

2014

UNIVERZA NA PRIMORSKEM
FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN
INFORMACIJSKE TEHNOLOGIJE

ZAKLJUČNA NALOGA

ZAKLJUČNA NALOGA

OPREDELITEV STRUKTUR IN FUNKCIJ IZOTIPOV
HUMANEGA SERUMSKEGA AMILOIDA A Z
BIOINFORMATIČNIMI PODATKOVNIMI BAZAMI IN
ORODJI

KOPITAR

LEON KOPITAR

UNIVERZA NA PRIMORSKEM
FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN
INFORMACIJSKE TEHNOLOGIJE

Zaključna naloga

**Opredelitev struktur in funkcij izotipov humanega serumskega amiloida
A z bioinformatičnimi podatkovnimi bazami in orodji**

(Structural And Functional Characterization of Human Serum Amyloid A Isotypes Using
Bioinformatic Databases and Tools)

Ime in priimek: Leon Kopitar
Študijski program: Bioinformatika
Mentor: doc. dr. Snežna Sodin Šemrl

Koper, avgust 2014

Ključna dokumentacijska informacija

Ime in PRIIMEK: Leon KOPITAR

Naslov zaključne naloge:

Opredelitev Struktur in Funkcij Izotipov Humanega Serumskega Amiloida A z Bioinformatičnimi Podatkovnimi Bazami in Orodji

Kraj : Koper

Leto: 2014

Število listov: 59

Število slik: 25

Število tabel: 16

Število referenc: 52

Mentor: doc. dr. Snežna Sodin Šemrl

Ključne besede: Serumski amiloid A, bioinformatična orodja, strukture in funkcije proteinov, SAA izotipi

Izvleček:

Serumski amiloid A je eden izmed glavnih akutno faznih proteinov. Tako s strukturnega kot funkcionalnega vidika proteina je še veliko neznank in nejasnosti. V zaključni nalogi s pomočjo bioinformatičnih podatkovnih baz in orodij poskušamo pridobiti čimveč podatkov in informacij. Določimo strukturne razlike izotipov serumskega amiloida A1-4 na vseh nivojih, proteinske interakcije SAA1-4 in njihovo celično lokalizacijo. Poleg tega primerjamo aminokislinske sestave SAA1-4 in njihove post-translacijske spremembe. Povezave strukturnih karakteristik in potencialnih funkcij SAA1-4 so preliminarno nakazane.

Key word documentation

Name and SURNAME: Leon KOPITAR

Title of the final project paper:

Structural and Functional Characterization of Human Serum Amyloid A Isotypes Using Bioinformatic Databases and Tools

Place: Koper

Year: 2014

Number of pages: 59

Number of figures: 25

Number of tables: 16

Number of references: 52

Mentor: Assist. Prof. Snežna Sodin Šemrl, PhD

Keywords: Serum amyloid A, bioinformatic tools, structures and functions of proteins, SAA isotypes

Abstract:

Serum amyloid A is one of the two major acute phase proteins. From the structural and functional aspects of the protein there remains quite a few unclear and unresolved issues. In the final thesis we try to gain as much information as possible using bioinformatic databases and tools. We determine structural differences of SAA isotypes at each level, protein interactions of SAA1-4 and their localization. We also compare the amino acid composition of SAA1-4, as well as their post-translational modifications. Connections between structural characteristics and potential functions of SAA1-4 are preliminarily explored.

Zahvala

Največja zahvala gre mentorici doc. dr. Snežni Sodin-Šemrl za prijazno, potrpežljivo in strokovno usmerjanje pri izdelavi zaključne naloge.

Kadar sem bil v dilemi, mi je bila vselej pripravljena podrobno razjasniti kakršnokoli nejasnost.

Zahvaljujem se doc. dr. Martinu Milaniču za strukturni pregled naloge in podajanju komentarjev.

Zahvaljujem se Katji Lakoti in Katjuši Mrak-Poljšak za podana strokovna mnenja in popravke.

Zahvaljujem se tudi družini za podporo in pomoč v vseh letih študija.

Kazalo vsebine

1	UVOD	1
1.1	Serumski amiloid A	1
1.1.1	Lastnosti in pomembnost proteina.....	1
1.1.2	Sinteza serumskega amiloida A.....	2
1.2	Izotipi SAA	2
2	NAMEN DELA	6
3	METODE	7
4	REZULTATI IN DISKUSIJA	11
4.1	Primarne strukture izotipov serumskega amiloida A.....	11
4.1.1	Primarna struktura SAA1	11
4.1.2	Primarna struktura SAA2	12
4.1.3	Primarna struktura SAA3	13
4.1.4	Primarna struktura SAA4	14
4.1.5	Primerjave aminokislinskih zaporedij izotipov SAA (med seboj)	15
4.2	Sekundarne strukture izotipov serumskega amiloida A	19
4.2.4	Sekundarna struktura SAA1	19
4.2.5	Sekundarna struktura SAA2	22
4.2.6	Sekundarna struktura SAA3	24
4.2.7	Sekundarna struktura SAA4	26
4.2.8	Primerjave sekundarnih struktur izotipov SAA.....	28
4.3	Terciarnne strukture izotipov serumskega amiloida A	29
4.3.1	Sekundarne strukture pomembne za terciarno strukturo	30
4.4	Kvartarne strukture serumskega amiloida A	32
4.5	Funkcionalna napoved SAA iz struktur.....	32
4.5.1	Primerjava aminokislinske sestave izotipov SAA.....	32
4.5.2	Primerjava post-translacijskih sprememb izotipov SAA	34
4.5.3	Proteinske interakcije SAA1-4	36
4.5.4	Lokalizacija SAA1-4	38
4.5.5	Potencialne funkcije SAA1-4	40
5	ZAKLJUČEK.....	43
6	LITERATURA IN VIRI	44

Kazalo preglednic

Tabela 1: Aminokislinska sestava SAA1	11
Tabela 2: Aminokislinska sestava SAA2	12
Tabela 3: Aminokislinska sestava SAA3	13
Tabela 4: Aminokislinska sestava SAA4	14
Tabela 5: Deleži podobnosti in identičnosti izotipov SAA med seboj (v %).....	19
Tabela 6: Deleži sekundarnih struktur v SAA1 v odstotkih (%).....	21
Tabela 7: Deleži sekundarnih struktur v SAA2 v odstotkih (%).....	23
Tabela 8: Deleži sekundarnih struktur v serumskem amiloidu A3 v odstotkih (%)	25
Tabela 9: Deleži sekundarnih struktur v serumskem amiloidu A4 v odstotkih (%)	27
Tabela 10: Delež sekundarnih struktur za posamezen izotip serumskega amiloida A v odstotkih (%)	28
Tabela 11: Primerjava aminokislinskih sestav za posamezen izotip SAA.....	33
Tabela 12: Postranslacijske modifikacije posameznih izotipov SAA.	36
Tabela 13: Proteinske interakcije z izotipi SAA1-4	37
Tabela 14: Podatki o lokalizacijah SAA1-4 s pomočjo portalov/podatkovnih baz in napovedovalnih orodij	39
Tabela 15: Mesta cepitve v izotipih SAA, verjetnosti transporta SAA do MT in mesta levcinsko bogatih jedr izvoznih signalov v zaporedju SAA	40
Tabela 16: Že znane in potencialne funkcije SAA1-4.....	41

Kazalo slik in grafikonov

Slika 1: Ekspresija mRNA SAA1 in SAA2 v človeških tkivih.....	3
Slika 2: Ekspresija mRNA SAA3 v človeških tkivih.....	4
Slika 3: Ekspresija mRNA SAA4 v človeških tkivih.....	4
Slika 4: Lokalna poravnava izotipov SAA1 in SAA2.....	15
Slika 5: Lokalna poravnava izotipov SAA1 in SAA3.....	15
Slika 6: Lokalna poravnava izotipov SAA1 in SAA4.....	16
Slika 7: Lokalna poravnava izotipov SAA2 in SAA3.....	16
Slika 8: Lokalna poravnava izotipov SAA2 in SAA4.....	17
Slika 9: Lokalna poravnava izotipov SAA3 in SAA4.....	17
Slika 10: Poravnava zaporedij SAA1-4 ena pod drugo.....	18
Slika 11: Odsek zaporedja SAA4 (CYLFGNSS) prisoten v proteinih različnih organizmov	18
Slika 12: Podatki o sekundarni strukturi humanega SAA1 s spletnim orodjem YASPIN ..	20
Slika 13: Podatki o sekundarni strukturi humanega SAA1 s spletnim orodjem SOPMA...	21
Slika 14: Podatki o sekundarni strukturi humanega SAA2 s spletnim orodjem YASPIN ..	22
Slika 15: Podatki o sekundarni strukturi humanega SAA2 s spletnim orodjem SOPMA...	23
Slika 16: Podatki o sekundarni strukturi humanega SAA3 s spletnim orodjem YASPIN ..	24
Slika 17: Podatki o sekundarni strukturi humanega SAA3 s spletnim orodjem SOPMA...	25
Slika 18: Podatki o sekundarni strukturi humanega SAA4 s spletnim orodjem YASPIN ..	26
Slika 19: Podatki o sekundarni strukturi humanega SAA4 s spletnim orodjem SOPMA...	27
Slika 20: Pojavnost orodij za napovedovanje sekundarnih struktur.....	28
Slika 21: Model monomera človeškega SAA1.....	29
Slika 22: Sekundarna struktura SAA1 z grafično razvidnimi alfa vijačnicami.....	30
Slika 23: Sekundarna struktura SAA2 z grafično razvidnimi alfa vijačnicami.....	31
Slika 24: Sekundarna struktura SAA3 z grafično razvidnimi alfa vijačnicami.....	31
Slika 25: Sekundarna struktura SAA4 z grafično razvidnimi alfa vijačnicami.....	32

Seznam kratic

AA amiloidoza	Amiloid A amiloidoza
AK	Aminokislina
AP-1	<i>Activator protein-1</i> / aktivatorski protein-1
APCS	Gen za Serumski amiloid protein (SAP)
A-SAA	Akutno fazni serumski amiloid A
ATPazo	Adenozin trifosfataza
BAT2(D6S51)	HLA-B (humani levkocitni antigen B) povezan transkript 2
Cconf	Zanesljivost napovedane strukture naključnih zank na mestu aminokislin
CD analize	Analize cirkularnega dihiroizma
CD55	<i>decay accelerating factor for complement</i> / pospeševalni faktor za razgradnjo komplementa
COL4A1	Kolagen 4 alfa1
COL4A2	Kolagen 4 alfa2
COL4A3	Kolagen 4 alfa3
COL4A4	Kolagen 4 alfa4
Conf.	Zanesljivost napovedi na določenem mestu zaporedja aminokislin
CRP	C – reaktivni protein
C-SAA	Konstitutivni serumski amiloid A
CTR9	<i>RNA polymerase-associated protein</i> / protein asociiran z RNA polimerazo
Da	Enota atomske mase ali Dalton
DNK	Deoksiribonukleinska kislina
Econf.	Zanesljivost napovedane beta strukture na mestu zaporedja
ER	Endoplazemski retikulum
FPRL-1/ALX	Formil peptidni receptor podobni-1/ <i>Lipoxin A₄ receptor</i>
GA	Golgijev aparat
GAGs	Glukozaminoglikani
GalNAc	N-acetil-galaktozamin
HCAEC	Človeške endotelijske celice koronarne arterije
Hconf.	Zanesljivost napovedane strukture alfa vijačnic na mestu zaporedja
HDL	Lipoprotein z visoko gostoto
HS	Heparin/Heparan
ID številka	Identifikacijska številka
IL-1	Interlevkin 1
IL-10	Interlevkin 10
IL-12p40	Interlevkin 12 p40
IL-1R	Interlevkin 1 receptor
IL-1 β	Interlevkin 1 β

IL-23p19	Interlevkin 23 p19 (<i>interleukin 23, alpha subunit p19</i>)
IL-6	Interlevkin 6
IL-8	Interlevkin 8
JNK	c-Jun N-terminalna kinaza
KCNC1	<i>Potassium Voltage-Gated Channel Subfamily C Member 1</i> / Kalijev Napetosti Varovan kanal pod-družine C1
kDA	1000 enot atomske mase ali kilo Dalton
LAMA1	Laminin alfa1
LDHA	Laktatna dehidrogenaza A
LDHC	Laktatna dehidrogenaza C
MAPK p42/44	Mitogenom-aktivirana proteinska kinaza podenot p42/44
mRNA	Informacijska RNA molekula
MT	Mitohondrij
NF-kB	Jedrni faktor kapa B
NLS	Jedrni lokalizacijski signal
NUDT9P1	Nukleozidni difosfat s skupino X- motiv 9 psevdogen 1
P0DJI8	Uniprot identifikacijska številka serumskega amiloida A 1
P0DJI9	Uniprot identifikacijska številka serumskega amiloida A 2
P22614	Uniprot identifikacijska številka serumskega amiloida A 3
P35542	Uniprot identifikacijska številka serumskega amiloida A 4
p38 MAPK	Mitogensko-aktivirana proteinska kinaza p38
p44/42	<i>Extracellular signal-regulated protein kinases 1</i> (p44, ERK1) and 2 (p42) (ERK2)/proteinske kinase regulirane od ekstracelularnega signala tudi p42/44 MAPK
RA	Revmatoidni artritis
RAGE	Receptor za napredno glikacijo končnih produktov (<i>receptor for advanced glycation endproducts</i>)
SAA	Serumski amiloid A
SAA1_HUMAN	Humani serumski amiloid A1
SAA2_HUMAN	Humani serumski amiloid A2
SAA3_HUMAN	Humani serumski amiloid A3
SAA4_HUMAN	Humani serumski amiloid A4
SCARB1	<i>Scavenger receptor class B member 1</i> /Receptor čistilec razred B tip 1 / SR-BI / ABCA1 (<i>ATP-binding cassette subfamily B1</i>)
TANIS	SAA receptor, a hebrew word for fasting /hebrejsko »postiti se«
TGF- β	Transformirajoči rastni faktor β
THP-1	Človeška monocitna celična linija
TLR-2	<i>Toll-like receptor2</i> /Toll-podoben receptor2
TLR-4	<i>Toll-like receptor4</i> / Toll-podoben receptor4
TNF- α	Tumor nekrozni faktor 1
UBC	Ubikvitin C
VIMP/SELS	Valosin containing protein (VCP)-Interacting Membrane

VKORC1 Protein/Selenoprotein S
Vitamin K epoxide reductase (VKOR) C1/encim odgovoren za
recikliranje vitamina K

1 UVOD

Diplomska naloga primerja različne izotipe humanega vnetnega biomarkerja in akutno faznega proteina Serumskega amiloida A (SAA), iz vidika strukturnih lastnosti, kakor tudi potencialnih funkcij z bioinformatičnimi podatkovnimi bazami in orodji.

1.1 Serumski amiloid A

1.1.1 Lastnosti in pomembnost proteina

Serumski amiloid A (SAA) je akutno fazni protein [22, 11], z visoko stopnjo polimorfizmov [51]. Je del družine apolipoproteinov, ki se vežejo na lipoproteine z visoko gostoto (angl. HDL – high density lipoprotein). Poleg C-reaktivnega proteina (CRP) je glavni pokazatelj vnetja pri ljudeh. Vendar je SAA v primerjavi z CRP veliko bolj občutljiv in hitrejši odzivni pokazatelj. SAA spada pod glavne pozitivno akutno fazne proteine, za katere je značilno, da se njihova koncentracija v krvi v času akutne faze vnetja poveča za najmanj 25 %, pa vse do 1000-krat normalne vrednosti [22, 11, 23, 18, 24, 41], kar je značilno za oba CRP in SAA [22]. SAA je tudi prekursor amiloida A (AA), glavne proteinske komponente prevladujočih zunajceličnih amiloidskih vlaken reaktivne sistemske amiloidoze AA [32]. Če je koncentracija SAA v krvi za dlje časa povišana lahko pride do amiloidoze [35]. SAA je globularni protein z hidrofili in hidrofobnimi lastnostmi.

Med drugim je tudi pokazatelj mnogih bolezni, kot so kardiovaskularna obolenja [27]. Odkrili so, da naj bi bil SAA lokaliziran v neposredni bližini mikrotubulov, v citosolu, v veziklih in tubulih ter znotraj jedra človeških endotelijskih celic koronarne arterije (HCAEC), kar lahko nakazuje njegove različne potencialne vloge. Zaradi lokacije v nanotubulih obstaja možnost, da je SAA vpleten tudi v komunikacijo med celicami [27]. Poleg naloge vnetnega pokazatelja, sodeluje tudi v prenosu holesterola in v metabolizmu oziroma presnovi [33]. Poleg tega pa SAA pripisujejo, da naj bi imel vplivno vlogo pri vzdrževanju homeostaze [41]. Zadnji dokazi nakazujejo, da naj bi bil SAA vpleten v karcinogenezo, kar pa pomeni, da se z razvojem rakastega obolenja (tumorja) povečuje tudi koncentracija SAA [41].

SAA je pomemben predvsem za diagnostične namene. Z izmerjeno koncentracijo proteina v telesu bolnika je tako mogoče postaviti diagnozo vnetne bolezni, ali je vnetje prisotno ali ni, poleg tega lahko sledimo tudi razvoju bolezni oziroma potencialnem izboljšanju vnetnih bolezni [23, 27].

1.1.2 Sinteza serumskega amiloida A

Kot odziv na bakterijske infekcije (okužbe), travme, opekline in druge vrste poškodb na telesu pride do posledice, kot je vnetje. Ta pa deluje kot mehanizem za zaščito [23, 41]. Pri akutni fazi vnetja se lahko koncentracija SAA poveča do 1000-krat v času enega do dveh dneh (24-48 ur) po travmi, poškodbi ali okužbi. Takrat je raven SAA tudi najvišja [24] in sicer približno 10.000 mg/L. Če upoštevamo, da je fiziološka koncentracija SAA v krvi navidezno zdravih krvodajalcev manjša od 10,0 mg/L [35], lahko dolgoročne zvišane koncentracije SAA predstavljajo breme za organizem. Prav tako hitro kot se koncentracija poveča, se lahko tudi zniža [32]. V roku 1-3 tednov pa se koncentracija SAA povrne na fiziološko raven, pod pogojem, da je šlo za akutno bolezen [22]. Pri kroničnih ireverzibilnih boleznih pa se koncentracija SAA nikoli ne povrne na začetno raven [11].

Primarno se akutno fazni proteini sintetizirajo v hepatocitih. Poleg hepatocitov so lahko vir akutno faznih proteinov tudi makrofagi, fibroblasti, epiteljske celice, maščobne celice oziroma različna ekstrahepatična tkiva, npr. rakave celične linije [23]. Sama sinteza SAA in izločanje je regulirano s strani citokinov [24]. Ti so znotrajcelični signalni polipeptidi, proizvedeni med samim vnetnim stanjem, v katerem tudi sodelujejo. So glavni stimulatorji pri nastanku akutno faznih proteinov [22]. Glavni citokini, ki so vpleteni v akutno fazni odziv so predvsem interleukin 6 (IL-6), ki je glavni stimulator pri sintezi akutno faznih proteinov, interleukin 1 β (IL-1 β) ter tumor nekrozn faktor α (TNF- α). Poleg njih pa so še interferon γ , transformirajoči rastni faktor β (TGF- β) in interleukin 8 (IL-8). TNF- α je glavni stimulator IL-1 pri revmatoidnem artritisu (RA), IL-6 inhibira ekspresijo TNF- α in hkrati je glavni stimulator ekspresije genov akutno faznih proteinov [22, 23]. Ti citokini so proizvedeni s strani nevtrofilov, limfocitov, fibroblastov, endotelijskih celic, epiteljskih celic, predvsem pa s strani monocitov in makrofagov [22, 23]. IL-6, IL-1 β ter TNF- α pa so tudi glavni produkti monocitov in makrofagov [23].

SAA je sintetiziran tudi v HCAEC [27]. V študiji so pokazali, da se SAA znotraj HCAEC kolokalizira z mikrotubuli, medtem ko pa niso opazili nobene kolokalizacije SAA z aktinskimi filamentami, z vimentinom je bila opažena le majhna kolokalizacija z SAA [27].

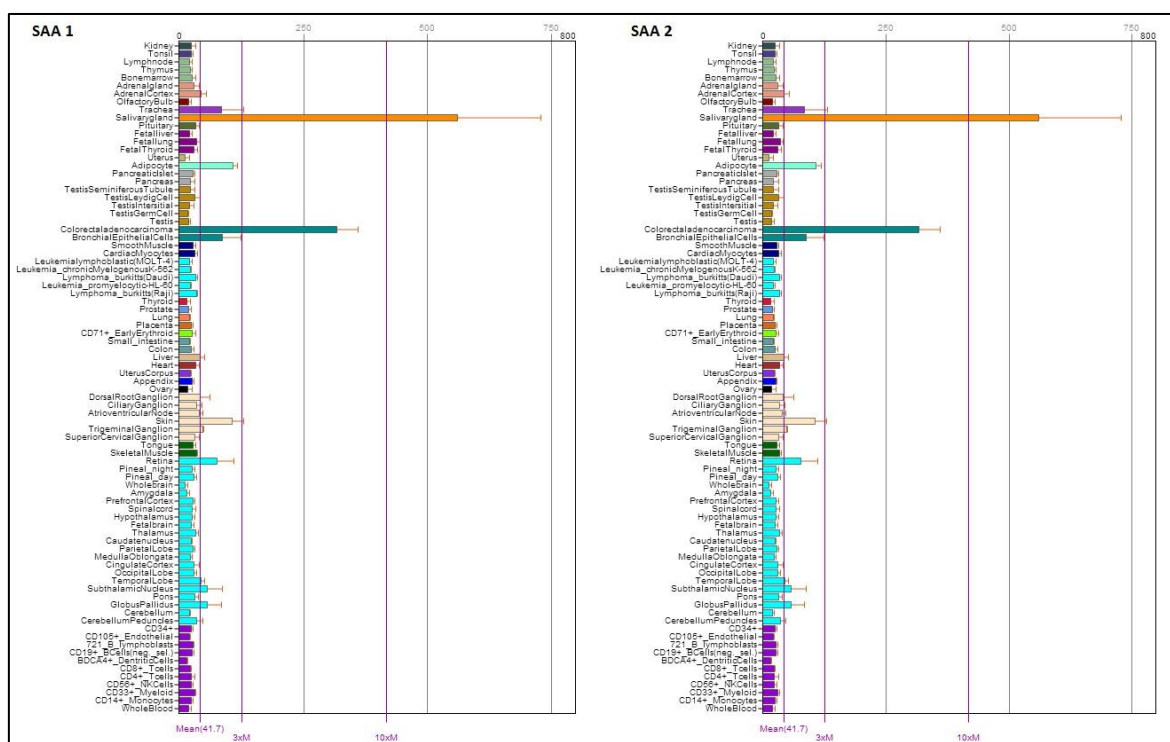
1.2 Izotipi SAA

Pri ljudeh najdemo štiri gene serumskega amiloida A (SAA1, SAA2, SAA3, SAA4). Vsi geni se nahajajo na kromosomu 11 [18].

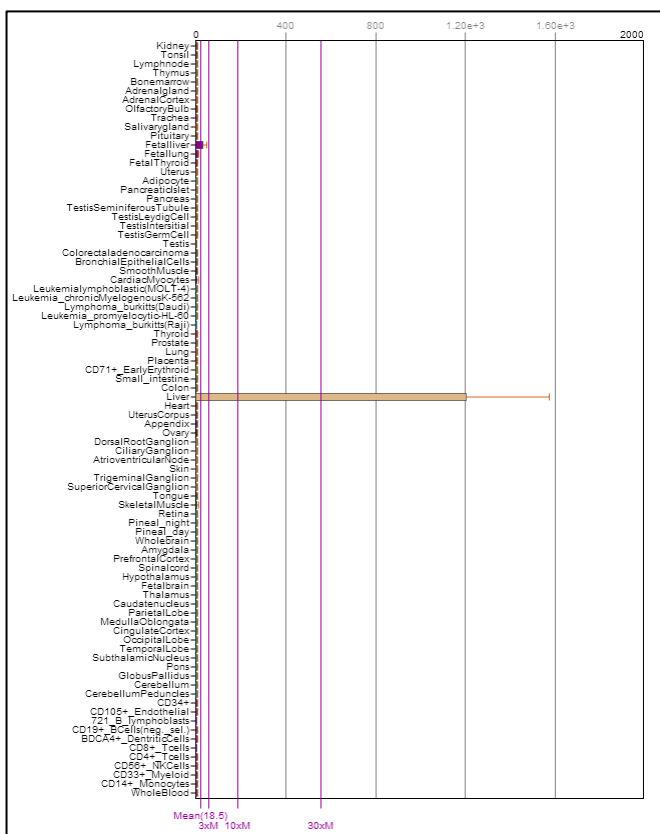
Gena SAA1 in SAA2 kodirata za akutno fazni serumski amiloid A (A-SAA). Inducirana akutno fazna proteina serumski amiloid A-1 (SAA1) in serumski amiloid A2 (SAA2) sta si sicer, tako v kodirajočih kot nekodirajočih regijah, 95% homologna [41, 42]. Ravno zaradi take podobnosti sta v nekaterih člankih označena kar z oznako SAA1/2 [33].

Serumski amiloid A1, ki je v nekaterih virih poimenovan tudi kot dominantni izotip, sestoji iz vsaj 5 različnih kombinacij alelov, pri čemer so le trije namenjeni za sintezo različnih proteinov (SAA1.1, SAA1.2, SAA1.3), medtem ko jih ima homologni izotip SAA2 tudi tri, pri čemer je SAA2.2 različen od ostalih dveh [41, 36]. Po podatkih iz internetne strani Uniprot (uniprot.org) je proteinsko zaporedje človeškega serumskega amiloida A1 (SAA1_HUMAN, P0DJ18) dolžine 122 aminokislin (AA), velikosti 13.532 daltonov (Da) oz. 13,5 kilodaltonov (kDa) [14]. Medtem ko pa je proteinsko zaporedje človeškega serumskega amiloida A2 (SAA2_HUMAN, P0DJ19) dolžine 122 aminokislin (AK) z velikostjo 13.527 daltonov (Da) oz. 13,5 kilodaltonov (kDa) [15].

Glavni vir sinteze akutno faznih proteinov (A-SAA) so jetra. Vendar pa so v številnih študijah pokazali, da se sintetizira tudi v drugih ekstrahepatskih tkivih/celicah. Lep primer je izražena informacijske ribonukleinske kisline (mRNA) akutno faznega serumskega amiloida A v sinoviocitih RA in pa tudi v monocitsko/makrofagskih celičnih linijah (THP-1 celicah, endotelijskih celicah, gladkih mišičnih celicah aterosklerotične lezije) [12]. Akutno fazna proteina (A-SAA) sta bila v študijah lokalizirana tudi v tkivih prsi, trebuha, črevesja, trebušne slinavke, žlezah slinavkah, raku debelega črevesa, maščobnih celicah, koži, vzdušnici (traheja), bronhialnih epitelnih celicah in drugod (Slika 1). Študija je bila izvedena na vzorcih 79 človeških tkiv [42, 5].

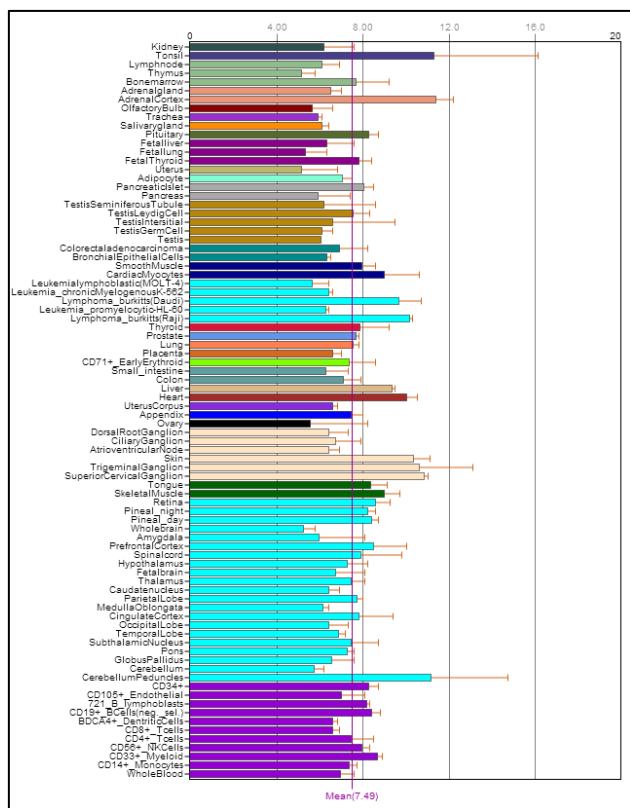


Slika 1: Ekspresija mRNA SAA1 in SAA2 v človeških tkivih (normalizirane vrednosti izražanja). Najbolj intenzivna v žlezah slinavkah, raku debelega črevesa, maščobnih celicah, koži, vzdušnici (traheja), bronhialnih epitelnih celicah (Vir: SAA1, <http://biogps.org/#goto=genereport&id=6288>; SAA2, <http://biogps.org/#goto=genereport&id=6289>)



Slika 2: Ekspresija mRNA SAA3 v človeških tkivih (Vir: <http://bioGPS.org/#goto=genereport&id=6290>)

Izotip SAA3 je bil do nedavnega poznan kot psevdogen, zaradi dodatne baze na drugi kodirajoči regiji zaporedja (2. ekson), po novjših raziskavah pa se je mnenje spremenilo, saj so izsledili koncentracije v materinem mleku (v mlečnih žlezah) [18, 27]. Relativno je enakomerno izražen v vseh pomembnejših človeških tkivih, najbolj izstopa izraženost v mandljih, skorji nadledvičnih žlez, trigeminalnih ganglijih (trivejni živec), cerebralnem peclju (Slika 2). Aminokislinsko zaporedje SAA izotipa 3 je prav tako dolgo 122 aminokislin (AK) [17].



Slika 3: Ekspresija mRNA SAA4 v človeških tkivih (Vir: <http://bioGPS.org/#goto=genereport&id=6291>)

Gen SAA4 pa kodira konstitutivni serumski amiloid A (C-SAA) [33]. Njegovo proteinsko zaporedje je dolgo 130 aminokislin (AK) z velikostjo 14.747 daltonov (Da) oz. 14,7 kilodaltonov (kDa) [18]. V mirujočih tkivih je izražen enakomerno porazdeljena (Slika 3). Z »Northern blot« analizo so v predhodno tudi dokazali, da so jetra med homeostazo glavni in edini vir SAA4 [18]. Med akutno faznim odzivom je serumski amiloid A-4 induciran minimalno [52], oziroma se njegova koncentracija ne poveča kot odgovor na stimulacijo, temveč ostaja na ravni med 80 in 140 mg/L [36].

Izotipi humanega SAA so znani že nekaj časa, vendar primerjava glede njihovih strukturnih značilnosti in potencialnih funkcij še ni bila opredeljena in zbrana z bioinformatičnimi podatkovnimi bazami in orodji.

2 NAMEN DELA

Serumski amiloid A (SAA) je eden glavnih vnetnih pokazateljev pri človeku. SAA genska družina vključuje inducibilna izotipa SAA1 in SAA2, SAA3 (predhodno imenovan kot pseudogen, vendar nedavno odkrit v mlečnih žlezah) in konstitutivno ekspresiran SAA4. Različni izotipi imajo različne vloge tako sistemsko kakor tudi znotraj celično, ki so v veliki meri še neznane.

Namen naloge je opredeliti z bioinformatičnimi podatkovnimi bazami in orodji primarno, sekundarno, terciarno strukturo izotipov SAA, jih primerjati in pridobiti vpogled v njihove morebitne funkcije glede na različne celične lokalizacije, proteinske interakcije in post-translacijske modifikacije.

SPECIFIČNI CILJI naloge so DOLOČITI:

1. Primarno strukturo SAA1, SAA2, SAA3 in SAA4 ter jih primerjati po zaporedju, kakor tudi po aminokislinski sestavi.
2. Sekundarno strukturo SAA1, SAA2, SAA3 in SAA4 in primerjati njihove deleže α vijačnic, β struktur in naključnih zank.
3. Terciarno strukturo SAA1, SAA2, SAA3 in SAA4.
4. Potencialne funkcije SAA1, SAA2, SAA3 in SAA4 glede na
 - a) celične lokalizacije
 - b) proteinske interakcije
 - c) post-translacijske modifikacije

3 METODE

Z namenom okarakterizirati strukturne in funkcijske lastnosti humanih SAA smo uporabili več različnih portalov, podatkovnih baz in bioinformatičnih orodij.

Portali		
Ime	Povezava	Specifike portala
PubMed NCBI portal	http://ncbi.nlm.nih.gov/pubmed/	<i>PubMed vsebuje več kot 23 milijonov citatov literature iz Biomedicine (MEDLINE, znanstvenih revij, spletnih knjig)</i>
ExPASy portal	http://expasy.org/	<i>ExPASy je SIB Bioinformatični portal virov, ki omogoča dostop do znanstvenih podatkovnih baz in programskih orodij v različnih področjih znanosti kot so proteomika, genomika, filogenija, sistemska biologija, populacijska genetika, transkriptomika.</i>
BioGPS (Gene Portal System)	http://biogps.org/	<i>Portal za opis in karakterizacijo genov in proteinskih funkcij.</i>
OBRC (Online Bioinformatics Resources Collection)	http://www.hsls.pitt.edu/obrc/	<i>Portal, ki ti omogoča iskanje molekularnih podatkovnih baz in programskih orodij.</i>
Bioinformatične podatkovne baze		
Ime	Povezava	Opis
Uniprot	http://uniprot.org/	<i>Prosto dostopna, visokokakovostna baza podatkov proteinskih zaporedij in njenih informacij.</i>
STRING	http://string-db.org/	<i>Podatkovna baza znanih in napovednih proteinskih interakcij. Podatke vsebuje iz direktnih (fizičnih) in indirektnih (funkcionalnih) vidikov (preko genskega konteksta, eksperimentov, ko-ekspresije in drugih že znanih raziskav).</i>
GenBank	http://ncbi.nlm.nih.gov/genbank/	<i>Podatkovna baza genskih zaporedij za primerjalna iskanja in karakterizacijo zaporedij preko Nacionalnega centra za biotehnološko informacijo (NCBI).</i>
RCSB PDB	http://rcsb.org/	<i>Podatkovna baza z informacijami o 3-dimenzionalnih strukturah bioloških molekul, vključno s proteini.</i>
Wiki-Pi	http://severus.dbmi.pitt.edu/wiki-pi/	<i>Podatkovna baza, ki vsebuje informacije o interakcijah med proteini.</i>
HPRD (Human Protein Reference Database)	http://hprd.org/	<i>Podatkovna baza človeških proteinov in njihovih lastnosti.</i>
Bioinformatična orodja		
Postranslacijske modifikacije		
Ime	Povezava	Opis

NetCGlyc 1.0 Server	http://www.cbs.dtu.dk/services/NetCGlyc/	Namenjeno je napovedovanju C-manozilacijskih mest v proteinih sesalcev. C-manozilacija je vrsta glikozilacije, kjer je ogljikov hidrat (manozni sladkor) vezan na triptofan.
NetNGlyc 1.0 Server	http://www.cbs.dtu.dk/services/NetNGlyc/	Namenjeno je napovedovanju N-glikozilacijskih mest v proteinih človeka za preučevanje dela zaporedij Asn-Xaa-Ser/Thr.
NetOGlyc 4.0 Server	http://www.cbs.dtu.dk/services/NetOGlyc/	Namenjeno je napovedovanju O-glikozilacijskih mest mucinskega tipa GalNac.
Myristoylator	http://web.expasy.org/myristoylator/	Namenjeno napovedovanju N-terminalnega miristoiliranega glicina.
NMT – The MYR Predictor	http://mendel.imp.ac.at/myristate/SUPLpredictor.htm	Namenjeno napovedovanju N-terminalne N-miristoilacije proteinov.
SUMOPlot™ Analysis Program	http://www.abgent.com/sumoplot	Namenjen napovedovanju in ocenjevanju mest sumoilacije v proteinih.
NetPhos 2.0 Server	http://www.cbs.dtu.dk/services/NetPhos/	Napovedovanje fosforilacijskih mest v serinih, treoninih in tirozinih evkariontskih proteinov.
The Sulfinator	http://web.expasy.org/sulfinator/	Namenjena je napovedovanju tirozinskih sulfatacijskih mest v proteinskih sekvencah. Tirozinska sulfatacija je pomembna post-translacijska modifikacija proteinov, ki gredo skozi izločitvene poti celic.

Sekundarna in terciarna struktura

Ime	Povezava	Opis
YASPIN	http://ibi.vu.nl/programs/yaspinwww/	Je HNN (Hidden Neural Network) oziroma orodje skrite nevronske mreže, kombinacija skritega Markovega modela in nevronske mreže. Namenjeno je napovedovanju sekundarnih struktur proteinov. Deluje s pomočjo algoritma PSI-BLAST, ki izdelava PSSM (Position-Specific scoring matrix) matriko, s katero nato dobimo napoved sekundarne strukture.
SOPMA	http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=NPSA/npsa_sopma.html	SOPMA je optimizirana določevalna metoda sekundarnih struktur s poravnami, temelji pa na metodi homologije Levina in njegovih sodelavcev.
Phyre2	http://sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index	Phyre2 je orodje za napovedovanje 3D strukture proteinov s pomočjo skritega Markovega modela. Je naslenjik orodij 3D-pssm in Phyre.
RaptorX	http://raptorx.uchicago.edu/StructurePrediction/predict/	Je strežnik namenjen za napovedovanje strukture proteinov. Napoveduje sekundarno, terciarno strukturo, površine dostopne molekulam topila in neurejene regije.
PredictProtein	http://predictprotein.org	Je spletni strežnik, ki vključuje naslednje podatke (kot so sekundarne strukture, površine dostopne molekulam topila, transmembranske vijačnice, lokalizacije, mesta disulfidnih vezi, med proteinske vezave), ki se jih pridobi s pomočjo proteinskega zaporedja.

ArgusLab	http://www.arguslab.com/arguslab.com/ArgusLab.html	<i>ArgusLab je program za molekulsko modeliranje.</i>
Lokalizacija proteina		
Ime	Povezava	Opis
Hum-mPLoc2.0	http://www.csbio.sjtu.edu.cn/bioinf/hum-multi-2/	<i>Namenjena napovedovanju znotraj celične lokalizacije človeških proteinov vključno s tistimi proteini ki imajo multiple lokalizacije.</i>
CELLO v.2.5: subCELLular LOcalization predictor	http://cello.life.nctu.edu.tw/	<i>CELLO je več razredna metoda podpornih vektorjev in klasifikacijski sistem. Za napovedovanje znotraj celične lokalizacije uporablja 4 vrste shem: - aminokislinsko sestavo - dipeptidno sestavo - porazdelilno aminokislinsko sestavo - sestavo, ki temelji na fizikalno kemijskih lastnosti aminokislin S kombinacijo rezultatov sestav dobimo oceno podatkov.</i>
PSORT II Prediction	http://psort.hgc.jp/form2.html	<i>Je računalniški program/orodje za napovedovanje proteinskih lokalizacijskih mest znotraj celic.</i>
TargetP 1.1 Server	http://www.cbs.dtu.dk/services/TargetP/	<i>Namenjen je subcelularnem napovedovanju lokalizacije proteinov evkariontov. Lahko zaznava N-terminalne sekvence: - Kloroplastne tranzitne peptide (cTPs) - mitohondrijske tarčne peptide (mTPs) - Signalne peptide izločevalnih celičnih poti (SP) Za napoved cepitvenih mest cTPs in SP uporablja ChloroP in SignalP orodji.</i>
NetNES 1.1 Server	http://www.cbs.dtu.dk/services/NetNES/	<i>Namenjena za napovedovanje levčinsko bogatih jedrnih izvoznih signalov (NES) v proteinih evkariontov.</i>
MitoProt II v1.101	http://ihg.gsf.de/ihg/mitoprot.html	<i>MitoProt izračuna N-terminalno proteinsko regijo, ki lahko podpre mitohondrijsko tarčno sekvenco in cepitveno mesto.</i>
Poravnava zaporedij		
Ime	Povezava	Opis
Clustal Omega	http://ebi.ac.uk/Tools/msa/clustalo/	<i>Orodje za poravnavo več zaporedij</i>
NCBI Blast (Basic Local Alignment Tool)	http://blast.ncbi.nlm.nih.gov/Blast.cgi	<i>Je osnovno iskalno orodje lokalne poravnave. Uporablja se za lokalno poravnavo zaporedij.</i>
Potencialne funkcije		
Ime	Povezava	Opis
ProtFun 2.2	http://cbs.dtu.dk/services/ProtFun/	<i>Orodje, ki napoveduje funkcije proteinov s pomočjo preko njihovih zaporedij. Uporablja</i>

		<i>ab initio metodo.</i>
HNB – Protein Function Prediction	http://dag.embl-heidelberg.de/hnb_cgi/show_input_page.pl?taskId=_notask&UserId=_nouser	<i>Orodje za napovedovanje potencialnih funkcij proteinov. Uporablja programa SMARTin miniPEDANT za prepoznavo funkcionalnih področij in orodje STRING za prepoznavo funkcionalnih povezav z drugimi proteini.</i>
CombFunc	http://sbg.bio.ic.ac.uk/~mwass/combfunc/	<i>Orodje za napovedovanje potencialnih funkcij proteinov. Prepozna prisotna konzervirana mesta proteinov v poravnavi z ostalimi proteini z enakim GO anotacijami in jim na tak način določi potencialne funkcije.</i>

4 REZULTATI IN DISKUSIJA

Rezultati izhajajo iz opredelitve posameznih izotipov SAA in iz njihove skupne primerjave, tako iz vidika primarnih, sekundarnih in terciarnih struktur kakor tudi potencialnih funkcij z bioinformatičnimi orodji.

4.1 Primarne strukture izotipov serumskega amiloida A

Primarna struktura je najnižji nivo strukture proteinov. Iz nje je mogoče razbrati dolžino proteinov in njihovo aminokislinsko sestavo.

4.1.1 Primarna struktura SAA1

Primarna struktura serumskega amiloida A1 (SAA1) je pridobljena iz spletne strani Uniprot (www.uniprot.org) [45].

Oznaka vzorca: SAA1_HUMAN
Uniprot ID številka: P0DJ18
Dolžina: 122

10	20	30	40	50	60
MKLLTGLVFC	SLVLGVSSRS	FFSFLGEAFD	GARDMWRAYS	DMREANYIGS	DKYFHARGNY
70	80	90	100	110	120
DAAKRGPGGV	WAAEAISDAR	ENIQRFFGHG	AEDSLADQAA	NEWGRSGKDP	NHFRPAGLPE
122					
KY					

Tabela 1: Aminokislinska sestava SAA1

Aminokislina	Št. AK
Alanin (A)	16
Cistein (C)	1
Asparaginska kislina (D)	9
Glutaminska kislina (E)	7
Fenilalanin (F)	9
Glicin (G)	14
Histidin (H)	3
Izolevcin (I)	3
Lizin (K)	5
Levcin (L)	8
Metionin (M)	3
Asparagin (N)	5
Prolin (P)	4
Glutamin (Q)	2
Arginin (R)	10
Serin (S)	10
Treonin (T)	1
Valin (V)	4
Triptofan (W)	3
Tirozin (Y)	5

Zaporedje vzorca človeškega serumskega amiloida A1 (SAA1_HUMAN) z Uniprot ID številko P0DJ18 je sestavljeno iz vseh 20 aminokislin. Največji delež v zaporedju zavzmeta aminokislini alanin (13,1%) in glicin (11,5%), sledijo jima še arginin in serin z vsak po 8,2%, asparaginska kislina (7,4%), fenilalanin (7,4%), levcin (6,6%) (Tabela 1).

4.1.2 Primarna struktura SAA2

Primarna struktura serumskega amiloida A2 (SAA2) je pridobljena iz spletne strani Uniprot [46].

Oznaka vzorca: SAA2_HUMAN
 Uniprot ID številka: P0DJI9
 Dolžina: 122

10	20	30	40	50	60
MKLLTGLVFC	SLVLSVSSRS	FFSFLGEAFD	GARDMWRAYS	DMREANYIGS	DKYFHARGNY
70	80	90	100	110	120
DAAKRGPGGG	WAAEVISNAR	ENIQRLTGRG	AEDSLADQAA	NKWGRSGRDP	NHFRPAGLPE
122					
KY					

Tabela 2: Aminokislinska sestava SAA2

Aminokislina	Št. AK
Alanin (A)	16
Cistein (C)	1
Asparaginska kislina (D)	8
Glutaminska kislina (E)	6
Fenilalanin (F)	7
Glicin (G)	13
Histidin (H)	2
Izolevcin (I)	3
Lizin (K)	5
Levcin (L)	9
Metionin (M)	3
Asparagin (N)	6
Prolin (P)	4
Glutamin (Q)	2
Arginin (R)	12
Serin (S)	11
Treonin (T)	2
Valin (V)	4
Triptofan (W)	3
Tirozin (Y)	5

Zaporedje vzorca človeškega serumskega amiloida A2 (SAA2_HUMAN) z Uniprot ID številko P0DJI9 je sestavljeno iz vseh 20 aminokislin. Največji delež v zaporedju zavzameta aminokislini alanin (13,1 %) in glicin (10,7 %), sledijo jima še arginin (9,8 %), serin (9,0 %), in tako dalje (Tabela 2).

4.1.3 Primarna struktura SAA3

Primarna struktura serumskega amiloida A3 (SAA3), različica številka 78, je pridobljena iz arhiva spletne strani Uniprot, saj je bila najnovejša različica (številka 79) zbrisana dne 06.13.2012 [47].

Oznaka vzorca: SAA3_HUMAN
 Uniprot ID številka: P22614
 Dolžina: 122

10	20	30	40	50	60
MKLSTGIIFC	SLVLGVSSQG	WLTFLKAAGQ	GAKDMWRAYS	DMKEANYKKS	DKYFHARGNY
70	80	90	100	110	120
DAVQRGPGGV	WATEVISDAR	ENVQRLTGDH	AEDSLAGQAT	NKWGQSGKDP	NHFRPAGLPE
122					
KY					

Tabela 3: Aminokislinska sestava SAA3

Aminokislina	Št. AK
Alanin (A)	13
Cistein (C)	1
Asparaginska kislina (D)	8
Glutaminska kislina (E)	5
Fenilalanin (F)	4
Glicin (G)	14
Histidin (H)	3
Izolevcin (I)	3
Lizin (K)	10
Levcin (L)	8
Metionin (M)	3
Asparagin (N)	5
Prolin (P)	4
Glutamin (Q)	6
Arginin (R)	6
Serin (S)	9
Treonin (T)	5
Valin (V)	6
Triptofan (W)	4
Tirozin (Y)	5

Zaporedje vzorca človeškega serumskega amiloida A3 (SAA3_HUMAN) z Uniprot ID številko P22614 je sestavljeno iz vseh 20 aminokislin. Največji delež v zaporedju zavzameta aminokislini glicin (11,5 %) in alanin (10,66 %). Sledijo lizin (8,2 %), serin (7,4 %), in tako dalje (Tabela 3).

4.1.4 Primarna struktura SAA4

Primarna struktura serumskega amiloida A4 (SAA4) je pridobljena iz spletne strani Uniprot [48].

Oznaka vzorca: SAA4_HUMAN
 Uniprot ID številka: P35542
 Dolžina: 130

10	20	30	40	50	60
MRLFTGIVFC	SLVMGVTSES	WRSFFKEALQ	GVGDMGRAYW	DIMISNHQNS	NRYLYARGNY
70	80	90	100	110	120
DAAQRGPGGV	WAAKLISRSR	VYLQGLIDCY	LFGNSSTVLE	DSKSNEKAE	WGRSGKDPDR
122					
FRPDGLPKKY					

Tabela 4: Aminokislinska sestava SAA4

Aminokislina	Št. AK
Alanin (A)	8
Cistein (C)	2
Asparaginska kislina (D)	8
Glutaminska kislina (E)	6
Fenilalanin (F)	6
Glicin (G)	14
Histidin (H)	1
Izolevcin (I)	5
Lizin (K)	7
Levcin (L)	10
Metionin (M)	4
Asparagin (N)	6
Prolin (P)	4
Glutamin (Q)	4
Arginin (R)	11
Serin (S)	13
Treonin (T)	3
Valin (V)	7
Triptofan (W)	4
Tirozin (Y)	7

Zaporedje vzorca človeškega serumskega amiloida A4 (SAA4_HUMAN) z Uniprot ID številko P35542 je sestavljeno iz vseh 20 aminokislin. Največji delež v zaporedju zavzameta aminokislini glicin (10,8 %) in serin (10 %). Sledijo arginin (8,5 %), levcin (7,7 %), in tako dalje (Tabela 4).

4.1.5 Primerjave aminokislinskih zaporedij izotipov SAA (med seboj)

Za poravnavo dveh zaporedij, sem uporabil orodje BLAST (Osnovno iskalno orodje lokalne poravnave), kjer je vselej prvo imenovani izotip izbran kot poizvedovalno zaporedje (angl. Query), drugo pa kot predmet (angl. Subject).

4.1.5.1 Primerjava zaporedij SAA1-SAA2

Zaporedji SAA1 in SAA2 (Slika 4) sta si popolnoma enaki v 93 % (Razlikujeta se na mestih 15, 70, 75, 78, 86, 87, 89, 102, 108), homologni pa v 95 % (Razlikujeta se na mestih 15, 70, 75, 86, 87, 89).

Score	Expect	Method	Identities	Positives	Gaps
206 bits(525)	5e-74	Compositional matrix adjust.	113/122(93%)	116/122(95%)	0/122(0%)
Query 1		MKLLTGLVFCSLVLGVSSRSFFSFLGEAFDGDARDMWRAYSDMREANYIGSDKYFHARGNY			60
Sbjct 1		MKLLTGLVFCSLVL VSSRSFFSFLGEAFDGDARDMWRAYSDMREANYIGSDKYFHARGNY			60
Query 61		DAAKRGGVWAAEAISDARENIQRFFGHGAEDSLADQAANEWGRSGKDPNHFRPAGLPE			120
Sbjct 61		DAAKRGGVWAAEISARENIQR G GAEDSLADQAAN+WGRSG+DPNHFRPAGLPE			120
Query 121		KY 122			
Sbjct 121		KY 122			

Slika 4: Lokalna poravnava izotipov SAA1 (P0DJ18) in SAA2 (P0DJ19)

4.1.5.2 Primerjava zaporedij SAA1-SAA3

Zaporedji SAA1 in SAA3 (Slika 5) sta si popolnoma enaki v 76 % (Razlike se pojavijo na mestih 4, 7, 8, 19-23, 26, 27, 29, 30, 33, 43, 48, 49, 63, 64, 73, 75, 83, 86, 87, 89, 90, 97, 100, 102, 105), homologni pa v 85 % (Razlike se pojavijo na mestih 4, 20, 22, 26, 27, 29, 30, 48, 49, 63, 73, 75, 86, 87, 89, 90, 97, 100).

Score	Expect	Method	Identities	Positives	Gaps
201 bits(510)	9e-72	Compositional matrix adjust.	93/122(76%)	104/122(85%)	0/122(0%)
Query 1		MKLLTGLVFCSLVLGVSSRSFFSFLGEAFDGDARDMWRAYSDMREANYIGSDKYFHARGNY			60
Sbjct 1		MKL TG++FCSLVLGVSS+ + +FL A GA+DMWRAYSDM+EANY SDKYFHARGNY			60
Query 61		DAAKRGGVWAAEAISDARENIQRFFGHGAEDSLADQAANEWGRSGKDPNHFRPAGLPE			120
Sbjct 61		DA +RGGVWVA E ISDAREN+QR G AEDSLA QA N+WG+SGKDPNHFRPAGLPE			120
Query 121		KY 122			
Sbjct 121		KY 122			

Slika 5: Lokalna poravnava izotipov SAA1 (P0DJ18) in SAA3 (P22614)

4.1.5.3 Primerjava zaporedij SAA1-SAA4

Zaporedji SAA1 in SAA4 (Slika 6) sta si popolnoma enaki v 5 % (Razlike se pojavijo na mestih 2, 4, 7, 14, 17, 19, 21, 22, 25, 26, 29, 30, 32, 33, 36, 40, 42- 45, 47- 49, 51, 52, 54, 55, 64, 74, 75, 78, 79, 81- 83, 85-91, 94-99, 103-107, 109, 119, 120, 124, 128), homologni

pa v 72 % (Razlike se pojavijo na mestih 4, 19, 22, 25, 26, 29, 30, 32, 33, 36, 40, 43, 44, 48, 49, 54, 75, 78, 81, 82, 85-91, 95, 97- 99, 103, 107, 109, 120, 124).

Score	Expect	Method	Identities	Positives	Gaps
141 bits(355)	3e-48	Compositional matrix adjust.	72/130(55%)	94/130(72%)	8/130(6%)
Query 1	MKLLTGLVFCSLVGVSSRSFFSFLGEAFDGMWRAYSMDREANYIGSDKYFHARGNY				60
Sbjct 1	M+L TG+VFCSLV+GV+S S+ SF EA G DM RAY D+ +N+ S++Y +ARGNY				60
Query 61	DAAKRGGPGGWAAEAIISDARENIQ-----RFFGHGA---EDSLADQAANEWGRSGKDPNH				112
Sbjct 61	DAA+RGPGGWAA+ IS +R +Q FG+ + EDS +++ A EWGRSGKDP+				120
Query 113	FRPAGLPEKY 122				
Sbjct 121	FRPDGLPKKY 130				

Slika 6: Lokalna poravnava izotipov SAA1 (P0DJ18) in SAA4 (P35542)

4.1.5.4 Primerjava zaporedij SAA2-SAA3

Zaporedji SAA2 in SAA3 (Slika 7) sta si popolnoma enaki v 76 % (Razlike se pojavijo na mestih 4, 7, 8, 15, 19-23, 26, 27, 29, 30, 33, 43, 48, 49, 63, 64, 70, 73, 78, 83, 89, 90, 97, 100, 105, 108), homologni pa v 86 % (Razlike se pojavijo na mestih 4, 15, 20, 22, 26, 27, 29, 30, 48, 49, 63, 70, 73, 89, 90, 97, 100).

Score	Expect	Method	Identities	Positives	Gaps
201 bits(512)	5e-72	Compositional matrix adjust.	93/122(76%)	105/122(86%)	0/122(0%)
Query 1	MKLLTGLVFCSLVSVSSRSFFSFLGEAFDGMWRAYSMDREANYIGSDKYFHARGNY				60
Sbjct 1	MKL TG++FCSLVL VSS+ + +FL A GA+DMWRAYSMD+EANY SDKYFHARGNY				60
Query 61	DAAKRGGPGGAWAAEVISNARENIQRLTGRGAEDSLADQAANKWGRSGRDPNHFRPAGLPE				120
Sbjct 61	DA +RGGG WA EVIS+AREN+QRLTG AEDSLA QA NKWG+SG+DPNHFRPAGLPE				120
Query 121	KY 122				
Sbjct 121	KY 122				

Slika 7: Lokalna poravnava izotipov SAA2 (P0DJ19) in SAA3 (P22614)

4.1.5.5 Primerjava zaporedij SAA2-SAA4

Zaporedji SAA2 in SAA4 (Slika 8) sta si popolnoma enaki v 52 % (Razlike se pojavijo na mestih 2, 4, 7, 14, 15, 17, 19, 21, 22, 25, 26, 29, 30, 32, 33, 36, 40, 42- 45, 47-49, 51, 52, 54, 55, 64, 70, 74, 75, 78, 79, 81-83, 85, 87-99, 103-107, 109, 110, 116, 119, 120, 124, 128), homologni pa v 70 % (Razlike se pojavijo na mestih 4, 15, 19, 22, 25, 26, 29, 30, 32, 33, 36, 40, 43, 44, 48, 49, 54, 70, 78, 81, 82, 85, 87-94, 96-99, 103, 107, 109, 120, 124).

Score	Expect	Method	Identities	Positives	Gaps
135 bits(339)	9e-46	Compositional matrix adjust.	67/130(52%)	91/130(70%)	8/130(6%)
Query 1	MKLLTGLVFCSLVLSVSSRSFFSFLGEAFDGDARDMWRAYSMDREANYIGSDKYFHARGNY				60
Sbjct 1	M+L TGI+VFCSLV+ V+S S+ SF EA G DM RAY D+ +N+ S++Y +ARGNY				60
Query 61	DAAKRGGPGGAWAAEVISNARENIQRL-----TGRGAEDSLADQAANKWGRSGRDPNH				112
Sbjct 61	DAA+RGGPGG WAA++IS +R +Q L + EDS +++ A +WGRSG+DP+				120
Query 113	FRPAGLPEKY	122			
Sbjct 121	FRPDGLPKKY	130			

Slika 8: Lokalna poravnava izotipov SAA2 (P0DJ19) in SAA4 (P35542)

4.1.5.6 Primerjava zaporedij SAA3-SAA4

Zaporedji SAA3 in SAA4 (Slika 9) sta si popolnoma enaki 52 % (Razlike se pojavijo na mestih 2, 4, 8, 14, 17, 19, 20, 22, 23, 25, 27, 29, 32, 33, 36, 40, 42-45, 47-49, 51, 52, 54, 55, 63, 73-75, 78, 79, 81-83, 85, 87-99, 103-110, 113, 119, 120, 124, 128), homologni pa v 70 % (Razlike se pojavijo na mestih 4, 20, 22, 25, 27, 29, 32, 33, 36, 40, 43, 44, 49, 54, 63, 73, 78, 81, 82, 85, 87-89, 91-99, 103, 105, 107-109, 120, 124).

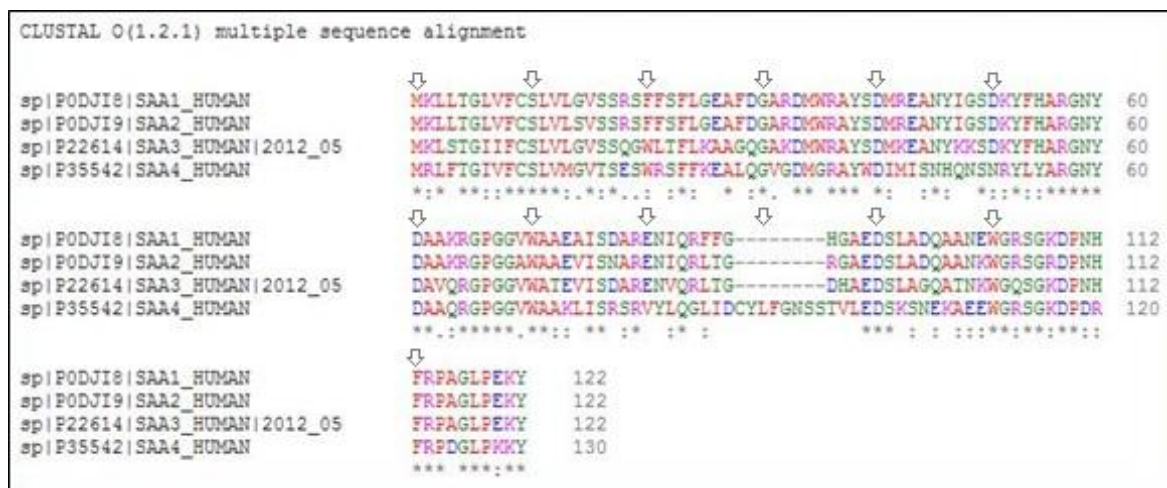
Score	Expect	Method	Identities	Positives	Gaps
139 bits(350)	2e-47	Compositional matrix adjust.	67/130(52%)	91/130(70%)	8/130(6%)
Query 1	MKLSTGIIFCSLVLGVSSQGWLTFLKAAGQGAQDMWRAYSMDKEANYKKSDDKYFHARGNY				60
Sbjct 1	M+L TGI+VFCSLV+GV+S+ W +F K A QG DM RAY D+ +N++ S++Y +ARGNY				60
Query 61	DAVQRGGPGGVWATEVISDARENVQRLTGDH-----AEDSLAGQATNKWGQSGKDPNH				112
Sbjct 61	DA QRGGPGGVWA ++IS +R +Q L + EDS + + +WG+SGKDP+				120
Query 113	FRPAGLPEKY	122			
Sbjct 121	FRPDGLPKKY	130			

Slika 9: Lokalna poravnava izotipov SAA3 (P22614) in SAA4 (P35542)

4.1.5.7 Primerjava zaporedij SAA1, SAA2, SAA3, SAA4

S poravnavo aminokislinskih (AK) zaporedij izotipov SAA dobimo izboljšani pregled nad mesti podobnosti in razlikami med zaporedji. Večje je AK zaporedje v primerjavi z ostalimi AK zaporedji, manjša je stopnja identičnosti večjega zaporedja z ostalimi. Izotop SAA4 je od ostalih treh izotipov večji za 8 aminokislin. Rezultati poravnave več zaporedij (*job: clustalo-I20140619-150455-0410-43679841-oy*) nazorno pokažejo, da odsek zaporedja SAA4, dolžine 8 aminokislin (od 88. do 95. mesta; CYLFGNSS), ni prisoten pri izotipih SAA1, SAA2 in SAA3 (Slika 10). Odsotnost odseka zaporedja aminokislin bi lahko pomenila, da je v preteklosti prišlo do razvoja izotipov iz enega samega izotipa. Ker pa so delecije pogostejše od insercij, substitucij in podovjitev, saj je zanje potrebno veliko manj energije, lahko sklepamo, da je izotip daljšega zaporedja AK (v tem primeru SAA4) najstarejši izotip znotraj družine SAA.

Sicer pa je že znano, da so proteini SAA visoko ohranjeni skozi evolucijo, saj so prisotni tako v sesalcih, kot v vretenčarjih in nevretenčarjih kot so morske kumare. Njihova aminokislinska zaporedja SAA so si od 61 do 80 % identična kar lahko kaže na to, da si delijo skupno strukturo [39].

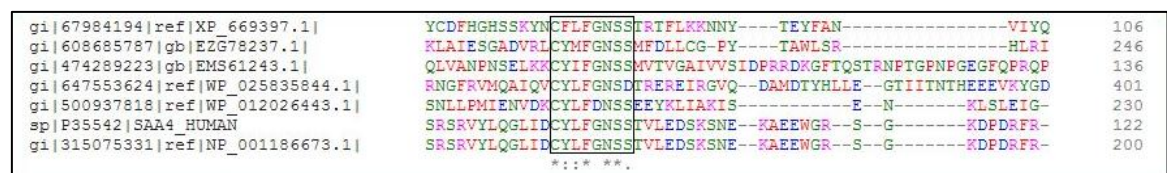


Slika 10: Poravnava zaporedij SAA1-4 ena pod drugo (Vir: <http://www.ebi.ac.uk/Tools/msa/clustalo>; job: clustalo-I20140619-150455-0410-43679841-oy)

NCBI Blast orodje izstopajoči odsek SAA4 zaporedja »CYLFGNSS« poravna z odseki proteinov SAA2 prekursorski protein (*Homo sapiens*; NP_001186673.1) s 100 % identičnostjo in pokritostjo. Človeški SAA2 prekursor, predhodnik zrelega SAA2 proteina, je poravnan na mestih 167-174.

Medtem, ko pa so hipotetična proteina (angl. »hypothetical proteins«) iz plazmodija (*Plasmodium berghei strain ANKA*; XP_669397.1 in flavobakterije (*Flavobacterium johnsoniae*; WP_012026443.1), transkripcijski faktor 27 z MADS zaporedjem (angl. »MADS-boxtranscription factor 27«) iz pšenice (*Triticum urartu*; EMS61243.1), diacilglicerol aciltransferaza (angl. »diacylglycerol acyltransferase«) iz gregarina (*Gregarina niphandrodes*; EZG78237.1), zaznani z 88 % identičnostjo in 100% pokritostjo (8/8) odseka SAA4 pri rezultatu poravnave z vsaj 25,7.

Omeniti velja še ATPazo (angl. »ATPase«) iz bakterije (*Bacteroides stercorisoris*; WP_025835844.1), ki z rezultatom poravnave 26,9 izstopa, kljub temu da je zaznan s 87 % pokritostjo (7/7) in 100 % identičnostjo odseka »CYLFGNS«, brez serina (S) na 8. mestu označenega odseka (Slika 11). Rezultati (job: clustalo-I20140617-110537-0291-24470089-oy) tako nakazujejo na zapleten evlucijski razvoj, skozi katerega je šla tudi družina SAA.



Slika 11: Odsek zaporedja SAA4 (CYLFGNSS) v proteinih različnih organizmov (Vir: <http://www.ebi.ac.uk/Tools/msa/clustalo>; job: clustalo-I20140617-110537-0291-24470089-oy)

Po rezultatih spletnega orodja PredictProtein je odsek »DKYFHARGN« najbolj konservativen (štiri mesta določena z rezultatom 8/9, preostalih pet pa z 9/9).

SAA1 in SAA2 sta si po zaporedju najbolj podobna, saj je pri njiju prisotna kar 93 % identičnost. Po podobnosti se nato od njiju najmanj razlikuje SAA3, ki je s 86 % homolognostjo s SAA2 sicer za 1 % bolj homologen v primerjavi s SAA1. Pri identičnosti v relacijah SAA1/SAA3 in SAA2/SAA3 ni opažene razlike, obe relaciji sta 76 % identični. SAA4 se z drugimi izotipi po primarni strukturi v družini SAA najbolj razlikuje. Njegovo zaporedje je še najbolj podobno zaporedju SAA1, z 72 % homologijo in 55 % identičnostjo, medtem pa v primerjavi z izotipoma SAA2 in SAA3 predstavlja najmanjšo podobnost v družini SAA (52 %) (Tabela 5).

Tabela 5: Deleži podobnosti in identičnosti izotipov SAA

		Identičnost			
		%	SAA1	SAA2	SAA3
podobnost / homologija	SAA1		93	76	55
	SAA2	95		76	52
	SAA3	85	86		52
	SAA4	72	70	70	

4.2 Sekundarne strukture izotipov serumskega amiloida A

Sekundarna struktura predstavlja drugi nivo strukture proteinov. Prikazuje nam sestavo primarne strukture (alfa vijačnice, beta strukture in naključne zanke). Nastane zaradi ionskih, vodikovih, kovalentnih vezi med aminokislinami.

4.2.4 Sekundarna struktura SAA1

Yaspinovo napovedovanje sekundarne strukture (Slika 12) prikazuje, da je po večini SAA1 sestavljen iz alfa vijačnic in naključnih zank. Če upoštevamo zanesljivost napovedi, označena z oznako Conf. z veljavnimi vrednostmi od 1 (nizka verjetnost) do 9 (visoka verjetnost) in privzamemo, da bomo upoštevali le tiste napovedi, kjer je zanesljivost napovedi večja ali enaka 7, in hkrati posledično tudi zanesljivost napovedane strukture (Econf, Hconf ali Cconf) večja ali enaka 7, pridemo do naslednjih ugotovitev. Aminokislinskih mest, ki so del naključnih zank je 14, beta strukture predstavljajo 3 mesta, alfa vijačnice pa 10 od skupno 27 mest. Za zanesljivost večjo ali enako od 8 pa velja naključnih zank (6), beta struktur (0), alfa vijačnic (5), kar pa je prestrog kriterij za določanje deleža struktur.

```
YASPIN PREDICTION RESULTS
Name: P0DJIB|(SAA1_HUMAN)
Length: 122

*Key
* AA: Target sequence
* Pred: Predicted secondary structure (H=helix, E=strand, -=coil)
* Conf: Confidence (0=low, 9=high)
*Hconf: Confidence of helix predictions
*Econf: Confidence of strand predictions
*Cconf: Confidence of coil predictions

AA: MKLLTGLVFCSLVGLGVSSRSFFSFLGEAFDGDARDMWRAYSMDREANYIGSKYFHARGNY
Pred: --EEEEEEEEEEEEEE-----HHHHH--HHHHHHHHHHHHHH-----EEEE-----
Conf: 911335566767766033561100117642762858682385724201156306222551

Hconf: 00000000000000000000003016999990099999999999999142400000000004
Econf: 05899999999999998000043000000000000000000000000000000899980000
Cconf: 9410000000000000199964530000099000000000000000857599100019995

AA: DAAKRGPGVWAAEAISDARENIQRFFGHGAEDSLADQAANEWGRSGKDPNHFRPAGLPE
Pred: -----EEEEEE---HHHHHHH-----HHHHHHHHH-----
Conf: 101016777514544121357742010578101014238761424166876325534887

Hconf: 252510000000000001999999750005736998999999000000000000000000
Econf: 00000000999999400000000000000000000000000000000000000000000000
Cconf: 74748999900000059800000002499942630010000009999999999999999999

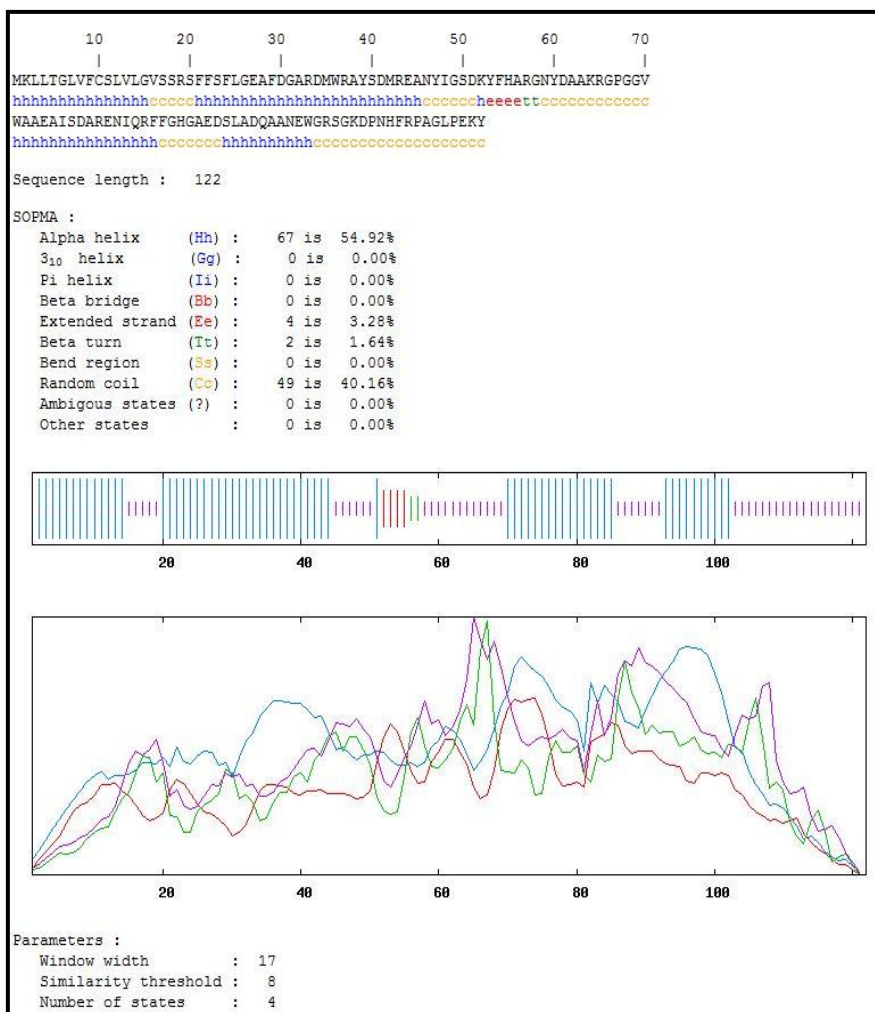
AA: KY
Pred: --
Conf: 79

Hconf: 00
Econf: 00
Cconf: 99
```

Slika 12: Napoved sekundarne strukture humanega SAA1 s spletnim orodjem YASPIN

Opombe: AA = aminokislina; Hconf = verjetnost alfa vijačnice ; Econf = verjetnost beta strukture; Cconf = verjetnost naključne zanke; Conf = Verjetnost pravilne napovedi (0 = najnižja, 9 = najvišja)

Za večjo zanesljivost sem poleg orodja YASPIN uporabil še orodje za napovedovanje sekundarne strukture proteinov SOPMA (Slika 13). Rezultati (*ID: 5d1f66d70b91*) prikazuje, da naj bi struktura vsebovala 67 mest, ki predstavljajo del alfa vijačnic, 49 mest za naključne zanke, 6 pa beta strukture od skupno 122 mest.



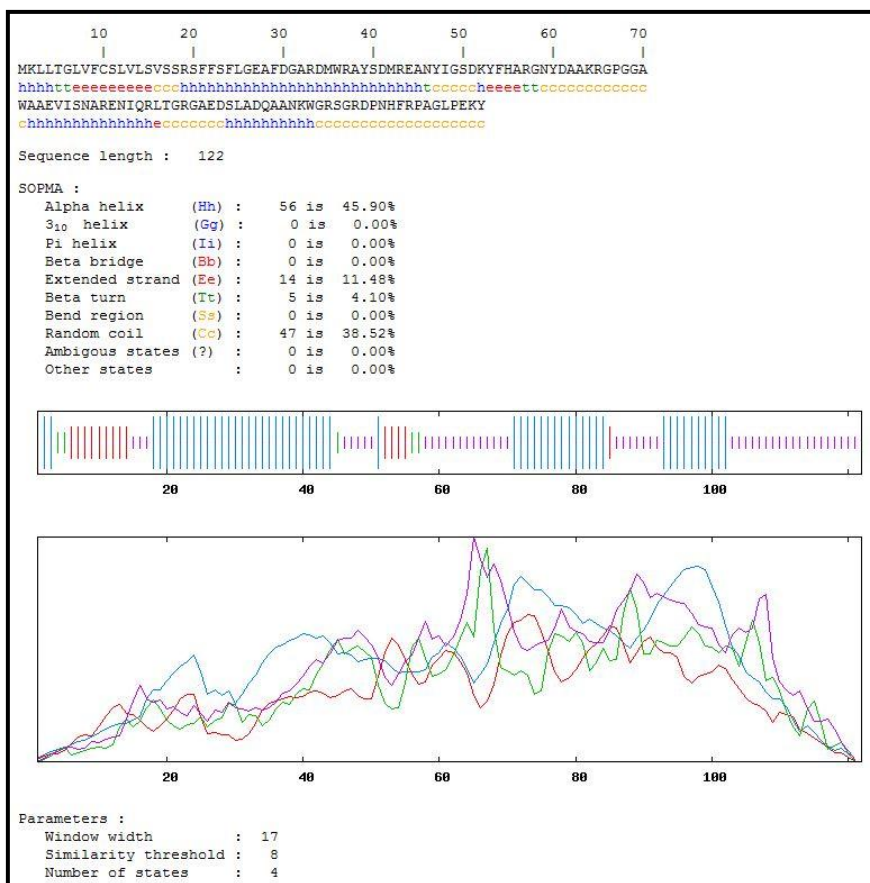
Slika 13: Napoved sekundarne strukture humanega SAA1 s spletnim orodjem SOPMA (Vir: http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html; ID: 5d1f66d70b91)

Poleg tega sem uporabil še orodja Phyre2, RaptorX in PredictProtein. Rezultate vseh spletnih orodij sem združil in izračunal povprečja za vsako sekundarno strukturo posebej. Ti pa nam prikazujejo, da je izotip SAA1 sestavljen večinoma iz alfa vijačnic (56,4 %) in naključnih zank (36,8 %). Beta struktura predstavlja pa le dobrih 6,6 % celotnega človeškega serumskega amiloida A1 (Tabela 6).

Tabela 6: Deleži sekundarnih struktur v SAA1 v odstotkih (%)

SAA1	Alfa vijačnica (α)	Beta struktura (β)	Naključna zanka	Dodatni podatki
YASPIN*	37,0	11,1	51,9	
PHYRE2	78,0	0,0	22,0	27 % neurejenost
RAPTORX	63,0	3,0	33,0	1 % neurejenost
PREDICTPROTEIN	49,2	13,9	36,9	
SOPMA	54,9	4,9	40,2	
Povprečje	56,4	6,6	36,8	

Opombe: *Upoštevana le tista mesta, kjer je zanesljivost napovedi (Conf) in zanesljivost določene strukture, alfa strukture (Hconf), naključne zanke (Cconf), beta strukture (Econf) na določenem mestu preseгла vrednost 6



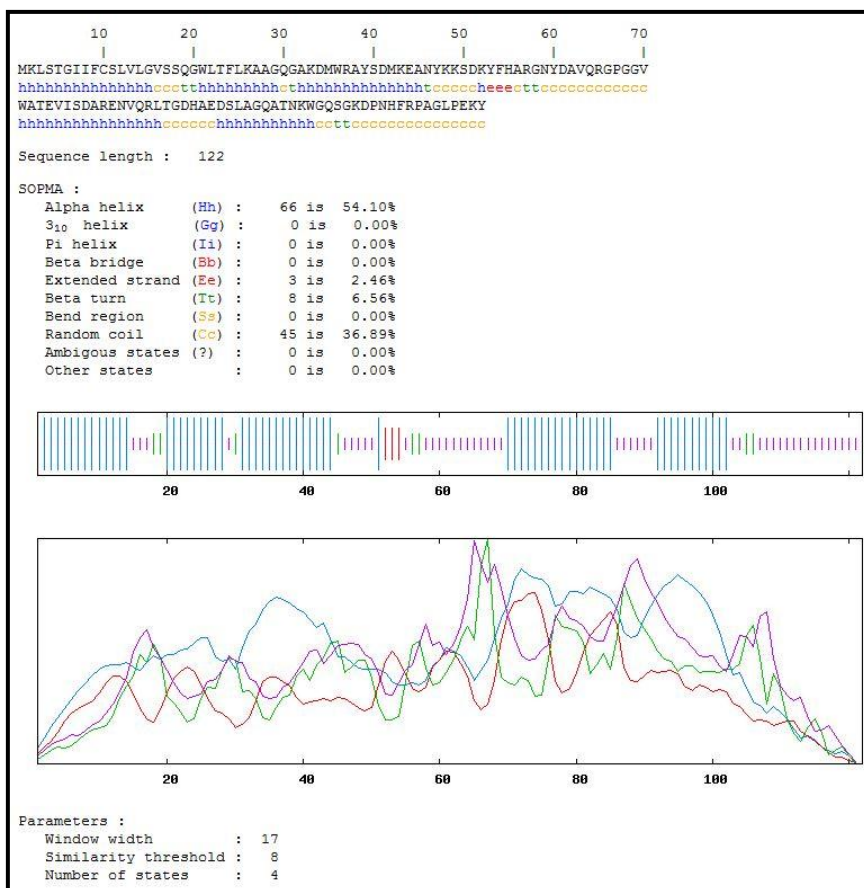
Slika 15: Napoved sekundarne strukture humanega SAA2 s spletnim orodjem SOPMA (Vir: http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html; ID: c8852f19a8ef)

Poleg tega sem uporabil še orodja Phyre2, RaptorX in PredictProtein. Rezultate vseh spletnih orodij sem združil in izračunal povprečja za vsako sekundarno strukturo posebej. Ti pa nam prikazujejo, da je izotip SAA2 sestavljen večinoma iz alfa vijačnic (50,8 %) in naključnih zank (35,1 %). Beta struktura predstavlja pa le dobrih 11,9 % celotnega človeškega SAA2 (Tabela 7).

Tabela 7: Deleži sekundarnih struktur v SAA2 v odstotkih (%)

SAA2	Alfa vijačnica (α)	Beta struktura (β)	Naključna zanka	Dodatni podatki
YASPIN*	34,5	17,2	38,5	
PHYRE2	75,0	0,0	25,0	25 % neurejenost
RAPTORX	50,0	13,0	36,0	2 % neurejenost
PREDICTPROTEIN	48,4	13,9	37,7	
SOPMA	45,9	15,6	38,5	
Povprečje	50,8	11,9	35,1	

Opombe: *Upoštevana le tista mesta, kjer je zanesljivost napovedi (Conf) in zanesljivost določene strukture, alfa strukture (Hconf), naključne zanke (Cconf), beta strukture (Econf) na določenem mestu presegla vrednost



Slika 17: Napoved sekundarne strukture humanega SAA3 s spletnim orodjem SOPMA (Vir: http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html; ID: e42a7d73a72b)

Poleg tega sem uporabil še orodja Phyre2, RaptorX in PredictProtein. Rezultate vseh spletnih orodij sem združil in izračunal povprečja za vsako sekundarno strukturo posebej. Ti pa nam prikazujejo, da je izotip SAA3 sestavljen večinoma iz alfa vijačnic (52,8 %) in naključnih zank (36,1 %). Beta struktura predstavlja pa le dobrih 10,7 % celotnega človeškega SAA3 (Tabela 8).

Tabela 8: Deleži sekundarnih struktur v serumskem amiloidu A3 v odstotkih (%)

SAA3	Alfa vijačnica (α)	Beta struktura (β)	Naključna zanka	Dodatni podatki
YASPIN *	36,0	12,0	52,0	
PHYRE2	80,0	0,0	20,0	25 % neurejenost 13 % trans. vijač.**
RAPTORX	52,0	13,0	33,0	1 % neurejenost
PREDICTPROTEIN	41,9	19,7	38,5	
SOPMA	54,1	9,0	36,9	
Povprečje	52,8	10,7	36,1	

Opombe: *Upoštevana le tista mesta, kjer je zanesljivost napovedi (Conf) in zanesljivost določene strukture, alfa strukture (Hconf), naključne zanke (Cconf), beta strukture (Econf) na določenem mestu preseglja vrednost 6; ** nanaša na delež transmembranskih vijačnic

4.2.7 Sekundarna struktura SAA4

Orodje za določevanje sekundarne strukture Yaspin prikazuje, da je po večini serumski amiloid A4 sestavljen iz alfa vijačnic in naključnih zank (Slika 18). Ob predpostavki, da je zanesljivost napovedi, označena z oznako Conf. z veljavnimi vrednostmi od 1 (nizka verjetnost) do 9 (visoka verjetno) in če upoševamo le tiste napovedi, kjer je zanesljivost napovedi večja ali enaka 7, in hkrati posledično tudi zanesljivost napovedane strukture (Econf, Hconf ali Cconf) večja ali enaka 7, pridemo do naslednjih ugotovitev. Aminokislin, ki so del alfa vijačnic je 10, kar predstavlja 37,0 % delež, beta struktur 4 (14,8 %), naključnih zank pa 13 (48,2 %).

```
YASPIN PREDICTION RESULTS

Name: P35542 (SAA4_HUMAN)|
Length: 130

*Key
* AA: Target sequence
* Pred: Predicted secondary structure (H=helix, E=strand, -=coil)
* Conf: Confidence (0=low, 9=high)
*Hconf: Confidence of helix predictions
*Econf: Confidence of strand predictions
*Cconf: Confidence of coil predictions

  AA: MRLFTGIVFCSLVMGVTSESWSRFFKEALQGVGDMGRAYWDIMISNHQNSNRYLYARGNY
  Pred: -EEEEEEEEEEEEEEEE---HHHHHHHHH--HHHHHHHHHHHHHHH-----EEEE----
  Conf: 910336577756765034560000227632873858683384734201135036222561

Hconf: 00000000000000000000000030199999990099999999999999254500000000005
Econf: 06899999999999998000066000000000000000000000000000000899980000
Cconf: 9310000000000000199963200000099000000000000000000745499100019994

  AA: DAAQRGPGGVWAAKLISRVRVYLQGLIDCYLFGNSSTVLEDSKSNEKAEWGRSGKDPDR
  Pred: -----EEEEEE---HHHHHHHHHHHHH-----HHHHHHHHH-----
  Conf: 101016777514533121225646850100005456771011143388603342668662

Hconf: 47472000000000000019999999976452000000632599999998000000000
Econf: 0000000009999995000000000000003130000000100000000000000000000
Cconf: 52527999900000049800000000023124999999356400000000199999999

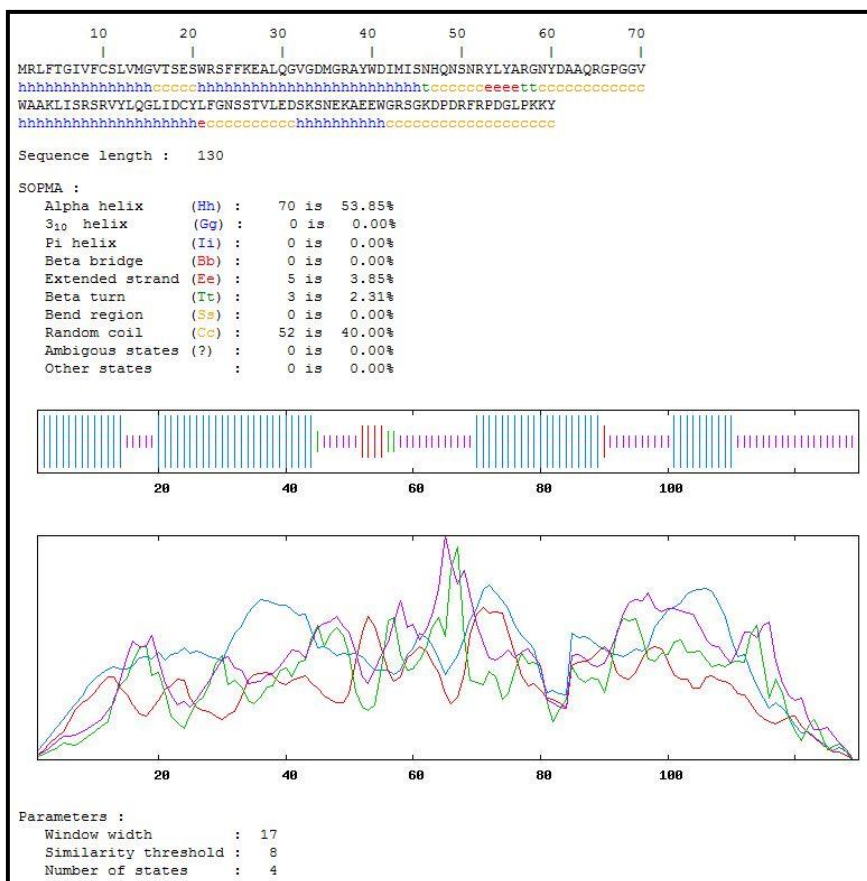
  AA: FRPDGLPKKY
  Pred: -----
  Conf: 1443488779

Hconf: 0000000000
Econf: 1000000000
Cconf: 8999999999
```

Slika 18: Napoved sekundarne strukture humanega SAA4 s spletnim orodjem YASPIN

Opombe: AA = aminokislina; Hconf = verjetnost alfa vijačnice ; Econf = verjetnost beta strukture; Cconf = verjetnost naključne zanke; Conf = Verjetnost pravilne napovedi (0 = najnižja, 9 = najvišja)

Internetno orodje SOPMA (ID: 4d51c0843422) pa nakazuje da naj bi struktura vsebovala 70 aminokislinskih mest, ki predstavljajo del alfa vijačnic, kar je 53,9 % delež celotnega zaporedja. 52 aminokislinskih mest predstavlja naključne zanke, 8 pa beta stukture od skupno 130 mest (6,2 %) (Slika 19).



Slika 19: Napoved sekundarne strukture humanega SAA4 s spletnim orodjem SOPMA (Vir: http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html; ID: 4d51c0843422)

Poleg tega sem uporabil še orodja *Phyre2* (SAA1, JobID: c6f99544b2edb7cb; SAA2, JobID: bc2fea1481d232ac; SAA3, JobID: 97de244bda9faa46; SAA4, JobID: 9a9dcd14075cf1bd), *RaptorX* (SAA1, JobID: 36134250; SAA2, JobID: 15019717; SAA3, JobID: 39254074; SAA4, JobID: 38943712) in *PredictProtein* (SAA1, JobID: 477570; SAA2, JobID: 477571; SAA3, JobID: 477581; SAA4, JobID: 477585). Rezultate vseh spletnih orodij sem združil in izračunal povprečja za vsako sekundarno strukturo posebej. Ti pa nam prikazujejo, da je izotip SAA4 sestavljen večinoma iz alfa vijačnic (50,5 %) in naključnih zank (39,2 %). Beta struktura predstavlja pa le dobrih 10,1 % celotnega človeškega serumskega amiloida A4 (Tabela 9).

Tabela 9: Deleži sekundarnih struktur v serumskem amiloidu A4 v odstotkih (%)

SAA4	Alfa vijačnica (α)	Beta struktura (β)	Naključna zanka	Dodatni podatki
YASPIN*	37,0	14,8	48,2	
Phyre2	72,0	0,0	28,0	30 % neurejenost
RaptorX	48,0	12,0	39,0	1 % neurejenost
PredictProtein	41,5	17,7	40,8	
SOPMA	53,9	6,2	40,0	
Povprečje	50,5	10,1	39,2	

Opomba: *Upoštevana le tista mesta, kjer je zanesljivost napovedi (Conf) in zanesljivost določene strukture, alfa strukture (Hconf), naključne zanke (Cconf), beta strukture (Econf) na določenem mestu preseгла vrednost 6

4.2.8 Primerjave sekundarnih struktur izotipov SAA

Vsi izotipi SAA so si po deležu posameznih sekundarnih struktur zelo podobni (Tabela 10). SAA1 sicer izstopa po deležu alfa vijačnice za približno 3-6 % (56,4 %) z razliko ostalih izotipov (v povprečju deleži alfa vijačnic znašajo ~ 51.4 %), zato pa vsebuje najmanjši delež beta struktur (6,6 %), ki je za vsaj 3,5 % nižji v primerjavi z ostalimi (≥ 10.1 %). Najnižji delež alfa vijačnic vsebujeta SAA2 in SAA4 in sicer 50.8 % oziroma 50.5 %. SAA2 glede na deleže sekundarnih struktur vsebuje najvišji delež beta struktur (11.95 %). V primerjavi z njegovim najbolj podobnim izotipom SAA1 (95 % homologija), ima nekaj manj kot dvakrat višji delež prisotnih beta struktur ($SAA1(\beta) = 6.6$ %; $SAA2(\beta) = 12.0$ %; $\Delta(SAA1(\beta), SAA2(\beta)) = 5.4$ %).

Tabela 10: Deleži sekundarnih struktur za posamezen izotip serumskega amiloida A v odstotkih (%)

2° struktura SAA izotipi	Alfa vijačnica (α)	Beta struktura (β)	Naključna zanka	Σ
SAA1	56,4	6,6	36,8	99,8
SAA2	50,8	12,0	35,2	98,0
SAA3	52,8	10,7	36,1	99,6
SAA4	50,5	10,1	39,2	99,8

Za ocenitev SAA struktur je bilo uporabljenih več orodij (Slika 20). Z namenom, da bi določili relevantnost podatkov smo kot kriterij uporabili citiranost orodij za določanje sekundarnih struktur, ki je bila pridobljena s pomočjo Google Učenjaka (<http://scholar.google.si>) za vsako orodje posebej.

SOPMA: (članek citirani 791-krat, vir iz leta 1995) [19]
 Phyre2: (članek citiran 1832-krat, vir iz leta 2009) [31]
 Yaspin: (članek citiran 161-krat, vir iz leta 2005) [23]
 RaptorX: (članki citirani največ 45-krat, viri med leti 2011-13) [24, 25, 26, 27]
 PredictProtein: (članek citiran 1 krat, vir iz leta 2014) [28]

Slika 20: Pojavnost orodij za napovedovanje sekundarnih struktur

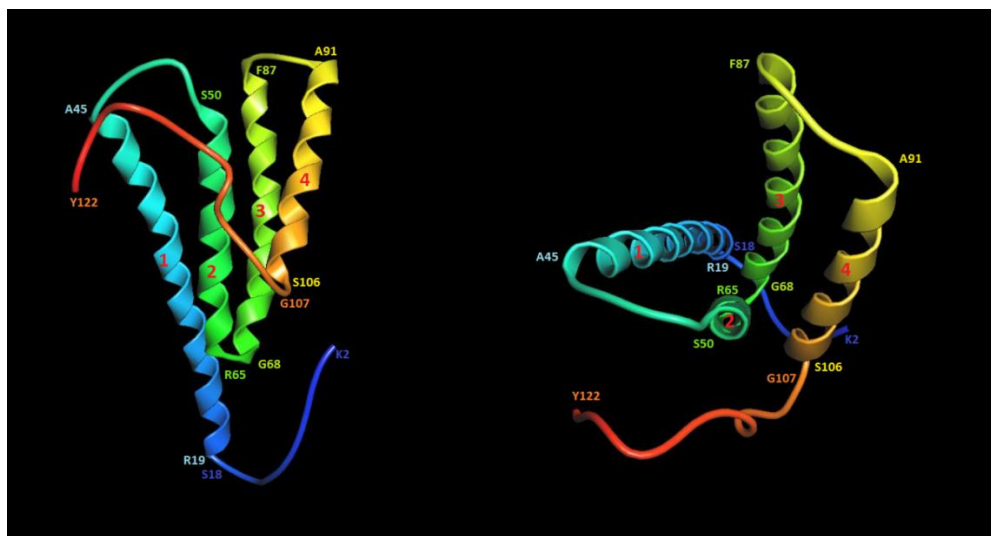
V nedavni raziskavi je bil SAA1.1 protein kristaliziran in njegova struktura določena [39]. SAA1.1 obstaja kot heksamer s podenotami, ki kažejo 4 vijačno konformacijo. Ista skupina je določila vezavo heparina in HDL na SAA1.1 ter amiloidogene epitope. Stevens F.L. je poročal o analizi primarne strukture humanega SAA ter nakazal verjetno homologijo z N-domeno hemocianinov členonožcev, ki kažejo na to, da približno 80 % proteina sestavljajo snopi alfa vijačnic vključno s (preostalim) C-terminalnim delom, ki je potencialno neurejen [50]. Prvoten SAA1.1 po zadnjih raziskavah naj ne bi vseboval beta struktur [20], kar se

približno ujema z napovedanimi rezultati (78 % alfa vijačnic in 0 % napovedanih beta struktur) najbolj citiranega uporabljenega orodja Phyre2, ki je bil po podatkih Google Učenjaka citiran že preko 1800 krat (Slika 20), kljub temu da obravnavana zaporedja SAA vsebujejo signalne peptide. Medtem ko ostala orodja predvidevajo v povprečju bistveno manjši delež v primeru alfa vijačnic (51,0 %), oziroma večji delež beta struktur (8,2 %). Slednji rezultati se bolje ujemajo s starejšimi študiji, starim 10 in več let, kjer analize cirkularnega dikroizma (CD) fuzijskega proteina, v kateri je bil humani SAA1.1 povezan s stafilokokno nukleazo nakazujejo, da naj bi alfa vijačnice predstavljali 33-44% vseh sekundarnih struktur v SAA [50]. Poleg tega pa naj bi v tem obdobju Wang in sodelavci odkrili približno 50% prisotnost strukture alfa vijačnice v njihovih analizah CD [37].

4.3 Terciarne strukture izotipov serumskega amiloida A

SAA1 je sestavljen iz snopa štirih alfa vijačnic, vključno s signalnim peptidom (K2-S18) in dolgim C-terminalnim repom (G107-Y122). Nizi ob obeh modelih predstavljajo začetna in končna mesta signalnega peptida, vijačnice 1, vijačnice 2, vijačnice 3, vijačnice 4 in C-terminalnega repa. Začetne črke v nizih predstavljajo aminokislino, številke pa njena zaporedna mesta v proteinu (Slika 21).

Alfa vijačnici N-konca (vijačnici 1 in 3) naj bi bili amiloidogena peptida SAA1.1. Ker pa sta skriti v heksameru, naj bi to pomenilo, da je naravni protein SAA nepatogen, kar se tiče amiloidnega potenciala. Monomer SAA1.1 se po strukturi razlikuje od monomera citokinov, vendar pa je zelo podoben N-terminalni domeni apolipoproteina E (ApoE), z razliko da sta položaja alfa vijačnic (vijačnici 3 in 4) pri SAA1.1 in ApoE zamenjani. Monomer SAA1.1 je storžaste oblike, njegov C-terminalni rep poteka po širini snopa štirih alfa vijačnic in tako tvori solne mostove in vodikove vezi s tremi alfa vijačnicami (vijačnice 1, 2 in 4). Tako da je C-terminalni rep ključnega pomena pri stabilizaciji monomera [39].



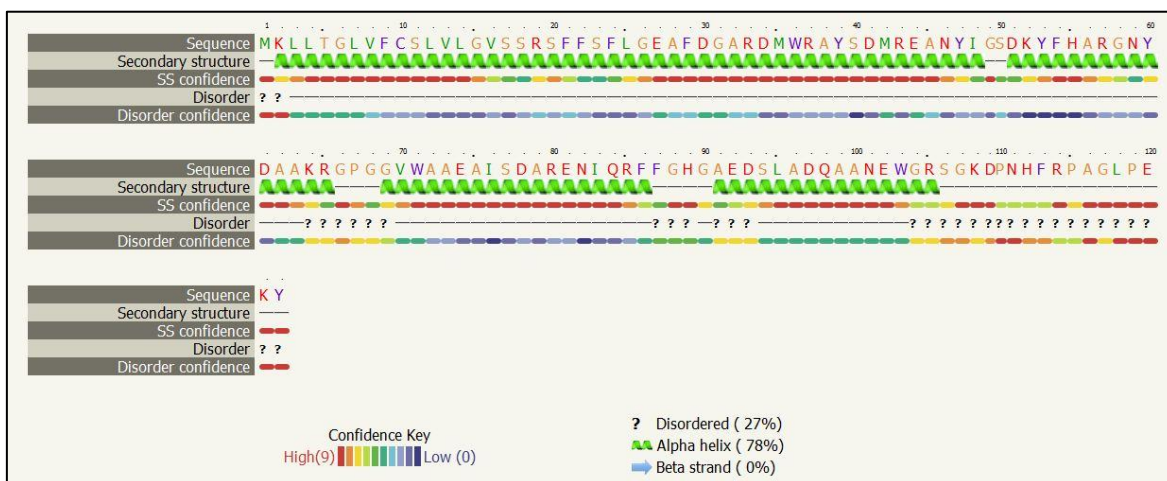
Slika 21: Model monomera človeškega SAA1 (Phyre ID: c4ip8B_) obdelan s pomočjo programa ArgusLab, prikazan s strani (na sliki levo) in od zgoraj (na sliki desno).

4.3.1 Sekundarne strukture pomembne za terciarno strukturo

SAA1

Monomer SAA1 je sestavljen iz snopa štirih alfa vijačnic [39]. Po rezultatih orodja Phyre 2 (*Job: c6f99544b2edb7cb*), prva vijačnica oziroma vijačnica 1 je dolžine 30 AK (od 19. do 48. mesta v zaporedju) (Slika 22). J. Lu in sodelavci so označili od 19. do 45. mesta v zaporedju, kar nanese, da je vijačnica 1 dolžine 27 AK [39]. Vijačnica 2 poteka od 51. mesta pa do 65. mesta v zaporedju (dolžine 15 AK), kar se ne razlikuje veliko v primerjavi z nedavno analizo (od 50. do 65. mesta v zaporedju; dolžine 16) [39], vijačnica 3 poteka od 69. do 86. mesta v zaporedju (dolžine 18 AK), kar se razlikuje za dolžino dveh aminokislin (od 68. do 87. mesta v zaporedju; dolžine 20) [39], vijačnica 4 poteka od 91. do 105 (dolžine 15 aminokislin), kar se razlikuje od raziskave J. Lu in sodelavcev za dolžino ene aminokislina (od 91. do 106. mesta v zaporedju; dolžine 16) [39].

Signalni peptid poteka od 2-18 (dolžine 17 AK), C-terminalni rep pa poteka od 107-122 (dolžine 16).

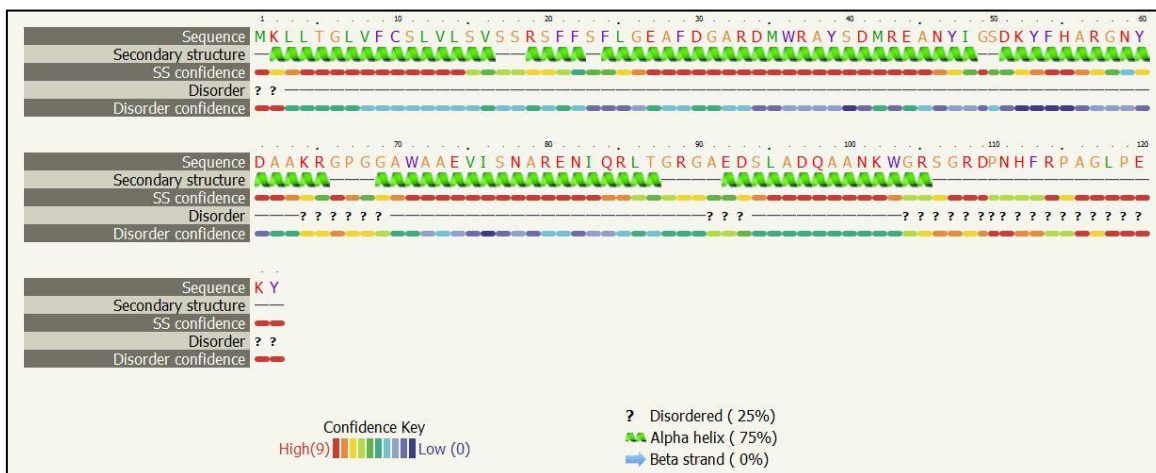


Slika 22: Sekundarna struktura SAA1 z grafično razvidnimi alfa vijačnicami (Phyre 2; Job: c6f99544b2edb7cb)

Ker skupina Lu et al. ni obravnavala SAA2-4, nadaljnje primerjave niso bile možne.

SAA2

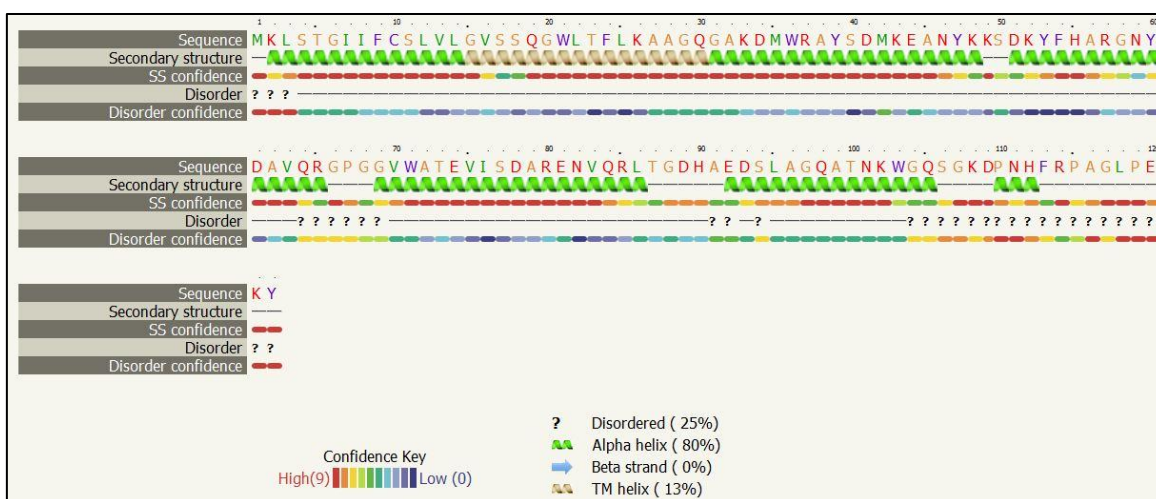
SAA2 naj bi bila po napovedih orodja Phyre 2 (*Job: bc2fea1481d232ac*) sestavljen iz snopa šestih vijačnic, pri čemer je prva vijačnica signalni peptid (Slika 23).



Slika 23: Sekundarna struktura SAA2 z grafično razvidnimi alfa vijačnicami (Phyre 2; Job: bc2fea1481d232ac)

SAA3

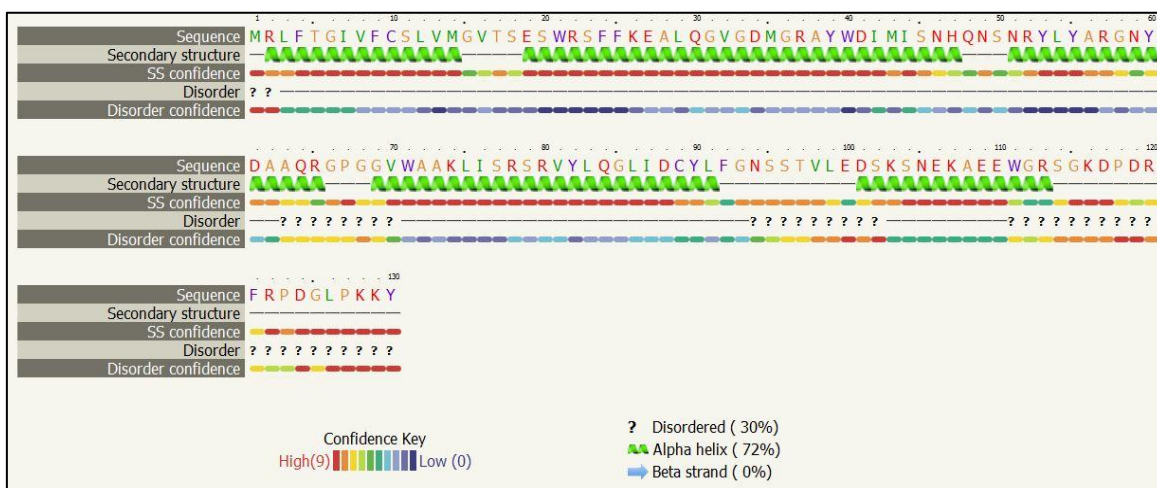
SAA3 naj bi bil po napovedih orodja Phyre 2 (*Job: 97de244bda9faa46*) sestavljen iz snopa petih vijačnic (Slika 24), pri tem naj bi prva vijačnica vsebovala tudi signalni peptid in pa poleg alfa vijačnice tudi transmembransko vijačnico od 15. do 30. mesta v aminokislinskem zaporedju (dolžine 16 AK).



Slika 24: Sekundarna struktura SAA3 z grafično razvidnimi alfa vijačnicami (Phyre 2; Job: 97de244bda9faa46)

SAA4

SAA4 naj bi bil po napovedih orodja Phyre 2 (*Job: 9a9dcd14075cf1bd*) sestavljen iz snopa petih vijačnic (Slika 25), pri čemer je prva vijačnica najverjetneje signalni peptid.



Slika 25: Sekundarna struktura SAA4 z grafično razvidnimi alfa vijačnicami (Phyre 2; *Job: 9a9dcd14075cf1bd*)

4.4 Kvartarne strukture serumskega amiloida A

SAA lahko oblikuje heksamere, sestavljen iz šestih monomerov, kjer je vsak sestavljen iz snopa štirih alfa vijačnic z dodatkom dolgega C-konca. C-terminalni rep snopa alfa vijačnic je razpotegnjen. Vsebuje dve pozitivno nabiti gruči, ki sta vključeni v SAA vezavo s heparinom ter visoko sulfatnim glukozaminom. Ena gruča se nahaja na sredi heksamera, druga pa na vrhu, kjer je mesto vezave HDL, kjer pa lahko predstavlja oviro za vezavo sam heparin [39].

4.5 Funkcionalna napoved SAA iz struktur

Glede na to, da sta SAA1/2 inducibilna proteina, SAA4 pa je konstitutivno izražen, smo iskali med njima tudi razlike struktur, ki bi nam lahko napovedovale določene funkcionalne razlike.

4.5.1 Primerjava aminokislinske sestave izotipov SAA

Da bi lahko ugotovili, kako se primerjajo število in delež posameznih aminokislin primarnih struktur SAA izotipov je bila narejena primerjava aminokislinske sestave SAA1, SAA2, SAA3 in SAA4 (Tabela 11).

Tabela 11: Primerjava aminokislinskih sestav za posamezen izotip SAA

SAA Izotipi	SAA1	SAA2	SAA3	SAA4
Aminokislina	n (%)	n (%)	n (%)	n (%)
Alanin (A)	16 (13,1)	16 (13,11)	13 (10,7)	8 (6,2)
Cistein (C)	1 (0,8)	1 (0,8)	1 (0,8)	2 (1,5)
Asparaginska kislina (D)	9 (7,4)	8 (6,6)	8 (6,6)	8 (6,2)
Glutaminska kislina (E)	7 (5,7)	6 (4,9)	5 (4,1)	6 (4,6)
Fenilalanin (F)	9 (7,4)	7 (5,7)	4 (3,3)	6 (4,6)
Glicin (G)	14 (11,5)	13 (10,7)	14 (11,5)	14 (10,8)
Histidin (H)	3 (2,5)	2 (1,6)	3 (2,5)	1 (0,8)
Izolevcin (I)	3 (2,5)	3 (2,5)	3 (2,5)	5 (3,9)
Lizin (K)	5 (4,1)	5 (4,1)	10 (8,2)	7 (5,4)
Levcin (L)	8 (6,6)	9 (7,4)	8 (6,6)	10 (7,7)
Metionin (M)	3 (2,5)	3 (2,5)	3 (2,5)	4 (3,1)
Asparagin (N)	5 (4,1)	6 (4,9)	5 (4,1)	6 (4,6)
Prolin (P)	4 (3,3)	4 (3,3)	4 (3,3)	4 (3,1)
Glutamin (Q)	2 (1,6)	2 (1,6)	6 (4,9)	4 (3,1)
Arginin (R)	10 (8,2)	12 (9,8)	6 (4,9)	11 (8,5)
Serin (S)	10 (8,2)	11 (9,0)	9 (7,4)	13 (10,0)
Treonin (T)	1 (0,8)	2 (1,6)	5 (4,1)	3 (2,3)
Valin (V)	4 (3,3)	4 (3,3)	6 (4,9)	7 (5,4)
Triptofan (W)	3 (2,5)	3 (2,5)	4 (3,3)	4 (3,1)
Tirozin (Y)	5 (4,1)	5 (4,10)	5 (4,1)	7 (5,4)
Skupaj	122 (100,2)	122 (100,0)	122 (100,3)	130 (100,3)

Opombe: n = število aminokislin, % = odstotek aminokislin

Najpogostejši aminokislini pri vseh izotipih SAA sta alanin in glicin (pri SAA4 tudi serin), naslednje najpogostejše so arginin, asparaginska kislina in serin. Po aminokislinski sestavi najbolj izstopa SAA3, ki ima manj serinov in argininov ter več lizinov, glutaminov, treoninov in valinov kot SAA1/2. SAA4 ima od vseh izotipov SAA4 največ aminokislin potencialno vpletenih v fosforilacijo serinov (13) ter tirozinov (7). Prav tako ima tudi dva cisteina, ki bi lahko bila vpletena v izgradnjo di-sulfidnega mostu.

4.5.2 Primerjava post-translacijskih sprememb izotipov SAA

Po sintezi polipeptidov ali proteinov lahko pride do dodatnih sprememb proteina, ki imajo lahko velik vpliv na njegovo delovanje ter lokalizacijo znotraj ali zunaj celice, kot so fosforilacija, glikozilacija, tirozin sulfatacija, sumoilacija, miristoilacija, manozilacija.

4.5.2.1 Fosforilacija

Je najbolj razširjena oblika post-translacijske modifikacije, saj naj bi bilo vselej do 30 % proteinov v celicah sesalcev fosforiliranih. Reverzibilna fosforilacija proteina vpliva na presnovo, rast, delitev, diferenciacijo, membranski transport, krčenje mišic, imuniteto, učenje, spomin in tako dalje [15]. Za napovedovanje fosforiliranih mest sem uporabil orodje NetPhos 2.0 Server [6].

4.5.2.2 Glikozilacija

Glikozilacija aminokislin je lahko O-vezana ali N-vezana. Več kot polovica vseh proteinov in veliko lipidov v bioloških sistemih je glikoziliranih [20].

N-glikozilacija

Ta post-translacijska modifikacija je bistvenega pomena za preživetje celic in razvoj celičnega cikla. Je metabolični proces, ki deluje tako, da spreminja/zamenjuje ustrezna mesta asparagina v proteinih z oligosaharidnimi strukturami ter tako vpliva na lastnosti in bioaktivnosti proteinov. N-povezani ogljikovi hidrati imajo pomembno vlogo pri zvijanju proteinov, konformaciji, stabilnosti in pri usmerjanju proteina znotraj celice, kakor tudi zunajcelično [34]. N-glikozilacija se začne v endoplazemskem retikulum (ER) in kasneje zavzame mesto v Golgijevem aparatu (GA) [20].

O-glikozilacija (mucinski tip GalNac)

O-glikozilacija je običajno stimulirana znotraj golgijevega aparata, večinoma s pomočjo N-acetil-galaktozaminil-transferaze, ki prenaša N-acetil-galaktozamine (GalNac) na stranske verige serinov ali treoninov [20].

Ima pomembno vlogo pri aktivnosti in transportu proteinov, še posebej pri tistih, ki se izločajo in potujejo skozi celične membrane. O-Glikani tipa mucin naj bi bili vezani na

znotrajcelični protein Orp150, ki je vpleten v jedrno lokalizacijsko zaporedje [49]. Za določanje O-glikozilacijskih mest sem uporabil orodje NetOGlyc 4.0 Server [28].

C-manozilacija

Je novejši tip glikozilacije v primerjavi z N- in O-glikozilacijo. C-manozilacija vsebuje α -manopiranozil, ki je kovalentno pripet na triptofanov indolni C2 ogljikov atom proteina. Prisotnost C-manozilacije naj ne bi imela vpliva na encimsko aktivnost [21]. Za določevanje manozilacijskih mest sem uporabil orodje NetCGlyc 1.0 Server [29].

4.5.2.3 Tirozinska sulfatacija

Je post-translacijska modifikacija, kjer se sulfatna skupina veže na tirozin. Glavna naloga je krepitev interakcij med proteini. Poleg tega pa je potrebna pri procesih koagulacije krvi (zazan je bil na fibrinogenu, Faktorjih V, VIII, IX), pri vezavi levkocitov na endotelij, vezavi hormonov na receptorje, vezavi liganda na kemokinske receptorje ter pri interakcijah proteinov v ekstracelularni matriks [2]. Za določevanje mest tirozinske sulfatacije sem uporabil orodje The Sulfinator [4].

4.5.2.4 N-terminalna miristoilacija

Je ko-translacijska modifikacija vključena v proteinske interakcije in pomembno vlogo ima tudi pri utrjevanju polipeptidov fosfolipidni dvosloj. Sestavljen je iz 14 ogljikovih nasičenih maščobnih kislin, miristinske kisline preko amidne vezi do NH₂-terminalnega glicina tarčnega proteina. Je zelo specifična modifikacija, pomembna za delovanje veliko regulatornih proteinov in encimov. N-miristoilacija je konstitutivna in ireverzibilna modifikacija [1]. Za določevanje mest miristoilacije sem uporabil orodji Myristolator in NMT [8].

4.5.2.5 Sumoilacija

Je post-translacijska modifikacija, kjer se SUMO protein kovalentno veže na protein. Prisoten je v procesih kot so jedrno citosolni transport, proces transkripcijske regulacije, apoptozi (kontrolirana celična smrt), stabilnosti proteina, odziv na stres [25]. Je vključena oziroma se pojavlja v signalnih poteh, znotraj jedra in tudi v ostalih delih celice. [3].

Rezultati zgoraj opisanih posameznih post-translacijskih modifikacijskih sprememb SAA1-4 so prikazani v tabeli (Tabela 12).

Tabela 12: Post-translacijske modifikacije izotipov SAA.

Posttranslacijske Modifikacije	SAA1	SAA2	SAA3	SAA4	Uporabljeno orodje	
Fosforilacija	Serin	40, 106	40, 94, 106	40, 50, 106	20, 50, 102, 104, 114	NetPhos 2.0 Server
	Treonin	/	87	87	/	
	Tirozin	47, 53	47, 53	47, 53	/	
N-glikozilacija	/	/	/	94 (0.6)	NetNGlyc 1.0 Server	
O-glikozilacija (mucinski tip GalNac)	106	106	106	114	NetOGlyc 4.0 Server	
Tirozin sulfatacija	/	/	/	/	The Sulfinator	
N-terminal miristoilacija	6, 58, 68, 90	6, 58, 69, 90	6, 58, 68, 97	6, 58, 68, 85, 93	- Myristoylator - NMT - ProteinPredict	
C-manozilacija	/	/	/	/	NetCGlyc 1.0 Server	
Sumoilacija	64 (0.6) 108 (0.6) 52 (0.2)	64 (0.6) 52 (0.2)	108(0.6) 49(0.5) 52(0.2)	116(0.6) 107(0.5)	SUMOplot™ Analysis Program	

Opombe: Številke označujejo mesto modifikacije v določenem izotipu, v oklepaju pa je označena verjetnost modifikacije na tem mestu

Največ fosforiliranih mest vsebujeta SAA2 in SAA3, vsaka po 6. Najmanj pa SAA1 in sicer 4. Treoninska fosforilacija je prisotna le v izotipih SAA2 in SAA3, medtem ko je tirozinska fosforilacija prisotna pri vseh izotipih razen SAA4. Pri SAA4 so bila zaznana le mesta serinske fosforilacije. Pri SAA1, SAA2 in SAA3 se vsa mesta fosforilacije razporejena med 40. in 160. mestom, pri SAA4 pa so fosforilacijska mesta razporejena že od 20. mesta pa vse do 114. mesta v zaporedju. N-glikozilacija je prisotna le pri SAA4 in sicer na 94. mestu AK zaporedja s 60 % verjetnostjo, medtem ko pa je bila O-glikozilacija zaznan pri vseh izotipih SAA1-4. Prisotnost tirozinske sulfatacije in C-manozilacije ni bila zaznana pri nobenem izotipu SAA1-4. Miristoilacijska mesta so bila zaznana pri vseh izotipih SAA le z spletnim orodjem ProteinPredict. Za izotipa SAA1 in SAA2 so zaznana skoraj popolnoma enaka miristoilacijska mesta. Največ miristoilacijskih mest ima protein SAA4. Največ mest sumoilacije ima izotip SAA1 (2 mesti), če upoštevamo le tista, ki so napovedana z 60 % verjetnostjo ali več. Ostali trije izotipi SAA2-4 pa imajo pri enakem pogoju vsak po eno mesto sumoilacije. Najmanj mest sumoilacije ima SAA2, če upoštevamo le tista mesta sumoilacije, ki so bila napovedana s 50 % verjetnostjo ali več.

4.5.3 Proteinske interakcije SAA1-4

Proteinske interakcije lahko veliko povedo o funkcijah, pa tudi o lokaciji proteina glede na protein, s katerim sodeluje v interakciji.

SAA1 sodeluje v interakcijah, po podatkih podatkovnih baz in analizah spletnih orodij z največ različnimi proteini (COL4A1, COL4A2, COL4A3, COL4A4, FPRL-1, LAMA1, VIMP/SELS, VKORC1), ti so bili določeni z večino orodij. BAT2 (D6S51), ENSG00000232432, UBC in SAA2 so bili določeni z poskusi pri visoki verjetnost (vsaj 71,1 % verjetnost). Preko »Textmining« metode so bili zaznana še proteini SCARB1 (85,9 %), CRP (80,6 %), CTR9 (80,0 %) in LDHC s 67,5 % verjetnostjo (Tabela 13).

SAA2 zagotovo sodeluje v interakcijah s proteini COL4A1, LAMA1 IN APOA1. SAA1, CRP in APCS so bili zaznani le s pomočjo podatkovne baze String. Poleg naštetih so v podatkovni bazi HighPredict zaznani tudi COL4A2, COL4A3, COL4A4, FPRL-1 in VIMP, vendar slednji niso zelo zanesljivi, saj so bili določeni s pomočjo proteina SAA1/2 (P02735), kar pa je veljal za predhodnika SAA1 (P0DJ18) in SAA2 (P0DJ19).

SAA3 je vključen v signalnih poteh metabolizma selena. SAA4 sodeluje v interakcijah s proteini COL4A1, NUDT9P1, SELS, LDHC, LDHA in KCNC1.

Tabela 13: Proteinske interakcije z izotipi SAA1-4

SAA1 Vežan protein	SAA1 Vir	SAA2 Vežan protein	SAA1 Vir	SAA3 Vežan protein	SAA Vir	SAA4 Vežan protein	SAA1 Vir
COL4A1	wiki pi, Ncbi, HPRD (in Vitro)	COL4A1	Wiki pi HPRD (in Vitro) String Experiments (0,619)	Selenium pathway	NCBI	COL4A1	wiki pi Ncbi HPRD (in Vitro) HitPredict (High confidence) String Experiment (0,634)
COL4A2	wiki pi, Ncbi, HPRD (in Vitro)	*COL4A2	HitPredict (High confidence)			NUDT9P1	wiki pi ncbi
COL4A3	wiki pi, Ncbi, HPRD (in Vitro)	*COL4A3	HitPredict (High confidence)			SELS	HitPredict (High confidence)
COL4A4	wiki pi, Ncbi, HPRD (in Vitro)	*COL4A4	HitPredict (High confidence)			LDHC	String Textmining (0,732)
FPRL-1	wiki pi, Ncbi, HPRD (in Vitro)	*FPRL-1	HitPredict (High confidence)			LDHA	String Textmining (0,609)
LAMA1	wiki pi, Ncbi, HPRD (in Vitro)	LAMA1	wiki pi HPRD (in Vitro) String Experiments (0,619)			KCNC1	String Textmining (0,679)
VIMP (SELS)	wiki pi, Ncbi, HPRD (in Vitro) String Experiments (0,674)	*VIMP	HitPredict (High confidence)				
VKORC1	wiki pi ncbi	APOA1	wiki pi ncbi String Experiments (0,789)				
SCARB1	String Textmining (0,859)	SAA1	String Coexpression, Homology (0,782)				
CRP	String Textmining (0,806)	CRP	String Textmining (0,788)				
CTR9	String Textmining (0,800)	APCS	String Textmining (0,671)				
BAT2 (D6S51)	String Experiments (0,786)						
SAA2	String Coexpression, Homology (0,782)						
ENSG00000232432	String Experiments (0,718)						

UBC	String Experiments (0,711)
LDHC	String Textmining (0,675)

Opombe: * Potencialni proteini ki sodelujejo v interakciji s SAA2, saj so bili določeni preko skupnega proteina z oznako SAA1/2 (P02735)

Glukozaminoglikani (GAGs) kot sta sulfaten heparin/heparan (HS) so povezani z reaktivno amiloidozo. Z analizami strukturnih mutacij so odkrili pozitivno nabita mesta v GAG in vsaj dve vezavni mesti HS na SAA. N-konec monomera SAA1.1 (vijačnica 1) naj bi bil pomemben del pri vezavi HDL na SAA [13].

Potencialni receptorji za SAA so FPRL-1/ALX, TLR-2, TLR-4, SR-BI/ABCA1, TANIS, CD55, RAGE. FPRL-1/ALX naj bi imel pomembno vlogo pri vnetju in sprostitvi IL-6, IL-8. Reguliral naj bi NF-KB, aktivnost AP-1 vezave DNK. Pomembno vlogo ima pri zaščiti nevrofilcev pri vnetjih in reumatskih sinoviocitih pred apoptozo ter sodeluje v regulaciji migracije levkocitov pri vnetju. Receptor TLR-2 je vpleten v aktivacijo NF-KB. Poveča ekspresijo dveh protivnetnih citokinov IL-10 in IL-1R receptorju agonistu, poveča ekspresijo IL-23p19, zmanjša pa ekspresijo IL-12p40 in TNF- α . Receptor TLR-4 ne stimulira nobene sinteze v makrofagih. Povzroča fosforilacijo p42/44 in p38 MAPK. Receptor SR-BI/ABCA1 povzroča povečane ravni fosforilacije MAPK p42/44, p38 in JNK. Vpliva na povečan vnos holesterilnih estrov. TANIS je izražen pri sladkorni bolezni tipa 2. CD55 je v velikih količinah ekspresiran v celicah malignih tumorjev. Medtem ko receptor RAGE aktivira NF-KB, MAPK, p42/44 in p38, kar vodi do povečane ekspresije monocitnega tkivnega faktorja [33].

4.5.4 Lokalizacija SAA1-4

Za vlogo proteina je pomembna njegova lokalizacija znotraj celice, kakor tudi zunajcelično.

S pomočjo večine portalov, orodij in podatkovnih baz sem prišel do zaključka, da je lokalizacija akutno faznih izotipov SAA1, SAA2 in SAA3 najverjetneje zunajcelična. Orodje Hum-mPLoc2 [13, 14], Orodje CELLO [12] prikazuje SAA1-4 kot družino, ki ima zunajcelično lokalizacijo, podobno velja za podatkovno bazo HPRD, vendar za razliko ta ne vsebuje nobenih podatkov za izotip SAA3.

CELLO sicer nakazuje možnosti lokalizacije SAA1-2 v mitohondriju, pri SAA1-2 z malo manjšo vrednostjo kot je napovedana zunajcelična lokalizacija. Orodje Psort2 izotipom SAA1-3 prav tako napoveduje zunajcelično lokalizacijo, pri čemer pripisuje zunajcelični lokalizaciji SAA1 največjo verjetnost (77,8%). Podatkovna baza BioGPS in portal NCBI za SAA1 navajata poleg zunajcelične lokalizacije tudi lokalizacijo z lipoproteini visokih gostot delcev, zunajceličnim vezikularnim eksosomom in endocitnim lumnom vezikla, medtem ko za SAA2 navajata le lokalizacijo z lipoproteini visokih gostot delcev.

Lokalizacijo SAA4 več kot polovica portalov/orodij/podatkovnih baz prikazuje kot zunajcelično. Orodje CELLO daje lokalizaciji v mitohondriju (vrednost 1,3) malenkostno prednost pred zunajcelično lokalizacijo (1,2). Orodje Psort2 navaja kot najbolj verjetno lokalizacijo v citoplazmi (30,4 %), sledita pa jima lokalizacija v mitohondriju (26,1 %) in zunajcelična (21,7 %). Portal NCBI oziroma podatkovna baza BioGPS navajata poleg zunajcelične lokalizacije tudi lokalizacijo z lipoproteini visokih gostot delcev.

Pri vseh štirih izotipih SAA je bilo z orodjem TargetP [9] določeno, da so SAA1-4 del izločevalne poti in da z vsaj 88,6 % verjetnostjo vsebujejo signalni peptid (Tabela 14).

Tabela 14: Podatki o lokalizacijah SAA1-4 s pomočjo portalov/podatkovnih baz in napovedovalnih orodij

Orodje/portal/ Baza podatkov	SAA1	SAA2	SAA3	SAA4
CELLO v.2.5	- zunajcelična (1.699) - mitohondrij (1.374)	- zunajcelična (1.614) - mitohondrij (1.441)	- zunajcelična (2.626) - mitohondrij (1.102)	- mitohondrij (1.330) -Zunajcelična (1.234)
Hum-mPLoc2	- zunajcelična	- zunajcelična	- Zunajcelična	-Zunajcelična
Psort2	- zunajcelična, vključno s celično steno (0.778)	- zunajcelična, vključno s celično steno (0.667)	- Zunajcelična, vključno s celično steno (0.667)	- citoplazma (0.304) - mitohondrij (0.261) - Zunajcelična, vključno s celično steno (0.217)
TargetP	- Izločevalna pot / vsebuje signalni peptid (0.907)	- Izločevalna pot / vsebuje signalni peptid (0.904)	- Izločevalna pot / vsebuje signalni peptid (0.938)	- Izločevalna pot/vsebuje signalni peptid (0.886)
HPRD	- zunajcelična	- zunajcelična	/	- Zunajcelična
BioGPS (Cellular Component)	- zunajcelična regija, - lipoprotein visoke gostote delcev, - Zunajcelični vezikularni eksosom - Endocitni lumen vezikla	- Lipoprotein visoke gostote delcev	/	- Zunajcelična regija, - Lipoprotein visoke gostote delcev
NCBI	- Zunajcelična regija, - lipoprotein visoke gostote delcev, - Zunajcelični vezikularni eksosom - Endocitni lumen vezikla	- lipoprotein visoke gostote delcev	/	- Zunajcelična regija, - Lipoprotein visoke gostote delcev

Kolokalizacija SAA v HCAEC je bila opažena v povezavi z mikrotubuli, redkeje z vimentin filamenti, medtem ko kolokalizacija z aktinskimi filamentami ni bila zaznana. Prav tako ni bilo zaznane nobene kolokalizacije med SAA in endomembranami/znotrajceličnimi membranami. SAA je bil lociran tudi v citosolu, znotraj jedra, v nanotubulih in eksosomih. Ekspresija mRNA SAA je pokazala sledi SAA v ateroskleroznih plakih, kar nakazuje na potencialno vlogo v aterogenezi. Leta 2007 je bila ekspresija mRNA SAA zaznana v humanih endotelnih celicah koronarne arterije (HCAEC), kjer so pokazali, da so celice občutljive in visoko odzivne na SAA, ki predstavlja tudi signal za lastno sintezo. Kasneje so pokazali, da je SAA večinoma lokalizirana v osrednjem delu celice [35].

SAA4 je edini izotip pri katerem je bilo zaznано mesto cepitve zaporedja aminokislin. Največjo verjetnost da je protein prenesen do mitohondrijev (MT) po napovedih orodja MitoProt II [16] imata SAA1 in SAA2. Prenos SAA2 do mitohondrija je 54,7 %, SAA1 pa 42,7%, medtem ko pa imata SAA3 (18,8 %) in SAA4 (22,8 %) zanemarljivo majhni verjetnosti. SAA1-3 imajo po napovedih orodja NetNES 1.1 [17] levcinska bogata jedra izvoznih signalov na začetku aminokislinskega zaporedja, na signalnih peptidih (mesta od 7 do 14), SAA4 pa v drugi polovici aminokislinskega zaporedja (mesta od 75 do 83) (Tabela 15).

Tabela 15: Mesta cepitve, mesta levcinsko bogatih jeder izvoznih signalov v zaporedju SAA in verjetnosti transporta do MT

	SAA1	SAA2	SAA3	SAA4	Orodje
Mesto cepitve (AK zapored.)	/	/	/	12 (MRLFTGIVFCS)	MitoProt II v1.101
Verjetnost transporta do MT	(0,4266)	(0,5466)	(0,1875)	(0,2281)	
Levcinsko bogata jedra izvoznih signalov	7-14	7-14	7-14	75-83	NetNES 1.1 Server

4.5.5 Potencialne funkcije SAA1-4

Določene funkcije akutno faznih proteinov SAA1/2 so že dalj časa znane. Pomembna sta pri metabolizmu in transportu lipidov, kemotaksi, regulaciji procesa vnetja in drugih [24]. Portal NCBI navaja še, da je SAA1 funkcionalno prisoten v imunskem odzivu, limfocitni kemotaksi, makrofagni kemotaksi, regulaciji vnetnega odziva, nevtrofilski kemotaksi, aktivaciji trombocitov, v pozitivni regulaciji celične adhezije in drugih (Tabela 16). Medtem, ko za SAA2 navaja funkcijo v akutno faznem odzivu. Orodje ProtFun 2.2 [7, 10] napoveduje funkcionalnost SAA1/2 v celični membrani, medtem ko pa daje funkciji transportu in vezavi manjšo vrednost. Genska ontologija: SAA1 - rastni faktor, SAA2 - imunski odziv.

SAA3 in SAA4 naj bi po podatkih orodja ProtFun 2.2 prav tako imela funkcijsko vlogo v celični membrani in pri transportu in vezavi molekul. Genska ontologija: SAA3 - strukturni protein.

Protein SAA4 je vpleten v transport lipidov med lipoproteini [5]. Genska ontologija: SAA4 - strukturni protein.

SAA2/4 naj bi bila po dobljenih rezultatih orodja ProtFun 2.2 encima, medtem ko SAA1/3 pa ne. Orodji HNB in CombFunc nista napovedali nobenih potencialnih funkcij.

Tabela 16: Že znane in potencialne funkcije SAA1-4

SAA1	Vir	SAA2	Vir	SAA3	Vir	SAA4	Vir		
Akutno fazni odziv	ncbi	Akutno fazni odziv	ncbi	Funkcijska kategorija: Celična membrana	ProtFun 2.2	Akutno fazni odziv	ncbi		
Notranji imunski odziv		Metabolizem lipidov		Funkcijska kategorija: Transport in vezava		Vpleten v transport lipidov med lipoproteini	[5]		
Limfocitna kemotaksa		Transport lipidov		Genska ontologija: Strukturni protein		Funkcijska kategorija: Celična membrana			
Makrofagna kemotaksa						Funkcijska kategorija: Transport in vezava			
Negativna regulacija vnetnega odziva			Regulacija procesa vnetja	[6]			Genska ontologija: Strukturni protein		
Nevtrofilska kemotaksa			Funkcijska kategorija: Celična membrana					Je encim	ProtFun 2.2
Aktivacija trombocitov			Funkcijska kategorija: Transport in vezava						
Pozitivna regulacija celične adhezije			Genska ontologija: Imunski odziv						
Pozitivna regulacija izločanja citokinov									
Pozitivna regulacija koncentracije citosolno kalcijevih ionov									
Pozitivna regulacija izločanja interleukina 1									
Regulacija izločanja proteinov		ncbi							
Metabolizem lipidov		[6]							
Transport lipidov									
chemotaksa									

Regulacija procesa vnetja				
Funkcijska kategorija: Celična membrana	ProtFun 2.2			
Funkcijska kategorija: Transport in vezava				
Genska ontologija: Rastni faktor				
Ni encim				

5 ZAKLJUČEK

SAA genska družina je pri človeku sestavljena iz štirih izotipov: SAA1, SAA2, SAA3 in SAA4, ki so izraženi kot svojevrstni proteini z različnimi vlogami pri vnetju, anti-mikrobnih procesih, poškodbah in podobno.

Najbolj podobna, tako strukturno kot funkcijsko sta si SAA1 in SAA2, ki sta tudi oba inducibilna in visoko zvišana med akutno fazo. SAA3 je bil, v preteklosti, mišljen kot psevdogen, ki ni izražen na proteinskem nivoju, vendar je bil SAA3 protein že zaznan v mlečnih žlezah. SAA4 se strukturno najbolj razlikuje od ostalih in domneva se, da je njegova vloga bolj oskrbovalna v celicah, glede na njegovo nizko konstitutivno izraženo znotraj različnih tipov celic.

Znotraj diplomske naloge smo prišli do naslednjih zaključkov z uporabo različnih bioinformatičnih baz in orodij:

1. SAA1 in SAA2 sta si po zaporedju najbolj podobna. Od njiju se najmanj razlikuje SAA3. SAA4 se od drugih izotipov po primarni strukturi v družini SAA najbolj razlikuje. Njegovo zaporedje je še najbolj podobno zaporedju SAA1, z 72 % homologijo in 55 % identičnostjo.
2. SAA1-4 so si po deležu posameznih sekundarnih struktur zelo podobni. SAA1 sicer izstopa po deležu alfa vijačnic za približno 3-6 %, zato pa vsebuje najmanjši delež beta struktur.
3. Post-translacijske spremembe kažejo, da je treoninska fosforilacija prisotna le v SAA2 in SAA3, medtem ko je tirozinska fosforilacija prisotna pri vseh SAA razen SAA4. Vsi serini so pri SAA4 tudi fosforilirani.
4. Vsi SAA proteini so O-glikozilirani, medtem ko je SAA4 tudi N-glikoziliran.
5. SAA1 in SAA3 imata 3 mesta sumoilacije, medtem ko imata SAA2 in SAA4 le dve mesti.
6. Proteinske interakcije kažejo, da sta si SAA1 in 2 zelo podobna glede vezave na druge proteine. Oba se vežeta z adhezivnimi molekulami kot so kolageni (Col4A1-4) in laminin-1 (LAMA1), receptorji FPRL-1 in VIMP ter serumskim proteinom CRP.
7. SAA1 se veže tudi na druge proteine VKORC1, SCARB1, CTR9, BAT2 (D6S51), SAA2, ENSG00000232432, UBC ter LDHC.
8. SAA2 se veže tudi na serumske proteine (APOA1, SAA1, APCS).
9. SAA4 sodeluje v interakcijah z adhezivno molekulo (COL4A1), s receptorjem SELS in ostalimi proteini (NUDT9P1, LDHC, LDHA in KCNC1).
10. Glede lokalizacije so SAA1-3 vsi prevalentno zunajcelični proteini, z verjetnostjo tudi mitohondrijske lokalizacije, medtem ko ima SAA4 manjšo verjetnost zunajcelične lokalizacije kot SAA1-3.

6 LITERATURA IN VIRI

1. D. Aggujaro, N. Borgese, P. Carrera et al., A role for N-Myristoylation in Protein Targeting: NADH-Cytochrome b_5 Reductase Requires Myristic Acid for Association with Outer Mitochondrial But Not ER Membranes, *J Cell Biol.* 135(6) (1996), 1501-1513. 39
2. M. R. Al-Ubaidi in Y. Kanan, Tyrosine O Sulfation : An Overview, *JSM Biotechnol Bioeng.* (2013), 1-5. 37
3. M. Andreou in N. Tavermarakis, SUMOylation and cell signalling, *Biotechnol J.* (2009), 1740-1752. 44
4. A. Bairoch, E. Gasteiger, F. Monigatti et al., The Sulfinator: predicting tyrosine sulfation sites in protein sequences, *Bioinformatics* 18(5) (2002), 769-770. 38
5. S. Batalov, A. L. Su, T. Wiltshire et al., A gene atlas of the mouse and human protein-encoding transcriptomes, *Proc Natl Acad Sci. USA* 101(16) (2004), 6062-6067. 16
6. N. Blom, S. Brunak in S. Gammeltoft, Sequence and structure-based prediction of eukaryotic protein phosphorylation sites, *J Mol Biol.* 294(5) (1999), 1351-1362. 32
7. N. Blom, R. Gupta, L. Juhl Jensen et al., Ab initio prediction of human orphan protein function from post-translational modifications and localization features, *J Mol Biol.* 319 (2002), 1257-1265. 51
8. G. Bologna, S. Duvaud, C. Yvon et al., N-Terminal myristoylation predictions by ensembles of neural networks, *Proteomics* 4(6) (2004), 1626-1632. 40
9. S. Brunak, O. Emanuelsson, H. Nielsen et al., Predicting subcellular localization of proteins based on their N-terminal amino acid sequence, *J Mol Biol.* 300(4) (2000), 1005-1016. 48
10. S. Brunak, L. J. Jensen in H. H. Starfeldt, Prediction of human protein function according to Gene Ontology categories, *Bioinformatics* 19 (2003), 635-642. 52
11. O. Buzan, M. Krank, K. Lakota et al., Acute Phase Proteins in Prototype Rheumatic Inflammatory Diseases, v: F. Veas (ur.), *Acute Phase Proteins – Regulation and Functions of Acute Phase Proteins*, InTech, Rijeka, 2011, 303-306. 2

12. Y. C. Chen, C. H. Lu, C. S. Yu et al., Prediction of protein subcellular localization, *Proteins* 64(3) (2006), 643-651. 47
13. K. C. Chou in H. B. Shen, A top-down approach to enhance the power of predicting human protein subcellular localization: Hum-mPLoc 2.0, *AnalBiochem.* 394(2) (2009), 269-274. 45
14. K. C. Chou in H. B. Shen, Hum-mPLoc: an ensemble classifier for large-scale human protein subcellular location prediction by incorporating samples with multiple sites, *Biochem Biophys Res Commun.* 355(4) (2007), 1006-1011. 46
15. J. Ciesla, T. Fraczyk in W. Rode, Phosphorylation of basic amino acid residues in proteins: important but easily missed, *Acta Biochim Pol.* 58(2) (2011), 137-148. 31
16. M. G. Claros in P. Vincens, Computational method to predict mitochondrially imported proteins and their targeting sequences, *Eur J Biochem.* 241(3) (1996), 779-786. 49
17. T. la Cour, L. Kiemer, A. Mølgaard et al., Analysis and prediction of leucine-rich nuclear export signals, *Protein Eng Des Sel.* 17(6) (2004), 527-536. 50
18. M. C. de Beer, M. S. Kindy, T. Yuan et al., Characterization of constitutive human serum amyloid A protein (SAA4) as an apolipoprotein, *J Lipid Res.* 36 (1995), 526-534. 5
19. G. Deleage in C. Geourjon, SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments, *Comput Appl Biosci.* 11(6) (1995), 681-684. 21
20. R. A. Dwek, P. M. Rudd, P. P. Van der Steen et al., Concepts and Principles of O-Linked Glycosylation, *Crit Rev Biochem Mol Biol.* 33(3) (1998), 151-208. 33
21. A. Furmanek in J. Hofsteenge, Protein C-mannosylation: Facts and questions, *Acta Biochim Pol.* 47(3) (2000), 781-789. 41
22. C. Gabay in I. Kushner, Acute-Phase Proteins and Other Systemic Responses to Inflammation, *N Engl J Med.* 360(6) (1999), 448-454. 1
23. C. Gabay, I. Kushner, Acute-phase Proteins, *Encyclopedia of Life Sciences* (2001), 1-6. 4

24. W. Gaastra, N. Upragarin, W. J. M. Landman et al., Extrahepatic production of acute phase serum amyloid A, *Histol Histopathol.* 20(4) (2005), 1295-1307. 6
25. R. T. Hay, SUMO: A History of Modification, *Mol Cell* 18 (2005), 1-12. 43
26. M. Hecht, L. Kajan, E. Kloppman et al., PredictProtein---an open resource for online prediction of protein structural and functional features, *Nucl Acids Res.* (2014), 1-7. 28
27. L. E. Jensen in A. S. Whitehead, Regulation of serum amyloid A protein expression during the acute-phase response, *Biochem J.* 334(3) (1998), 489-503. 10
28. H. J. Joshi, C. Steentoft, S. Y. Vakhrushev et al., Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology, *EMBO J.* 32(10) (2013), 1478-1488. 36
29. K. Julenius, NetCGlyc 1.0: Prediction of mammalian C-mannosylation sites, *Glycobiology* 17(8) (2007), 868-876. 42
30. M. Källberg, H. Wang, S. Wang et al., Template-based protein structure modeling using the RaptorX web server, *Nat Protoc.* 7(8) (2012), 1511-1522. 24
31. L. A. Kelley in M. J. E. Sternberg, Protein structure prediction on the web: a case study using the Phyre server, *Nat Protoc.* 4(3) (2009), 163-371. 22
32. F. A. Khan in M. F. Khan, Inflammation and acute phase response, *Int J Appl Biol Pharmaceut Technol.* 1(2) (2010), 312-321. 8
33. A. Kovacevic, E. Malle in S. Sodin-Šemrl, Serum amyloid A: An acute-phase protein involved in tumour Pathogenesis, *Cell Mol Life Sci.* 66(1) (2009), 9-26. 11
34. M. A. Kukuruzinska in K. Lennon, Protein N-Glycosylation: Molecular Genetics and Functional Significance, *Crit Rev Oral Biol Med.* 9(4) (1998), 415-448. 34
35. K. Lakota, K. Mrak-Poljšak, N. Resnik et al., Colocalization of Serum Amyloid A with Microtubules in Human Coronary Artery Endothelial Cells, *J.Biomed Biotechnol.* (2011), 1-8. 9

36. K. Lakota, K. Mrak-Poljšak, B. Rozman et al., Serum Amyloid A and Its Potential Physiological / Pathological Functions - An Overview of Patents, *Recent Pat Endocr Metab Immune Drug Discov.* 4(2) (2010), 89-99. 13
37. H. A. Lashuel, T. Walz, L. Wang et al., Murine apolipoprotein serum amyloid A in solution forms a hexamer containing a central channel, *Proc Natl Acad Sci. USA* 99(25) (2002), 15947-15952. 30
38. K. Lin, V. A. Simossis, W. R. Taylor et al., A Simple and Fast Secondary Structure Prediction Method using Hidden Neural Networks, *Bioinformatics* 21(2) (2005), 152-159. 23
39. J. Lu, Y. Yu, I. Zhu et al., Structural mechanism of serum amyloid A-mediated inflammatory amyloidosis, *Proc Natl Acad Sci. USA* 111(14) (2014), 5189-5194. 20
40. J. Ma, S. Wang, F. Zhao et al., Protein threading using context-specific alignment potential, *Bioinformatics* 29(13) (2013), i257-i265. 25
41. J. F. Morrow, C. G. Peltzman, R. S. Stearman et al., Induction of hepatic synthesis of serum amyloid A protein and actin, *Proc Natl Acad Sci. USA* 78(8) (1981), 4718-4722. 7
42. P. Murphy, R. O'Hara, A. S. Whitehead et al., Acute-phase serum amyloid A production by rheumatoid arthritis synovial tissue, *Arthritis Res.* 2(2) (2000), 142-144. 12
43. J. Peng in J. Xu, A multiple-template approach to protein threading, *Proteins* 79(6) (2011), 1930-1939. 26
44. J. Peng in J. Xu, RaptorX: exploiting structure information for protein alignment by statistical inference, *Proteins* 79(S10) (2011), 161-171. 27
45. Proteinsko zaporedje SAA1 (2014, 14. maj). Najdeno 15. maja 2014 na spletnem naslovu <http://uniprot.org/uniprot/P0DJJ8>. 14
46. Proteinsko zaporedje SAA2 (2014, 14. maj). Najdeno 15. maja 2014 na spletnem naslovu <http://uniprot.org/uniprot/P0DJJ9>. 15

47. Proteinsko zaporedje SAA3 (2012, 16. maj). Najdeno 15. maja 2014 na spletnem naslovu <http://uniprot.org/uniprot/P22614.fasta?version=78>. 17
48. Proteinsko zaporedje SAA4 (2014, 16. april). Najdeno 15. maja 2014 na spletnem naslovu <http://uniprot.org/uniprot/P35542>. 18
49. J. M. Rhodes, B. Tam in L. G. Yu, Effect of O-Glycosylation and Sialylation Inhibitors on Classical NLS-Dependent Nuclear Protein Import in HT29-MTX Human Colon Cancer Cells, *Glycobiology Insights*. 2 (2010), 75-82. 35
50. J. Stevens, Hypothetical structure of human serum amyloid A protein, *Amyloid*. 11(2) (2004), 71-80. 29
51. E. Tîrziu, Acute-phase proteins in immune response, *Lucrări științifice medicină veterinară* 42 (2009), 329-339. 3
52. M. Uhlar in A. S. Whitehead, Serum amyloid A, the major vertebrate acute-phase reactant, *Eur J Biochem*. 265(2) (1999), 501-523. 19