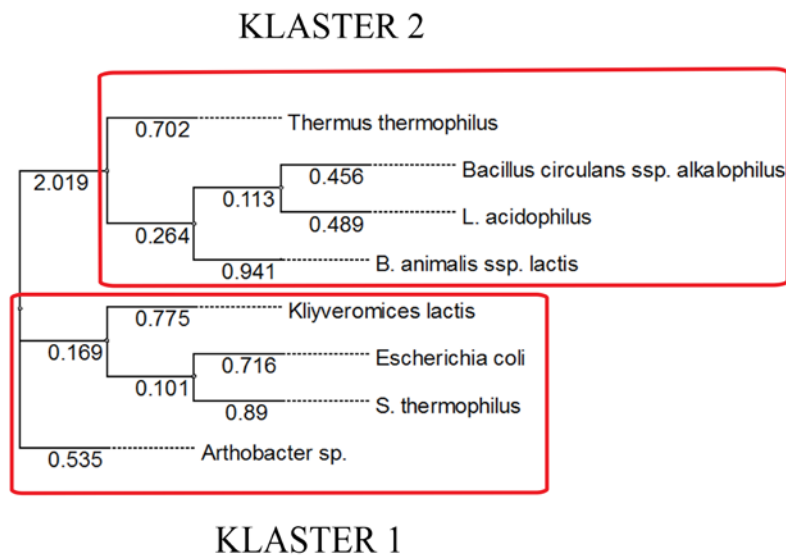
Svi eksperimenti su rađeni u tri ponavljanja i vrednosti su iskazane kao srednja vrednost. Statistička analiza i grafički prikaz rezultata rađeni su u softverskom paketu Origin 6.1 i Excel 2007. Eksperimentalni podaci su fitovani u empirijski model primenom Origin-a. Model koji najbolje opisuje sistem je exp. Dec.2 funkcija: $S(t) = S_0 + A_1 e^{-x/t^1} + A_2 e^{-x/t^2}$, gde je S koncentracija laktoze (g/100g) koja se menja u vremenu t, dok su parametri A₁ i A₂ vrednosti asimptota krive S(t).

4. Rezultati i diskusija

4.1. Klaster i filogenetska analiza beta-galaktozidaza koje pripadaju subfamilijama GH-2 i GH-42

Klaster analiza na osnovu aminokiselinskih sekvenci za beta-galaktozidaze sprovedena je sa osam sojeva mikroorganizama (*E. coli*, *Thermus thermophilus*, *Arthobacter* sp., *Kluyveromyces lactis*, *Bacillus circulans* ssp. *alkalophilus*, *Streptococcus thermophilus*, *Lactobacillus acidophilus* and *Bifidobacterium animalis* ssp. *lactis* BB12). Klaster analiza je vršena za sojeve čije beta-galaktozidaze želimo da istražimo, kao i za pripadnike njihovih subfamilija čija trodimenzionalna struktura je ekperimentalno određena primenom kristalografije rendgenskim zracima. Na osnovu sličnosti, analizirane sekvence se mogu podeliti u dva klastera (slika 53). Ovakva podela je očekivana obzirom da članovi istih klasera pripadaju istoj subfamiliji glukozid

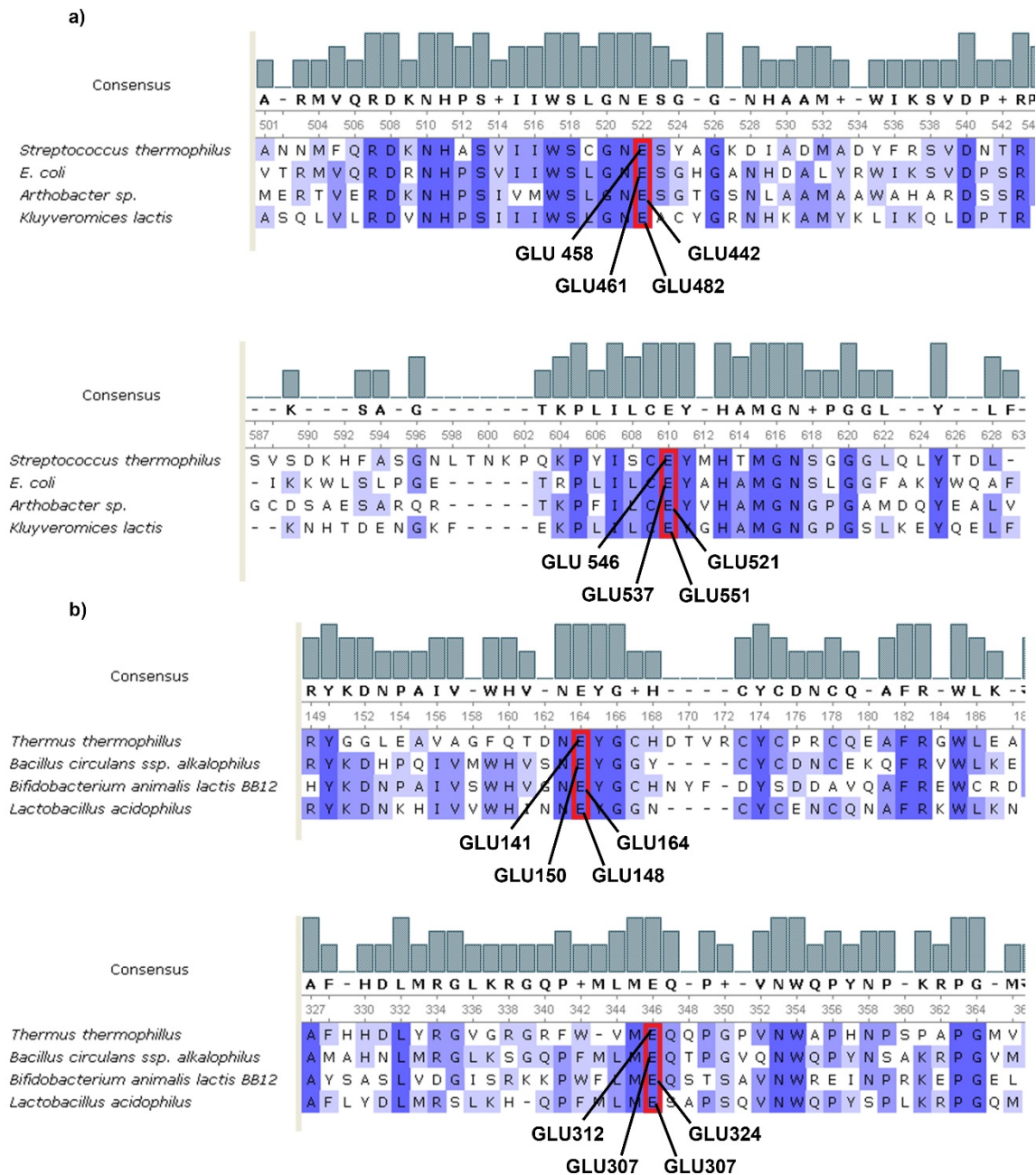
hidrolaza. Članovi klastera 1 pripadaju GH-2 subfamiliji, dok članovi klastera 2 pripadaju GH-42 subfamiliji.



Slika 53. Klaster analiza odabranih beta-galactozidaza na osnovi sličnosti aminokiselinskih sekvenci; klaster 1 – GH-2 subfamilija, klaster 2 – GH-42 subfamilija

Poravnanje sekvenci predstavlja prvi korak u određivanju aktivnih rezidua kod ispitivanih sekvenci. Eksperimentalno utvrđene aktivne rezidue beta-galaktosidaza *E. coli* (GLU461 i GLU537) i *Bacillus circulans* ssp. *alcalophilus* (GLU150 i GLU307) mogu se iskoristiti za određivanje aktivnih rezidua ispitivanih sekvenci (Gebler i sar. 1992; Hidaka i sar. 2002). Poravnanjem sekvenci dobijene su potencijalne aktivne rezidue (slika 50). U klaster 1 spadaju beta-galaktosidaze iz *S. thermophilus*, *E. coli*, *Arthobacter* sp. i *K. lactis*. Njihova približna ukupna sličnost sekvenci iznosi 35-40%, s tim što je značano veća za rezidue koje grade aktivna mesta (slika 50a). Klaster 2 uključuje beta-galaktosidaze iz *L. acidophilus*, *B. animalis* ssp. *lactis* BB12 i *B. circulans*, sa ukupnom sličnosti sekvenci od 30-46% (slika 54b). Kod klastera 2 takođe zabeležena značajno veća sličnost sekvenci u regionima koji grade aktivna mesta. Kao što je poznato, sve do sada ispitane beta-galaktosidaze imaju dve visokokonzervativne rezidue glutamata koje deluju kao aktivne rezidue prilikom hidrolize supstrata. One tokom razgradnje supstrata vrše funkciju proton donora i nukleofila. Na osnovu klaster analize i poravnanja sekvenci katalitički aktivne rezidue beta-galaktosidaza ispitivanih sojeva mogle bi biti

GLU458/GLU546 kod *S. thermophilus*-a, GLU148/GLU307 kod *L. acidophilus*-a i GLU164/GLU324 kod *B. animalis* ssp. *lactis*-a.



Slika 54. Poravnanje sekvenci ispitivanih beta-galaktosidaza u okolini potencijalnih aktivnih mesta a) klaster 1; b) klaster 2.

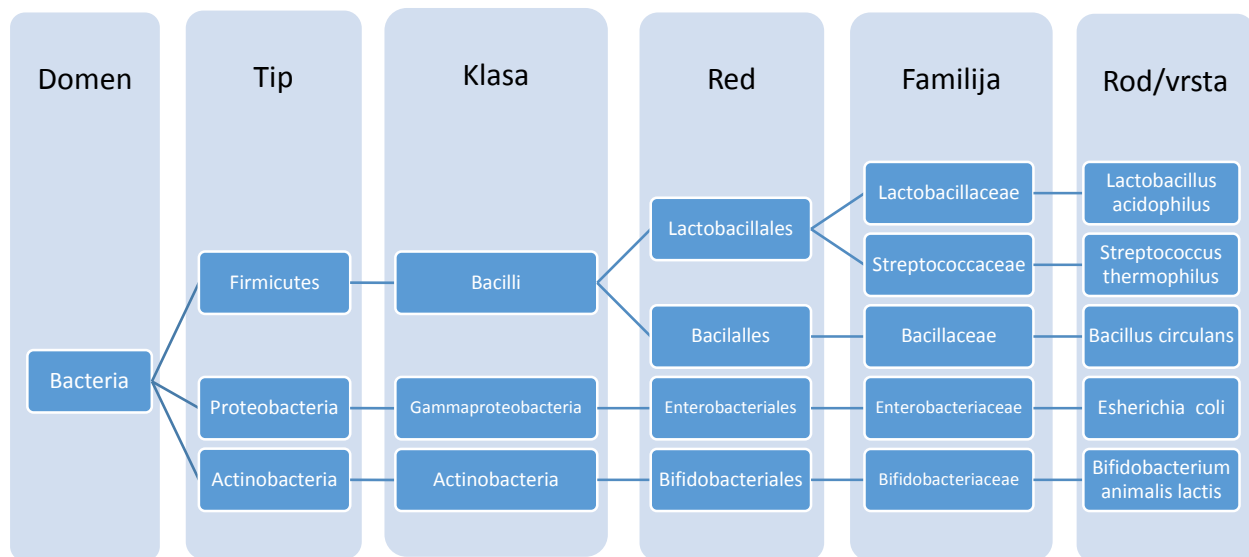
Pretraživanje BlastP programom, pokazalo je da beta-galaktozidaze *S. thermophilus*, *L. acidophilus* i *B. animalis* ssp. *lactis* imaju zadovoljavajući stepen sličnosti sa nekoliko beta-galaktozidaza čija je struktura eksperimentalno određena kristalografijom rendgenskim zracima, da bi se njihova struktura mogla predvideti homologim modelovanjem. BlastP programom je vršeno pretraživanje PDB baze podataka (tabela 6). E vrednost predstavlja pokazatelj stepena sličnosti, na osnovu kojeg se može zaključiti da li su sekvence homologe. Što je vrednost parametra E manja, to je veća verovatnoća da su sekvence homologe. Za vrednosti $E < 1e^{-50}$ uzima se sa sigurnošću da su sekvence homologe. Za vrednosti $1e^{-50} < E < 0,01$ sekvence se smatraju homologim. Ukoliko je vrednost E od 0,01 do 10 poklapanje nije značajno, ali je moguća daleka homologija, dok se za vrednosti E preko 10 smatra da je dobijena vrednost rezultat slučajnosti i da nije moguće detektovati srodnost sekvenci datom metodom. Dobijene E vrednosti potvrđuju homologiju između ispitivanih sekvenci. Da bi homologni model koji se želi konstruisati odgovarao realnom stanju u prirodi, odnosno da bi imao dovoljnu pouzdanost za dalja istraživanja, potrebno je da sličnost sa šablon sekvencom bude visoka. Dosadašnja iskustva su pokazala da sličnost sekvenci ispod 20% daje nepouzdan rezultate i da je visoko odstupanje od realnog stanja u prirodi. Sličnost sekvenci preko 30% daje zadovoljavajuće rezultate i model se može smatrati pouzdanim. Vrednosti preko 40% se smatraju visokopouzdanim za homologno modelovanje. Na osnovu vrednosti dobijenih BlastP pretraživanjem, može se zaključiti da dobijeni rezultati spadaju u zadovoljavajuće i visokopouzdan i da bi dobijeni modeli trebalo da odgovaraju realnom stanju u prirodi.

Tabela 6. BlastP rezultati pretraživanja PDB baze podataka

Željena sekvenca	PDB sekvenca	Pokrivenost sekvenci	E vrednost	Sličnost sekvenci
<i>S. thermophilus</i>	1DP0A (<i>Escherichia coli</i>)	98%	$2e^{-163}$	33%
	1YQ2 A (<i>Arthrobacter</i> sp.)	97%	$4e^{-135}$	31%
	3OB8 A (<i>Kluyveromyces lactis</i>)	91%	$5e^{-101}$	30%
<i>L. acidophilus</i>	3TTS A (<i>Bacillus circulans</i>)	97%	0	47%
	1KWG A (<i>Thermus thermophilus</i>)	77%	$2e^{-70}$	30%
<i>B. animalis</i> ssp. <i>lactis</i>	3TTS A (<i>Bacillus circulans</i>)	96%	$2e^{-119}$	33%
	1KWG A (<i>Thermus thermophilus</i>)	93%	$3e^{-65}$	29%

Filogenetska analiza ispitivanih vrsta (slika 55) pokazala je da *S. thermophilus* i *L. acidophilus* pripadaju istom redu, dok *B. animalis* ssp. *lactis* pripada drugom tipu u odnosu na

njih. *Bacillus circulans* čija beta-galaktozidaza ima visoku homologiju sa beta-galaktozidazama iz *L. acidophilus*-ai *B. animalis* ssp. *lactis*-a pripada istom redu kao *L. acidophilus*, dok sa druge strane pripada različitom tipu nego *B. animalis* ssp. *lactis*.



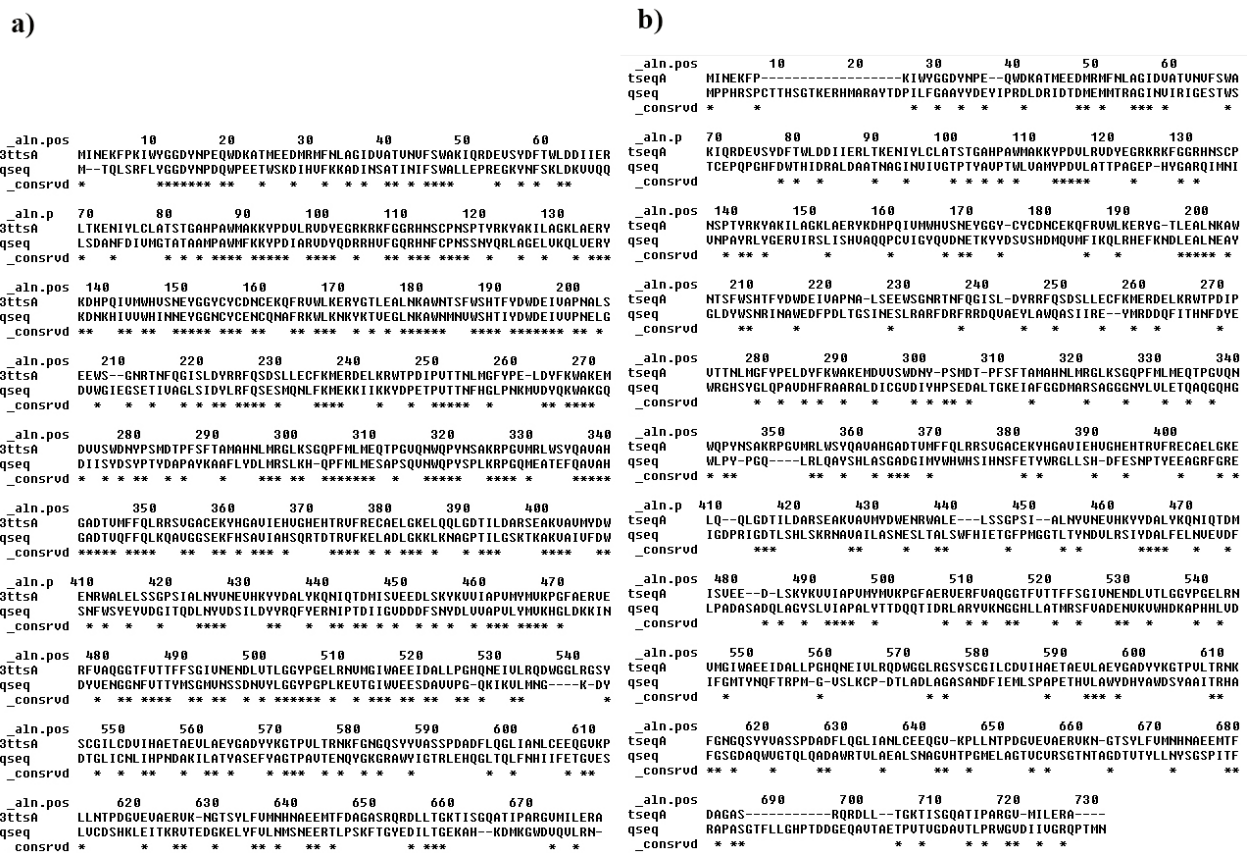
Slika 55. Filogenetska analiza ispitivanih vrsta

Beta-galaktozidaze koje pripadaju klasteru 1 (GH-2) imaju visokokonzervativne regione u okolini aktivnih rezidua, dok su članovi klastera 2 (GH-42) pokazali manju konzervativnost rezidua (slika 54), iako i kod njih rezidue proton donora i nukleofila ostaju visokokonzervativne. *L. acidophilus* i *B. animalis* ssp. *lactis* su probiotske bakterije koje normalno nastanjuju ljudski gastrointestinalni trakt. Uzimajući u obzir filogenetsku analizu (raličiti tipovi), BlastP rezultati (vrlo homologe sekvence) ukazuju na horizontalni tranfer gena tokom evolucije *L. acidophilus*-a i *B. animalis* *lactis*-a. Dalje analize su potrebne da bi se napravila kompletna filogenetska slika. Ipak, dobijeni rezultati ukazuju da evolutivno bliske vrste ne moraju uvek imati visok stepen homologije beta-galaktozidaze, i obrnuto, evolutivno udaljene vrste mogu imati visok stepen homologije sekvenci.

4.2. Konstrukcija modela metodom modelovanja kompletne strukture proteina

Prilikom konstrukcije homologih modela program Modeller pravi poravnanje šablon i ispitivane sekvence na osnovu kojeg uzima koordinate podudarajućih regiona kostura strukture

kao i podudarajućih sekundarnih struktura i konstruiše homologe modele. Na slici 56 prikazana su dobijena poravnanja sekvenci za modelovane beta-galaktozidaze *L. acidophilus* *B. animalis* ssp. *lactis*. Oučljivo je da je beta-galaktozidaza *L. acidophilus* manje promenjena u odnosu na njenu šablon sekvencu, što utiče na preciznost modela koji se konstruiše. Međutim, iako beta-galaktozidaza *B. animalis* ssp. *lactis*-a ima manje konzervativnih rezidua, njihov raspored je duž cele sekvence, što je ključno za dobijanje ispravnog modela (pokrivenost modela je velika, po Blast-u 96%). Konzervativne rezidue pored funkcionalne imaju i ulogu u stabilizaciji strukture enzima, usled čega je važno da budu prisutne u svim njegovim delovima.



Slika 56. Poravnanja sekvenci modelovanih beta-galaktozidaza a) *L. acidophilus*; b) *B. animalis* ssp. *lactis*

Kao rezultat Modeller daje pet potencijalno odgovaraju realnom stanju u prirodi. Od njih je potrebno odabrati odgovarajući koji će se dalje koristiti za analizu. Kao presudni parametar za odabir najboljeg modela uzima se vrednost DOPE scor-a. Što je vrednost

DOPE scor-a manja to je veća pouzdanost modela. U tabeli 7 prikazane su DOPE score vrednosti za konstruisane modele. Odabrani modeli obeleženi su crvenom bojom.

Tabela 7. DOPE score vrednosti konstruisanih modela

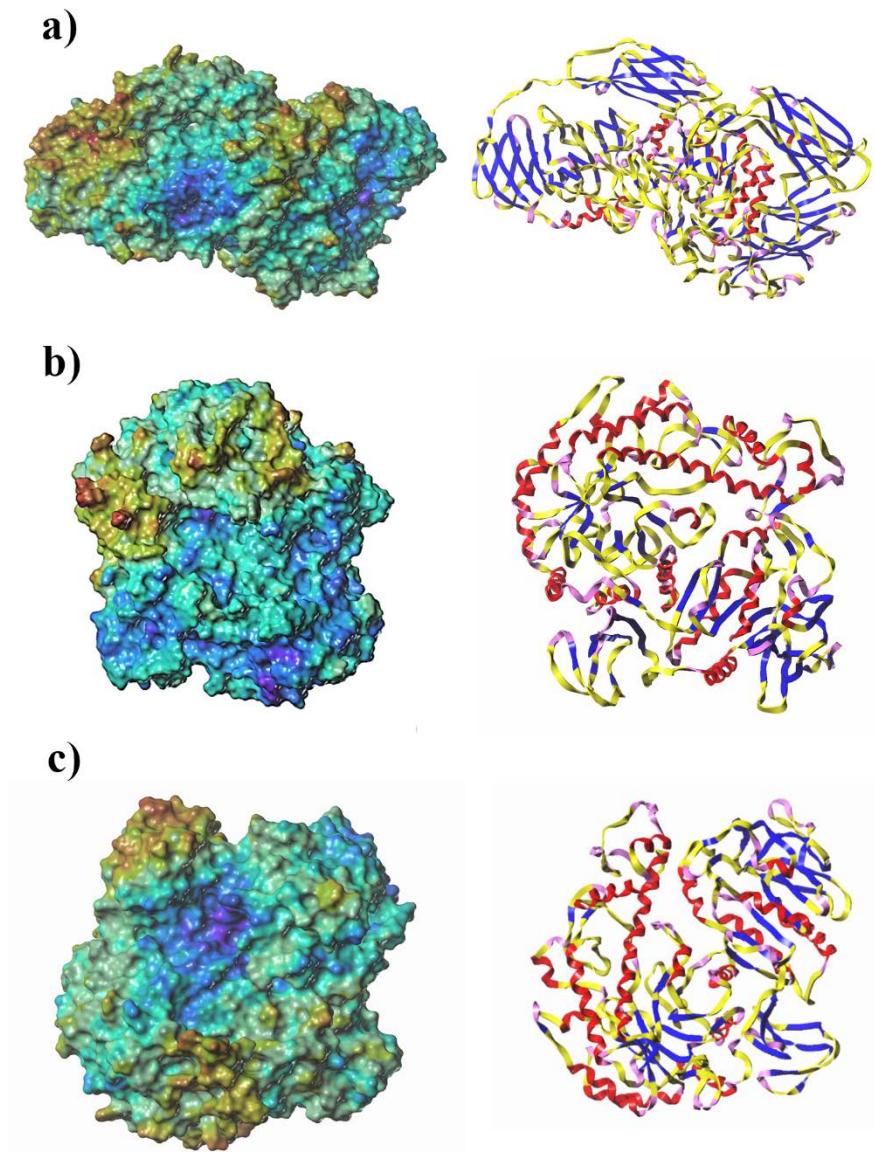
Sekvenca	<i>S. thermophilus</i>	<i>L. acidophilus</i>	<i>B. animalis</i> ssp. <i>lactis</i>
Qseq.B99990001.pdb	-107754,86	-78720,09	-78101,89
Qseq.B99990002.pdb	-107714,64	-79126,94	-78995,02
Qseq.B99990003.pdb	-108335,30	-79394,22	-78605,78
Qseq.B99990004.pdb	-107245,07	-77698,19	-79242,48
Qseq.B99990005.pdb	-107314,00	-78624,20	-78624,20
Vreme CPU (s)	3156,80	1504,92	1496,61

Konstruisani su homologni modeli beta-galaktozidaza *S. thermophilus*, *L. acidophilus*, *B. animalis* ssp. *lactis* (slika 57). Monomer beta-galaktozidaze *S. thermophilus*-a se sastoji od 5 domena: Domen 1 – jelly-roll barell (šećer vezujući domen sa reziduama 42-218), Domeni 2 i 4 su slični fibronektin III domenu (Domen 2 je prepoznat od strane Conserved Domain Database kao fibronektin sličan domen sa reziduama 220-322, dok Domen 4 nije prepoznat, usled malog stepena homologije, oko 110 rezidua), jedan veliki beta sendvič domen (Domen 5, sa reziduama 749-1020) i centralni TIM barel domen koji ima funkcionalnu ulogu u razgradnji supstrata (Domen 3, sa reziduama 324-634).

Beta-galaktozidaze *L. acidophilus*-a i *B. animalis* ssp. *lactis*-a imaju sličnu ukupnu strukturu, što je očekivano obzirom da su konstruisani primenom iste šablon strukture.

Monomer beta-galaktozidaze iz *L. acidophilus*-a i *B. animalis* ssp. *lactis*-a se sastoji iz 3 domena. Domen 1 predstavlja katalitički domen sa TIM barelom i obuhvata rezidue 12-384 *L. acidophilus*-a i 28-402 *B. animalis* ssp. *lactis*-a. Kao i kod beta-galaktozidaze *B. circulans*-a petlje koje povezuju C terminalne krajeve alfa heliksa i N terminalne krajeve beta ploča su kratke čineći N terminalne krajeve TIM barela relativno glatkim. Petlje koje povezuju C terminalne krajeve beta ploča i N terminalne krajeve alfa heliksa su duže i često sadrže elemente sekundarne strukture. Takođe, u njihovoj strukturi prisutne su i petlje za koje se na osnovu homologije sa beta-galaktozidazom *B. circulans*-a može pretpostaviti da imaju važnu ulogu u formiranju kvarterne strukture učestvujući u interakcijama između monomera. Domen 2 (rezidue 399-601 i 417-539) ima strukturu sastavljenu od centralno postavljenih sedam uspravnih beta ploča i pet alfa heliksa. Takođe i za njega se može pretpostaviti da ima strukturnu funkciju formiranjem

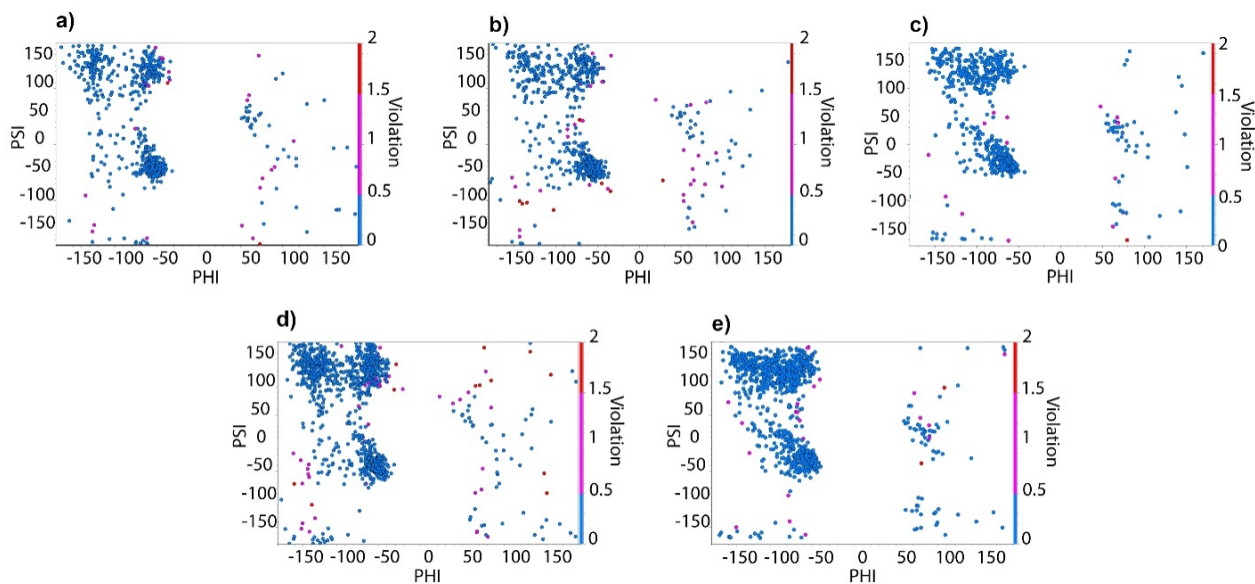
interfejsa sa drugim monomerom stabilišući na taj način enzim. Treći domen (rezidue 610-667) je najmanji i ima antiparalelnu beta-sendvič strukturu od dve beta ploče koje grade strukturu sa izgledom rolata. Pretpostavlja se da je i njegova uloga strukturna.



Slika 57. Struktura modelovanih beta-galaktozidaza: a) *S. thermophilus*, b) *L. acidophilus*, c) *B. animalis ssp. lactis*

4.3. Validacija konstruisanih modela

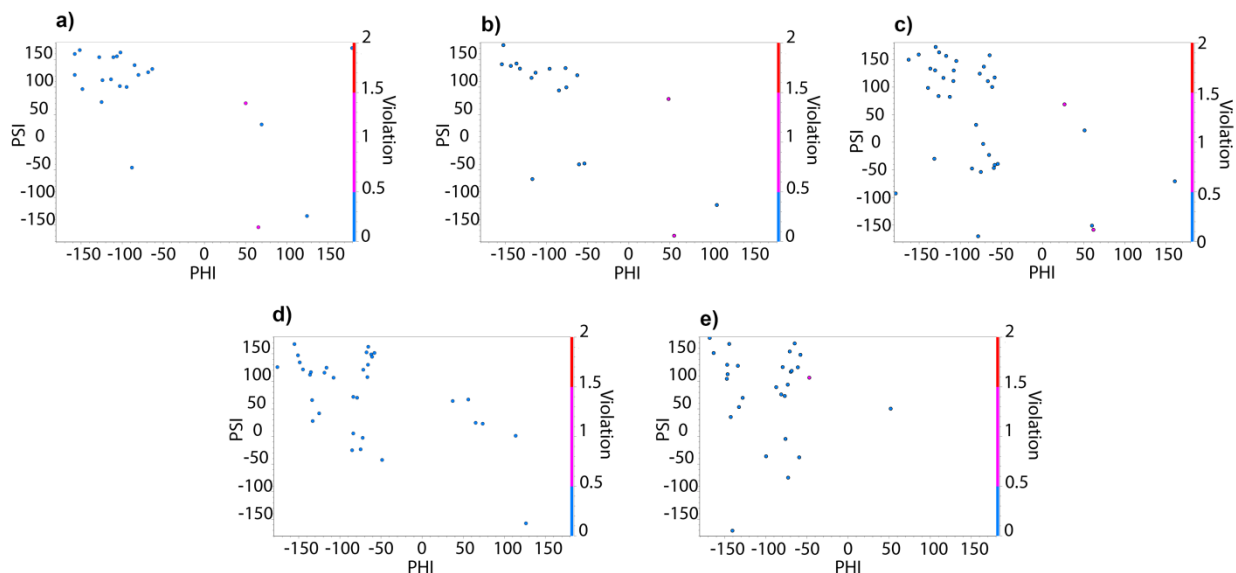
Iako se pomoću DOPE score-a bira model za dalju analizu, potrebno je izvršiti validaciju ispravnosti konstruisanih modela. Za proveru tri modelovana enzima (*S. thermophilus*, *L. acidophilus*, *B. animalis* ssp. *lactis*) konstruisani su Ramačandranovi plotovi (slika 58). Takođe, plotovi su konstruisani i za dve šablon sekvence korišćene za homologo modelovanje. Ramačandranov plot ukazuje na fizičku mogućnost modelovane strukture putem merenja uglova veze između aminokiselina. Ramačandranovi plotovi pokazuju da se kod beta-galaktozidaze *L. acidophilus*-a samo dve aminokiseline nalaze u nedozvoljenom regionu, na osnovu čega se može zaključiti da beta-galaktozidaza iz *L. acidophilus* ima najveći strukturalni kvalitet u poređenju sa ostalim modelovanim enzimima. Ovakvi rezultati za konstruisane modele nisu iznenađujući, imajući u vidu da sekveca beta-galaktozidaze *L. acidophilus*-a ima najveću sličnost (47%) sa svojom šablon sekvencom (*B. circulans*). Druga dva modela (*S. thermophilus*, *B. animalis* ssp. *lactis*) pokazala su nešto slabiji kvalitet (osam nedozvoljenih aminokiselina kod beta-galaktozidaze *B. animalis* ssp. *lactis*-a i 12 kod *S. thermophilus*-a), ali su i kod njih većina rezidua u povoljnim regionima, što ukazuje na ispravno konstruisane modele uzimajući u obzir strukturalnu konformaciju. Šablon strukture imaju *B. circulans* i *E. coli* imaju po dve aminokiseline u nedozvoljenim regionima.



Slika 58. Ramačandranovi plotovi za konstruisane modele i šablon strukture: a) *L. acidophilus*, b) *B. animalis* ssp. *lactis*, c) *B. circulans*, d) *S. thermophilus*, e) *E. coli*

Detaljnija analiza konformacija aktivnih mesta Ramačandranovim plotom pokazala je da sve rezidue koje učestvuju u građenju aktivnog mesta imaju povoljne uglove veza, dok se svega

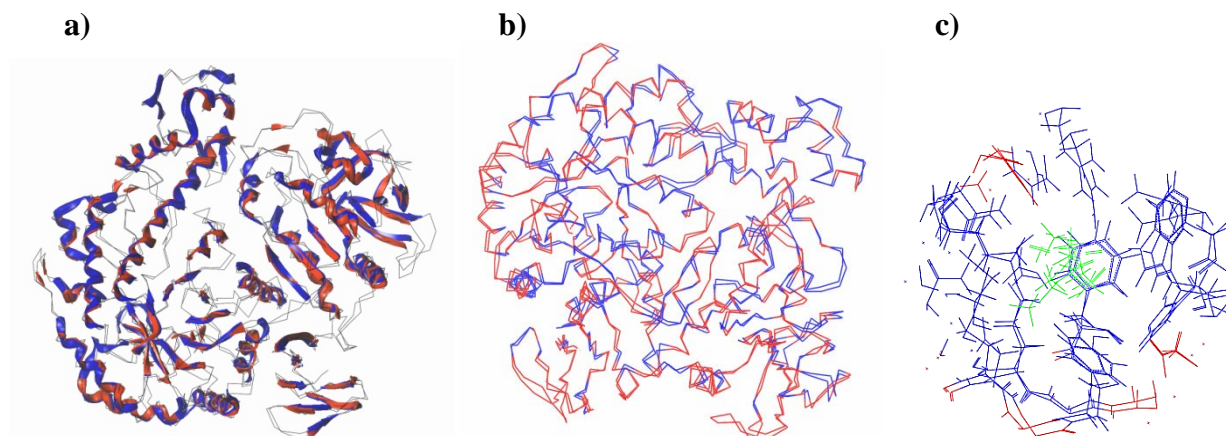
po dve aminokiseline kod beta-galaktozidaza *L. acidophilus*-a i *B. animalis* ssp. *lactis*-a nalaze na graničnim pozicijama u plotu, isto koliko i kod šablon strukture *B. circulans*-a, što je potvrda validnosti modela i od izuzetnog je značaja za njihovu dalju upotrebu za doking simulaciju (slika 59). Sa druge strane sve residue beta-galaktozidaze *S. thermophilus*-a su u povoljnim regionima.



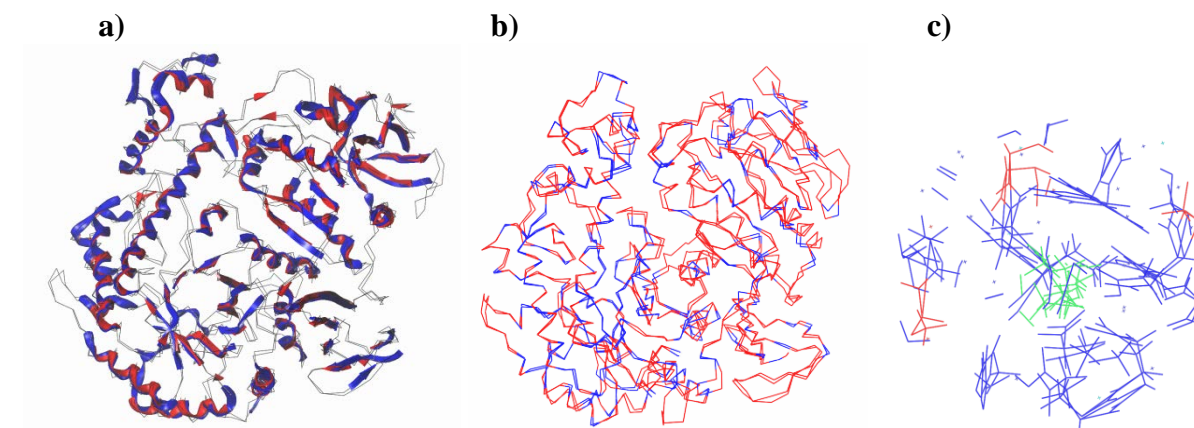
Slika 59. Ramačandranovi plotovi aktivnih mesta konstruisanih modela i šablon struktura: a) *L. acidophilus*, b) *B. animalis* ssp. *lactis*, c) *B. circulans*, d) *S. thermophilus*, e) *E. coli*

Modelovani enzimi su dalje validovani strukturalnim superpozicioniranjem za njihovim šablon strukturama (slika 60, slika 61 i slika 62). Svi modeli su pokazali visok stepen sličnosti trodimenzionalnih struktura sa šablon strukturama. Odstupanja su izražena kao koren kvadrata odstupanja (RMSD – Root Mean Square Deviation). Vrednosti su bile prilično male i iznosile: za *S. thermophilus* 2.6Å za C α trag, dok je za proteinski kostur vrednost bila 2.6Å. Kod *B. animalis* ssp. *lactis*-a odstupanje za C α trag je iznosilo 1.5 Å, dok je za kostur bilo 1.6Å, *L. acidophilus* je imao najmanja odstupanja C α = 1.1 i kostur = 1.2Å. Na slikama b može se videti da i pored velikih razlika u sekvencama proteina (plava boja – podudaranje sekvenci, crvena boja – različite sekvence) ukupna struktura monomera je ostala očuvana. To ukazuje da je evolutivni pritisak veći na strukturalnoj organizaciji proteina, odnosno da je za njegovu funkciju važnija struktura nego aminokiselinska sekvencen. Iz ovog pravila su izuzete pojedine rezide koje imaju ulogu u stabilizaciji strukture enzima, vezivanju supstrata, kao i njegovoj katalizi. Takve rezidue ostaju konzervativne kroz evoluciju enzima. Slike c pokazuju visoku konzervativnost rezidua u regionima

aktivnog mesta, što potvrđuje prethodno iznešenu tvrdnju. Globalno posmatrano, modelovane strukture imaju visoku sličnost sa šablon strukturama. Ipak, postoje promene u reziduama u okolini aktivnog mesta što može uticati na energiju vezivanja supstrata, kao i na pK vrednosti aktivnih rezidua, a time in a aktivnost enzima.

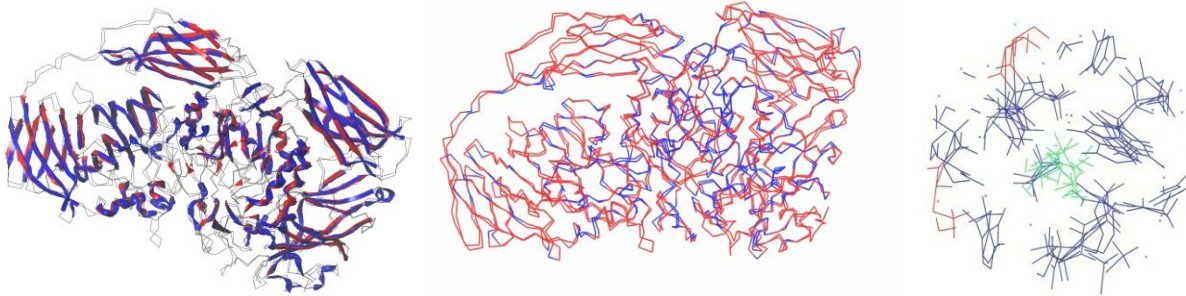


Slika 60. Poravnanje struktura beta-galaktozidaza *B. animalis* ssp. *lactis*-a i *B. circulans*-a: a) prepoklapanje struktura monomera (plavo – šablon sekvenca, crveno – ispitivana sekvenca), b) sličnosti i razlike sekvenci monomera (plava boja – podudaranje sekvenci, crvena boja – različite sekvence), c) sličnosti i razlike sekvenci u okolini aktivnog mesta



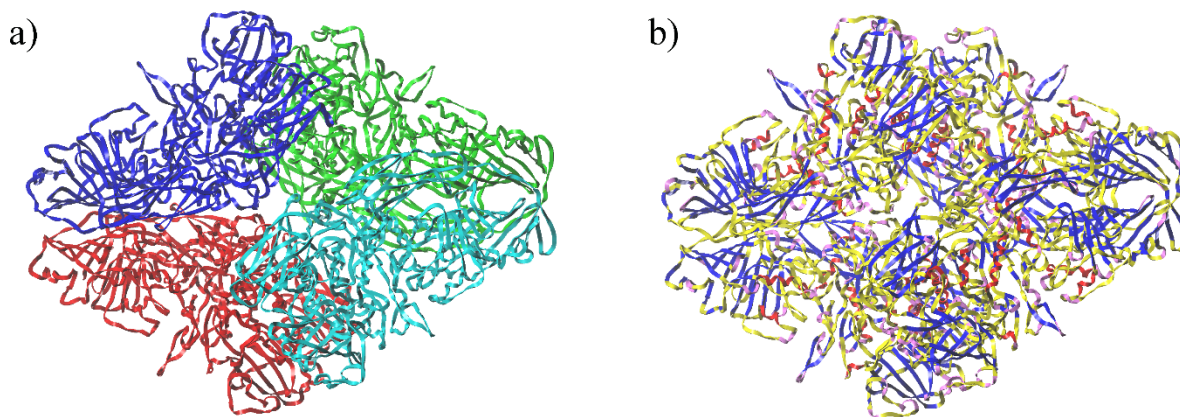
Slika 61. Poravnanje struktura beta-galaktozidaza *L. acidophilus*-a i *B. circulans*-a: a) prepoklapanje struktura monomera (plavo – šablon sekvenca, crveno – ispitivana sekvenca), b) sličnosti i razlike sekvenci monomera (plava boja – podudaranje sekvenci, crvena boja – različite sekvence), c) sličnosti i razlike sekvenci u okolini aktivnog mesta

a) b) c)



Slika 62. Poravnanje struktura beta-galaktozidaza *S. thermophilus*-a i *E. coli*,: a) prepoklapanje struktura monomera (plavo – šablon sekvenca, crveno – ispitivana sekvenca), b) sličnosti i razlike sekvenci monomera (plava boja – podudaranje sekvenci, crvena boja – različite sekvence), c) sličnosti i razlike sekvenci u okolini aktivnog mesta

Utvrđeno je da je beta-galaktozidaza *S. thermophilus*-a tetramerni protein sastavljen od četiri identične subjedinice (Rhim i sar. 2010). Usled toga, moguće je predvideti njegovu kvarternu strukturu na osnovu konstruisanih modela i šablon strukture beta-galaktozidaze *E. coli* (slika 63). Kvarтерна struktura je kao i sekundarna pokazala visok stepen sličnosti sa šablon strukturom.

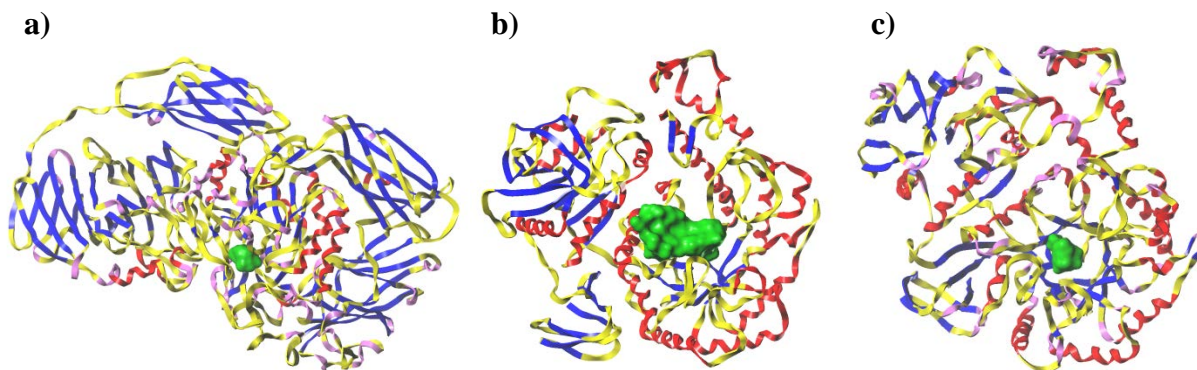


Slika 63. Kvarterna struktura beta-galaktozidaze *S. thermophilus*-a: a) subjedinice tetramera; b) sekundarna struktura tetramera

Broj subjedinica kvarterne strukture beta-galaktozidaza *L. acidophilus*-a i *B. animalis* ssp. *lactis*-a nije eksperimentalno utvrđen, usled čega nije moguće predvideti kvarternu strukturu navedenih beta-galaktozidaza.

4.4. Analiza interakcije enzima sa supstratom (doking analiza)

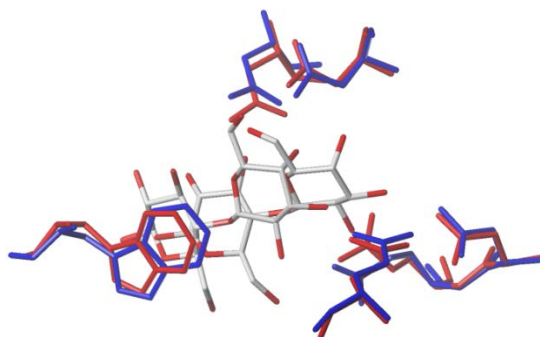
Iako su ispitivane beta-galaktozidaze oligomerne strukture (trimeri i tetrameri, u zavisnosti od subfamilije kojoj pripadaju), simulacije molekularnog dokinga su izvođene sa monomernom strukturom, obzirom da je ustanovljeno da u direktnom vezivanju supstrata učestvuju samo pojedinačni monomeri (McCarter i Stephen Withers 1994). Na slici 64 prikazane su pozicije aktivnih mesta u modelovanim beta-galaktozidazama. Za konstrukciju aktivnih mesta uziman je prostor koji uključuje sve tri potencijalne aktivne rezidue sa dodatnih 5Å prostora, sa ciljem da bi se uključile sve rezidue koje bi mogle učestvovati u vezivanju laktoze.



Slika 64. Pozicije aktivnih mesta u modelovanim beta-galaktozidazama a) *S. thermophilus*; b) *L. acidophilus*; c) *B. animalis ssp. lactis*

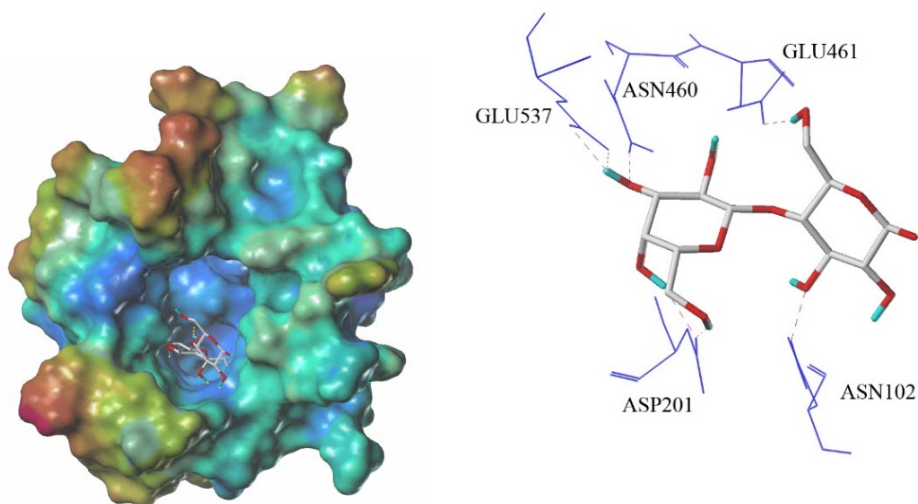
Obzirom da poravnanje sekvenci nije dovoljno pouzdano da bi se predvidele rezidue koje učestvuju u katalizi laktoze, simulacije molekularnog dokinga su izvedene sa ciljem da predvide rezidue uključene u vezivanje laktoze i galaktoze kod modelovanih beta-galaktozidaza. Takođe, obzirom da je objavljena samo jedna kristalna struktura beta-galaktozidaze sa laktozom u aktivnom mestu, analiza molekularnog dokinga laktoze kao supstrata je sprovedena i sa beta-galaktozidazama koje imaju određenu trodimenzionalnu strukturu. Kao pozitivna kontrola korišćena je kristalna struktura *E. coli* u kompleksu sa laktozom (PDB:1JYN), gde je sproveden identični protokol za doking kao i kod ostalih enzima (Herrchen i Legler 1984; Gebler i sar. 1992; Juers i sar. 2001). Primenom razvijenog protokola, simulacija dokinga laktoze uspešno izvedena sa nativnom strukturom beta-galaktozidaze *E. coli*. Molekul laktoze je imao istu orijentaciju u enzimu sa odstupanjem od eksperimentalnih podataka od $\text{RMSD} = 0,787\text{Å}$, čime je

potvrđeno da su metode korišćene u razvijenom protokolu odgovarajuće za dato ispitivanje (slika 65).



Slika 65. Preklapanje struktura dokinga beta-galaktozidaza *E.coli* sa ekperimentalno određenim položajem laktoze u enzimu

Molekularni doking nativne beta-galaktozidaze *E. coli*, sproveden je sa ciljem da se ispituju početne faze vezivanja laktoze. Rezultati su pokazali da se obe aktivne rezidue GLU461 i GLU537 beta-galaktozidaze *E. coli* (1DP0) nalaze u povoljnom položaju za građenje vodoničnih veza sa laktozom, i da se obe nalaze u blizini mesta hidrolize laktoze (slika 66).

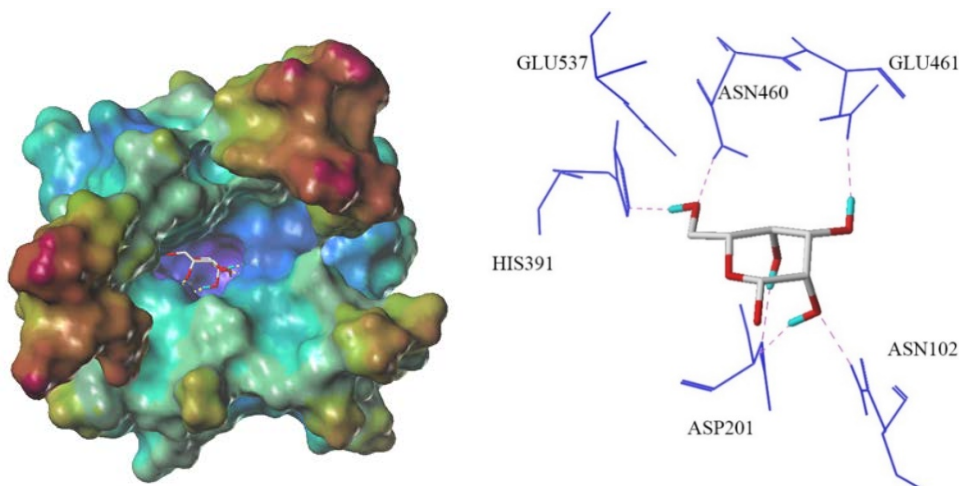


Slika 66. Doking beta-galaktozidaze *E. colisa* laktozom kao supstratom

Iz rezultata se vidi da se supstrat veže za aktivno mesto koje se nalazi u regionu visoke elektronegativnosti, što je na slici predstavljeno plavom bojom. Ovakvi rezultati su u skladu sa

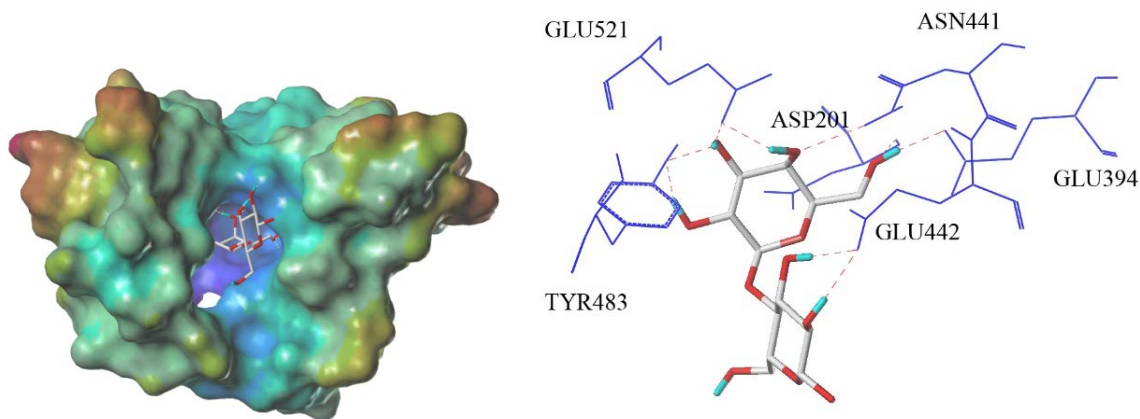
mehanizmom delovanja TIM brel katalitičkog domena gde reakcija započinje delovanjem nukleofila na vezu razgradnje supstrata. Iako se prvobitno smatralo da je GLU461 nukleofilna rezidua beta-galaktoze *E. coli*, eksperimentalno je utvrđeno da je Glu527 katalitički nukleofil, koji gradi kovalentnu vezu sa laktozom kao supstratom (Gebler i sar. 1992). Ostale rezidue koje vežu laktozu su ASN102, ASP201 i ASN460.

Simulacija dokinga sa molekulom galaktoze sprovedena je sa svim enzimima sa ciljem ispitivanja uspostavljanja veza nakon razgradnje laktoze i određivanja rezidua koje učestvuju u procesu transgalaktozilacije. Simulacija dokinga beta-galaktozidaze *E. coli* sa galaktozom pokazala je da nukleofilna rezidua GLU537 nije u mogućnosti da gradi vezu sa produktom reakcije (slika 67). Na njeno mesto kao supstrat vezujuća rezidua javlja se HIS391, što je čini važnim faktorom za reakciju koji treba uzeti u obzor prilikom budućih ispitivanja ili modifikacija enzima. Ostale supstrat vezujuće residue su nepromenjene.



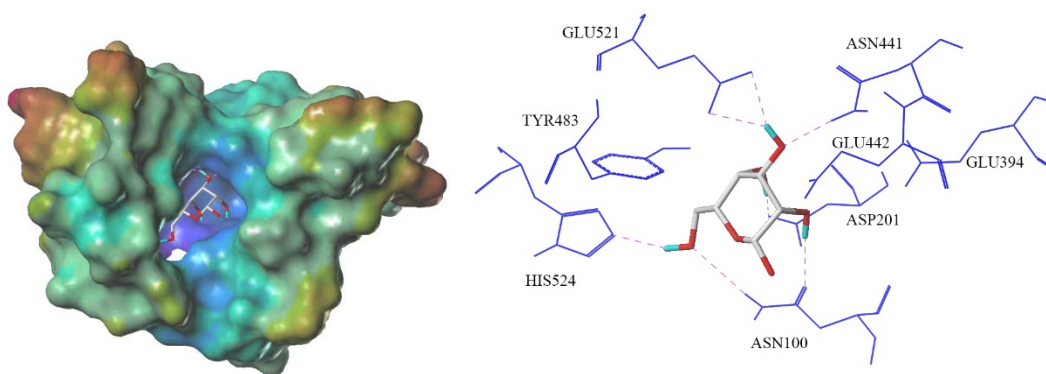
Slika 67. Doking beta-galaktozidaze *E. coli* sa galaktozom kao supstratom

Analizom beta-galaktozidaze *Arthrobacter* sp. (1YQ2) takođe su utvrđene aktivne rezidue glutaminske kiseline (GLU442/GLU521), što je u saglasnosti sa prethodno objavljenim rezultatima (Skalova i sar. 2005). Ostale rezidue koje interaguju sa laktozom su ASP201, GLU394 ASN441 i TYR483 (slika 68). Aktivno mesto se i kod beta-galaktoze *Arthrobacter* sp. nalazi u regionu visoke elektronegativnosti.



Slika 68. Doking beta-galaktozidaze *Arthrobacter* sp. sa laktozom kao supstratom

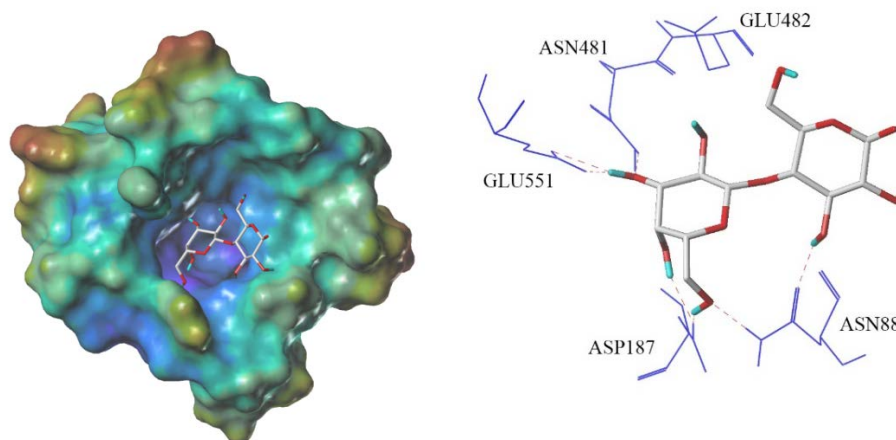
Kao i kod *E. coli*, jedna aktivna rezidua beta-galaktozidaze *Arthrobacter* sp. nije u mogućnosti da gradi vodonične veze sa galaktozom kao supstratom – GLU442 (slika 69). Galaktoza se ne nalazi u povoljnom položaju za građenje vodonične veze sa reziduom TYR438 niti sa GLU394, ali se nalazi u povoljnom položaju za vezu sa reziduom HIS524. Kao i kod *E. coli*, i u ovom slučaju se pojavljuje histidin kao supstrat vezujuća rezidua galaktoze. Za razliku od interakcije sa laktozom, kod vezivanja galaktoze javlja se rezidua ASN100. Rezidua homologa ASN100 učestvuje kod beta-galaktozidaze *E. coli* (ASN102) u vezivanju laktoze i galaktoze.



Slika 69. Doking beta-galaktozidaze *Arthrobacter* sp. sa galaktozom kao supstratom

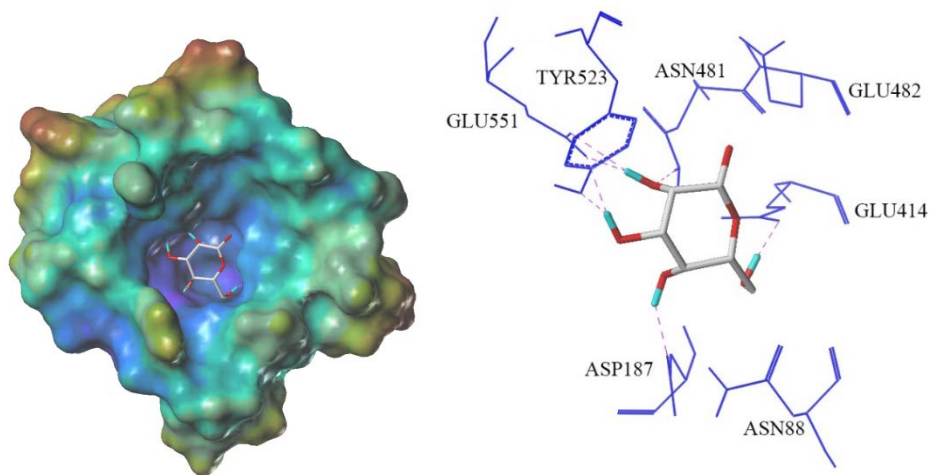
Interesantan je podatak, da prema našim doking rezultatima proton donor rezidua (GLU482) beta-galaktozidaze *K. lactis*-a (3OBA) (slika 70), nije u mogućnosti da gradi

vodonične veze sa laktozom. Nasuprot tome, nukleofilna rezidua GLU551 se nalazi na rastojanju koje dozvoljava uspostavljanje vodoničnih veza, što je u saglasnosti sa literaturnim podacima (Pereira-Rodriguez i sar. 2012). Laktoza interaguje i sa reziduama ASN88, ASP187 i ASN481.



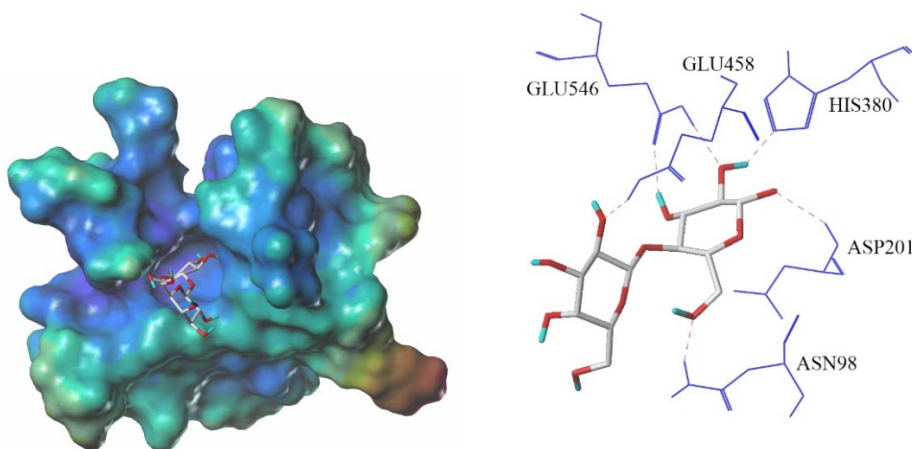
Slika 70. Doking beta-galaktozidaze *K. lactis*-a sa laktozom kao supstratom

Proton donor GLU482 nije u povoljnom položaju da gradi vodonične veze ni sa galaktozom kao supstratom (slika 71). Za razliku od laktoze, u povoljnom položaju za vezivanje sa galaktozom nalazi se TYR523. ASN88, koji ima mogućnost građenja dve vodonične veze sa laktozom, nema takvih mogućnosti za vezivanje galaktoze. Ostale supstrat vezujuće rezidue su nepromenjene.



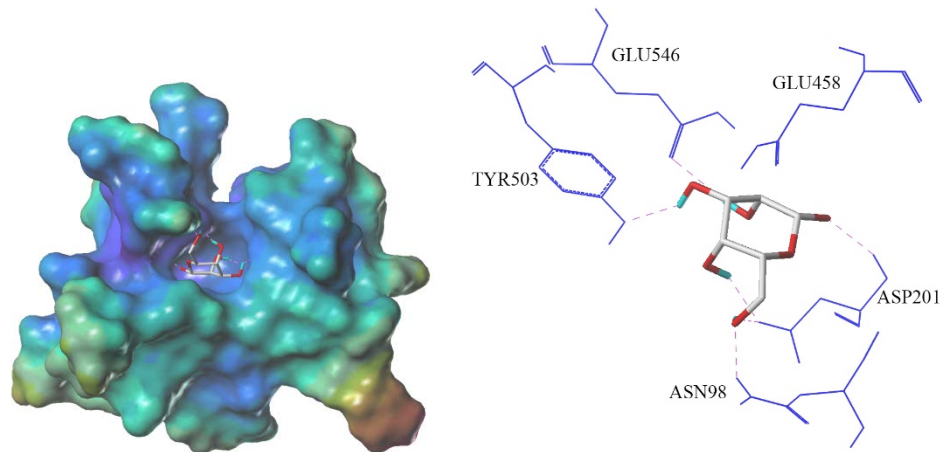
Slika 71. Doking beta-galaktozidaze *K. lactis*-a sa galaktozom kao supstratom

Na slici 72 prikazana je simulacija dokinga beta-galaktozidaze *S. thermophilus*-a sa laktozom. Uočljivo je da obe predložene katalitičke rezidue (GLU458 i GLU546) u interakciji sa našim modelom beta-galaktozidaze *S. thermophilus*-a (klaster 1) imaju mogućnost da grade vodonične veze sa laktozom. Rastojanja između aktivnih rezidua i laktoze iznose 1,80Å za GLU458 i 2,01Å za GLU546. Njihovo međusobno rastojanje je 4,26Å, što je u saglasnosti sa tipom glikozid hidrolaza kojem pripadaju (neinvertujuće glikozidaze sa TIM barel domenom). Ostale rezidue koje učestvuju u vezivanju laktoze (sa rastojanjima od liganda) su ASN98 (1,98Å), ASP201 (2,34Å) i HIS380 (1,76Å).



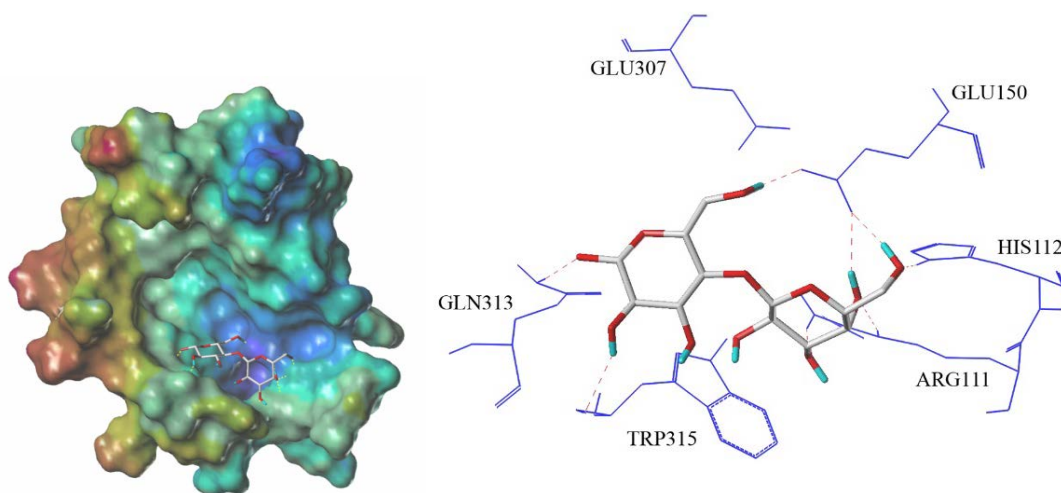
Slika 72. Doking beta-galaktozidaze *S. thermophilus*-a sa laktozom kao supstratom

Ispitivanja interakcije modelovane beta-galaktozidaze *S. thermophilus*-a sa galaktozom kao supstratom pokazala su da predložena aktivna rezidua GLU458 nije u mogućnosti da sa njom gradi vodonične veze (slika 73). GLU458 se nalazi na rastojanju od 2,74Å od galaktoze. Takođe, ni rezidua HIS380 takođe nema mogućnost vezivanja sa galaktozom. Druga predložena aktivna rezidua GLU546 ima mogućnost vezivanja sa galaktozom. Kao i kod beta-galaktozidaze *K. lactis*-a javlja se rezidua tirozina (TYR503) koja se nalazi u povoljnom položaju za interakciju sa galaktozom sa rastojanjem od 2,05Å.



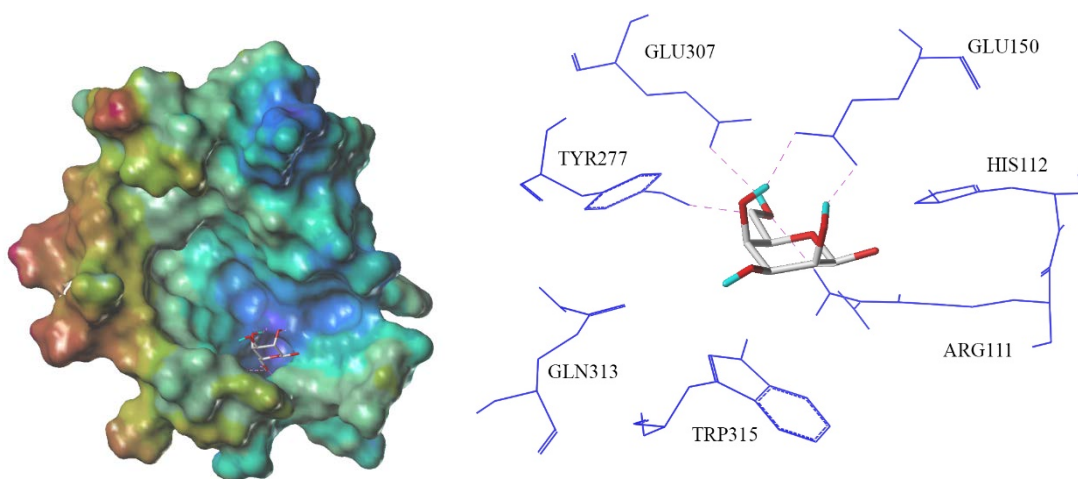
Slika 73. Doking beta-galaktozidaze *S. thermophilus*-a sa galaktozom kao supstratom

Simulacija molekularnog dokinga izvedena je i sa beta-galaktozidazom iz *B. circulans* ssp. *alkalophilus* (PDB:3TTS), čija struktura je korišćena kao šablon za homologo modelovanje beta-galaktozidaza iz klastera 2. Važno je napomenuti da je struktura beta-galaktozidaze *B. circulans* ssp. *alkalophilus*-a određena bez prisustva liganda. Simulacija dokinga laktoze je otkrila blizak položaj aktivnih rezidua GLU150 i GLU307 međuatomske vezi na kojoj se razgrađuje laktoza (slika 74). Prema našoj simulaciji predviđena nukleofilna rezidua (GLU307) nije u mogućnosti da gradi vodonične veze sa laktozom. Kao rezidue koje vezuju laktozu identifikovane su ARG111, HIS112, GLN313 i TRP317.



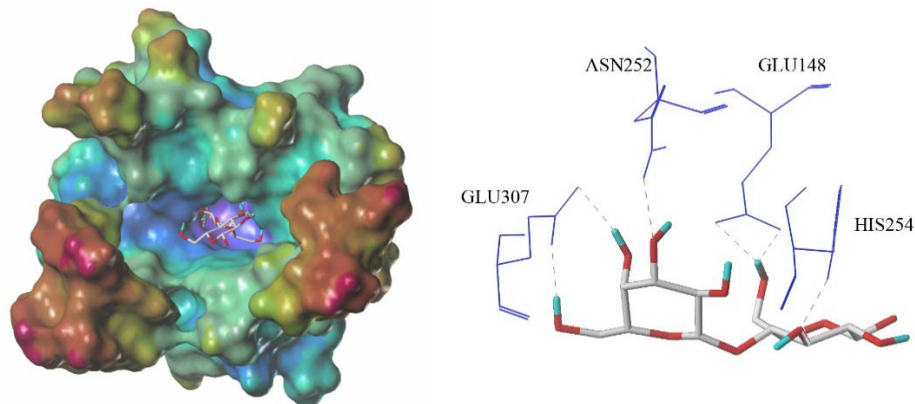
Slika 74. Doking beta-galaktozidaze *B. circulans* ssp. *alkalophilus*-a sa laktozom kao supstratom

Simulacija dokinga sa galaktozom pokazala je da aktivna rezidua GLU307 ima mogućnost da sa njom reaguje, što nije bio slučaj kod ispitivanja laktoze kao supstrata (slika 75). Za razliku od laktoze, kao supstrat vezujuća rezidua prepoznata je TYR313. Nasuprot tome, sa galaktozom nemaju mogućnost interakcije GLN313, TRP315 i HIS112. Poređenjem ovih rezultata sa rezultatima kristalografije Maksimainen i sar. (2012), može se uočiti da se podudaraju residue koje reaguju sa galaktozom, dok sa druge strane nije potvrđena njihova pretpostavka da je TYR277 vezivno mesto za laktozu, koji na prikazanom modelu nema interakciju sa laktozom.



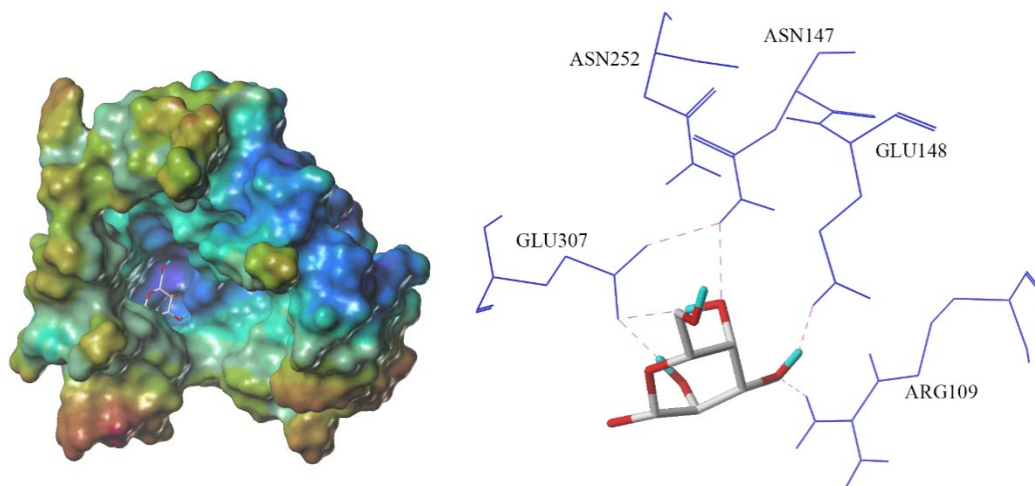
Slika 75. Doking beta-galaktozidaze *B. circulans* ssp. *alkalophilus* galaktozom kao supstratom

Simulacija molekularnog dokinga model beta-galaktozidaze *L. acidophilus*-a (klaster 2) ukazuje da enzim poseduje iste aktivne residue kao šablon struktura (beta-galaktozidaza *B. circulans* ssp. *alkalophilus*-a) (slika 76). Proton donor rezidua GLU148 je očigledno sposobna da gradi vodonične veze sa laktozom i pozicionirana je veoma blizu veze na kojoj se razgrađuje laktoza (2,08Å). Za razliku od šablon struktura (3TTS) nukleofilna rezidua GLU307 takođe ima mogućnost da gradi vodonične veze sa laktozom (2,15Å). Razdaljina između aktivnih rezidua je 5,1Å. Ostale residue koje grade vodonične veze sa laktozom su ASN252 (2,53Å) i HIS254 (2,05Å).



Slika 76. Doking beta-galaktozidaze *L. acidophilus*-a sa laktozom kao supstratom

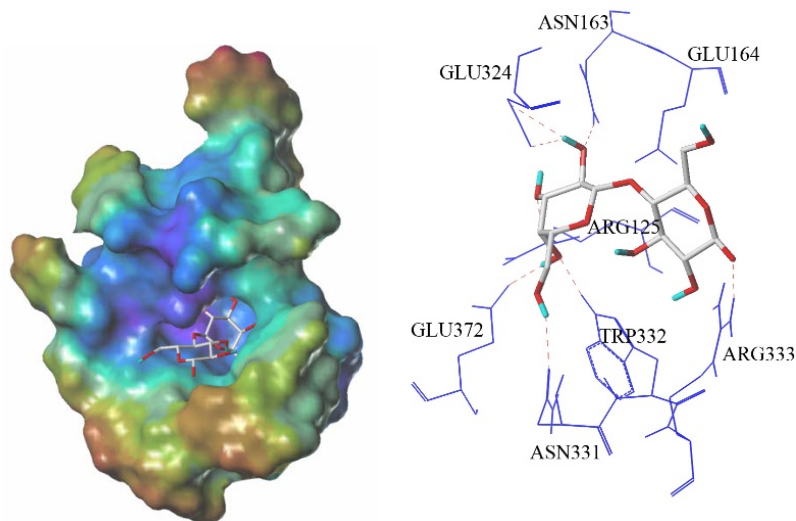
Doking galaktoze u model beta-galaktozidaze *L. acidophilus*-a pokazao je da obe predložene aktivne rezidue imaju mogućnost da grade vodonične veze sa supstratom (slika 77). Rezidua GLU148 se nalazi na rastojanju od 1,95Å, dok je rezidua GLU307 na rastojanju od 1,88Å. Uočljivo je da ASN252 (3,7Å) nema mogućnost vezivanja sa galaktozom vodoničnim vezama. Kao supstrat vezujuća rezidua javlja se ARG109 sa rastojanjem od 2,06Å.



Slika 77. Doking beta-galaktozidaze *L. acidophilus*-a sa galaktozom kao supstratom

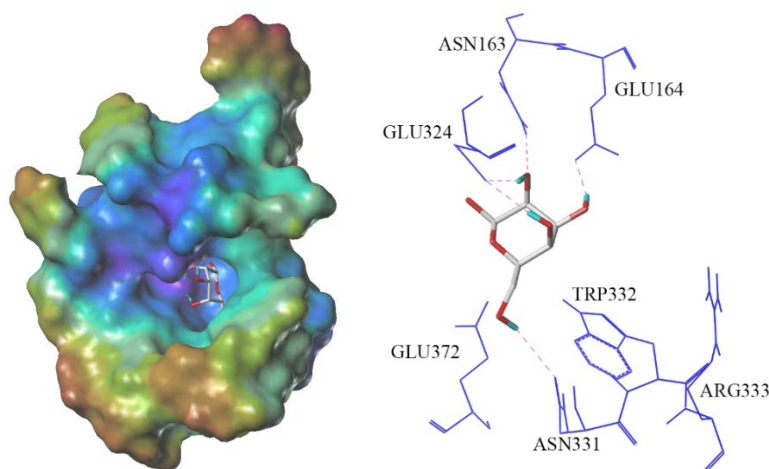
Prema simulaciji sprovedenoj u ovoj disertaciji predviđena aktivna rezidua GLU164 beta-galaktozidaze *B. animalis* ssp. *lactis* ne gradi vodonične veze sa laktozom, od koje je udaljena 2,48Å (slika 78). Nasuprot tome, druga predviđena aktivna rezidua GLU324 gradi vodonične veze i nalazi se na rastojanju od 1,94Å. Razdaljina između rezidua iznosi 5,45 Å, što je kao i u

prethodnim modelima odgovarajuće za neinvertujuće glikozidaze i TIM barel domen. Ostale vezujuće rezidue (sa rastojanjima od laktoze) su ASN163 (2,14Å), ASN331 (2,12Å), TRP332 (1,86Å), ARG333 (2,02Å) and GLU372 (1,94Å).



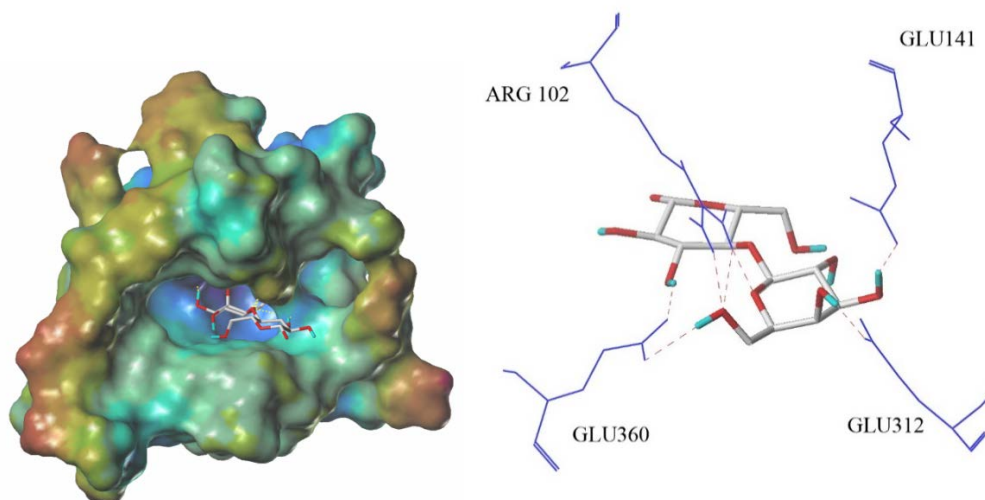
Slika 78. Doking beta-galaktozidaze *B. animalis* ssp. *lactis*-a sa laktozom kao supstratom

Za razliku od laktoze predložena aktivna rezidua GLU164 ima mogućnost građenja vodoničnih veza sa galaktozom kao supstratom (slika 79). Rastojanje GLU164 od galaktoze iznosi 1,65Å. Aktivna rezidua GLU324 (1,95Å) ima mogućnost građenja vodoničnih veza sa galaktozom, kao što je slučaj i sa laktozom kao supstratom. Za razliku od laktoze veze ne mogu graditi TRP332 (2,63Å), ARG333(7,45Å) i GLU372(3,81Å).



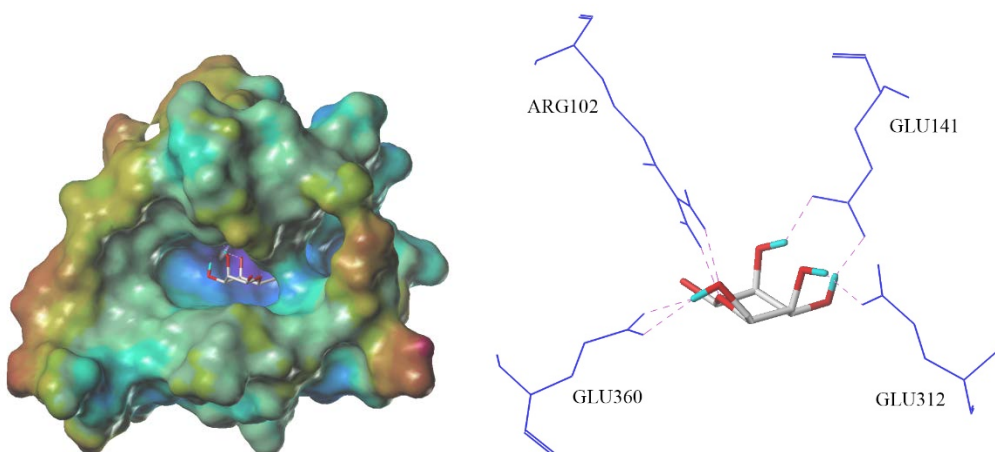
Slika 79. Doking beta-galaktozidaze *B. animalis* ssp. *lactis*-a sa galaktozom kao supstratom

Analizom aktivnih rezidua beta-galaktozidaze *T. thermophilus*-a nisu zabeležena značajnija odstupanja od ostalih članova klastera 2. Kao aktivne rezidue predviđene su GLU141 i GLU312, dok su ostale vezujuće rezidue ARG102 i GLU360 (slika 80).



Slika 80. Doking beta-galaktozidaze *T. thermophilus*-a sa laktozom kao supstratom

Ispitivanja simulacija dokinga sa galaktozom kao supstratom su pokazala da iste residue učestvuju u njenom vezivanju kao i u slučaju laktoze kao supstrata (aktivne rezidue GLU141 i GLU312, kao i supstrat vezujuće rezidue ARG102 i GLU360) (slika 81).



Slika 81. Doking beta-galaktozidaze *T. thermophilus*-a sa galaktozom kao supstratom

Dobijeni rezultati ukazuju da simulacije dokinga sprovedene u ovom istraživanju sa velikom preciznošću identifikuju rezidue uključene u katalizu i vezivanje laktoze kao supstrata, što ih čini pogodnim za ispitivanje interakcija kod konstruisanih modela gde ne postoje eksperimentalni podaci i gde je potrebno po prvi put predložiti takve rezidue.

Energije vezivanja gore opisanih simulacija dokinga prikazane su u tabeli 8. Iz tabele se vidi da beta-galaktozidaza *K. lactis*-a ima najveći afinitet za laktozu, dok beta-galaktozidaza *B. circulans* ssp. *alkalophilus*-a ima najmanji afinitet.

Tabela 8. Ukupna vrednost vezivanja (total score) za laktozu u odabranim beta-galaktozidazama primenom Surflex-Dock programa

Organizam	Total Score -log(K _d)
<i>E. coli</i>	6.444
<i>Arthrobacter</i> sp.	6.918
<i>K. lactis</i>	8.056
<i>B. circulans</i> ssp. <i>alkalophilus</i>	6.200
<i>T. thermophilus</i>	7.223
<i>S.thermophilus</i>	7.232
<i>B. animalis</i> ssp. <i>lactis</i>	7.163
<i>L. acidophilus</i>	6.761

Analiziranjem molekularnog dokinga mogu se dobiti veoma važne informacije o reziduama koje učestvuju u vezivanju supstrata za enzim, kao i katalitički aktivnim reziduama. Doking analiza ima poseban značaj jer klaster analizom i poravnanjem sekvenci nije moguće sa sigurnošću utvrditi sve aktivne rezidue koje učestvuju u vezivanju supstrata, u ovom slučaju laktoze. Doking analiza u ovom radu je potvrdila aktivne rezidue kod beta-galaktozidaze *E. coli* (1DP0) predložene u UNIPROT bazi podataka. Iako proton donor rezidua GLU482 beta-galaktozidaze *K. lactis*-a ne gradi vodonične veze sa laktozom, pozicionirana je veoma blizu veze na kojoj se razgrađuje laktoza, što je ključan preduslov za njenu katalitičku ulogu. Obe katalitičke rezidue (GLU 458 i GLU546) beta-galaktozidaze *S. thermophilus*-a grade vodonične veze sa laktozom i obe se nalaze blizu mesta razgradnje laktoze, što je u saglasnosti sa njihovom katalitičkom ulogom u reakciji (slika 72). Iako je rastojanje između aktivnih rezidua nešto manje nego što je prosek za neinvertujuće glikozid hidrolaze, i dalje se nalazi u prihvatljivom opsegu. Neinvertujuće glikozid hirolaze imaju prosečno rastojanje između aktivnih rezidua glutaminskih

kiselina od 5,5Å (McCarter i Stephen Withers 1994; Wang i sar. 1994). Ovi rezultati ukazuju da su napravljeni model i simulacija donikga beta-galaktozidaze *S. thermophilus*-a adekvatno konstruisani i pogodni za dalju analizu. Na osnovu dobijenih rezultata, pored aktivnih rezidua i rezidue koje učestvuju u vezivanju supstrata su konzervativne unutar klastera 1. Tako su na primer iste rezidue ASN i ASP prisutne u većini prokazanih enzima iz klastera 1. Proton donorska rezidua GLU150 *B. circulans* ssp. *alkalophilus*-a (3TTS) gradi jake vodonične veze sa laktozom, čime su potvrđeni literaturni podaci za navedenu beta-galaktozidazu (slika 74) (Maksimainen i sar. 2012). Iste aktivne rezidue kod beta-galaktozidaza *L. acidophilus*-a i *B. circulans* ssp. *Alkalophilus*-a je u saglanosti sa rezultatima poravnanja sekvenci (slika 54) kao i sa njihovom pripadnošću subfamiliji GH-42. Obzirom da pripadaju različitim tipovima i ovaj podatak ide u prilog gore iznetoj hipotezi o horizontalnom transferu gena u njihovom evolutivnom razvoju. Razdaljine aktivnih glutaminskih rezidua i kod beta-galaktozidaze *L. acidophilus*-a potvrđuje njenu pripadnost neinvertujućim glikozid hidrolazama. Iako predložena aktivna rezidua GLU164 ne gradi vodonične veze sa laktozom, pozicionirana je u blizini veze na kojoj se razgrđuje laktoza, što ukazuje da ima istu ulogu kao i kod ostalih članova klastera 2 i GH-42 subfamilije. Rezidua GLU324 vezuje laktozu, što uzimajući u obzir našu filogenetsku analizu ukazuje na njenu katalitičku aktivnost. Razdaljina između aktivnih glutamatskih rezidua beta-galaktozidaze *B. animalis* ssp. *lactis*-a je takođe u okviru prosečnih vrednosti za neinvertujuće glikozid hidrolaze. Iz dobijenih doking rezultata je uočljivo da su aktivne rezidue glutamata izrazito konzervativne kod pripadnika istih klastera. Takođe, konzervativnost je prisutna i kod rezidua koje učestvuju u vezivanju laktoze za enzim. Aktivne rezidue glutamata su u svim simulacijama pozicionirane blizu veze na kojoj se razgrađuje laktoza.

Na osnovu dobijenih rezultata, može se zaključiti da su konstruisani modeli pogodni za dalja istraživanja. Takva istraživanja mogla bi uključiti usmerenu evoluciju sa ciljem povećanja stabilnosti i aktivnosti enzima. Iako dokingom predviđne energije vezivanja nisu pouzdan parametar za izražavanje efikasnosti enzima, one ukazuju na stabilnost kompleksa enzima sa supstratom i brzinom reakcije. Takođe, predviđene energije vezivanja za laktozu za modelovane strukture mogu biti upoređene sa dobijenim rezultatima za beta-galaktozidaze sa eksperimentalno određenim strukturama (tabela 8), što opet ukazuje da su konstruisani modeli precizni. Interesantan je podatak da beta-galaktozidaza *K. lactis*-a ima najveću vrednost energije vezivanja, što može ukazivati na visoku efikasnost enzima. Dobijeni rezultati pokazuju da je u

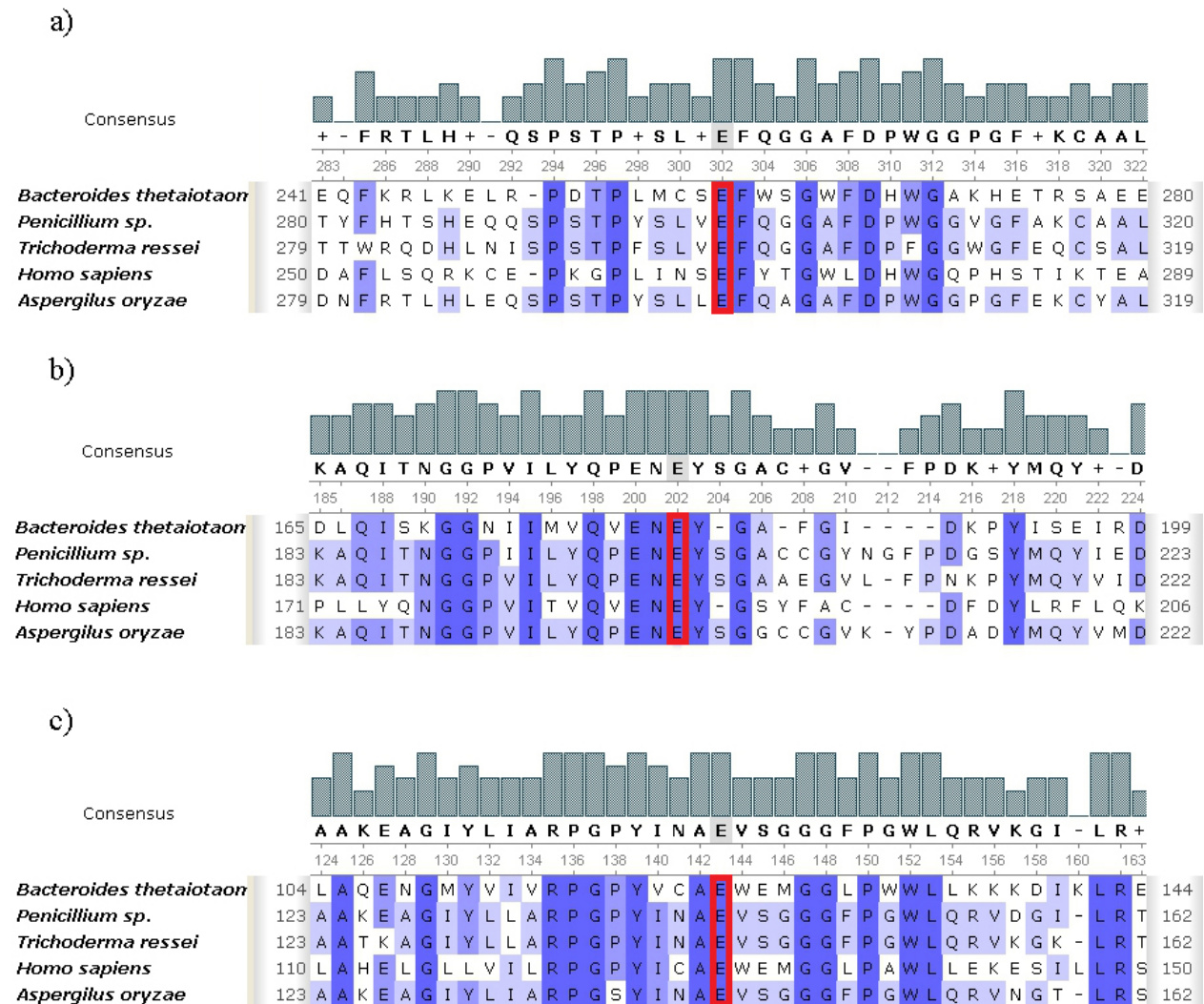
ovoj disertaciji predstavljen protokol za modelovanje komercijalno važnih beta-galaktozidaza i analiziranje njihovih interakcija sa laktozom kao supstratom. Ovi rezultati treba da pruže okvir za molekularna istraživanja vezivanja i razgradnje laktoze, kao i za buduće dizajniranje komercijalno interesantnih beta-galaktozidaza.

4.5. Beta-galaktozidaza *B. thetaiotaomicron*-a

Poseban segment ove disertacije čini ispitivanje beta-galaktozidaze *B. thetaiotaomicron*-a jer se ne koristi u tehnologiji mlečnih proizvoda, već predstavlja bakteriju koja nastanjuje ljudski crevni trakt gde čini dominantnu vrstu sa udelom od 6% od ukupnih bakterija i 12% od bakterija roda *Bacterioides* u ljudskim crevima.

Analiza poravnanja celih sekvenci otkrila je da beta-galaktozidaza *B. thetaiotaomicron*-a ima više od 70% sličnosti sekvenci sa beta-galaktozidazama *Penicillium sp*, *T. ressei* i *A. oryzae*, dok sličnost sa beta-galaktozidazom *H. sapiens*-a iznosi nešto manje od 60%. Sve navedene beta-galaktozidaze pripadaju subfamiliji GH-35, u koju po CAZY bazi podataka spadaju neinvertujuće glikozid hidrolaze (Cantarel i sar. 2009).

Najbliža homologa beta-galaktozidaza koja ima eksperimentalno određenog aktivnog nukleofila jeste beta-galaktozidaza *Homo sapiens*-a, gde je utvrđeno da je GLU268 nukleofilna rezidua (McCarter i sar. 1997). Na osnovu slučnosti sekvenci (slika 82a) može se zaključiti da je nukleofilna rezidua beta-galaktozidaze *B. thetaiotaomicron*-a GLU259. Analiza konzervativnih regiona u okolini potencijalne nukleofilne rezidue pokazala je da je 15,3% rezidua konzervativno kod svih 5 sekvenci, 10,2% kod 4 od 5 sekvenci, dok je 35% konzervativno kod 3 od 5 sekvenci. Iako je niska konzervativnost okolnih rezidua, uočljiva je konzervativnost aktivne rezidue koja je uključena u razgradnju laktoze.

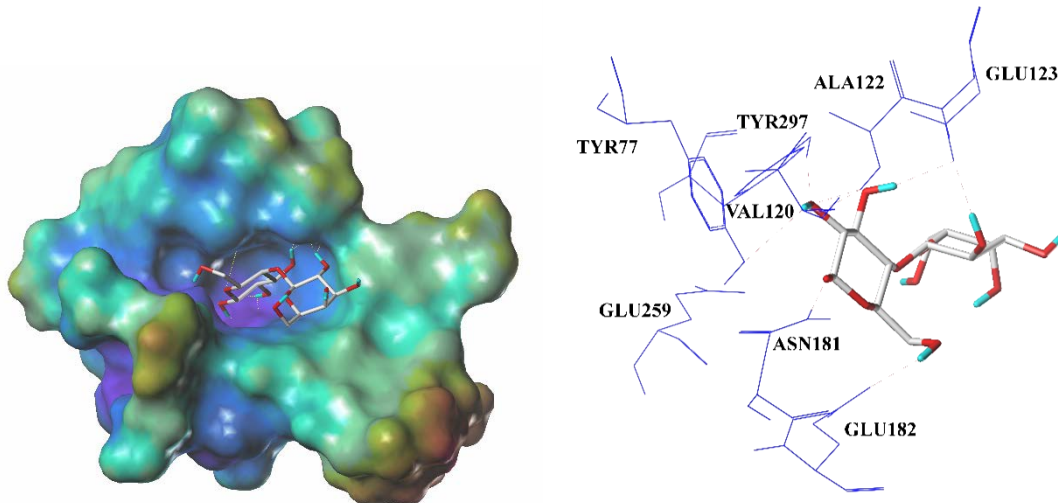


Slika 82. Predložene aktivne rezidue beta-galaktozidaze *B. thetaiotaomicron*-a: a) region u okolini nukleofilne rezidue, b) region u okolini GLU182 potencijalnog proton donora c) region u okolini GLU123 potencijalnog proton donora

Obzirom da proton donor rezidua nije eksperimentalno utvrđena kod homologih sekvenci, potrebno je predvideti je, kako bi bolje razumeli mehanizam delovanja enzima i razgradnju laktoze. Kao potencijalne proton donor rezidue javljaju se dve rezidue glutaminske kiseline: GLU182 i GLU123.

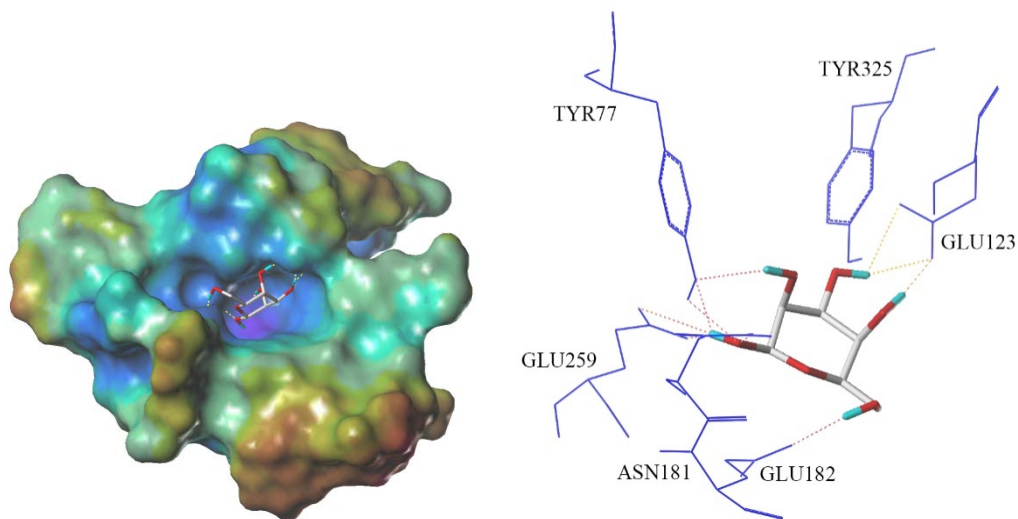
Analiza konzervativnih regiona u okolini GLU182 potencijalne proton donor rezidue pokazala je da je 25% rezidua konzervativno kod svih 5 sekvenci, 33,3% kod 4 od 5, dok je 25% konzervativno kod 3 od 5 sekvenci. Uočljiva je velika konzervativnost rezidua u neposrednom okruženju GLU182 rezidue. Druga potencijalna proton donor rezidua GLU123 ima nešto manje konzervativne rezidue u svom neposrednom okruženju, dok u širem okruženju ima konzervativno 38,9% rezidua kod svih 5 sekvenci, 10,25% kod 4 od 5 i 43,6% kod 3 od 5 sekvenci.

Obzirom da na osnovu dobijenih poravnanja sekvenci nije moguće sa velikom sigurnošću predvideti proton donora, za ispitivane beta-galaktozidaze su analizirane i simulacije dokinga. Analizirano je aktivno mesto koje uključuje sve tri potencijalne aktivne rezidue sa dodatnih 5Å prostora, sa ciljem da bi se uključile sve rezidue koje bi mogle učestvovati u vezivanju laktoze. Nakon dokinga laktoze u aktivno mesto analizirana je njena konformacija i potencijalne aktivne i vezivne rezidue (slika 83). Eksperimentalno utvrđena nukleofilna rezidua GLU259 je udaljena 5,26Å od laktoze i prema dobijenim rezultatima ne gradi vodonične veze sa njom. Obe potencijalne proton donor rezidue su pozicionirane blizu laktoze i veze na kojoj se ona razgrađuje, (GLU 123 4,39Å, GLU182 5,60Å). Kao što je poznato rastojanje između proton donora i nukleofila kod neinvertujućih glukozid hidrolaza iznosi oko 5,5Å (Wang i sar. 1994; McCarter i sar. 1997). Merenjem razdaljina eksperimentalno utvrđenog nukleofila i potencijalnih proton donora utvrđeno je da je GLU259 udaljen od GLU182 5,47Å, dok rastojanje GLU 259 od GLU123 iznosi 8,0Å. Uzimajući u obzir razdaljinu rezidua od laktoze, kao i njihovu međusobnu razdaljinu, može se pretpostaviti da ulogu proton donora u beta-galaktozidazi *B. thetaiotaomicron*-a ima rezidua GLU182. Ostale rezidue koje učestvuju u vezivanju laktoze su TYR77, ASN181 i TYR325. I u ispitivanju beta-galaktozidaza subfamilije GH-35 uočeno je da evolutivno udaljene vrste imaju visoku homologiju sekvenci, što ukazuje na horizontalni transfer gena, posebno ukoliko se uzme u obzir činjenica da *B. thetaiotaomicron* nastanjuje ljudski crevni trakt.



Slika 83. Doking beta-galaktozidaze *B. thetaiotaomicron*-a sa laktozom kao supstratom

Doking analiza galaktoze ukazala je da sve tri glutaminske residue imaju mogućnost da sa njom grade vodonične veze (slika 84). Za razliku od laktoze, rezidua TYR325 nema mogućnost vezivanja galaktoze. Ostale galaktoza vezujuće residue su ASN181 i TYR77.



Slika 84. Doking beta-galaktozidaze *B. thetaiotaomicron*-a sa galaktozom kao supstratom

4.6. Kinetika razgradnje laktoze beta-galaktozidazom *K. lactis*-a

Obzirom da se beta-galaktozidaza *K. lactis*-a komercijalno upotrebljava prilikom proizvodnje delaktoziranih mlečnih proizvoda, ispitana je njena kinetika razgradnje laktoze. Kinetika je praćena primenom permeata kao izvora laktoze koji predstavlja nus proizvod mlečne industrije prilikom proizvodnje feta sira. Analiza hemijskog sastava vršena je kako sirovina, tako i produkta realizovanog procesa. Hemijski sastav mleka, permeata i hidrolizovanog permeta dati su u tabeli 9.

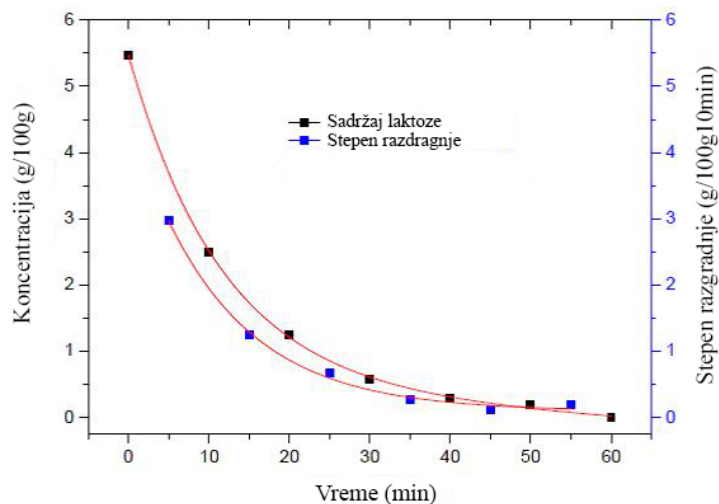
Tabela 9. Hemijski sastav i energetska vrednost permeata mleka i hidrolizovanog permeata

Analiza	Mleko	Permeat	Hidrolizovani permeat
pH	6,75	6,46	6,28
Suva materija (g/100g)	12,67	5,55	5,41
Mlećna mast (g/100g)	3,7	<0,1	0,00
Ukupni proteini (g/100g)	3,13	0,20	0,18
Laktoza (g/100g)	5,47	5,72	0,00
Pepeo (g/100g)	0,76	0,48	0,51
Energetska vrednost (kJ/100g)	288,88	87,27	84,32

Hemijski sastav mleka je u skladu sa Pravilnikom o kvalitetu proizvoda od mleka i starter kultura Ministarstva poljoprivrede i zaštite životne sredine Republike Srbije (Sl glasnik RS. 33/2010, i 69/2010). pH vrednost u mleku je iznosila 6,75, što je u okviru optimalnih vrednosti (pH = 6,5 – 7,0) za beta-galaktozidazu *Kluyveromyces lactis*-a (Kim i sar. 2004; Demirhan i sar. 2010). Uočljivo je da se sadržaj laktoze povećao u permeatu u odnosu na mleko. To je posledica prolaska laktoze kroz membranu prilikom ultrafiltracije, dok se većina proteina, mlećna mast i drugi veliki makromolekuli zadržavaju, čime se povećava krajnja koncentracija laktoze u permeatu. Usled navedenog dolazi i do smanjenja sadržaja suve materije i pepela. Energetska vrednost se takodje smanjuje, kako kod permeata, tako i dodatno kod hidrolizovanog permeata usled ragradnje beta-galaktozidne veze prilikom koje se oslobađa energija.

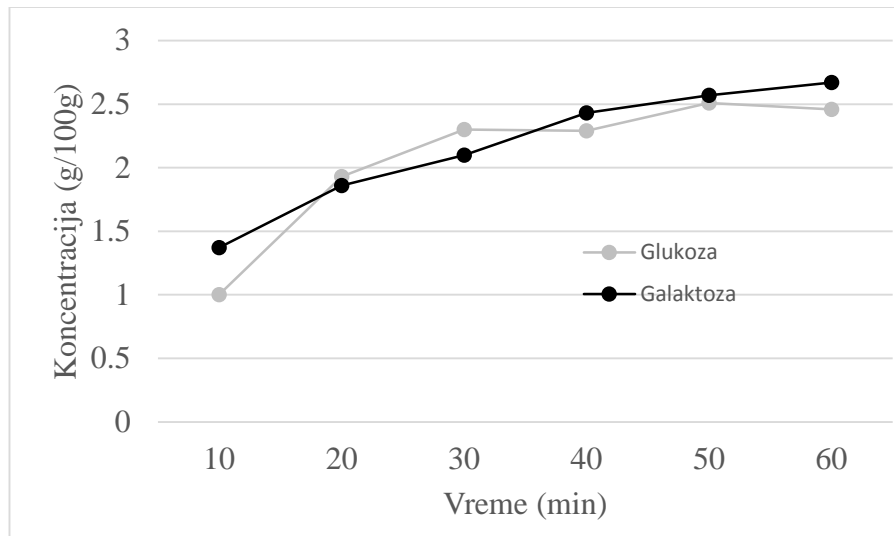
Na slici 85 prikazano je smanjenje sadržaja laktoze tokom hidrolize u permeatu na temperaturi od 40°C i pri koncentraciji enzima od 0,1% tokom 60 minuta. Celokupna laktoza se

razgradila nakon 60 minuta od trenutka dodatka enzima. Analiza krive razgradnje laktoze ukazuje da je razgradnja najveća na početku ukazujući da razgradnja laktoze dostiže maksimalnu brzinu u prvih 10 minuta. Najveća brzina se dostiže usled velike koncentracije laktoze koje ima dovoljno da se angažuje veći broj molekula enzima. Vremenom se koncentracija laktoze smanjuje, što dovodi do smanjene pristupačnosti laktoze enzimu i smanjenja brzine njene razgradnje. Razgradnja laktoze je pokazala eksponencijalni model razgradnje sa veoma visokom R^2 vrednosti od 0,99.



Slika 85. Kinetika hidrolize laktoze u permeatu dodavanjem 0,1% enzima u toku 60 minuta

Ispitane promene koncentracije glukoze i galaktoze tokom 60 min hidrolize prikazane su na slici 86. Najveći porast sadržaja glukoze u početnim fazama reakcije dobijen je primenom koncentracije enzima od 0,5%, što je posledica najveće koncentracije enzima. Krajnja koncentracija glukoze iznosila je 2,8g/100g. Najmanji porast sadržaja glukoze zabeležen je prilikom primene 0,1% enzima. Krajnja koncentracija bila je 2,4g/100g. Sadržaj galaktoze je takođe pokazivao rast tokom vremena primenom svih koncentracija enzima što je očekivano za produkt enzimske reakcije.



Slika 86. Sadržaj glukoze i galaktoze u hidrolizovanom permeate u toku 60 min

Primenom koncentracija enzima od 0,3% i 0,5%, ukupni sadržaj laktoze je razgrađen već nakon prvih 10 min i iznosio je 0 g/100g.

Iz dobijenih rezultata se može zaključiti da beta-galaktozidaza *K. lactis*-a ima visoku efikasnost prilikom razgradnje laktoze što je čini pogodnom za komercijalnu upotrebu u industriji mleka. Kao što je pokazano doking simulacijom, beta-galaktozidaza *K. lactis*-a ima najveću vrednost energije vezivanja sa laktozom, što je ukazuje na potencijalnu vezu energije vezivanja i efikasnosti razgradnje laktoze. Ovi rezultati mogu imati veliku primenu u industriji mleka za razgradnju laktoze prilikom proizvodnje delaktoziranog mleka ili drugih proizvoda bez laktoze. Takođe, primenom metoda usmerene evolucije i racionalnog dizajna moguće je poboljšati efikasnost datih enzima i učiniti ih još atraktivnijim za industriju mleka.

5. Zaključci

U ovoj disertaciji predstavljen je protokol za modelovanje komercijalno važnih beta-galaktozidaza i analiziranje njihovih interakcija sa laktozom kao supstratom. Ovi rezultati treba da pruže okvir za molekularna istraživanja vezivanja i razgradnje laktoze, kao i za buduće dizajniranje komercijalno interesantnih beta-galaktozidaza.

Na osnovu sličnosti analizirane sekvence se mogu podeliti u dva klastera. Ovakva podela je očekivana obzirom da članovi istih klasera pripadaju istoj subfamiji glikozid hidrolaza. Članovi klastera 1 pripadaju GH-2 subfamiji, dok članovi klastera 2 pripadaju GH-42 subfamiji. Poravnanjem sekvenci dobijene su potencijalne aktivne rezidue. Na osnovu klaster analize i poravnanja sekvenci katalitički aktivne rezidue beta-galaktozidaza ispitivanih sojeva mogle bi biti GLU458/GLU546 kod *S. thermophilus*-a, GLU148/GLU307 kod *L. Acidophilus*-a i GLU164/GLU324 kod *B. animalis* ssp. *lactis*-a.

Pretraživanje BlastP programom, pokazalo je da beta-galaktozidaze *S. thermophiles* imaju visok stepen sličnosti sa 1DP0A (*Escherichia coli*), dok beta-galaktozidaze *L. acidophilus* i *B. animalis* ssp. *lactis* imaju visok stepen sličnosti sa 3TTS A (*Bacillus circulans*). Uzimajući u obzir filogenetsku analizu (raličiti tipovi), BlastP rezultati (visok stepen homologije između sekvenci) ukazuju na horizontalni tranfer gena tokom evolucije *L. acidophilus*-a i *B. animalis* *lactis*-a.

Konstruisani su homologni modeli beta-galaktozidaza *S. thermophilus*, *L. acidophilus*, *B. animalis* ssp. *lactis*. Monomer beta-galaktozidaze *S. thermophilus*-a se sastoji od 5 domena, od kojih centralni TIM barel domen ima funkcionalnu ulogu u razgradnji supstrata. Beta-galaktozidaze *L. acidophilus*-a i *B. animalis* ssp. *lactis*-a imaju sličnu ukupnu strukturu, što je očekivano obzirom da su konstruisani primenom iste šablon strukture. Monomer beta-galaktozidaze iz *L. acidophilus*-a i *B. animalis* ssp. *lactis*-a se sastoji iz 3 domena sa katalitičkim TIM barelom domenom.

Za proveru tri modelovana enzima (*S. thermophilus*, *L. acidophilus*, *B. animalis* ssp. *lactis*) konstruisani su Ramačandranovi plotovi. Na osnovu Ramačandranovih plotova može se zaključiti da su modelovane beta-galaktozidaze imaju povoljne uglove veza, posebno u regionu aktivnog mesta što je od ključnog značaja za njihovu primenu u daljim istraživanjima.

Modelovani enzimi su dalje validovani strukturalnim superpozicioniranjem sa njihovim šablon strukturama.

Molekularni doking sa betagalaktozidazama iz *Arthrobacter* sp., *K. lactis*, *S. thermophilus*, *B. circulans* ssp. *alkalophilus*, *L. acidophilus*, *B. animalis* ssp. *lactis*, *T. thermophiles* i *B. thetaiotaomicron* je sproveden sa ciljem da se ispituju početne faze vezivanja laktoze. Iz rezultata simulacije molekularnog dokinga svih beta-galaktozidaza vidi se da se supstrat veže za aktivno mesto koje se nalazi u regionu visoke elektronegativnosti. Ovakvi rezultati su u skladu sa mehanizmom delovanja TIM barel katalitičkog domena gde reakcija započinje delovanjem nukleofila na vezu razgradnje supstrata. Simulacije dokinga su potvrdile aktivna mesta predložena poravnanjem sekvenci i otkrile rezidue koje učestvuju u vezivanju laktoze i galaktoze kao supstrata.

Na osnovu sličnosti sekvenci može se zaključiti da je nukleofilna rezidua beta-galaktozidaze *B. thetaiotaomicron*-a GLU259. Analizom konzervativnih regiona, kao i doking simulacije, može se zaključiti da ulogu proton donora u beta-galaktozidazi *B. thetaiotaomicron*-a ima rezidua GLU182.

Hemijski sastav mleka za ispitivanje aktivnosti beta-galaktozidaze *Kluyveromyces lactis*-a je u skladu sa Pravilnikom o kvalitetu proizvoda od mleka i starter kultura Ministarstva poljoprivrede i zaštite životne sredine Republike Srbije. Celokupna laktoza se razgradila u permeatu na temperaturi od 40°C i pri koncentraciji enzima od 0,1% nakon 60 minuta od trenutka dodatka enzima. Maksimalna brzina se postiže u toku prvih 10 minuta. Iz dobijenih rezultata se može zaključiti da beta-galaktozidaza *K. lactis*-a ima visoku efikasnost prilikom razgradnje laktoze što je čini pogodnom za komercijalnu upotrebu u industriji mleka.

Rezultati dobijeni u ovoj diertaciji mogu imati veliki aplikativni značaj u industriji mleka gde se primenjuje enzim beta-galaktozidaza za razgradnju laktoze prilikom proizvodnje delaktoziranog mleka ili drugih proizvoda bez laktoze.

6. Literatura

- Abagyan, R. A., Batalov, S. (1997). "Do aligned sequences share the same fold?" J Mol Biol **273**(1): 355-368.
- Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M., Pearl, L. H. (1997). "Crystal structure of the beta-glycosidase from the hyperthermophilic archeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability." J Mol Biol **271**(5): 789-802.
- Alexandrov, N. N., Luethy, R. (1998). "Alignment algorithm for homology modeling and threading." Protein Sci **7**(2): 254-258.
- Altermann, E., Russell, W. M., Azcarate-Peril, M. A., Barrangou, R., Buck, B. L., McAuliffe, O., Souther, N., i sar. (2005). "Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM." Proc Natl Acad Sci U S A **102**(11): 3906-3912.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J. (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-3402.
- Andersen, J. M., Barrangou, R., Abou Hachem, M., Lahtinen, S., Goh, Y. J., Svensson, B., Klaenhammer, T. R. (2011). "Transcriptional and functional analysis of galactooligosaccharide uptake by lacS in *Lactobacillus acidophilus*." Proc Natl Acad Sci U S A **108**(43): 17785-17790.
- Anfinsen, C. B. (1973). "Principles that govern the folding of protein chains." Science **181**(4096): 223-230.
- Anfinsen, C. B., Haber, E., Sela, M., White, F. H., Jr. (1961). "The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain." Proc Natl Acad Sci U S A **47**: 1309-1314.
- Armen, R. S., Chen, J., Brooks, C. L., 3rd (2009). "An Evaluation of Explicit Receptor Flexibility in Molecular Docking Using Molecular Dynamics and Torsion Angle Molecular Dynamics." J Chem Theory Comput **5**(10): 2909-2923.
- Bairoch, A., Apweiler, C., Wu, H. (2005). "The Universal Protein Resource (UniProt)." Nucleic Acids Res **33**: D154-159.
- Baldi, P., Chauvin, Y., Hunkapiller, T., McClure, M. A. (1994). "Hidden Markov models of biological primary sequence information." Proc Natl Acad Sci U S A **91**(3): 1059-1063.

Barrangou, R., Azcarate-Peril, M. A., Duong, T., Conners, S. B., Kelly, R. M., Klaenhammer, T. R. (2006). "Global analysis of carbohydrate utilization by *Lactobacillus acidophilus* using cDNA microarrays." Proc Natl Acad Sci U S A **103**(10): 3816-3821.

Barrangou, R., Briczinski, E. P., Traeger, L. L., Loquasto, J. R., Richards, M., Horvath, P., Coute-Monvoisin, A. C., i sar. (2009). "Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and BI-04." J Bacteriol **191**(13): 4144-4151.

Bates, P. A., Jackson, R. M., Sternberg, M. J. (1997). "Model building by comparison: a combination of expert knowledge and computer automation." Proteins Suppl **1**: 59-67.

Becerra M, Cerdán E, González Siso, M. I. (1998). "Micro-scale purification of β -galactosidase from *Kluyveromyces lactis* reveals that dimeric and tetrameric forms are active." Biotechnol Techniques **12**: 253–256.

Benedetti, E., Morelli, G., Nemethy, G., Scheraga, H. A. (1983). "Statistical and energetic analysis of side-chain conformations in oligopeptides." Int J Pept Protein Res **22**(1): 1-15.

Benner, S. A., Gerloff, D. (1991). "Patterns of divergence in homologous proteins as indicators of secondary and tertiary structure: a prediction of the structure of the catalytic domain of protein kinases." Adv Enzyme Regul **31**: 121-181.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Wheeler, D. L. (2005). "GenBank." Nucleic Acids Res **33**(Database issue): D34-38.

Berg, O. G., Winter, R. B., von Hippel, P. H. (1981). "Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory." Biochemistry **20**(24): 6929-6948.

Berger, J.-L., Lee, B. H., Lacroix, C. (1997). "Purification, properties and characterization of a high-molecular-mass β -galactosidase isoenzyme from *Thermusaquaticus* YT-I." Biotechnology and Applied Biochemistry **25**(1): 29-41.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., i sar. (2000). "The Protein Data Bank." Nucleic Acids Res **28**(1): 235-242.

Beveridge, D. L., DiCapua, F. M. (1989). "Free energy via molecular simulation: applications to chemical and biomolecular systems." Annu Rev Biophys Biophys Chem **18**: 431-492.

Bhat, T. N., Sasisekharan, V., Vijayan, M. (1979). "An analysis of side-chain conformation in proteins." Int J Pept Protein Res **13**(2): 170-184.

Blundell, T. L., Sibanda, B. L., Sternberg, M. J., Thornton, J. M. (1987). "Knowledge-based prediction of protein structures and the design of novel molecules." Nature **326**(6111): 347-352.

Bohm, H. J. (1992). "LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads." J Comput Aided Mol Des **6**(6): 593-606.

- Bonneau, R., Strauss, C. E., Baker, D. (2001). "Improving the performance of Rosetta using multiple sequence alignment information and global measures of hydrophobic core formation." Proteins **43**(1): 1-11.
- Boomsma, W., Hamelryck, T. (2005). "Full cyclic coordinate descent: solving the protein loop closure problem in C α space." BMC Bioinformatics **6**: 159.
- Bower, M. J., Cohen, F. E., Dunbrack, R. L., Jr. (1997). "Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool." J Mol Biol **267**(5): 1268-1282.
- Brenner, S. E. (2001). "A tour of structural genomics." Nat Rev Genet **2**(10): 801-809.
- Brenner, S. E., Chothia, C., Hubbard, T. J. (1998). "Assessing sequence comparison methods with reliable structurally identified distant evolutionary relationships." Proc Natl Acad Sci U S A **95**(11): 6073-6078.
- Brooks, B. R., Brooks, C. L., MacKerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., i sar. (2009). "CHARMM: The Biomolecular Simulation Program." J Comput Chem **30**(10): 1545-1614.
- Browne, W. J., North, A. C., Phillips, D. C., Brew, K., Vanaman, T. C., Hill, R. L. (1969). "A possible three-dimensional structure of bovine alpha-lactalbumin based on that of hen's egg-white lysozyme." J Mol Biol **42**(1): 65-86.
- Bursulaya, B. D., Totrov, M., Abagyan, R., Brooks, C. L., 3rd (2003). "Comparative study of several algorithms for flexible ligand docking." J Comput Aided Mol Des **17**(11): 755-763.
- Candela, M., Maccaferri, S., Turrone, S., Carnevali, P., Brigidi, P. (2010). "Functional intestinal microbiome, new frontiers in prebiotic design." International Journal of Food Microbiology **140**(2-3): 93-101.
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B. (2009). "The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics." Nucleic Acids Res **37**(Database issue): D233-238.
- Canutescu, A. A., Dunbrack, R. L., Jr. (2003). "Cyclic coordinate descent: A robotics algorithm for protein loop closure." Protein Sci **12**(5): 963-972.
- Canutescu, A. A., Dunbrack, R. L., Jr. (2005). "MolDE: a homology modeling framework you can click with." Bioinformatics **21**(12): 2914-2916.
- Carić, M., Milanović, S., Vucelja, D., (2000). Standard methods of milk and dairy products analyses. Novi Sad, Yugoslavia:, Prometej & Faculty of Technology.

- Chen, W., Chen, H., Xia, Y., Zhao, J., Tian, F., Zhang, H. (2008). "Production, purification, and characterization of a potential thermostable galactosidase for milk lactose hydrolysis from *Bacillus stearothermophilus*." J Dairy Sci **91**(5): 1751-1758.
- Chothia, C., Lesk, A. M. (1986). "The relation between the divergence of sequence and structure in proteins." EMBO J **5**(4): 823-826.
- Cieslinski, H., Kur, J., Bialkowska, A., Baran, I., Makowski, K., Turkiewicz, M. (2005). "Cloning, expression, and purification of a recombinant cold-adapted beta-galactosidase from antarctic bacterium *Pseudoalteromonas* sp. 22b." Protein Expr Purif **39**(1): 27-34.
- Cincilla, G., Vidal, D., Pons, M. (2009). "An improved scoring function for suboptimal polar ligand complexes." J Comput Aided Mol Des **23**(3): 143-152.
- Coombs, J. M., Brenchley, J. E. (1999). "Biochemical and phylogenetic analyses of a cold-active beta-galactosidase from the lactic acid bacterium *Carnobacterium piscicola* BA." Appl Environ Microbiol **65**(12): 5443-5450.
- Crippen, G. (2013). Distance Geometry for Realistic Molecular Conformations. Distance Geometry. A. Mucherino, C. Lavor, L. Liberti, N. Maculan, Springer New York: 315-328.
- Čurić, M. (1997). Uloga starter kultura u formiranju arome fermentisanih napitaka, . Kvalitet mleka i fermentisanih proizvoda. Beograd, Jugoslovenski mlekarski simpozijum: 38-40.
- Davies, G., Henrissat, B. (1995). "Structures and mechanisms of glycosyl hydrolases." Structure **3**(9): 853-859.
- Davis, I. W., Murray, L. W., Richardson, J. S., Richardson, D. C. (2004). "MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes." Nucleic Acids Res **32**(Web Server issue): W615-619.
- Davis, I. W., Raha, K., Head, M. S., Baker, D. (2009). "Blind docking of pharmaceutically relevant compounds using RosettaLigand." Protein Sci **18**(9): 1998-2002.
- Dayhoff, M. O., Barker, W. C., Hunt, L. T. (1983). "Establishing homologies in protein sequences." Methods Enzymol **91**: 524-545.
- Degraeve, P., Lemay, P. (1997). "High pressure-induced modulation of the activity and stability of *Escherichia coli* (iacZ) P-galactosidase: Potential applications." Enzyme and Microbial Technology **20**: 550-557
- Demirhan, E., Apar, D., Özbek, B. (2010). "A modelling study on hydrolysis of whey lactose and stability of β -galactosidase." Korean Journal of Chemical Engineering **27**(2): 536-545.
- Di Nola, A., Roccatano, D., Berendsen, H. J. (1994). "Molecular dynamics simulation of the docking of substrates to proteins." Proteins **19**(3): 174-182.

Donate, L. E., Rufino, S. D., Canard, L. H., Blundell, T. L. (1996). "Conformational analysis and clustering of short and medium size loops connecting regular secondary structures: a database for modeling and prediction." Protein Sci **5**(12): 2600-2616.

Drenth, J. (2007). Crystallizing a Protein. Principles of Protein X-Ray Crystallography, Springer New York: 1-20.

Drenth, J. (2007). Theory of X-ray Diffraction by a Crystal. Principles of Protein X-Ray Crystallography, Springer New York: 64-108.

Drenth, J. (2007). X-ray Sources and Detectors. Principles of Protein X-Ray Crystallography, Springer New York: 21-44.

Dunbrack, R. L., Jr. (1999). "Comparative modeling of CASP3 targets using PSI-BLAST and SCWRL." Proteins Suppl **3**: 81-87.

Dunbrack, R. L., Jr. (2002). "Rotamer libraries in the 21st century." Curr Opin Struct Biol **12**(4): 431-440.

Dunbrack, R. L., Jr., Cohen, F. E. (1997). "Bayesian statistical analysis of protein side-chain rotamer preferences." Protein Sci **6**(8): 1661-1681.

Dunbrack, R. L., Jr., Karplus, M. (1994). "Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains." Nat Struct Biol **1**(5): 334-340.

Durbin, R., .. et al. (1998). Biological Sequence Analysis, Cambridge University Press.

Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic Acids Res **32**(5): 1792-1797.

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U., i sar. (2007). "Comparative protein structure modeling using MODELLER." Curr Protoc Protein Sci **Chapter 2**: Unit 2 9.

Fan, H., Mark, A. E. (2004). "Refinement of homology-based protein structures by molecular dynamics simulation techniques." Protein Sci **13**(1): 211-220.

Fechteler, T., Dengler, U., Schomburg, D. (1995). "Prediction of protein three-dimensional structures in insertion and deletion regions: a procedure for searching data bases of representative protein fragments using geometric scoring criteria." J Mol Biol **253**(1): 114-131.

Fernandez-Fuentes, N., Querol, E., Aviles, F. X., Sternberg, M. J., Oliva, B. (2005). "Prediction of the conformation and geometry of loops in globular proteins: testing ArchDB, a structural classification of loops." Proteins **60**(4): 746-757.

- Fersht, A. (1985). Enzyme Structure and Mechanism. New York, , W.H. Freeman & Company, .
- Fischer, D., Eisenberg, D. (1996). "Protein fold recognition using sequence-derived predictions." Protein Sci **5**(5): 947-955.
- Fiser, A., Do, R. K., Sali, A. (2000). "Modeling of loops in protein structures." Protein Sci **9**(9): 1753-1773.
- Fogolari, F., Tosatto, S. C. (2005). "Application of MM/PBSA colony free energy to loop decoy discrimination: toward correlation between energy and root mean square deviation." Protein Sci **14**(4): 889-901.
- Fong, P., McNamara, J. P., Hillier, I. H., Bryce, R. A. (2009). "Assessment of QM/MM Scoring Functions for Molecular Docking to HIV-1 Protease." J Chem Inf Model **49**(4): 913-924.
- Fowler, A. V., Zabin, I. (1983). "Purification, structure, and properties of hybrid beta-galactosidase proteins." J Biol Chem **258**(23): 14354-14358.
- Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., i sar. (2004). "Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy." J Med Chem **47**(7): 1739-1749.
- Frishman, D., Argos, P. (1995). "Knowledge-based protein secondary structure assignment." Proteins **23**(4): 566-579.
- Fuller, R. (1989). "Probiotics in man and animals." J Appl Bacteriol **66**(5): 365-378.
- Gänzle, M. G., Haase, G., Jelen, P. (2008). "Lactose: Crystallization, hydrolysis and value-added derivatives." International Dairy Journal **18**(7): 685-694.
- Gebler, J. C., Aebersold, R., Withers, S. G. (1992). "Glu-537, not Glu-461, is the nucleophile in the active site of (lac Z) beta-galactosidase from Escherichia coli." J Biol Chem **267**(16): 11126-11130.
- Gehlhaar, D. K., Moerder, K. E., Zichi, D., Sherman, C. J., Ogden, R. C., Freer, S. T. (1995). "De novo design of enzyme inhibitors by Monte Carlo ligand generation." J Med Chem **38**(3): 466-472.
- Gill, S. R., Pop, M., Deboy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S., Gordon, J. I., i sar. (2006). "Metagenomic analysis of the human distal gut microbiome." Science **312**(5778): 1355-1359.
- Goldberg, D. E. (1989). Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley Longman Publishing Co., Inc.

- Goodsell, D. S., Morris, G. M., Olson, A. J. (1996). "Automated docking of flexible ligands: applications of AutoDock." J Mol Recognit **9**(1): 1-5.
- Goodsell, D. S., Olson, A. J. (1990). "Automated docking of substrates to proteins by simulated annealing." Proteins **8**(3): 195-202.
- Guex, N., Peitsch, M. C. (1997). "SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling." Electrophoresis **18**(15): 2714-2723.
- Guillotin, L., Lafite, P., Daniellou, R. (2014). "Unraveling the Substrate Recognition Mechanism and Specificity of the Unusual Glycosyl Hydrolase Family 29 BT2192 from *Bacteroides thetaiotaomicron*." Biochemistry **53**(9): 1447-1455.
- Guimarães, P. M. R., Teixeira, J. A., Domingues, L. (2010). "Fermentation of lactose to bio-ethanol by yeasts as part of integrated solutions for the valorisation of cheese whey." Biotechnol Adv **28**(3): 375-384.
- Guyonnet, D., Woodcock, A., Stefani, B., Trevisan, C., Hall, C. (2009). "Fermented milk containing *Bifidobacterium lactis* DN-173 010 improved self-reported digestive comfort amongst a general population of adults. A randomized, open-label, controlled, pilot study." J Dig Dis **10**(1): 61-70.
- Hamilton-Miller, J. M. T. (1999). "Probiotics: a Critical Review." Journal of Antimicrobial Chemotherapy **43**(6): 849.
- Havel, T. F., Kuntz, I. D., Crippen, G. M. (1983). "The combinatorial distance geometry method for the calculation of molecular conformation. I. A new approach to an old problem." J Theor Biol **104**(3): 359-381.
- Havel, T. F., Kuntz, I. D., Crippen, G. M. (1983). "The theory and practice of distance geometry." Bulletin of Mathematical Biology **45**(5): 665-720.
- Hendrickson, W. A. (1991). "Determination of macromolecular structures from anomalous diffraction of synchrotron radiation." Science **254**(5028): 51-58.
- Henikoff, S., Henikoff, J. G. (1992). "Amino acid substitution matrices from protein blocks." Proc Natl Acad Sci U S A **89**(22): 10915-10919.
- Henrissat, B., Bairoch, A. (1996). "Updating the sequence-based classification of glycosyl hydrolases." Biochem J **316** (Pt 2): 695-696.
- Herrchen, M., Legler, G. (1984). "Identification of an essential carboxylate group at the active site of lacZ beta-galactosidase from *Escherichia coli*." Eur J Biochem **138**(3): 527-531.
- Heuser, P., Wohlfahrt, G., Schomburg, D. (2004). "Efficient methods for filtering and ranking fragments for the prediction of structurally variable regions in proteins." Proteins **54**(3): 583-595.

- Hidaka, M., Fushinobu, S., Ohtsu, N., Motoshima, H., Matsuzawa, H., Shoun, H., Wakagi, T. (2002). "Trimeric crystal structure of the glycoside hydrolase family 42 beta-galactosidase from *Thermus thermophilus* A4 and the structure of its complex with galactose." J Mol Biol **322**(1): 79-91.
- Holzappel, W. H., Haberer, P., Geisen, R., Bjorkroth, J., Schillinger, U. (2001). "Taxonomy and important features of probiotic microorganisms in food and nutrition." Am J Clin Nutr **73**(2 Suppl): 365S-373S.
- Holzappel, W. H., Haberer, P., Snel, J., Schillinger, U., Huis in't Veld, J. H. (1998). "Overview of gut flora and probiotics." Int J Food Microbiol **41**(2): 85-101.
- Hoyoux, A., Jennes, I., Dubois, P., Genicot, S., Dubail, F., Francois, J. M., Baise, E., i sar. (2001). "Cold-adapted beta-galactosidase from the Antarctic psychrophile *Pseudoalteromonas haloplanktis*." Appl Environ Microbiol **67**(4): 1529-1535.
- Hu, J. M., Li, H., Cao, L. X., Wu, P. C., Zhang, C. T., Sang, S. L., Zhang, X. Y., i sar. (2007). "Molecular cloning and characterization of the gene encoding cold-active beta-galactosidase from a psychrotrophic and halotolerant *Planococcus* sp. L4." J Agric Food Chem **55**(6): 2217-2224.
- Huber, R. E., Hurlburt, K. L. (1986). "Reversion reactions of beta-galactosidase (*Escherichia coli*)." Arch Biochem Biophys **246**(1): 411-418.
- Hung, M. N., Lee, B. H. (2002). "Purification and characterization of a recombinant beta-galactosidase with transgalactosylation activity from *Bifidobacterium infantis* HL96." Appl Microbiol Biotechnol **58**(4): 439-445.
- Husain, Q. (2010). "Beta galactosidases and their potential applications: a review." Crit Rev Biotechnol **30**(1): 41-62.
- Itsaranuwat, P., Al-Haddad, K. S. H., Robinson, R. K. (2003). "The potential therapeutic benefits of consuming 'health-promoting' fermented dairy products: a brief update." International Journal of Dairy Technology **56**(4): 203-210.
- Jacob, F., Monod, J. (1961). "Genetic regulatory mechanisms in the synthesis of proteins." J Mol Biol **3**: 318-356.
- Jacobson, R. H., Zhang, X. J., DuBose, R. F., Matthews, B. W. (1994). "Three-dimensional structure of beta-galactosidase from *E. coli*." Nature **369**(6483): 761-766.
- Jain, A. N. (2003). "Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine." J Med Chem **46**(4): 499-511.

- Jermin, L. S., Olsen, G. J., Mengersen, K. L., Eastal, S. (1997). "Majority-rule consensus of phylogenetic trees obtained by maximumlikelihood analysis." Molecular Biology and Evolution **14**: 1296–1302
- Jobe, A., Bourgeois, S. (1972). "lac Repressor-operator interaction. VI. The natural inducer of the lac operon." J Mol Biol **69**(3): 397-408.
- Jones, D. T. (1999). "Protein secondary structure prediction based on position-specific scoring matrices." J Mol Biol **292**(2): 195-202.
- Jones, G., Willett, P., Glen, R. C., Leach, A. R., Taylor, R. (1997). "Development and validation of a genetic algorithm for flexible docking." J Mol Biol **267**(3): 727-748.
- Jones, T. A., Thirup, S. (1986). "Using known substructures in protein model building and crystallography." EMBO J **5**(4): 819-822.
- Juajun, O., Nguyen, T. H., Maischberger, T., Iqbal, S., Haltrich, D., Yamabhai, M. (2011). "Cloning, purification, and characterization of beta-galactosidase from *Bacillus licheniformis* DSM 13." Appl Microbiol Biotechnol **89**(3): 645-654.
- Juers, D. H., Heightman, T. D., Vasella, A., McCarter, J. D., Mackenzie, L., Withers, S. G., Matthews, B. W. (2001). "A structural view of the action of *Escherichia coli* (lacZ) beta-galactosidase." Biochemistry **40**(49): 14781-14794.
- Juers, D. H., Jacobson, R. H., Wigley, D., Zhang, X. J., Huber, R. E., Tronrud, D. E., Matthews, B. W. (2000). "High resolution refinement of beta-galactosidase in a new crystal form reveals multiple metal-binding sites and provides a structural basis for alpha-complementation." Protein Sci **9**(9): 1685-1699.
- Juers, D. H., Matthews, B. W., Huber, R. E. (2012). "LacZ beta-galactosidase: structure and function of an enzyme of historical and molecular biological importance." Protein Sci **21**(12): 1792-1807.
- Juers, D. H., Rob, B., Dugdale, M. L., Rahimzadeh, N., Giang, C., Lee, M., Matthews, B. W., i sar. (2009). "Direct and indirect roles of His-418 in metal binding and in the activity of beta-galactosidase (*E. coli*)." Protein Sci **18**(6): 1281-1292.
- Kabsch, W., Sander, C. (1983). "Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features." Biopolymers **22**(12): 2577-2637.
- Kim, C. S., Ji, E. S., Oh, D. K. (2003). "Expression and characterization of *Kluyveromyces lactis* beta-galactosidase in *Escherichia coli*." Biotechnol Lett **25**(20): 1769-1774.

- Kim, C. S., Ji, E. S., Oh, D. K. (2004). "A new kinetic model of recombinant beta-galactosidase from *Kluyveromyces lactis* for both hydrolysis and transgalactosylation reactions." Biochem Biophys Res Commun **316**(3): 738-743.
- Kirkpatrick, S., Gelatt, C. D., Jr., Vecchi, M. P. (1983). "Optimization by simulated annealing." Science **220**(4598): 671-680.
- Klaenhammer, T. R., Kullen, M. J. (1999). "Selection and design of probiotics." Int J Food Microbiol **50**(1-2): 45-57.
- Knegtel, R. M., Kuntz, I. D., Oshiro, C. M. (1997). "Molecular docking to ensembles of protein structures." J Mol Biol **266**(2): 424-440.
- Kochev, N. T., Paskaleva, V. H., Jeliaskova, N. (2013). "Ambit-Tautomer: An Open Source Tool for Tautomer Generation." Molecular Informatics **32**(5-6): 481-504.
- Koehl, P., Levitt, M. (1999). "A brighter future for protein structure prediction." Nat Struct Biol **6**(2): 108-111.
- Koolman, J., Röhm, K. H. (2005). Color Atlas of Biochemistry, Thieme.
- Krammer, A., Kirchhoff, P. D., Jiang, X., Venkatachalam, C. M., Waldman, M. (2005). "LigScore: a novel scoring function for predicting binding affinities." J Mol Graph Model **23**(5): 395-407.
- Krogh, A., Brown, M., Mian, I. S., Sjolander, K., Haussler, D. (1994). "Hidden Markov models in computational biology. Applications to protein modeling." J Mol Biol **235**(5): 1501-1531.
- Kuby, S. A., Lardy, H. A. (1953). "Purification and Kinetics of β -D-Galactosidase from *Escherichia coli*, Strain K-121,2." Journal of the American Chemical Society **75**(4): 890-896.
- Kumar, M., Bhasin, M., Natt, N. K., Raghava, G. P. (2005). "BhairPred: prediction of beta-hairpins in a protein from multiple alignment information using ANN and SVM techniques." Nucleic Acids Res **33**(Web Server issue): W154-159.
- Lang, P. T., Brozell, S. R., Mukherjee, S., Pettersen, E. F., Meng, E. C., Thomas, V., Rizzo, R. C., i sar. (2009). "DOCK 6: combining techniques to model RNA-small molecule complexes." RNA **15**(6): 1219-1230.
- Lazar, C., Kluczyk, A., Kiyota, T., Konishi, Y. (2004). "Drug evolution concept in drug design: 1. Hybridization method." J Med Chem **47**(27): 6973-6982.
- Leach, A. R. (2001). Molecular Modelling: Principles and Applications, second edition. Harlow, UK, Pearson Education EMA.

- Lee, J., Ametani, A., Enomoto, A., Sato, Y., Motoshima, H., Ike, F., Kaminogawa, S. (1993). "Screening for the Immunopotentiating Activity of Food Microorganisms and Enhancement of the Immune Response by *Bifidobacterium adolescentis* M101-4." Bioscience, biotechnology, and biochemistry **57**(12): 2127-2132.
- Levitt, M., Gerstein, M., Huang, E., Subbiah, S., Tsai, J. (1997). "Protein folding: the endgame." Annu Rev Biochem **66**: 549-579.
- Li, H., Tejero, R., Monleon, D., Bassolino-Klimas, D., Abate-Shen, C., Bruccoleri, R. E., Montelione, G. T. (1997). "Homology modeling using simulated annealing of restrained molecular dynamics and conformational search calculations with CONGEN: application in predicting the three-dimensional structure of murine homeodomain Msx-1." Protein Sci **6**(5): 956-970.
- Li, W., Xiang, X., Tang, S., Hu, B., Tian, L., Sun, Y., Ye, H., i sar. (2009). "Effective enzymatic synthesis of lactosucrose and its analogues by beta-D-galactosidase from *Bacillus circulans*." J Agric Food Chem **57**(9): 3927-3933.
- Liang, S., Grishin, N. V. (2002). "Side-chain modeling with an optimized scoring function." Protein Sci **11**(2): 322-331.
- Liu, J., Rost, B. (2002). "Target space for structural genomics revisited." Bioinformatics **18**(7): 922-933.
- Liu, J., Rost, B. (2004). "CHOP: parsing proteins into structural domains." Nucleic Acids Res **32**(Web Server issue): W569-571.
- Lo, S., Dugdale, M. L., Jeerh, N., Ku, T., Roth, N. J. (2010). "Studies of Glu-416 variants of b-galactosidase (*E. coli*) show that the active site Mg²⁺ is not important for structure and indicate that the main role of Mg²⁺ is to mediate optimization of active site chemistry." Protein. Journal **29**: 26392-.
- Loquasto, J. R., Barrangou, R., Dudley, E. G., Stahl, B., Chen, C., Roberts, R. F. (2013). "*Bifidobacterium animalis* subsp. *lactis* ATCC 27673 is a genomically unique strain within its conserved subspecies." Appl Environ Microbiol **79**(22): 6903-6910.
- Lorber, D. M., Shoichet, B. K. (1998). "Flexible ligand docking using conformational ensembles." Protein Sci **7**(4): 938-950.
- Lovell, S. C., Word, J. M., Richardson, J. S., Richardson, D. C. (2000). "The penultimate rotamer library." Proteins **40**(3): 389-408.
- Lowe, J. (2013). Introduction to X-ray crystallography. Cambridge.

Maksimainen, M., Hakulinen, N., Kallio, J. M., Timoharju, T., Turunen, O., Rouvinen, J. (2011). "Crystal structures of *Trichoderma reesei* beta-galactosidase reveal conformational changes in the active site." J Struct Biol **174**(1): 156-163.

Maksimainen, M., Paavilainen, S., Hakulinen, N., Rouvinen, J. (2012). "Structural analysis, enzymatic characterization, and catalytic mechanisms of beta-galactosidase from *Bacillus circulans* sp. *alkalophilus*." FEBS J **279**(10): 1788-1798.

Maksimainen, M. M., Lampio, A., Mertanen, M., Turunen, O., Rouvinen, J. (2013). "The crystal structure of acidic beta-galactosidase from *Aspergillus oryzae*." Int J Biol Macromol **60**: 109-115.

Marsden, P. M., Puvanendrapillai, D., Mitchell, J. B., Glen, R. C. (2004). "Predicting protein-ligand binding affinities: a low scoring game?" Org Biomol Chem **2**(22): 3267-3273.

Martínez-Villaluenga, C., Cardelle-Cobas, A., Corzo, N., Olano, A., Villamiel, M. (2008). "Optimization of conditions for galactooligosaccharide synthesis during lactose hydrolysis by β -galactosidase from *Kluyveromyces lactis* (Lactozym 3000 L HP G)." Food Chemistry **107**(1): 258-264.

McCarter, J. D., Burgoyne, D. L., Miao, S., Zhang, S., Callahan, J. W., Withers, S. G. (1997). "Identification of Glu-268 as the catalytic nucleophile of human lysosomal beta-galactosidase precursor by mass spectrometry." J Biol Chem **272**(1): 396-400.

McCarter, J. D., Stephen Withers, G. (1994). "Mechanisms of enzymatic glycoside hydrolysis." Curr Opin Struct Biol **4**(6): 885-892.

McDonough, F. E., Hitchins, A. D., Wong, N. P., Wells, P., Bodwell, C. E. (1987). "Modification of sweet acidophilus milk to improve utilization by lactose-intolerant persons." Am J Clin Nutr **45**(3): 570-574.

McMurry, J., Fay, R. (2003). Chemistry. New Jersey, USA, Prentice Hall.

Meile, L., Ludwig, W., Rueger, U., Gut, C., Kaufmann, P., Dasen, G., Wenger, S., i sar. (1997). "*Bifidobacterium lactis* sp. nov., a Moderately Oxygen Tolerant Species Isolated from Fermented Milk." Systematic and Applied Microbiology **20**(1): 57-64.

Merlitz, H., Burghardt, B., Wenzel, W. (2003). "Application of the stochastic tunneling method to high throughput database screening." Chemical Physics Letters **370**(1-2): 68-73.

Milanović, S. (1997). Fermentisani mlečni napici za specijalne namene. Kvalitet mleka i fermentisanih proizvoda. Zlatibor, Jugoslovenski mlekarski simpozijum: 49-55.

Miller, M. D., Kearsley, S. K., Underwood, D. J., Sheridan, R. P. (1994). "FLOG: a system to select 'quasi-flexible' ligands complementary to a receptor of known three-dimensional structure." J Comput Aided Mol Des **8**(2): 153-174.

Morgan, X. C., Segata, N., Huttenhower, C. (2013). "Biodiversity and functional genomics in the human microbiome." Trends Genet **29**(1): 51-58.

Müller-Hill, B. (1996). The Lac Operon: A Short History of a Genetic Paradigm. Berlin, New York, Walter de Gruyter.

Murray, R. K., Granner, D. K., Mayes, P. A., Rodwell, V. W. (2006). Harper's Illustrated Biochemistry, McGraw-Hill Medical Publishing Division.

Nakagawa, T., Fujimoto, Y., Ikehata, R., Miyaji, T., Tomizuka, N. (2006). "Purification and molecular characterization of cold-active beta-galactosidase from *Arthrobacter psychrolactophilus* strain F2." Appl Microbiol Biotechnol **72**(4): 720-725.

Nakanishi, K., Matsuno, R., Torii, K., Yamamoto, K., Kamikubo, T. (1983). "Properties of immobilized β -d-galactosidase from *Bacillus circulans*." Enzyme and Microbial Technology **5**(2): 115-120.

Nguyen, T. H., Splechna, B., Krasteva, S., Kneifel, W., Kulbe, K. D., Divne, C., Haltrich, D. (2007). "Characterization and molecular cloning of a heterodimeric beta-galactosidase from the probiotic strain *Lactobacillus acidophilus* R22." FEMS Microbiol Lett **269**(1): 136-144.

Nguyen, T. H., Splechna, B., Yamabhai, M., Haltrich, D., Peterbauer, C. (2007). "Cloning and expression of the beta-galactosidase genes from *Lactobacillus reuteri* in *Escherichia coli*." J Biotechnol **129**(4): 581-591.

Ohto, U., Usui, K., Ochi, T., Yuki, K., Satow, Y., Shimizu, T. (2012). "Crystal structure of human beta-galactosidase: structural basis of Gm1 gangliosidosis and morquio B diseases." J Biol Chem **287**(3): 1801-1812.

Ohtsu, N., Motoshima, H., Goto, K., Tsukasaki, F., Matsuzawa, H. (1998). "Thermostable β -Galactosidase from an Extreme Thermophile, *Thermus* sp. A4: Enzyme Purification and Characterization, and Gene Cloning and Sequencing." Bioscience, biotechnology, and biochemistry **62**(8): 1539-1545.

Okonechnikov, K., Golosova, O., Fursov, M., team, U. (2012). "Unipro UGENE: a unified bioinformatics toolkit." Bioinformatics **28**(8): 1166-1167.

Oostenbrink, C., Villa, A., Mark, A. E., van Gunsteren, W. F. (2004). "A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6." J Comput Chem **25**(13): 1656-1676.

Pan, Q., Zhu, J., Liu, L., Cong, Y., Hu, F., Li, J., Yu, X. (2010). "Functional identification of a putative beta-galactosidase gene in the special lac gene cluster of *Lactobacillus acidophilus*." Curr Microbiol **60**(3): 172-178.

- Pang, Y. P., Perola, E., Xu, K., Prendergast, F. G. (2001). "EUDOC: a computer program for identification of drug interaction sites in macromolecules and drug leads from chemical databases." J Comput Chem **22**(15): 1750-1771.
- Pardee, A. B., Jacob, F., Monod, J. (1959). "The Genetic Control and Cytoplasmic Expression of "Inducibility" in the Synthesis of 8-galactosidase by *E. coli* ." J Mol Biol **1**: 165-178.
- Park, J., Karplus, K., Barrett, C., Hughey, R., Haussler, D., Hubbard, T., Chothia, C. (1998). "Sequence comparisons using multiple sequences detect three times as many remote homologues as pairwise methods." J Mol Biol **284**(4): 1201-1210.
- Pauling, L., Corey, R. B. (1951). "Configurations of Polypeptide Chains With Favored Orientations Around Single Bonds: Two New Pleated Sheets." Proc Natl Acad Sci U S A **37**(11): 729-740.
- Peitsch, M. C. (1996). "ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling." Biochem Soc Trans **24**(1): 274-279.
- Pereira-Rodriguez, A., Fernandez-Leiro, R., Gonzalez-Siso, M. I., Cerdan, M. E., Becerra, M., Sanz-Aparicio, J. (2012). "Structural basis of specificity in tetrameric *Kluyveromyces lactis* beta-galactosidase." J Struct Biol **177**(2): 392-401.
- Peterson, R. W., Dutton, P. L., Wand, A. J. (2004). "Improved side-chain prediction accuracy using an ab initio potential energy function and a very large rotamer library." Protein Sci **13**(3): 735-751.
- Pivarnik, L. F., Senecal, A. G., Rand, A. G. (1995). "Hydrolytic and transgalactosylic activities of commercial beta-galactosidase (lactase) in food processing." Adv Food Nutr Res **38**: 1-102.
- Platzer, K. E., Momany, F. A., Scheraga, H. A. (1972). "Conformational energy calculations of enzyme-substrate interactions. II. Computation of the binding energy for substrates in the active site of -chymotrypsin." Int J Pept Protein Res **4**(3): 201-219.
- Przybylski, D., Rost, B. (2002). "Alignments grow, secondary structure prediction improves." Proteins **46**(2): 197-205.
- Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S., Medzhitov, R. (2004). "Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis." Cell **118**(2): 229-241.
- Ramachandran, G. N., Ramakrishnan, C., Sasisekharan, V. (1963). "Stereochemistry of polypeptide chain configurations." J Mol Biol **7**: 95-99.
- Rao, M. V. R., Dutta, S. M. (1981). "Purification and Properties of Beta-galactosidase from *Streptococcus thermophilus*." Journal of Food Science **46**(5): 1419-1423.

- Rarey, M., Kramer, B., Lengauer, T. (1999). "The particle concept: placing discrete water molecules during protein-ligand docking predictions." Proteins **34**(1): 17-28.
- Reddy, B. S., Rivenson, A. (1993). "Inhibitory effect of Bifidobacterium longum on colon, mammary, and liver carcinogenesis induced by 2-amino-3-methylimidazo[4,5-f]quinoline, a food mutagen." Cancer Res **53**(17): 3914-3918.
- Rhimi, M., Aghajari, N., Jaouadi, B., Juy, M., Boudebouze, S., Maguin, E., Haser, R., i sar. (2009). "Exploring the acidotolerance of beta-galactosidase from Lactobacillus delbrueckii subsp. bulgaricus: an attractive enzyme for lactose bioconversion." Res Microbiol **160**(10): 775-784.
- Rhimi, M., Boisson, A., Dejob, M., Boudebouze, S., Maguin, E., Haser, R., Aghajari, N. (2010). "Efficient bioconversion of lactose in milk and whey: immobilization and biochemical characterization of a β -galactosidase from the dairy Streptococcus thermophilus LMD9 strain." Research in Microbiology **161**(7): 515-525.
- Richards, F. M., Kundrot, C. E. (1988). "Identification of structural motifs from protein coordinate data: secondary structure and first-level supersecondary structure." Proteins **3**(2): 71-84.
- Rizzardini, G., Eskesen, D., Calder, P. C., Capetti, A., Jespersen, L., Clerici, M. (2012). "Evaluation of the immune benefits of two probiotic strains Bifidobacterium animalis ssp. lactis, BB-12(R) and Lactobacillus paracasei ssp. paracasei, L. casei 431(R) in an influenza vaccination model: a randomised, double-blind, placebo-controlled study." Br J Nutr **107**(6): 876-884.
- Rodriguez, A. P., Leiro, R. F., Trillo, M. C., Cerdan, M. E., Siso, M. I., Becerra, M. (2006). "Secretion and properties of a hybrid Kluyveromyces lactis-Aspergillus niger beta-galactosidase." Microb Cell Fact **5**: 41.
- Rohl, C. A., Strauss, C. E., Misura, K. M., Baker, D. (2004). "Protein structure prediction using Rosetta." Methods Enzymol **383**: 66-93.
- Rojas, A. L., Nagem, R. A., Neustroev, K. N., Arand, M., Adamska, M., Eneyskaya, E. V., Kulminskaya, A. A., i sar. (2004). "Crystal structures of beta-galactosidase from Penicillium sp. and its complex with galactose." J Mol Biol **343**(5): 1281-1292.
- Rost, B. (1995). "TOPITS: threading one-dimensional predictions into three-dimensional structures." Proc Int Conf Intell Syst Mol Biol **3**: 314-321.
- Rost, B. (2005). How to Use Protein 1- D Structure Predicted by PROFphd: 875-901.
- Rost, B., Sander, C., Schneider, R. (1994). "PHD--an automatic mail server for protein secondary structure prediction." Comput Appl Biosci **10**(1): 53-60.

Russell, R. B., Copley, R. R., Barton, G. J. (1996). "Protein Fold Recognition by Mapping Predicted Secondary Structures." J Mol Biol **259**(3): 349-365.

S.A., I., Bezkorovainy, A. (1993). "Inhibition of Escherichia coli by Bifidobacteria." Food Protein **56**: 713–715.

Sahasrabudhe, P. V., Tejero, R., Kitao, S., Furuichi, Y., Montelione, G. T. (1998). "Homology modeling of an RNP domain from a human RNA-binding protein: Homology-constrained energy optimization provides a criterion for distinguishing potential sequence alignments." Proteins **33**(4): 558-566.

Sali, A., Blundell, T. L. (1993). "Comparative protein modelling by satisfaction of spatial restraints." J Mol Biol **234**(3): 779-815.

Sanchez, R., Sali, A. (1997). "Evaluation of comparative protein structure modeling by MODELLER-3." Proteins Suppl **1**: 50-58.

Sandak, B., Nussinov, R., Wolfson, H. J. (1994). 3-D flexible docking of molecules. In Proc. 1st IEEE Workshop on Shape and Pattern Recognition in Computational Biology. Seattle: 41–54.

Sanders, M. E. (1998). "Overview of Functional Foods: Emphasis on Probiotic Bacteria." International Dairy Journal **8**(5–6): 341-347.

Sasisekharan, V., Ponnuswamy, P. K. (1970). "Backbone and side-chain conformations of amino acids and amino acid residues in peptides." Biopolymers **9**(10): 1249-1256.

Sasisekharan, V., Ponnuswamy, P. K. (1971). "Studies on the conformation of amino acids. X. Conformations of norvalyl, leucyl and aromatic side groups in a dipeptide unit." Biopolymers **10**(3): 583-592.

Schneider, R. (1989). Sekundärstrukturvorhersage von Proteinen unter Berücksichtigung von Tertiärstrukturaspekten
Department of Biology. Heidelberg, University of Heidelberg. **Bsc Thesis**.

Sen, T. Z., Jernigan, R. L., Garnier, J., Kloczkowski, A. (2005). "GOR V server for protein secondary structure prediction." Bioinformatics **21**(11): 2787-2788.

Shen, M. Y., Sali, A. (2006). "Statistical potential for assessment and prediction of protein structures." Protein Sci **15**(11): 2507-2524.

Shepherd, A. J., Gorse, D., Thornton, J. M. (1999). "Prediction of the location and type of beta-turns in proteins using neural networks." Protein Science **8**: 1045–1055.

Sibanda, B. L., Thornton, J. M. (1991). Conformation of β hairpins in protein structures: Classification and diversity in homologous structures. Methods Enzymol. J. L. John, Academic Press. **Volume 202**: 59-82.

- Simons, K. T., Kooperberg, C., Huang, E., Baker, D. (1997). "Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions." J Mol Biol **268**(1): 209-225.
- Skalova, T., Dohnalek, J., Spiwok, V., Lipovova, P., Vondrackova, E., Petrokova, H., Duskova, J., i sar. (2005). "Cold-active beta-galactosidase from *Arthrobacter* sp. C2-2 forms compact 660 kDa hexamers: crystal structure at 1.9A resolution." J Mol Biol **353**(2): 282-294.
- Sklenar, H., Etchebest, C., Lavery, R. (1989). "Describing protein structure: a general algorithm yielding complete helicoidal parameters and a unique overall axis." Proteins **6**(1): 46-60.
- Sonnenburg, J. L., Xu, J., Leip, D. D., Chen, C. H., Westover, B. P., Weatherford, J., Buhler, J. D., i sar. (2005). "Glycan foraging in vivo by an intestine-adapted bacterial symbiont." Science **307**(5717): 1955-1959.
- Sorensen, H. P., Porsgaard, T. K., Kahn, R. A., Stougaard, P., Mortensen, K. K., Johnsen, M. G. (2006). "Secreted beta-galactosidase from a *Flavobacterium* sp. isolated from a low-temperature environment." Appl Microbiol Biotechnol **70**(5): 548-557.
- Sousa, S. F., Ribeiro, A. J., Coimbra, J. T., Neves, R. P., Martins, S. A., Moorthy, N. S., Fernandes, P. A., i sar. (2013). "Protein-ligand docking in the new millennium--a retrospective of 10 years in the field." Curr Med Chem **20**(18): 2296-2314.
- Stappenbeck, T. S., Hooper, L. V., Gordon, J. I. (2002). "Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells." Proc Natl Acad Sci U S A **99**(24): 15451-15455.
- Sutcliffe, M. J., Haneef, I., Carney, D., Blundell, T. L. (1987). "Knowledge based modelling of homologous proteins, Part I: Three-dimensional frameworks derived from the simultaneous superposition of multiple structures." Protein Eng **1**(5): 377-384.
- Sutcliffe, M. J., Hayes, F. R., Blundell, T. L. (1987). "Knowledge based modelling of homologous proteins, Part II: Rules for the conformations of substituted sidechains." Protein Eng **1**(5): 385-392.
- Tamime, A. Y. (2008). Fermented Milks, Wiley.
- Tamime, A. Y., Robinson, R. K. (1999). Yoghurt: Science and Technology, CRC Press.
- Tamime, A. Y., Robinson, R. K. (2007). Tamime and Robinson's Yoghurt: Science and Technology, Elsevier Science.
- ten Brink, T., Exner, T. E. (2009). "Influence of protonation, tautomeric, and stereoisomeric states on protein-ligand docking results." J Chem Inf Model **49**(6): 1535-1546.

- Tenu, J. P., Viratelle, O. M., Garnier, J., Yon, J. (1971). "pH dependence of the activity of beta-galactosidase from *Escherichia coli*." Eur J Biochem **20**(3): 363-370.
- Torres, D. P. M., Gonçalves, M. d. P. F., Teixeira, J. A., Rodrigues, L. R. (2010). "Galacto-Oligosaccharides: Production, Properties, Applications, and Significance as Prebiotics." Comprehensive Reviews in Food Science and Food Safety **9**(5): 438-454.
- Tran, L. S., Szabo, L., Fulop, L., Orosz, L., Sik, T., Holczinger, A. (1998). "Isolation of a beta-galactosidase-encoding gene from *Bacillus licheniformis*: purification and characterization of the recombinant enzyme expressed in *Escherichia coli*." Curr Microbiol **37**(1): 39-43.
- Ullmann, A. (1992). "Complementation in beta-galactosidase: from protein structure to genetic engineering." Bioessays **14**(3): 201-205.
- Ullmann, A., Jacob, F., Monod, J. (1967). "Characterization by in vitro complementation of a peptide corresponding to an operator-proximal segment of the beta-galactosidase structural gene of *Escherichia coli*." J Mol Biol **24**(2): 339-343.
- van Gunsteren, W. F., Hünenberger, P. H., Mark, A. E., Smith, P. E., Tironi, I. G. (1995). "Computer simulation of protein motion." Computer Physics Communications **91**(1-3): 305-319.
- Vitkup, D., Melamud, E., Moulton, J., Sander, C. (2001). "Completeness in structural genomics." Nat Struct Biol **8**(6): 559-566.
- Vriend, G. (1990). "WHAT IF: a molecular modeling and drug design program." J Mol Graph **8**(1): 52-56, 29.
- Wallenfels, K., Weil, R. (1972). b-Galactosidase. The Enzymes. P. D. Boyer. New York, Academic Press., **7**: 617-663.
- Wallner, B., Elofsson, A. (2005). "All are not equal: a benchmark of different homology modeling programs." Protein Sci **14**(5): 1315-1327.
- Wang, J., Sheng, X., Zhao, Y., Liu, Y., Liu, C. (2012). "QM/MM investigation on the catalytic mechanism of *Bacteroides thetaiotaomicron* alpha-glucosidase BtGH97a." Biochim Biophys Acta **1824**(5): 750-758.
- Wang, J., Shoemaker, N. B., Wang, G. R., Salyers, A. A. (2000). "Characterization of a *Bacteroides mobilizable* transposon, NBU2, which carries a functional lincomycin resistance gene." J Bacteriol **182**(12): 3559-3571.
- Wang, Q., Graham, R. W., Trimbur, D., Warren, R. A. J., Withers, S. G. (1994). "Changing Enzymic Reaction Mechanisms by Mutagenesis: Conversion of a Retaining Glucosidase to an Inverting Enzyme." Journal of the American Chemical Society **116**(25): 11594-11595.

- Weber, I. T. (1990). "Evaluation of homology modeling of HIV protease." *Proteins* **7**(2): 172-184.
- Weber, I. T., Miller, M., Jaskolski, M., Leis, J., Skalka, A. M., Wlodawer, A. (1989). "Molecular modeling of the HIV-1 protease and its substrate binding site." *Science* **243**(4893): 928-931.
- Wexler, H. M. (2007). "Bacteroides: the good, the bad, and the nitty-gritty." *Clin Microbiol Rev* **20**(4): 593-621.
- Wong, P., Gladney, S., Keasling, J. D. (1997). "Mathematical model of the lac operon: inducer exclusion, catabolite repression, and diauxic growth on glucose and lactose." *Biotechnol Prog* **13**(2): 132-143.
- Xu, J., Bjursell, M. K., Himrod, J., Deng, S., Carmichael, L. K., Chiang, H. C., Hooper, L. V., et al. (2003). "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis." *Science* **299**(5615): 2074-2076.
- Yuriev, E., Agostino, M., Ramsland, P. A. (2011). "Challenges and advances in computational docking: 2009 in review." *J Mol Recognit* **24**(2): 149-164.
- Zvelebil, M. J., Barton, G. J., Taylor, W. R., Sternberg, M. J. (1987). "Prediction of protein secondary structure and active sites using the alignment of homologous sequences." *J Mol Biol* **195**(4): 957-961.