

# Evolucija genetskog koda

---

**Vuković, Borna Branimir**

**Undergraduate thesis / Završni rad**

**2021**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:217:034679>

*Rights / Prava:* [In copyright](#)/[Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2024-11-29**



*Repository / Repozitorij:*

[Repository of the Faculty of Science - University of Zagreb](#)



Sveučilište u Zagrebu  
Prirodoslovno-matematički fakultet  
Biološki odsjek

Borna Branimir Vuković

**Evolucija genetskog koda**  
Završni rad

Zagreb, 2021.

Ovaj rad je izrađen na Zavodu za biokemiju Prirodoslovno-matematičkog fakulteta u Zagrebu,  
pod voditeljstvom doc. dr. sc. Jasmine Rokov Plavec.

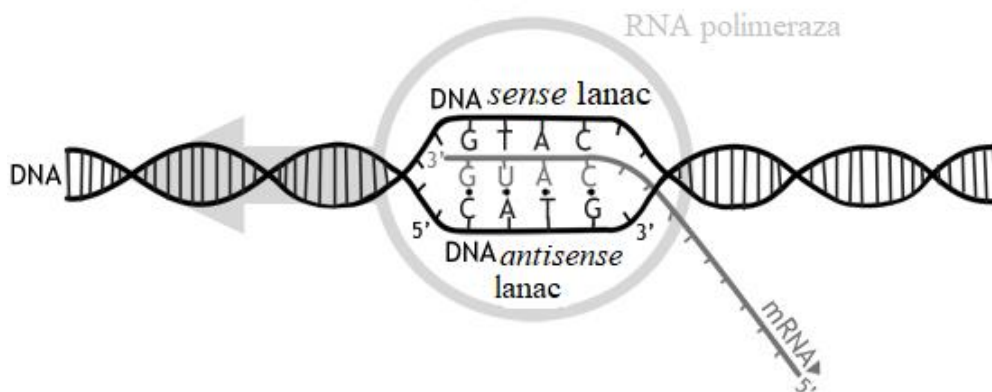
# SADRŽAJ

1. UVOD .....	1
1.1. Transkripcija i translacija .....	1
1.2. Značajke genetskog koda .....	3
1.2.1. Univerzalnost .....	3
1.2.2. Otpornost na mutacije .....	3
2. IZNIMKE KOJE POTVRĐUJU PRAVILO .....	5
3. POGLED U PROŠLOST – HIPOTEZE O EVOLUCIJI KODA .....	9
3.1. Hipoteza zamrznute slučajnosti .....	9
3.1.1. Rana evolucija i horizontalni prijenos gena .....	10
3.2. Rane i kasne aminokiseline .....	11
3.3. Stereokemijska hipoteza .....	12
3.4. Koevolucijska hipoteza .....	16
3.5. Hipoteza umanjivanja grešaka .....	19
3.6. Evolucija translacijske mašinerije .....	24
4. POGLED U BUDUĆNOST - PROŠIRIVANJE GENETSKOG KODA ..	29
5. ZAKLJUČAK .....	32
6. LITERATURA .....	33
7. SAŽETAK .....	38
8. SUMMARY .....	39
9. ŽIVOTOPIS .....	40

# 1. UVOD

## 1.1 Transkripcija i translacija

Genetski kod je informacija u obliku slijeda dušičnih baza u DNA koje kodiraju specifičan slijed aminokiselina. Posrednička molekula u tom procesu je *messenger* RNA – mRNA. Ona sadrži slijed dušičnih baza ekvivalentan kodirajućem, odnosno *sense* lancu DNA, a komplementaran lancu kalupa, odnosno *antisense* lancu s kojeg se prepisuje u procesu transkripcije (Slika 1.)



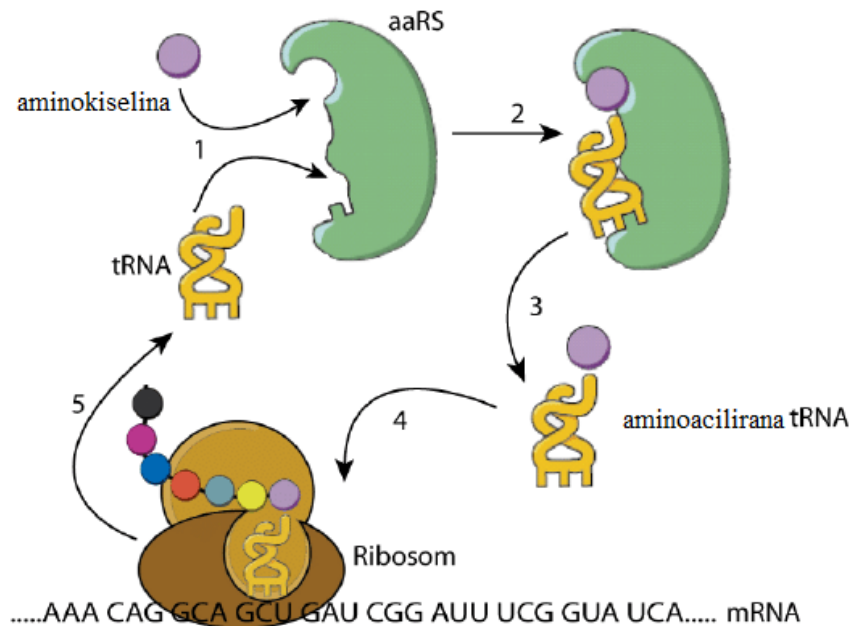
Slika 1. Shematski prikaz procesa transkripcije s označenim *sense* i *antisense* lancem DNA, te enzimom RNA-polimerazom koji sudjeluje u procesu. Preuzeto i prilagođeno prema: [www.pediaa.com/difference-between-sense-and-antisense-strand/](http://www.pediaa.com/difference-between-sense-and-antisense-strand/) (pristupljeno 30. 7. 2021.).

Procesom translacije se zatim slijed dušičnih baza mRNA prevodi u slijed aminokiselina budućeg proteina, čime je definirana primarna struktura proteina. Informacija na mRNA organizirana je u triplete dušičnih baza – kodone, gdje svaki kodon kodira jednu aminokiselinu. Osim toga postoje i tripleti dušičnih baza koji signaliziraju prekid translacije – stop kodoni. Stoga, kako bi se "jezik" od 4 slova dušičnih baza efikasno preveo u 20 aminokiselina zajedničkih svim živim bićima, genetski kod mora sadržavati minimalno  $4^3 = 64$  tripleta dušičnih baza (Alberts, 2015). Tripleti te njima kodirane pripadajuće aminokiseline prikazani su na Slici 2.

1. pozicija (5' kraj)	2. pozicija				3. pozicija (3' kraj)
↓	U	C	A	G	↓
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Slika 2. Genetski kod. Prikazani su kodoni na mRNA i pripadajuće aminokiseline koje kodiraju, uključujući i stop kodone (slova U, C, A i G označavaju redom dušične baze uracil, citozin, adenin i gvanin). Prva pozicija u kodonu (lijevi stupac) predstavlja 5'-kraj kodona na mRNA, a treća pozicija 3'-kraj kodona. Aminokiseline su označene kraticama, a puni nazivi odgovarajućih kratica na engleskom jeziku dostupni su na: [www.ddbj.nig.ac.jp/ddbj/code-e.html](http://www.ddbj.nig.ac.jp/ddbj/code-e.html) (pristupljeno 30. 7. 2021.). Preuzeto i prilagođeno prema: Alberts, 2015.

Molekule koje posreduju u procesu translacije su *transfer* RNA – tRNA. One u svojoj strukturi sadrže antikodon – triplet dušičnih baza komplementaran kodonu. Također, one na svom 3'-kraju vežu aminokiselinu koja odgovara kodonu prema pravilima organizacije genetskog koda prikazanim na Slici 2. Interakcijom tRNA, mRNA i ribosoma se zatim jednu po jednu aminokiselinu gradi primarna struktura proteina. Međutim, tRNA molekule ne vrše proces prepoznavanja odgovarajuće aminokiseline. Taj proces, nazvan aminoacilacija, kataliziraju enzimi aminoacil-tRNA-sintetaze (Slika 3). Upravo je interakcija između aminoacil-tRNA-sintetaza i pripadajućih tRNA ključna za pravilno čitanje genetskog koda (Nelson i Cox, 2017).



Slika 3. Pojednostavljeni prikaz translacije, s označenim aminokiselinama (raznobojni kružići), zeleno obojanom aminoacil-tRNA-sintetazom (aaRS), žuto obojanom tRNA, smeđe obojanim ribosomom te kodonima mRNA na dnu slike. Preuzeto i prilagođeno prema: Wang i sur., 2012.

## 1.2. Značajke genetskog koda

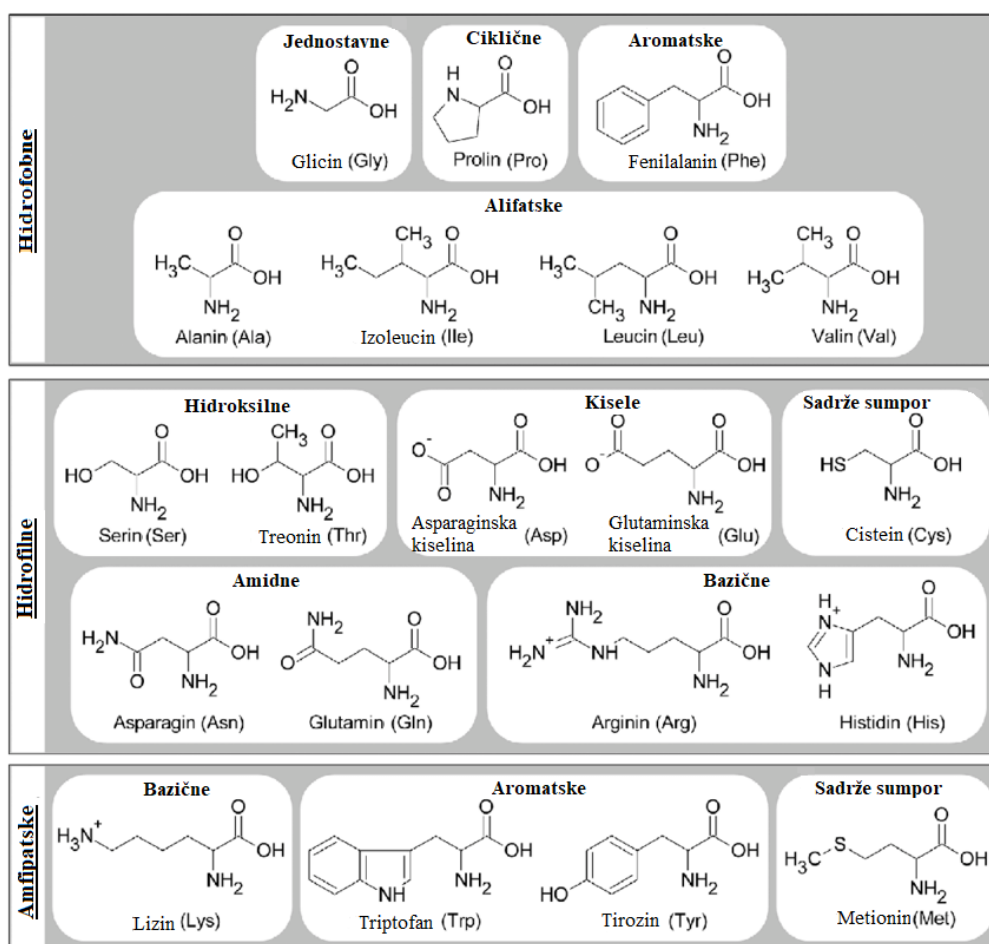
### 1.2.1. Univerzalnost

Prethodno opisani procesi transkripcije i translacije funkcioniraju gotovo na istim načelima u svim oblicima života na Zemlji (Koonin i Novozhilov, 2017). Tako je i genetski kod gotovo univerzalan u svim organizmima, od jednostaničnih prokariota do kompleksnih višestaničnih eukariota. Ta se značajka koda danas smatra činjenicom (Koonin i Novozhilov, 2017), što proizlazi iz brojnih istraživanja čiji su pioniri bili uključeni u otkriće genetskog koda i njegovih svojstava (Crick, 1968; Woese i sur., 1964). Moguća objašnjenja univerzalnosti, kao i ostalih svojstava koda, bit će opisana u ovom radu.

### 1.2.2. Otpornost na mutacije

Ranim istraživačima genetskog koda su i druga njegova svojstva osim univerzalnosti bila dobro poznata. Uz samu strukturu koda prikazanu na Slici 2 (Crick, 1966b), znalo se da 20 aminokiselina nisu nasumično kodirane kodonskim tripletima. Određeni su se uzorci mogli primijetiti unutar koda. Naime, na Slici 2 vidljivo je da kodoni oblika XYC i XYU, gdje su X i Y baze na prvoj i drugoj poziciji kodona, uvijek kodiraju istu aminokiselinu. Također, kodoni oblika XYA i XYG gotovo uvijek kodiraju istu aminokiselinu, uz iznimke metionina i triptofana (Crick, 1968). Takav uzorak osigurava otpornost koda na tranzicijske mutacije.

Naime, ukoliko se u trećoj bazi kodona dogodi mutacija pirimidinske dušične baze u pirimidinsku (C u U ili obratno) ili purinske u purinsku (A u G ili obratno), u najvećem broju slučajeva mutirani triplet će kodirati za istu aminokiselinu, uz navedene iznimke (Nelson i Cox, 2017). Također, aminokiseline sličnih kemijskih svojstava su kodirane bliskim kodonima, što je vidljivo uspoređujući Sliku 2 i Sliku 4, na kojoj su prikazane aminokiseline grupirane prema svojstvima.



Slika 4. Aminokiseline grupirane prema kemijskim svojstvima. Gornji red sadrži hidrofobne aminokiseline, srednji red hidrofilne, a donji amfipatske. Unutar redova navedena su svojstva pojedinih aminokiselina. Preuzeto i prilagođeno prema: Vnućec i sur., 2016.

Primjerice, hidrofobne aminokiseline su kodirane kodonima koji na drugoj poziciji sadrže uracil, a aromatske aminokiseline (Phe, Tyr i Trp) kodirane su kodonima koji na prvoj poziciji sadrže uracil (Crick, 1968). Stoga je genetski kod otporan i na mutacije u drugim pozicijama unutar kodona. Takvim rasporedom kodona osigurava se da supstitucijske mutacije u kodonu dovode do ugradnje aminokiseline sličnih svojstava ispravnoj kako bi se smanjio negativan učinak na promjenu konačne strukture i funkcije proteina (Koonin i Novozhilov, 2017).



## 2. IZNIMKE KOJE POTVRĐUJU PRAVILO

Genetski kod, iako univerzalan, produkt je prirodne selekcije te je podložan varijacijama. Ipak, varijacije su ograničene jer bi značajne promjene u kodu bile pogubne za strukturu proteina u stanici, što bi za posljedicu imalo smrt organizma ukoliko su zahvaćeni esencijalni proteini (Nelson i Cox, 2017). Uočena odstupanja od standardnog koda najvjerojatnije su rezultat specifičnih selekcijskih pritisaka na pojedine organizme, primjerice endosimbionte. Upravo je u mitohondrijima pronađeno najviše poznatih odstupanja (Sengupta i Higgs, 2015). Jedan od značajnijih mehanizama kojim se postižu varijacije u kodu je supresija stop kodona, čime se umjesto zaustavljanja sinteze proteina uvodi nova aminokiselina i lanac se produžuje. Osim toga, mogući su i rearanžmani koda u kojima kodon za jednu aminokiselinu kodira neku drugu. Nekoliko je hipoteza ponuđeno kako bi se ta odstupanja objasnila (Mühlhausen i sur., 2016; Kollmar i Mühlhausen, 2017). U ovom će radu biti predstavljena samo ona koju su predložili Mühlhausen i sur. (2016) jer objašnjava sva do sad otkrivena odstupanja. Naime, oni smatraju da se rearanžmani koda javljaju ukoliko se dogodi mutacija ili gubitak gena za neku tRNA ili terminacijski faktor transkripcije, čime se pripadajući kodon (za neku aminokiselinu ili stop kodon) "oslobađa". Taj kodon zatim mogu prepoznati druge tRNA molekule s odgovarajućom mutacijom u antikodonu kojom postaju komplementarni slobodnom kodonu. Povezivanje mutirane tRNA i odgovarajuće aminokiseline katalizira aminoacil-tRNA-sintetaza (aaRS). Stoga je bitno da ona može prepoznati mutiranu tRNA. Jedan od načina kako se to može dogoditi u prirodi je ukoliko mjesto prepoznavanja za tRNA na aaRS ne obuhvaća antikodon pripadajuće tRNA. Tada je moguće da sintetazi "promakne" mutacija u antikodonu, čime je rearanžman standardnog koda omogućen. Neki primjeri odstupanja od standardnog koda navedeni su u Tablici 1, no taj se popis širi kako raste broj podataka o genomima sve većeg broja organizama (Kollmar i Mühlhausen, 2017).

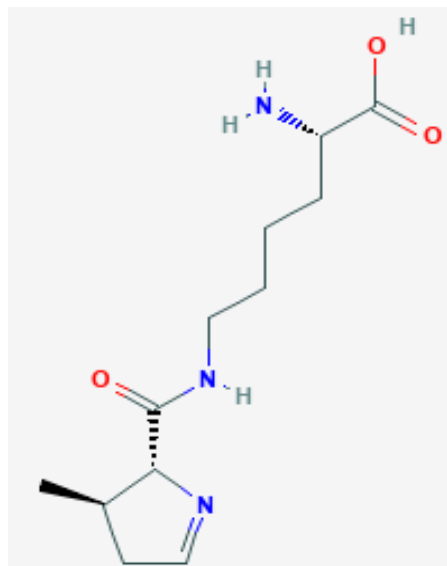
Tablica 1. Primjeri odstupanja od standardnog koda. S lijeve strane strelice u lijevom stupcu navedeni su međuodnosi aminokiselina i kodona u standardnom kodu, a s desne strane strelice navedeno je odstupanje. Slovo N u kodonu označava bilo koju od 4 dušične baze, a slovo R označava purinsku bazu – A ili G ([www.ddbj.nig.ac.jp/ddbj/code-e.html](http://www.ddbj.nig.ac.jp/ddbj/code-e.html), pristupljeno 30. 7. 2021.). Oznaka "Nedodijeljeno" označava slučajeve u kojima ne postoji odgovarajuća tRNA koja interagira s navedenim kodonom, stoga taj kodon više ne nosi informaciju za aminokiselinu. U zagradama u desnom stupcu naveden je pronađeni broj slučajeva pojedinog rearanžmana u odgovarajućim organizmima. Podaci u tablici preuzeti su iz: Sengupta i Higgs, 2015.

<b>Rearanžman koda</b>	<b>Lokacija (broj slučajeva)</b>
UUA: Leu → Stop	mitohondriji (1)
UCA: Ser → Stop	mitohondriji (1)
UAG: Stop → Leu	mitohondriji (2)
UAG: Stop → Ala	mitohondriji (1)
UGA: Stop → Trp	mitohondriji (12), <i>Ciliophora</i> (2), bakterije (3)
CUN: Leu → Thr	mitohondriji (1)
CUN: Thr → Nedodijeljeno	mitohondriji (1)
CGN: Arg → Nedodijeljeno	mitohondriji (5)
AUA: Ile → Met ili Nedodijeljeno	mitohondriji (3)
AAA: Lys → Asn	mitohondriji (2)
AAA: Lys → Nedodijeljeno	mitohondriji (1)
AGR: Arg → Ser	mitohondriji (1)
AGR: Ser → Stop	mitohondriji (1)
AGR: Ser → Gly	mitohondriji (1)
UAR: Stop → Gln	zelene alge (1), <i>Ciliophora</i> (4), <i>Diplomonadida</i> (1)
UGA: Stop → Cys	<i>Ciliophora</i> (1)
UGA: Stop → Nedodijeljeno	<i>Ciliophora</i> (2)
CUG: Leu → Ser	gljive (1)
CGG: Arg → Nedodijeljeno	bakterije (1)
AUA: Ile → Nedodijeljeno	bakterije (1)
AGA: Arg → Nedodijeljeno	bakterije (1)

Osim varijacija u kodonima za neke od 20 standardnih aminokiselina, među organizmima na Zemlji pojavila su se i proširenja koda aminokiselinama selenocisteinom i pirolizinom.

Selenocistein (Sec) je aminokiselina koja umjesto atoma sumpora u tiolnoj skupini cisteina sadrži atom selenija. Pronađen je u organizmima iz svih triju domena (*Archaea*, *Bacteria*, *Eukarya*) te sudjeluje u pojačavanju katalitičke aktivnosti selenoproteina zbog veće reaktivnosti u odnosu na cistein (Ambrogelly i sur., 2007). Povezivanje selenocisteina i pripadajuće tRNA katalizira seril-tRNA-sintetaza. Selenocistein stoga ne postoji kao slobodna aminokiselina, već se sintetizira iz serina. Detalji sinteze neće biti obuhvaćeni u ovom radu, no opisali su ih Yuan i sur., 2010. Za ugradnju selenocisteina u polipeptidni lanac potrebno je da pripadajuća tRNA<sup>Sec</sup> prepozna UGA stop kodon, nizvodno od kojeg se nalazi tzv. SECIS mRNA omča na koju se veže specijalizirani elongacijski faktor koji omogućuje vezanje pripadajuće tRNA<sup>Sec</sup> na ribosom (Ambrogelly i sur. 2007).

Pirolizin (Pyl) je aminokiselina dosad pronađena samo u nekoliko vrsta metanogenih arheja gdje je sastavni dio aktivnih mjesta enzima metiltransferaza koji kataliziraju proizvodnju metana (Ambrogelly i sur., 2007). Njena kemijska struktura prikazana je na Slici 5.



Slika 5. Kemijska struktura aminokiseline pirolizina (Pyl). Preuzeto iz: [www.pubchem.ncbi.nlm.nih.gov/compound/Pyrrolysine#section=2D-Structure](http://www.pubchem.ncbi.nlm.nih.gov/compound/Pyrrolysine#section=2D-Structure) (pristupljeno 1. 8. 2021.)

Za razliku od selenocisteina, povezivanje priolizina i pripadajuće tRNA<sup>Pyl</sup> katalizira aminoacil-tRNA-sintetaza specijalizirana za pirolizin – pirolizil-tRNA-sintetaza. To je prvi poznati slučaj sintetaze u prirodi specijalizirane za neku nestandardnu aminokiselinu (Yuan i sur., 2010).

Također, u proučavanju su arhejama pronađeni geni za biosintetski put pirolizina te on postoji kao slobodna aminokiselina u stani. Ugradnja pirolizina u polipeptidni lanac zahtijeva prepoznavanje UAG stop kodona od strane Pyl-tRNA<sup>Pyl</sup>. Zbog malog broja UAG stop kodona u metanogenim arhejama, kao i malog broja proteina koji sadrže Pyl, pretpostavlja se da u tim organizmima UAG nema funkciju stop kodona već kodona za Pyl. Unatoč tome, moguće je i da sekundarne strukture u mRNA, tzv. PYLIS sljedovi, služe kao signal za supresiju UAG stop kodona i ugradnju Pyl (Zhang i sur., 2005).

Ograničen opseg odstupanja od standardnog genetskog koda može se tumačiti kao iznimka koja potvrđuje pravilo. Naime, sva odstupanja proizlaze iz samog koda te u prirodi još nije pronađen genetski kod koji se potpuno razlikuje od standardnog. Stoga pronađene iznimke idu u prilog univerzalnosti koda podložnog ograničenom stupnju varijabilnosti (Nelson i Cox, 2017).

### 3. POGLED U PROŠLOST – HIPOTEZE O EVOLUCIJI KODA

Svojstva genetskog koda navedena u uvodu potaknula su brojna istraživanja kojima se nastojala objasniti njihova evolucija. Od početaka istraživanja genetskog koda do danas nije postignut jasan konsenzus po tom pitanju. U ovom će radu stoga biti predstavljene hipoteze koje su rezultat dugogodišnjih istraživanja genetskog koda (Koonin i Novozhilov, 2017), s naglaskom na činjenicu da one nisu međusobno isključive. Stoga je moguće da je genetski kod nastao kao rezultat međudjelovanja brojnih čimbenika tijekom evolucije.

#### 3.1. Hipoteza zamrznute slučajnosti

Zašto je genetski kod gotovo univerzalan? Znanstvenici su pokušavali naći odgovor na to pitanje od samog početka otkrića koda. Tako je Crick (1968) predstavio hipotezu zamrznute slučajnosti (eng. *frozen accident*). Prema toj hipotezi, svaka promjena u genetskom kodu bila bi letalna za organizam, ili bar pod negativnim selekcijskim pritiskom zato što kod određuje aminokiselinski slijed velikog broja važnih proteina. Ukoliko bi se kod poremetio, pojavile bi se brojne greške u slijedu aminokiselina te bi pravilno funkcioniranje tih proteina bilo onemogućeno. Genetski je kod stoga "zamrznut" u svom današnjem stanju, uz ograničene varijacije obrađene u prethodnom poglavlju.

Današnji je izgled koda, smatra Crick (1968), nastao slučajno u zajedničkom pretku svih organizama (eng. *Last Universal Common Ancestor – LUCA*) evolucijom iz primitivnijih oblika koda koji su kodirali manji broj aminokiselina i proteina. Primitivni se kod potencijalno, ali malo vjerojatno, sastojao od kodona s dvije dušične baze umjesto tri kao u današnjem obliku. Pretpostavku da je LUCA imao razvijen genetski kod dovoljno sofisticiran da bude "zamrznut" i otporan na daljnje modifikacije (Koonin i Novozhilov, 2009) podržavaju rezultati drugih istraživanja. Primjerice, Harris i sur. (2003) su pokazali da su brojni geni za proteine uključene u proces translacije konzervirani u trima domenama života, što ih je navelo na zaključak da su ti geni naslijeđeni od LUCA-e te da je LUCA sadržavao sustav za translaciju sličan onome u organizmima koji postoje danas.

Unatoč prednostima ove hipoteze, ona ne može objasniti specifično povezivanje kodona s pripadajućim aminokiselinama. Sam Crick (1968) ističe da ova hipoteza prepušta slučaju previše koraka u ranoj evoluciji koda te da ju je vrlo teško eksperimentalno provjeriti.

### 3.1.1. Rana evolucija i horizontalni prijenos gena

Crickova hipoteza također nije ponudila objašnjenje jedinstvenosti koda te nije razmotrila mogućnost većeg broja zamrznutih slučajnosti tijekom evolucije. Taj su problem istražili Vetsigian i sur. (2006), predloživši model rane evolucije koda. Prema njihovom modelu univerzalni genetski kod nastao je kao posljedica izmjene gena između prvih oblika života na Zemlji putem horizontalnog prijenosa. Horizontalni prijenos gena, za razliku od vertikalnog (s roditelja na potomstvo), podrazumijeva izmjenu DNA ili RNA između genetski nesrodnih organizama. Pritom genetski materijal donora može zamijeniti neke gene recipijenta ili u njega uvesti nove gene (Keeling i Palmer, 2008). Vetsigian i sur. (2006) smatraju da je u ranim fazama evolucije života postojalo više različitih kodova te da je horizontalni prijenos gena bio najvažniji čimbenik u uspostavljanju jedinstvenog koda zajedničkog svim organizmima. Matematičkim su modelima pokazali da u sustavu s nasumično generiranim kodovima u kojem nije uveden horizontalni prijenos gena kodovi postaju međusobno sve različitiji, a ukoliko se u sustav uvede horizontalni prijenos, kodovi konvergiraju prema jednom univerzalnom kodu. Njihove su rezultate potvrdila i novija istraživanja (Aggarwal i sur., 2016).

Ipak, horizontalni prijenos gena je najefikasniji među stanicama sa srodnijim genetskim kodovima, što je i pokazano ovim istraživanjem. Kako onda objasniti njegovu važnu ulogu u evoluciji ranih oblika života sa znatno različitim kodovima? Vetsigian i sur. (2006) smatraju da je horizontalni prijenos u takvim uvjetima bio omogućen jer novi donirani geni nisu ometali važne funkcije recipijentnih stanica s obzirom na to da su njihovi kodovi bili otporniji na greške od današnjeg. Naime, ti su kodovi bili manje sofisticirani i kodirali su manji broj proteina u usporedbi sa standardnim genetskim kodom stoga su greške u translaciji imale manje negativne učinke nego u današnjim organizmima. Horizontalni je prijenos tada pružao evolucijsku prednost jer su njime uvedene nove kombinacije u kod čime se povećavala efikasnost translacije i broj različitih proteina koji su translatirani. Skupine stanica u kojima je horizontalni prijenos gena bio prisutan u tolikoj mjeri da su im kodovi postali međusobno gotovo identični su stoga postale brojnije od drugih. Posljedično, donirale su veći broj svojih gena dalje i širile vlastitu verziju koda u druge populacije. Rezultat takvog slijeda događaja je univerzalni genetski kod koji je "zamrznut" te daljnji horizontalni prijenos na njega može imati negativan učinak zbog uvođenja grešaka u translaciji.

Horizontalni prijenos gena bi također mogao biti važan čimbenik pri nastajanju iznimaka od standardnog koda obrađenih u prethodnom poglavlju. Naime, pod specifičnim selekcijskim

pristicima kojima su izloženi pojedini organizmi, primjerice endosimbionti, promjene u kodu mogu pružiti evolucijsku prednost koja postaje fiksirana u populaciji putem genetičkog drifta. Ta im je prednost osigurana upravo horizontalnim prijenosom gena (Koonin i Novozhilov, 2017).

### **3.2. Rane i kasne aminokiseline**

Jedna od postavki hipoteze zamrznute slučajnosti je postojanje jednostavnijih kodova iz kojih se razvio univerzalni genetski kod. Jednostavniji su kodovi kodirali manji broj proteinogenih aminokiselina od današnjih 20. Stoga se nameće pitanje koje su aminokiseline bile uključene u sastav prvih proteina. Eksperimenti kojima se nastojao oponašati sastav prebiotske Zemljine atmosfere (Miller, 1953; 1957) dali su prve odgovore na to pitanje i postavili smjer budućih istraživanja. Prebiotsku atmosferu predstavljali su metan, amonijak, vodena para i vodik, a kao izvor energije korišteni su električni izboji pod pretpostavkom da su oni bili važni pokretači kemijskih reakcija u prebiotskim uvjetima. Aminokiseline dobivene u tim eksperimentima su Gly, Ala, Asp i Glu.

Nastavak istraživanja prebiotskih uvjeta na Zemlji donio je i nove pretpostavke o porijeklu proteinogenih aminokiselina. Tako su meteoriti i hidrotermalna vrela bili predloženi kao potencijalni izvori aminokiselina za prve oblike života (Higgs i Pudritz, 2009). Uspoređivanjem sastava i količine aminokiselina u meteoritima, hidrotermalnim vrelima i različitim pokusima koji oponašaju prebiotske uvjete Zemljine atmosfere, Higgs i Pudritz (2009) predložili su 10 konsenzusnih aminokiselina koje su pronašli u svim proučavanim podacima. Njih su nazvali ranima: Gly, Ala, Asp, Glu, Val, Ser, Ile, Leu, Pro i Thr (puni nazivi aminokiselina navedeni su na Slici 4). Druga su istraživanja rezultirala sličnim zaključcima (Zaia i sur., 2008). Ostalih 10 aminokiselina (Arg, Gln, Asn, His, Lys, Cys, Phe, Tyr, Met, Trp) nazvane su kasnima te Higgs i Pudritz (2009) smatraju da su one uklopljene u genetski kod tek nakon razvoja njihovih biosintetskih puteva. Rane aminokiseline naime imaju jednostavnije biosintetske puteve, te su osim toga mogle nastati iz anorganskih prekursora.

Trifonov (2004) je čak predložio kronološki poredak uklapanja proteinogenih aminokiselina u genetski kod spojivši 60 različitih kriterija. Prema tom istraživanju, redosljed uklapanja aminokiselina u kod je sljedeći: Gly, Ala, Asp, Val, Pro, Ser, Glu, Leu ili Thr, Arg, Ile ili Gln ili Asn, His, Lys, Cys, Phe, Tyr, Met, Trp. Prvih 9 aminokiselina te izoleucin (11. ili 12. po redu) pojavljuju se u pokusima koji oponašaju uvjete prebiotske atmosfere te u meteoritima i hidrotermalnim vrelima, što dodatno podržava podjelu na rane i kasne aminokiseline.

### 3.3. Stereokemijska hipoteza

Otkriće strukture molekule DNA potaknulo je pitanje o povezanosti između "jezika" četiriju dušičnih baza i "jezika" 20 aminokiselina. Gamow (1954) je predložio mehanizam ključ-brava između aminokiselina u polipeptidnom lancu i utora između dušičnih baza u lancu DNA. Prema tom bi mehanizmu slobodne aminokiseline ulazile u utore u DNA, na način da 20 različitih utora strukturno odgovara 20 aminokiselina. Tada bi se aminokiseline u susjednim utorima povezivale u polipeptidni lanac. Danas je poznato da translacija ne funkcionira na taj način te postoje dokazi da je već LUCA imao razvijenu translacijsku mašineriju (Harris i sur., 2003). Stoga takav mehanizam vjerojatno nije postojao ni u ranim stadijima evolucije koda. Unatoč tome je važan jer predstavlja prvu formulaciju stereokemijske hipoteze, prema kojoj povezanost kodona i aminokiselina u genetskom kodu slijedi direktno iz fizikalno-kemijskog afiniteta između aminokiselina i pripadajućih kodona ili antikodona (Koonin i Novozhilov, 2009).

Nastavak istraživanja stereokemijske hipoteze donio je brojne pretpostavke o interakcijama aminokiselina i kodona tijekom rane evolucije koda. Tako je primjerice Woese (1968) predložio model prema kojem su pozitivno nabijene aminokiseline interagirale s negativno nabijenom okosnicom polinukleotidnog lanca RNA. Primitivna je translacija prema tome bila neprecizna u usporedbi s današnjom te dvosmjerna – polipeptidi su mogli davati informaciju za sintezu polinukleotidnih lanaca i obratno. Eksperimenti koji su oponašali uvjete prebiotske atmosfere nisu rezultirali sintezom bazičnih aminokiselina koje nose pozitivan naboj, no Woese (1968) predlaže da su u primitivnoj translaciji sudjelovale jednostavnije aminokiseline koje su nosile pozitivan naboj te služile kao prekursori za neke od ranih proteinogenih aminokiselina. Njegov argument počiva na tome što su takve molekule pronađene u navedenim eksperimentima.

Woese i sur. (1966) su proučavali interakcije aminokiselina s piridinom, aromatskom molekulom koja sadrži polarni atom dušika zbog čega je strukturno i kemijski slična pirimidinskim (citozin, timin, uracil) i purinskim (adenin, gvanin) dušičnim bazama (Slika 6).





Slika 6. Usporedba kemijskih struktura piridina, pirimidina i purina. Polarni dušik prikazan je plavom bojom. Preuzeto i prilagođeno prema: [www.pubchem.ncbi.nlm.nih.gov](http://www.pubchem.ncbi.nlm.nih.gov) (pristupljeno 17. 8. 2021.)

Ovo je istraživanje rezultiralo uspostavljanjem polarnog uvjeta (eng. *polar requirement*) – kriterija podjele aminokiselina prema različitoj topljivosti u piridinu koji se i danas koristi u procjenama interakcija između aminokiselina i dušičnih baza (Yarus, 2021). Istraživanje je također pokazalo da aminokiseline interagiraju s piridinom preko polarnog dušika, ali i preko nepolarnog prstenastog dijela molekule. Najvažniji zaključci su da aminokiseline čiji se kodoni razlikuju u 3. poziciji (Slika 2) imaju gotovo identične vrijednosti polarnog uvjeta te da aminokiseline čiji kodoni sadrže uracil na 2. poziciji (hidrofobne aminokiseline) također imaju gotovo identične vrijednosti polarnog uvjeta. Zbog korelacije između vrijednosti polarnog uvjeta pojedinih aminokiselina i njihovih pozicija u genetskom kodu Woese i sur. (1966) zaključuju da su za uspostavu genetskog koda u današnjem obliku ključne bile direktne interakcije aminokiselina i pripadajućih kodona. Također su pokazali da aminokiseline mogu interagirati sa 3-4 molekule piridina istovremeno, što dodatno podržava zaključak da su aminokiseline interagirale s kodonskim tripletima. Unatoč brojnim pretpostavkama o direktnim stereokemijskim interakcijama između aminokiselina i kodona, eksperimenti u to vrijeme nisu jasno dokazali mogućnost takvih interakcija za svaku od 20 proteinogenih aminokiselina, već samo za neke (Crick, 1968). Stoga se smatralo da su za povezivanje kodona i aminokiselina ključne posredničke molekule – adaptori (Crick, 1958). Danas je poznato da ulogu adaptoru imaju molekule tRNA.

Nakon što se pokazalo da specifične sekundarne strukture mRNA – ribosklopke (eng. *riboswitch*) mogu vezati male molekule i regulirati ekspresiju gena (Nahvi i sur., 2002; Winkler i sur., 2002), znanstvenici su bili potaknuti na otkrivanje potencijalnih veznih mjesta za aminokiseline na RNA molekulama te nastavak istraživanja stereokemijske hipoteze i njenog značaja u evoluciji genetskog koda. Za takva su istraživanja korišteni aptameri – RNA molekule generirane *in vitro* koje imaju mogućnost vezanja za specifični ligand (Nelson i Cox, 2017).

Tako su Yarus i sur. (2009) analizirali sekvence 337 veznih mjesta na RNA aptamerima za 8 aminokiselina – Arg, Gln, His, Ile, Leu, Phe, Trp, Tyr. Pokazali su da aptameri interagiraju s polarnim dijelovima aminokiselina te ih na taj način vežu, a veza se učvršćuje specifičnim interakcijama pojedinog aptamera s bočnim ogrankom pripadajuće aminokiseline. Na taj način aptameri mogu razlikovati aminokiseline što je, smatraju Yarus i sur. (2009), bitno za početak procesa kodiranja. Zatim su analizirali postojanje kodonskih odnosno antikodonskih tripleta u blizini veznih mjesta na aptamerima, što bi ukazivalo na direktnu interakciju aminokiselina i kodona ili antikodona. Navedenih 8 aminokiselina povezano je s 24 pripadajuća kodona (Slika 2) i 24 antikodona. U ovom je istraživanju pokazano da se 3 od 24 kodona i 7 od 24 antikodona pojavljuju blizu veznih mjesta na aptamerima. Konkretno, vezna mjesta za Arg i Ile sadrže kodone i antikodone za te aminokiseline, a vezna mjesta za His, Phe, Trp i Tyr sadrže samo antikodone. Gln je pokazao mali afinitet vezanja za aptamere korištene u eksperimentu. Vezna mjesta za Leu nisu sadržavala pripadajuće kodone ili antikodone. Yarus i sur. (2009) zaključuju da su pronađeni tripleti uključeni u kod zbog svoje stereokemijske interakcije s pripadajućim aminokiselinama. Takav se zaključak slaže s hipotezom odbjeglog tripleta (eng. *escaped triplet*) koju su izložili Yarus i sur. (2005), prema kojoj je originalna funkcija određenih tripleta na RNA bila vezanje aminokiselina, no tijekom evolucije su "pobjegli" od te funkcije i poprimili novu – kodona i antikodona. Objašnjenje za ostatak kodona i antikodona koji nisu pronađeni u blizini veznih mjesta na aptamerima Yarus i sur. (2009) nude u kasnijem uključivanju tih tripleta u genetski kod neovisno o stereokemijskim interakcijama, o čemu će više riječi biti kasnije. Prema njihovom modelu, direktne interakcije između RNA i aminokiselina mogle su u ranim fazama evolucije koda služiti za kodiranje polipeptidnog lanca. Taj je proces, smatraju, kasnije optimiziran uvođenjem aminoacil-tRNA-sintetaza. Ipak, ovo istraživanje analizira uglavnom kasne aminokiseline, uz iznimke Leu i Ile, stoga ne nudi objašnjenje za povezanost ranih aminokiselina s njima pripadajućim kodonima odnosno antikodonima (Koonin i Novozhilov, 2017).

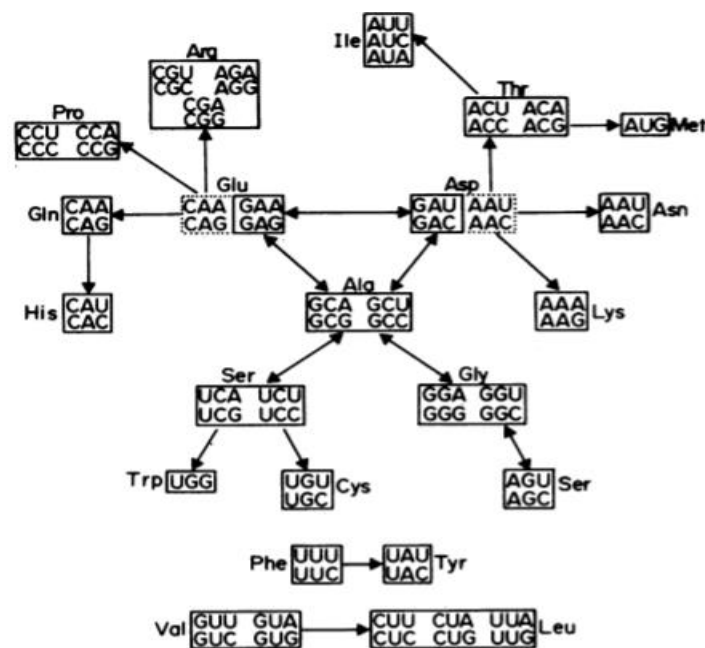
Eksperiment Yarusa i sur. (2009) pokazao je interakciju između aminokiselina i tripleta u *in vitro* uvjetima, što ne dokazuje da su takve interakcije bile prisutne *in vivo* tijekom evolucije koda. Stoga su Johnson i Wang (2010) proučavali interakcije ribosomskih proteina s ribosomskim RNA (rRNA). Ribosom je, smatraju, odličan model za dokazivanje interakcija između aminokiselina i tripleta *in vivo* zato što sadrži velik broj proteina i rRNA čije mu interakcije osiguravaju stabilnost. Također, tvrde Johnson i Wang (2010), ribosom se pojavio u vrijeme rane evolucije koda, prije LUCA-e, pa postoji velika vjerojatnost da sadrži relikte

stereokemijskih interakcija, ukoliko takve postoje, koje su bile važne u uspostavljanju genetskog koda u današnjem obliku. Njihovo je istraživanje pokazalo da unutar strukture ribosoma, aminokiseline Ile, Leu, Thr, Glu, Gly, Arg, Lys i Tyr interagiraju s tripletima na rRNA koji odgovaraju njima pripadajućim kodonima, a Phe, Trp, Asp, Ile, Gln, Tyr, Leu, His, Arg i Met interagiraju s pripadajućim antikodonskim tripletima. Aminokiseline koje interagiraju s antikodonima su, uz iznimke Ile i Leu, kasne. Johnson i Wang (2010) stoga zaključuju da su te aminokiseline uklopljene u kod putem interakcija s antikodonima koje su služile kao baza za kasniju evoluciju pripadajućih tRNA. Uklapanje ranih aminokiselina u kod, smatraju, nije ovisilo o stereokemijskim interakcijama s antikodonskim tripletima, ili su pak takve interakcije izgubljene tijekom evolucije zbog premještanja kodona, o čemu će biti riječi kasnije.

Proučavanje interakcija između proteina i RNA potaknulo je oživljenje hipoteze o primitivnoj translaciji koja proizlazi iz interakcije polipeptida i polinukleotida (Woese, 1968). Istraživanje Hlevnjaka i sur. (2012) pokazalo je da mRNA s većim udjelom pirimidinskih baza (C, U) u sekvenci kodiraju proteine čije aminokiseline imaju veću sklonost interakcije s pirimidinima, što je izmjereno putem polarnog uvjeta. Polyansky i Zagrovic (2013) su proširili to istraživanje proučavajući proteine koji u stanicama stvaraju komplekse s RNA, primjerice ribosomske proteine. Pokazali su da aminokiseline čiji kodoni sadrže više purinskih baza stvaraju više interakcija s purinskim bazama u odnosu na pirimidinske unutar kompleksa RNA i proteina. Isto tako, aminokiseline čiji su kodoni bogatiji pirimidinskim bazama stvaraju više interakcija s pirimidinskim bazama u RNA u odnosu na purinske. Najveća korelacija pokazana je za baze gvanin i citozin, stoga Polyansky i Zagrovic (2013) predlažu model rane evolucije koda u kojem su polinukleotidni lanci bogati bazama G i C služili kao predložak za sintezu polipeptidnih lanaca na temelju direktnih stereokemijskih interakcija. S obzirom da su Gly i Ala najzastupljenije rane aminokiseline (Miller, 1953; Trifonov, 2004), a njihovi su kodoni upravo GGN i GCN, moguće je da su u ranoj fazi evolucije koda bile zastupljenije RNA molekule s visokim udjelom G i C. Kasnijim uključivanjem baza A i U u genetski kod, smatraju Polyansky i Zagrovic (2013), počinju djelovati drugi evolucijski mehanizmi i stereokemijske interakcije više nisu glavni čimbenik evolucije koda. Ipak, ovo istraživanje nije pokazalo specifične interakcije aminokiselina s pripadajućim kodonima te je fokus bio na peptidima umjesto na slobodnim aminokiselinama. Stoga je upitno jesu li pokazane interakcije bile značajne za ranu evoluciju koda (Koonin i Novozhilov, 2017).

### 3.4. Koevolucijska hipoteza

Eksperimenti koji su proučavali interakcije između aminokiselina i tripleta (Yarus i sur., 2005; Yarus i sur., 2009) pokazali su da stereokemijske interakcije mogu objasniti uklapanje malog broja kodona u genetski kod. Kao objašnjenje za uklapanje ostalih kodona predložena je koevolucijska hipoteza. Prema toj hipotezi današnji raspored kodona u genetskom kodu nastao je kao rezultat povezanosti prebiotskih puteva sinteze aminokiselina. Wong (1975) je prvi predložio ovu hipotezu, iako je ideja da su metabolički srodne aminokiseline kodirane bliskim kodonima postojala i prije (Nirenberg i sur., 1963; Pelc, 1965). Naime, Wong (1975) smatra da su rane aminokiseline nastale abiotskim putem (Miller, 1953) služile kao prekursori u biosintetskim putevima kasnijih aminokiselina. Razvoj novih biosintetskih puteva i pojava novih aminokiselina događali su se u isto vrijeme kad i mutacije u kodonima ranih aminokiselina. Novi kodoni nastali tim mutacijama zatim su kodirali za novonastale aminokiseline (Slika 7). Genetski je kod stoga potencijalno koevoluirao s biosintetskim putevima aminokiselina.



Slika 7. Predloženi putevi koevolucije genetskog koda i biosintetskih puteva aminokiselina. Jednosmjerne strelice označavaju put od prekursora do produkta u biosintetskom putu određenih aminokiselina, a dvosmjerne strelice označavaju odnos u kojem se jedna aminokiselina može sintetizirati iz druge i obratno. Kvadratići omeđeni punom crtom označavaju današnji raspored kodona po aminokiselinama, a kvadratići omeđeni iscrtkanom crtom označavaju moguće kodone za pripadajuće aminokiseline u vrijeme rane evolucije koda. Preuzeto iz: Wong, 1975.

Mutacije u kodonima događale su se samo u jednoj bazi, što objašnjava današnji raspored genetskog koda u kojem bliski kodoni kodiraju aminokiseline sličnih svojstava (Slike 2 i 4) te otpornost koda na tranzicijske mutacije. Ipak, određene aminokiseline u odnosu prekursor-produkt kako je prikazano na Slici 7 kodirane su bliskim kodonima (s razlikom u jednoj bazi), a nemaju slična kemijska svojstva. To su primjerice parovi aminokiselina Ser-Trp, Thr-Ile, Gln-His, Ile-Met, Asn-Lys i Cys-Trp. Wong (1975) objašnjava te pojave kao očuvane ostatke premještanja kodona iz vremena rane evolucije koda. Također smatra da očuvanost takvih pojava u genetskom kodu odražava otpornost koda na promjene kako je predviđeno hipotezom zamrznute slučajnosti. Osim mutacijama postojećih kodona, nove aminokiseline svoje su kodone mogle steći i na drugi način. Na Slici 7 vidljivo je da su Glu i Asp u razdoblju rane evolucije koda potencijalno bili kodirani kodonima koji danas služe za kodiranje Gln odnosno Asn. Wong (1975) objašnjava tu pojavu premještanjem kodona (eng. *codon reassignment*). Naime, kako Glu služi kao prekursor u biosintetskom putu za Gln, kao i Asp za Asn, kodoni koji su prvotno kodirali Glu odnosno Asp pojavom novih aminokiselina bivaju "premješteni" i služe kao kodoni za Gln odnosno Asn.

Kako su kodoni ranih aminokiselina mogli "prijeći" na novonastale aminokiseline? Wong (1975) nudi nekoliko mogućih mehanizama, a svi uključuju adaptorske molekule – proto-tRNA. Naime, moguće je da su produkti novih biosintetskih puteva bili dovoljno strukturno slični prekursorima da su se mogli vezati za adaptore prekursora, ili se pretvorba prekursora u produkt dogodila dok je prekursor bio vezan za svoj adaptor. Također, moguće je da se intermedijer novonastalog biosintetskog puta povezao s adaptorom prekursora te se nakon toga formirao produkt – nova aminokiselina.

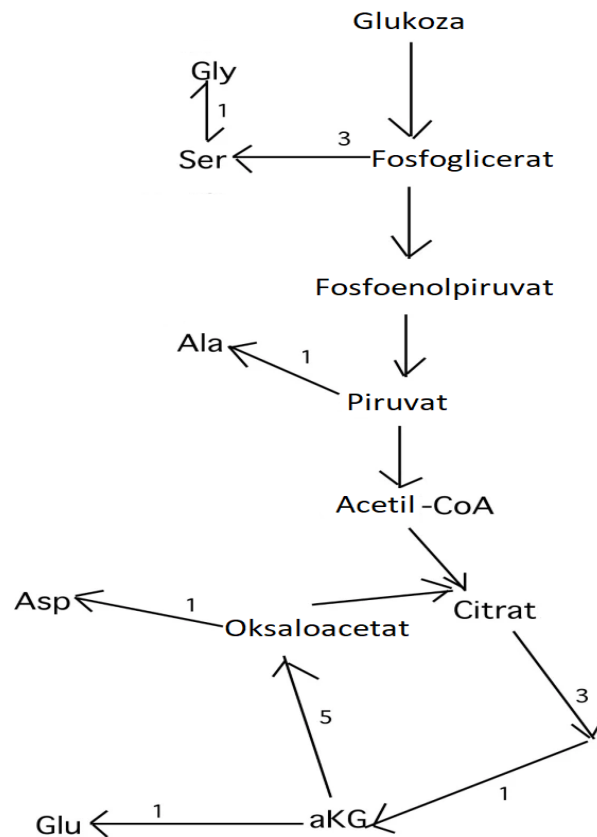
Koevolucijska hipoteza predlaže da su aminokiseline Glu, Asp, Ala, Ser i Gly služile kao centar iz kojeg je evoluiralo 11 drugih aminokiselina (Slika 7). Taj je prijedlog podržan rezultatima eksperimenata koji oponašaju uvjete prebiotske atmosfere (Miller, 1953; 1957), gdje su navedene aminokiseline bile najzastupljenije, te istraživanjima kronološkog poretka uklapanja aminokiselina u genetski kod (Trifonov, 2004). Koevolucijska hipoteza također nudi objašnjenje brojnosti kodona za pojedine aminokiseline. Naime, Wong (1975) smatra da broj kodona koji kodira za pojedinu aminokiselinu ovisi o tome koliko je rano ta aminokiselina uklopljena u genetski kod te koliko je aminokiselina iz nje nastalo evolucijom novih biosintetskih puteva. Primjerice, Met i Trp su svaki kodirani samo jednim kodonom, stoga su vjerojatno najkasnije uklopljeni u genetski kod, što potvrđuje istraživanje Trifonova (2004). S druge strane, Glu i Asp su među prvima uklopljeni u kod (Miller, 1957; Trifonov, 2004; Higgs

i Pudritz, 2009), a kodirani su svaki dvama kodonima. Prema Slici 7 vidljivo je da su iz njih evoluirale brojne aminokiseline, stoga su vjerojatno procesom premještanja kodona "izgubile" neke od svojih izvornih kodona i do danas "zadržale" samo dva. Aminokiseline Leu i Arg, pak, kodirane su sa 6 kodona, a uklopljene su u kod nakon Glu i Asp (Trifonov, 2004). Prema koevolucijskoj hipotezi, veći broj kodona za Leu i Arg u usporedbi s ranijim aminokiselinama rezultat je toga što iz njih nisu evoluirale druge aminokiseline, odnosno nisu služile kao prekursori novih biosintetskih puteva. Stoga su one "evolucijski inertne" (Wong, 1975).

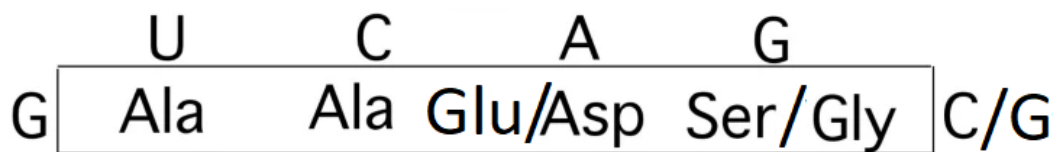
Unatoč objašnjenjima koje ova hipoteza nudi, moguće je da počiva na krivim pretpostavkama. Naime, putevi biosinteze aminokiselina koje predlaže Wong (1957) su prebiotski, dakle nastali su prije evolucije enzima koji sudjeluju u današnjim biosintetskim putevima aminokiselina. Stoga ostaje upitno jesu li biosintetski putevi evoluirali na način prikazan na Slici 7 (Koonin i Novozhilov, 2009). Primjerice, Johnson i Wang (2010) su pokazali da je Met "preuzeo" AUG kodon koji je prije kodirao za Ile te da je Phe "preuzeo" UUU i UUC kodone od Leu. Takvi rezultati odudaraju od predviđenih koevolucijskih odnosa prikazanih na Slici 7 te otvaraju mogućnost drukčijih puteva biosinteze aminokiselina u vrijeme rane evolucije koda. Osim toga, koevolucijska hipoteza kako ju originalno predlaže Wong (1975) ne objašnjava povezanost najranijih aminokiselina s njima pripadajućim kodonima, već to objašnjenje prepušta zamrznutoj slučajnosti ili stereokemijskim interakcijama.

Di Giulio (2008) proširuje koevolucijsku hipotezu i predlaže da su aminokiseline Glu, Asp, Ala, Ser i Gly bile dio prvog genetskog koda jer mogu biti sintetizirane iz intermedijera glikolize i Krebsovog ciklusa (Slika 8). Stoga su prve proteinogene aminokiseline mogle nastati iz prekursorskih molekula koje su imale ulogu u proizvodnji energije za stanicu. Povezanost između tih aminokiselina vidljiva je i u tome što njihovi kodoni počinju bazom G, uz iznimku Ser. Di Giulio (2008) pretpostavlja da su u ranoj fazi evolucije koda, kad su bile prisutne samo navedene aminokiseline, postojale proto-tRNA molekule povezane s tim aminokiselinama koje su prepoznavale sljedove na mRNA bogate G i C parovima baza. Postojanje takvih mRNA pretpostavljaju i druga istraživanja (Trifonov, 2004; Polyansky i Zagrovic, 2013). S obzirom da je raznolikost aminokiselina u toj fazi evolucije koda bila mala, kao i raznolikost sekvenci na mRNA, tada prisutne proto-tRNA su vjerojatno imale mali broj sličnih antikodonskih sekvenci te su prepoznavale slične kodone. Na temelju tih pretpostavki Di Giulio (2008) predlaže izgled genetskog koda za prvih 5 aminokiselina (Slika 9). Taj se kod naziva GNC odnosno GNG, ovisno o bazi na trećoj poziciji. Objašnjenje razlike između današnjih kodona za Ser te onih u GNC/GNG kodu Di Giulio (2008) nudi u činjenici da Ser može biti sintetiziran

iz Gly i obratno. Stoga je moguće da je u fazi GNC/GNG koda bio kodiran istim ili sličnim kodonima kao Gly, no daljnjom je evolucijom biosintetskih puteva te mutacijama i premještanjem kodona "stekao" nove kodone, dok su prvotni kodoni ostali vezani uz Gly.



Slika 8. Povezanost prvih aminokiselina uklopljenih u genetski kod s intermedijerima glikolize i Krebsovog ciklusa. Brojevi na strelicama označavaju broj koraka u biosintetskom putu potreban da od prekursora nastane produkt. Skraćenica aKG označava alfa-ketoglutarat. Preuzeto i prilagođeno prema: Di Giulio, 2008.



Slika 9. Pretpostavljeni izgled GNC/GNG genetskog koda za prvih 5 aminokiselina. Kao i kod standardnog genetskog koda, prva baza slijeva predstavlja 5'-kraj kodona na mRNA, a treća baza 3'-kraj kodona. Preuzeto i prilagođeno prema: Di Giulio, 2008.

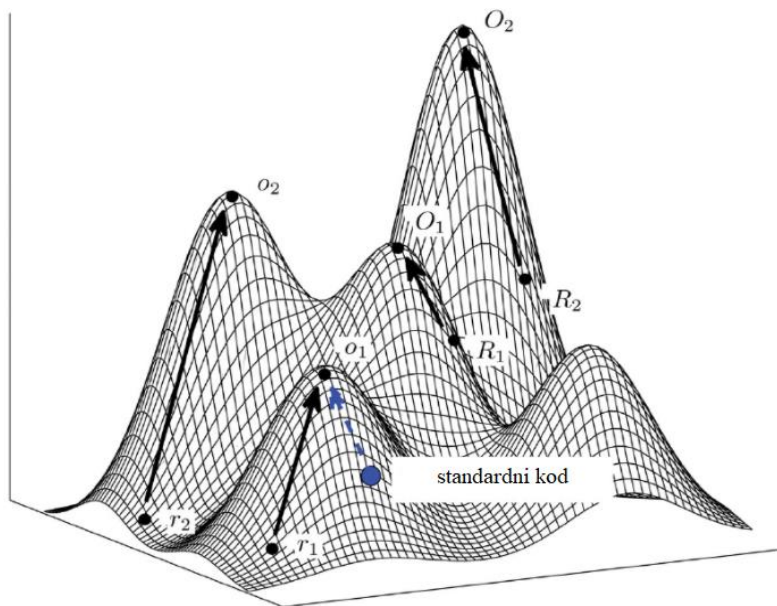
### 3.5. Hipoteza umanjivanja grešaka

Nastojeći objasniti otpornost genetskog koda na učinke mutacija, Woese (1965) je na temelju dotadašnjih istraživanja (Bryson i Vogel, 1965) predložio umanjivanje grešaka prilikom

translacije (eng. *error minimization*) kao glavni selekcijski pritisak tijekom evolucije koda. Današnji je izgled koda u kojem slični kodoni kodiraju aminokiseline sličnih svojstava, smatra Woese (1965), rezultat tog selekcijskog pritiska koji je usmjeravao početne verzije koda prema sve točnijoj translaciji. Ta je pretpostavka navela buduće istraživače da pokušaju izmjeriti otpornost genetskog koda na učinke mutacija i usporediti tu otpornost s drugim, nasumično stvorenim modelnim kodovima kako bi dokazali da je standardni kod najotporniji. Freeland i Hurst (1998) su u svojem istraživanju zaključili da je tek jedan u milijun nasumično stvorenih kodova efikasniji u suzbijanju učinaka mutacija od prirodnog genetskog koda. Njihov se rezultat poklapa s hipotezom umanjivanja grešaka te navodi na zaključak da je organizacija univerzalnog koda nastala prirodnom selekcijom. Međutim, s obzirom da je broj različitih potencijalnih genetskih kodova za 20 aminokiselina veći od  $10^{84}$  (Koonin i Novozhilov, 2009), očito je da je tijekom evolucije mogao nastati velik broj kodova s većom otpornošću od standardnog.

Istraživanje Novozhilova i sur. (2007) je pokazalo da genetski kod nije optimalan u suzbijanju mutacijskih učinaka te da postoje drugi, nasumično generirani kodovi koji su efikasniji u tom pogledu. Slika 10 prikazuje rezultat njihovog istraživanja u kojem su kodovi prikazani kao točke na reljefu, gdje više točke predstavljaju kodove koji su efikasniji u suzbijanju mutacijskih učinaka.





Slika 10. Reljefni prikaz evolucije nasumično generiranih kodova i standardnog genetskog koda iz regija manje efikasnosti suzbijanja mutacijskih učinaka (niže na reljefu) u regije veće efikasnosti (više na reljefu). Crne točke označene slovima O i R predstavljaju pozicije nasumično generiranih kodova, a plava točka poziciju standardnog genetskog koda. Strelice predstavljaju evolucijski put koda iz niže u višu točku. Iscrtkana plava strelica označava hipotetski evolucijski put standardnog koda prema maksimalnoj efikasnosti u suzbijanju učinaka mutacija. Preuzeto i prilagođeno prema: Novozhilov i sur., 2007.

Na Slici 10 vidljivo je da se standardni genetski kod nalazi na poziciji otprilike na pola puta do maksimalne efikasnosti u suzbijanju mutacija. Novozhilov i sur. (2007) smatraju da je "zaustavljanje" koda na toj poziciji rezultat međudjelovanja dvaju evolucijskih pritisaka – pozitivne selekcije prema većoj efikasnosti u suzbijanju učinaka mutacija te negativne selekcije prema promjenama u kodu, kako je predviđeno hipotezom zamrznute slučajnosti.

Crick (1968) predlaže da današnji izgled koda nije nastao pod selekcijskim pritiskom umanjivanja grešaka, već da je otpornost koda na učinke mutacija nastala kao nusproizvod evolucije. Buduća su istraživanja rezultirala razvojem modela koji podržavaju tu pretpostavku.

Massey (2006) predlaže "2-1-3" model evolucije koda prema kojem je u prvom stadiju evolucije koda samo druga pozicija kodona nosila informaciju za aminokiselinu. Zatim je, dodavanjem novih aminokiselina, kod proširen tako da su prva i druga pozicija bile bitne u kodiranju aminokiselina, a treća je pozicija kodona bila zadnja uključena u kodiranje aminokiselina (2-1-3 redosljed). Ovaj model objašnjava zašto mutacija druge pozicije kodona u današnjem genetskom kodu najčešće rezultira ugradnjom aminokiseline drukčijih svojstava od prvotne,

odnosno zašto je druga pozicija kodona najmanje otporna na učinke mutacija. Naime, kod u kojem samo druga pozicija nosi informaciju sadrži 4 različita kodona. Mutacija bilo koje baze na drugoj poziciji automatski rezultira drukčijim kodonom koji ne može kodirati aminokiselinu sličnih svojstava prvotnoj, već se na to mjesto u polipeptidni lanac ugrađuje aminokiselina drukčijih svojstava, što može imati negativan učinak na strukturu i funkciju proteina. Stoga takav kod nije efikasan u suzbijanju štetnih učinaka mutacija (Massey, 2006). Uvođenjem novih aminokiselina u genetski kod on se, prema ovom modelu, proširuje na 16 kodona određenih prvom i drugom pozicijom. Mutacija prve pozicije kodona u takvom kodu mogla je rezultirati ugradnjom aminokiseline sličnih svojstava prvotnoj. Massey (2006) stoga smatra da je rezultat prirodne selekcije bio kodiranje aminokiselina sličnih svojstava kodonima koji se razlikuju u prvoj poziciji, a aminokiselina drukčijih svojstava kodonima koji se razlikuju u prvoj i drugoj poziciji. Ovakav model može objasniti zašto je prva pozicija kodona otpornija na učinke mutacija od druge u današnjem standardnom kodu. Treća je baza kodona zadnja uključena u kodiranje aminokiselina jer antikodon na tRNA s tom bazom ostvaruje najslabije interakcije, stoga se ona naziva kolebljivom bazom (Crick, 1966a). Mutacija u toj bazi najčešće rezultira ugradnjom iste aminokiseline, stoga je ona najotpornija na učinke mutacija. Massey (2016) je istražio različite modele proširivanja genetskog koda, uključujući i "2-1-3" model, u kojima se novouvedenim aminokiselinama dodjeljuju kodoni koji su prethodno kodirali za aminokiseline sličnih svojstava. Različiti putevi proširivanja koda rezultirali su istim ishodom – kodovima koji imaju veću otpornost na učinke mutacija od standardnog. Massey (2016) stoga smatra da standardni genetski kod nije nastao kao rezultat selekcije prema umanjivanju grešaka jer bi u tom slučaju upravo on pokazivao najveću otpornost na učinke mutacija. Također, prema hipotezi umanjivanja grešaka pretpostavlja se da je selekcija prema optimalnoj otpornosti genetskog koda postignuta rearanžmanima kodona (Woese, 1965). No, visok stupanj rearanžmana kodona doveo bi do velikog broja promjena u strukturi proteina, što bi negativno utjecalo na funkciju proteina i preživljavanje stanica (Crick, 1968). Massey (2016) stoga zaključuje da je otpornost standardnog genetskog koda "primjer neutralne pojave povoljnog svojstva".

Higgs (2009) predlaže model 4 stupca (eng. *four column model*) kojim nastoji objasniti današnji izgled genetskog koda. Prema ovom modelu prvih 5 aminokiselina u kodu su Gly, Ala, Asp, Glu i Val te je baza na drugoj poziciji kodona najbitnija pri kodiranju. Stoga svi kodoni u istom stupcu kodiraju istu aminokiselinu (Slika 11) te se najraniji genetski kod na temelju ovog modela može podijeliti na 4 kodonska bloka odnosno 4 skupine kodona od kojih svaka skupina

kodira jednu aminokiselinu (potencijalno i dvije u slučaju kodonskog bloka za Asp/Glu). Higgs (2009) smatra da raspored aminokiselina današnjeg koda reflektira nekadašnji kod prikazan na Slici 11 jer aminokiseline raspoređene u isti stupac u današnjem kodu (Slika 2), odnosno one čiji kodoni imaju istu bazu na drugoj poziciji, najčešće imaju slična kemijska svojstva.

	U	C	A	G	
U	Val	Ala	Asp i/ili Glu	Gly	U C
					A G
C					U C
					A G
A					U C
					A G
G					U C
					A G

Slika 11. Predloženi izgled najranijeg genetskog koda, prema modelu 4 stupca. Kao i kod standardnog genetskog koda, prva baza slijeva predstavlja 5'-kraj kodona na mRNA, a treća baza 3'-kraj kodona. Preuzeto i prilagođeno prema: Higgs, 2009.

Kako su se nove aminokiseline uklapale u kod, ovaj model predviđa da su im "dodijeljivani" kodoni koji su prethodno kodirali aminokiseline sličnih svojstava. Time se naime postižu najmanji poremećaji u strukturi postojećih proteina kodiranih jednostavnijim verzijama koda, što ističe i Crick (1968). Primjerice, kodonski blok za Val (Slika 11) se mogao podijeliti na tri kodonska bloka za Val, Ile i Leu (Higgs, 2009). Redoslijed i način uklapanja novih aminokiselina u genetski kod predviđen ovim modelom poklapa se s predviđanjima koevolucijske hipoteze. Higgs (2009) zaključuje da je glavni selekcijski pritisak tijekom evolucije koda bio usmjeren prema povećanju raznolikosti proteina i njihovih funkcija, što je postignuto većim brojem aminokiselina u genetskom kodu. Rezultat takvog evolucijskog puta je kod otporan na greške prilikom translacije. Stoga model 4 stupca, kao i "2-1-3" model, ide u prilog ideji da je otpornost koda na učinke mutacija svojstvo koje se pojavilo kao nusproizvod drugih selekcijskih pritisaka.

Glavni problem u istraživanjima koja proučavaju otpornost genetskog koda na učinke mutacija je što ne postoji jasan konsenzus oko odabira najboljeg kriterija za procjenu otpornosti koda (Koonin i Novozhilov, 2009). Stoga svako istraživanje može koristiti drukčije kriterije, što otežava usporedbu dobivenih rezultata i donošenje zaključaka o otpornosti koda.

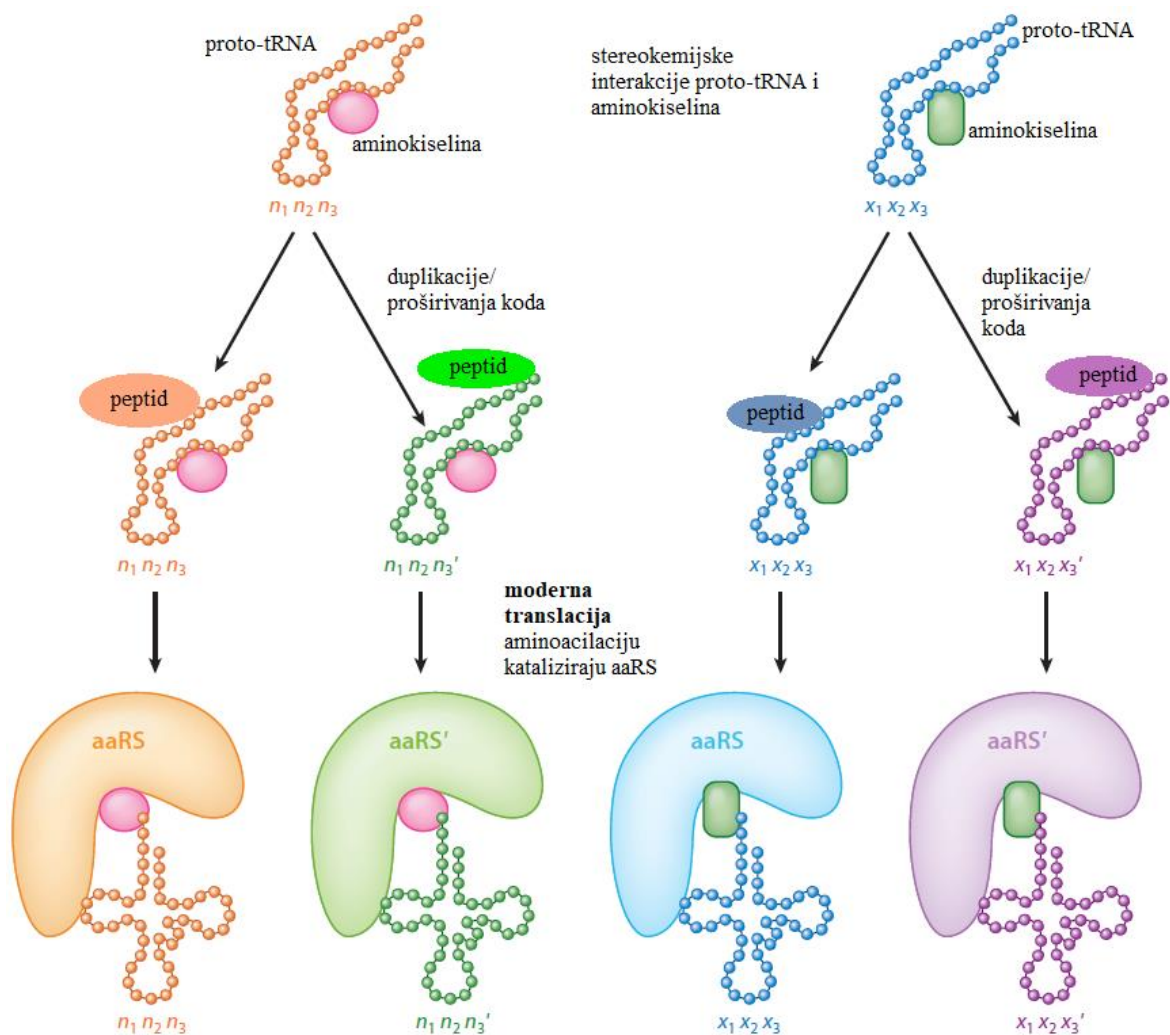
### **3.6. Evolucija translacijske mašinerije**

Pronalazak dokaza da je LUCA imao razvijen sustav za translaciju sličan onom u današnjim organizmima (Harris i sur., 2003) istaknuo je važnost istraživanja evolucije translacijske mašinerije za objašnjavanje porijekla i evolucijskog puta genetskog koda. Hohn i sur. (2006) su istraživali evoluciju tRNA molekula i aminoacil-tRNA-sintetaza (aaRS), najvažnijih molekula u procesu "čitanja" genetskog koda. Zaključili su da su dijelovi tRNA molekula ključni za interakcije s aminokiselinama i mRNA evoluirali prije sustava za aminoacilaciju kakav postoji danas. To implicira da su specifične strukture i funkcije aaRS evoluirale kasnije te nisu igrale ulogu u povezivanju aminokiselina s pripadajućim tRNA u ranoj fazi evolucije genetskog koda, prije pojave LUCA-e.

Međutim, to nameće problem nazvan paradoksom proteinske evolucije (Koonin i Novozhilov, 2017). Naime, za evoluciju kompleksnih proteinskih struktura ključna je visoka točnost translacije, a ona je ostvarena tek nakon uvođenja aaRS u translacijsku mašineriju. Kako je onda moguće da su specifične strukture aaRS evoluirale prije nego su same aaRS bile u mogućnosti osigurati dovoljno točnu translaciju? Objašnjenje je ponuđeno u okviru hipoteze RNA svijeta. Prema ovoj su hipotezi prije pojave sinteze proteina i bioloških procesa kataliziranih proteinskim enzimima, RNA molekule imale ulogu katalizatora vlastite replikacije (tzv. ribozimi), potičući vlastitu evoluciju (Robertson i Joyce, 2010). Znanstvenici su stoga zaključili da je i genetski kod nastao u periodu RNA svijeta te da su upravo RNA molekule povezivale aminokiseline s pripadajućim tRNA prije pojave aaRS (Koonin i Novozhilov, 2017). Eksperimenti kojima su se u laboratorijskim uvjetima uspjeli sintetizirati ribozimi s aminoacilirajućom aktivnošću (Lee i sur., 2000; Müller, 2015) potvrđuju takav zaključak.

Koonin (2017) predlaže model evolucije translacijskog sustava u skladu s hipotezom RNA svijeta. Prema tom modelu, prvi stadij u evoluciji koda temeljio se na stereokemijskim interakcijama između proto-tRNA i aminokiselina tako da su proto-tRNA imale ulogu vlastite aminoacilacije. Te interakcije nisu uključivale antikodone niti kodone, već su aminokiseline interagirale sa segmentima u terciarnoj strukturi proto-tRNA molekula. Time je osigurano

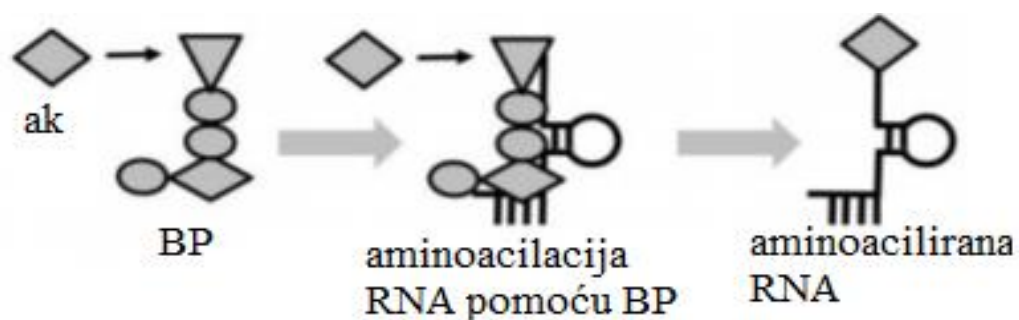
specifično prepoznavanje aminokiselina i pripadajućih tRNA. Stoga ovaj model nije u potpunosti u skladu sa stereokemijskom hipotezom, no podržava hipotezu zamrznute slučajnosti jer pretpostavlja da su navedenim interakcijama nasumično povezane aminokiseline i pripadajuće proto-tRNA te da je nastala povezanost "zamrznuta" tijekom evolucije. Iduća faza u evoluciji koda uključuje duplikacije tRNA molekula i uklapanje novih aminokiselina. Duplicirane tRNA bile su strukturno slične svojim prethodnicama, stoga vežu aminokiseline sličnih svojstava onima vezanima na prethodničke tRNA (Koonin, 2017). Posljedično, aminokiseline sličnih svojstava bivaju kodirane bliskim kodonima. Dakle, prema ovom je modelu otpornost genetskog koda na učinke mutacija nastala kao nusprodukt evolucije, kao i prema "2-1-3" modelu i modelu 4 stupca. U navedenoj fazi evolucije koda prethodno sintetizirani kratki peptidi služili su kao kofaktori u aminoacilaciji proto-tRNA. Daljnjim poboljšavanjem točnosti translacije i kompleksnosti peptidnih kofaktora nastupa posljednja faza evolucije koda u kojoj su aaRS evoluirale i preuzele ulogu katalize aminoacilacije (Slika 12). Koonin (2017) smatra da nakon što su aaRS preuzele ulogu specifičnog prepoznavanja aminokiselina, segmenti u tercijarnoj strukturi tRNA koji su nekoć imali tu ulogu nestaju te danas u strukturi tRNA ne postoje ostaci tih segmenata. Iako ovaj model nije eksperimentalno dokazan (Koonin, 2017), može služiti kao osnova za eksperimente kojima se nastoji rekonstruirati evolucijski put translacijskog sustava u skladu s hipotezom RNA svijeta.



Slika 12. Model evolucije translacijskog sustava koji predlaže Koonin (2017). Objašnjenje pojedinih koraka navedeno je u tekstu. Dijelovi proto-tRNA označeni s ' $n_1 n_2 n_3$ ' i ' $x_1 x_2 x_3$ ' predstavljaju antikodone. Preuzeto i prilagođeno prema: Koonin i Novozhilov, 2017.

Kunnev i Gospodinov (2018) predlažu drukčiji model evolucije translacijskog sustava koji ima određenih sličnosti s prethodno opisanim modelom u početnom stadiju. Naime, ovaj model pretpostavlja da su RNA molekule vršile vlastitu aminoacilaciju prije pojave aaRS te kao i model koji predlaže Koonin (2017) pretpostavlja da su aminoacilirajuće RNA u početku bile nasumično povezane s pripadajućim aminokiselinama, podržavajući hipotezu zamrznute slučajnosti. Aminoacilirane RNA su zatim mogle hibridizirati s drugima u blizini na način da se dijelovi sekundarnih struktura tih RNA povežani s aminokiselinama dovoljno približe da omoguće stvaranje peptidnih veza između aminokiselina. Peptidi koji na taj način nastaju bili bi sastavljeni od 2-5 aminokiselina te bi njihov sastav ovisio o dostupnosti aminokiselina u okolišu. Stoga se pretpostavlja da su bili sastavljeni od aminokiselina najzastupljenijih u

eksperimentima koji oponašaju uvjete prebiotske atmosfere (Miller, 1957). Idući stadij je selekcija onih peptida koji povećavaju stabilnost RNA molekula za koje su vezani. Naime, što su kompleksi RNA i peptida stabilniji, mogu dulje opstati u okolišu i veća je vjerojatnost da će nastaviti evoluirati. Malobrojne peptide s tim svojstvom Kunnev i Gospodinov (2018) nazivaju RNA-vezujućim peptidima (eng. *RNA-binding peptides* – RBP). Ipak, RBP-ovi nisu imali mogućnost čitanja informacija u sekvenci pripadajuće RNA, što naizgled zaustavlja daljnju evoluciju prema sve stabilnijim kompleksima RNA i peptida jer se nove aminokiseline ugrađuju u RBP-ove nasumično, bez selekcije za povećanu stabilnost. Ovaj model nudi rješenje pretpostavljajući da je selekcija išla u drugom smjeru – prema RBP-ovima koji su imali mogućnost interakcije s kemijski sličnim aminokiselinama te su ih mogli povezati s pripadajućim RNA. Takve RBP-ove Kunnev i Gospodinov (2018) nazivaju premošćujućim peptidima (eng. *bridge peptides* – BP). Prema tome, BP-ovi bi imali mogućnost aminoacilacije RNA (Slika 13), no taj proces ne bi bio nasumičan, već bi se na RNA vezale aminokiseline sličnih svojstava onima koje omogućavaju stabilne interakcije BP-a i RNA. Stoga bi nosintetizirani RBP-ovi sadržavali sve veći udio aminokiselina koje osiguravaju stabilnije povezivanje s pripadajućom RNA. Na taj je način ostvarena pozitivna povratna sprega kojom nastaju sve stabilniji kompleksi peptida i RNA. Sastav aminokiselina u novonastalim peptidima ovisio bi o sekvenci na pripadajućoj RNA molekuli, jer je stabilnost kompleksa osigurana specifičnim interakcijama između dušičnih baza i aminokiselina. Stoga je razvojem ovakvog sustava aminoacilacije nastao "protokod", odnosno po prvi je put sekvenca na RNA molekuli služila za određivanje aminokiselinskog slijeda peptida.



Slika 13. Predloženi put aminoacilacije RNA molekula u ranoj fazi evolucije translacijskog sustava. BP – premošćujući peptid, ak – vrsta aminokiseline koja osigurava stabilno povezivanje peptida i RNA. Aminokiselina u strukturi BP-a prikazana trokutom predstavlja aminokiselinu sličnih svojstava onoj označenoj s "ak" na slici. Preuzeto i prilagođeno prema: Kunnev i Gospodinov, 2018.

Daljnjom koevolucijom komponenata ovog sustava te povećanjem njihove kompleksnosti postepeno nastaje moderni sustav za translaciju zajednički svim organizmima na Zemlji. Specijalizacijom aminoaciliranih RNA molekula (Slika 13) nastaju tRNA koje su povezivane s pripadajućim aminokiselinama na način opisan Kooninovima (2017) modelom. BP-ovi su postupno evoluirali u aminoacil-tRNA-sintetaze koje su mogle razlikovati pojedine tRNA molekule i povezivati ih s pripadajućim aminokiselinama. Kompleksi RNA i peptida evoluirali su u ribozime, uključujući i ribosom, čija velika podjedinica sadrži peptidil-transferazni centar – mjesto formiranja peptidnih veza nastajućeg polipeptidnog lanca. Uključivanjem kompleksnih proteina u proces translacije onemogućene su promjene u uspostavljenom genetskom kodu jer bi imale nepovoljno djelovanje na funkciju tih proteina. Stoga je ovaj model u skladu s hipotezom zamrznute slučajnosti. Specifičnost ovog modela u odnosu na onaj koji predlaže Koonin (2017) je odstupanje od uvriježene hipoteze RNA svijeta. Umjesto toga pretpostavlja da je translacijski sustav evoluirao putem interakcija između RNA i peptida, odnosno u okviru RNA-peptidnog svijeta. Ovaj model, kao ni onaj koji predlaže Koonin (2017), nije eksperimentalno dokazan. Unatoč tome je važan jer nudi osnovu za eksperimente kojima bi se pokušala dokazati mogućnost evolucije translacijske mašinerije van okvira RNA svijeta.

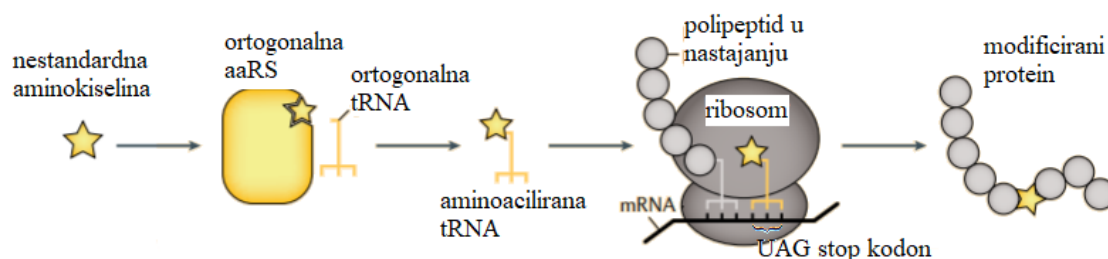


#### 4. POGLED U BUDUĆNOST – PROŠIRIVANJE GENETSKOG KODA

Standardni genetski kod je "zamrznut" u današnjem stanju (Crick, 1968) i podložan je organskim promjenama i proširenjima koji su rezultat specifičnih selekcijskih pritisaka na pojedine organizme. Ipak, u laboratorijskim uvjetima uspješno su provedena proširenja genetskog koda različitim metodama, od kojih će neke biti opisane u ovom poglavlju.

Genetski kod je proširen uklapanjem nestandardnih aminokiselina, sintetiziranih *in vitro*, u postojeći translacijski sustav stanica (de la Torre i Chin, 2020). Na taj način nestandardne aminokiseline postaju dio strukture postojećih proteina koji dobivaju nova svojstva. Primjerice, uklapanjem aminokiselina koje se pod utjecajem svjetlosnog zračenja povezuju s fluorescirajućim molekulama moguće je pratiti interakcije proteina *in vivo*. Ova i brojne druge primjene nestandardnih aminokiselina opisane su u radu Davisa i China (2012), no o njima neće biti riječi jer prelaze opus ovog rada.

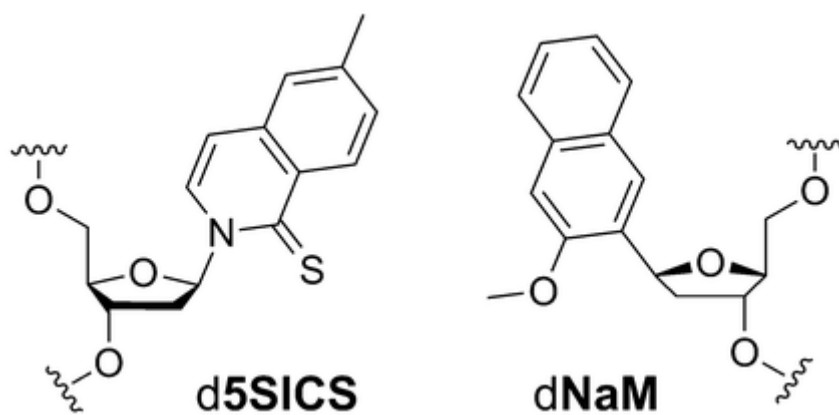
Najčešći način uklapanja nestandardnih aminokiselina ostvaruje se ugradnjom nove aminokiseline na mjesto u polipeptidnom lancu koje odgovara stop kodonu, najčešće UAG, u sekvenci mRNA (de la Torre i Chin, 2020). To se postiže pomoću ortogonalnih parova aaRS/tRNA (Slika 14). Naime, tRNA koja prepoznaje stop kodon i na njegovo mjesto veže neku aminokiselinu, tzv. supresorska tRNA, mora biti povezana s nestandardnom aminokiselinom od interesa kako bi se ona ugradila u polipeptidni lanac. Aminoacilaciju tRNA katalizira aaRS, stoga ona mora biti modificirana tako da specifično prepoznaje nestandardnu aminokiselinu i ne prepoznaje ni jednu drugu. Osim toga, mora specifično prepoznavati samo supresorsku tRNA od interesa, kako bi mogla uspješno povezati tRNA i nestandardnu aminokiselinu. Stoga tRNA treba biti modificirana tako da se omogući specifično vezanje samo za aaRS od interesa te istovremeno onemogući vezanje za bilo koju drugu aaRS, čime nastaje ortogonalni par. Ortogonalni parovi aaRS/tRNA najčešće su dobiveni modifikacijom parova aaRS i pripadajuće supresorske tRNA pronađenih u bakterijama ili arhejama.



Slika 14. Ugradanja nestandardnih aminokiselina u proteine na mjesto UAG stop kodona pomoću ortogonalnog para aaRS/tRNA. Preuzeto i prilagođeno prema: de la Torre i Chin, 2020.

Osim supresijom stop kodona, nestandardne aminokiseline mogu zamijeniti standardne u polipeptidnom lancu. Naime, većina aminokiselina u standardnom genetskom kodu kodirana je s više od jednog kodona (Slika 2). Različiti kodoni za istu aminokiselinu nazivaju se sinonimima. Znanstvenici su razvili metode kojima deaktiviraju dio sinonimnih kodona za određenu aminokiselinu na način da ih uklone sa ciljnih mRNA. Zatim se iz stanice uklanjaju tRNA koje se povezuju s tim kodonima. Tada je standardna aminokiselina kodirana samo dijelom svojih sinonimnih kodona. Ovaj se proces naziva kompresija sinonimnih kodona (eng. *synonymous codon compression*). Uklonjeni kodoni se zatim ponovo ugrađuju na željena mjesta u sekvenci ciljne mRNA, no ovaj put se za njih vežu modificirane tRNA koje nose nestandardnu aminokiselinu. Ugradnja nestandardne aminokiseline postignuta je pomoću dizajniranih ortogonalnih parova aaRS/tRNA (de la Torre i Chin, 2020).

Gore navedene metode koriste prirodno prisutne dušične baze na mRNA za ugradnju nestandardnih aminokiselina, što ograničava broj kodona koji potencijalno kodiraju nestandardne aminokiseline. Uvođenjem sintetskih baza u sekvencu DNA ili RNA može se znatno povećati taj broj. Primjerice, uvođenjem jednog para sintetskih baza standardni kod se proširuje s originalna 64 kodona na ukupno 216 (Blažej i sur., 2020). Najpoznatiji primjer sintetskih baza korištenih za proširivanje koda *in vivo* je par d5SICS-dNaM (Slika 15). Ove sintetske baze međusobno su povezane hidrofobnim interakcijama te su stabilizirane unutar strukture DNA interakcijama slaganja s ostalim bazama (Malyshev i sur., 2012).



Slika 15. Kemijska struktura sintetskog para baza d5SICS-dNaM. Preuzeto iz: Malyshev i sur., 2012.

Osim navedenih sintetskih baza, razvijene su brojne druge čija se povezanost temelji na hidrofobnim interakcijama (Dhami i sur., 2014) i vodikovim vezama koje se razlikuju od onih u Watson-Crickovom sparivanju (Ohtsuki i sur., 2001). Prilikom proširivanja genetskog koda

bitno je paziti na umanjivanje potencijalnih štetnih učinaka mutacija koje se mogu pojaviti dodavanjem novih baza. Stoga Błażej i sur. (2020) predlažu da se kodoni s novim bazama koji kodiraju nove aminokiseline dodaju u kod postepeno, zamjenom jedne po jedne prirodne baze sa sintetskom u postojećim tripletima. Time bi se omogućila otpornost novog koda na mutacije. Svoj prijedlog temelje na pretpostavci da su u ranoj evoluciji koda novouvedene aminokiseline "stjecale" kodone slične već postojećima, što je rezultiralo standardnim kodom otpornim na učinke mutacija. Ovime je istaknuta važnost proučavanja evolucijske prošlosti genetskog koda. Naime, spoznaje o principima rane evolucije koda mogu pomoći znanstvenicima u otkrivanju efikasnijih metoda proširivanja koda temeljenih na tim principima.

Metode proširivanja standardnog genetskog koda opisane u ovom pogavlju predstavljaju *in vitro* fazu evolucije koda. Razvojem novih metoda i sve efikasnijih načina za uklapanje nestandardnih aminokiselina i sintetskih baza u postojeći kod (Shandell i sur., 2021), ova će se faza zasigurno nastaviti u budućnosti. Nedvojbene su i brojne korisne primjene ovih metoda u medicini, genetičkim istraživanjima, istraživanjima strukture i funkcije proteina i mnogim drugim područjima biologije (Davis i Chin, 2012; Shandell i sur., 2021).

## 5. ZAKLJUČAK

Evolucijski put kojim je standardni genetski kod "stekao" svoja svojstva i današnji izgled ostaje neotkriven unatoč više od 60 godina istraživanja. Ipak, 4 najvažnije hipoteze razvijene tim istraživanjima ponudile su važne uvide u brojne mehanizme i procese važne za evoluciju koda. Hipoteza zamrznute slučajnosti ponudila je osnovu za kasnije modele evolucije koda. Stereokemijska hipoteza potaknula je značajan napredak u istraživanju interakcija između aminokiselina i RNA. Koevolucijska hipoteza je integrirala biosintetske puteve aminokiselina u evoluciju koda te pojasnila razliku između ranih i kasnih aminokiselina. Hipoteza umanjivanja grešaka dovela je do napretka u spoznajama o otpornosti koda na učinke mutacija te je ponudila osnovu za različite modele evolucije koda. Niti jedna od 4 navedene hipoteze ne može samostalno objasniti evoluciju svojstava standardnog genetskog koda. Međutim, zajednički nude mogućnost evolucijskog puta u kojem su međudjelovanja različitih selekcijskih pritisaka oblikovala današnji izgled koda (Koonin i Novozhilov, 2009; Yarus i sur., 2009). Eksperimenti koji povezuju ideje četiriju hipoteza s modelima evolucije translacijske mašinerije mogli bi u budućnosti polučiti značajan uspjeh u dokazivanju konkretnih evolucijskih puteva i selekcijskih pritisaka važnih za uspostavljanje današnjeg izgleda standardnog genetskog koda (Koonin i Novozhilov, 2017). Naposljetku, ulaskom genetskog koda u *in vitro* evolucijsku fazu povećava se količina informacija pohranjenih u kodu pomoću nestandardnih aminokiselina i sintetskih baza. Proširivanje genetskog koda stoga obećava brojne korisne primjene u različitim granama biologije. Otkrivanje principa evolucije standardnog koda moglo bi pridonijeti razvitku novih i efikasnijih metoda proširivanja koda. Time je istaknuta važnost razumijevanja prošlosti za ostvarivanje napretka u budućnosti.

## 6. LITERATURA

- Aggarwal N., Bandhu A. V., Sengupta S. (2016). Finite population analysis of the effect of horizontal gene transfer on the origin of an universal and optimal genetic code. *Physical Biology* 13: 036007.
- Alberts B. (2015). *Molecular biology of the cell*. New York, Garland Science.
- Ambrogelly A., Palioura S., Söll D. (2007). Natural expansion of the genetic code. *Nature Chemical Biology* 3: 29–35.
- Błażej P., Wnetrzak M., Mackiewicz D., Mackiewicz P. (2020). Basic principles of the genetic code extension. *Royal Society Open Science* 7: 191384.
- Bryson V., Vogel H. J. (1965). Evolving genes and proteins. *Science* 147: 68–71.
- Crick F. H. C. (1958). On protein synthesis. *Symposia of the Society for Experimental Biology* 12: 138–163.
- Crick F. H. C. (1966a). Codon—anticodon pairing: the wobble hypothesis. *Journal of Molecular Biology* 19: 548–555.
- Crick F. H. C. (1966b). The genetic code – yesterday, today, and tomorrow. *Cold Spring Harbor Symposia on Quantitative Biology* 31: 3–9.
- Crick F. H. C. (1968). The origin of the genetic code. *Journal of Molecular Biology* 38: 367–379.
- Davis L., Chin J. W. (2012). Designer proteins: applications of genetic code expansion in cell biology. *Nature Reviews Molecular Cell Biology* 13: 168–182.
- de la Torre D., Chin J. W. (2020). Reprogramming the genetic code. *Nature Reviews Genetics* 22: 169–184.
- Dhami K., Malyshev D. A., Ordoukhanian P., Kubelka T., Hocek M., Romesberg F. E. (2014). Systematic exploration of a class of hydrophobic unnatural base pairs yields multiple new candidates for the expansion of the genetic alphabet. *Nucleic acids research* 42: 10235–10244.
- Di Giulio M. (2008). An extension of the coevolution theory of the origin of the genetic code. *Biology Direct* 3: 37.

- Freeland S. J., Hurst L. D. (1998). The genetic code is one in a million. *Journal of Molecular Evolution* 47: 238–248.
- Gamow G. (1954). Possible relation between deoxyribonucleic acid and protein structures. *Nature* 173: 318–318.
- Harris J. K., Kelley S. T., Spiegelman G. B., Pace N. R. (2003). The genetic core of the universal ancestor. *Genome Research* 13: 407–412.
- Higgs P. G. (2009). A four-column theory for the origin of the genetic code: tracing the evolutionary pathways that gave rise to an optimized code. *Biology Direct* 4: 16.
- Higgs P. G., Pudritz R. E. (2009). A thermodynamic basis for prebiotic amino acid synthesis and the nature of the first genetic code. *Astrobiology* 9: 483–490.
- Hlevnjak M., Polyansky A. A., Zagrovic B. (2012). Sequence signatures of direct complementarity between mRNAs and cognate proteins on multiple levels. *Nucleic Acids Research* 40: 8874–8882.
- Hohn M. J., Park H. S., O'Donoghue P., Schnitzbauer M., Söll D. (2006). Emergence of the universal genetic code imprinted in an RNA record. *Proceedings of the National Academy of Sciences* 103: 18095–18100.
- Johnson D. B. F., Wang L. (2010). Imprints of the genetic code in the ribosome. *Proceedings of the National Academy of Sciences* 107: 8298–8303.
- Keeling P., Palmer J. (2008). Horizontal gene transfer in eukaryotic evolution. *Nature Reviews Genetics* 9: 605–618.
- Kollmar M., Mühlhausen S. (2017). Nuclear codon reassignments in the genomics era and mechanisms behind their evolution. *BioEssays* 39: 1600221.
- Koonin E. V. (2017). Frozen accident pushing 50: stereochemistry, expansion, and chance in the evolution of the genetic code. *Life* 7: 22.
- Koonin E. V., Novozhilov A. S. (2009). Origin and evolution of the genetic code: the universal enigma. *IUBMB Life* 61: 99–111.
- Koonin E. V., Novozhilov A. S. (2017). Origin and evolution of the universal genetic code. *Annual Review of Genetics* 51: 45–62.

- Kunnev D., Gospodinov A. (2018). Possible emergence of sequence specific RNA aminoacylation via peptide intermediary to initiate Darwinian evolution and code through origin of life. *Life* 8: 44.
- Lee N., Bessho Y., Wei K., Szostak J. W., Suga H. (2000). Ribozyme-catalyzed tRNA aminoacylation. *Nature Structural Biology* 7: 28–33.
- Malyshev D. A., Dhimi K., Quach H. T., Lavergne T., Ordoukhanian P., Torkamani A., Romesberg F. E. (2012). Efficient and sequence-independent replication of DNA containing a third base pair establishes a functional six-letter genetic alphabet. *Proceedings of the National Academy of Sciences* 109: 12005–12010.
- Massey S. E. (2006). A sequential “2-1-3” model of genetic code evolution that explains codon constraints. *Journal of Molecular Evolution* 62: 809–810.
- Massey S. E. (2016). The neutral emergence of error minimized genetic codes superior to the standard genetic code. *Journal of Theoretical Biology* 408: 237–242.
- Miller S. L. (1953). A production of amino acids under possible primitive Earth conditions. *Science* 117: 528–529.
- Miller S. L. (1957). The mechanism of synthesis of amino acids by electric discharges. *Biochimica et Biophysica Acta* 23: 480–489.
- Mühlhausen S., Findeisen P., Plessmann U., Urlaub H., Kollmar M. (2016). A novel nuclear genetic code alteration in yeasts and the evolution of codon reassignment in eukaryotes. *Genome Research* 26: 945–955.
- Müller S. (2015). Engineering of ribozymes with useful activities in the ancient RNA world. *Annals of the New York Academy of Sciences* 1341: 54–60.
- Nahvi A., Sudarsan N., Ebert M. S., Zou X., Brown K. L., Breaker R. R. (2002). Genetic control by a metabolite binding mRNA. *Chemistry & Biology* 9: 1043–1049.
- Nelson D. L., Cox M. M. (2017). *Lehninger principles of biochemistry*. New York, W. H. Freeman and Company.
- Nirenberg M., Jones O. W., Leder P., Clark B., Sly W., Pestka S. (1963). On the coding of genetic information. *Cold Spring Harbor Symposia on Quantitative Biology* 28: 549–557.

- Novozhilov A. S., Wolf Y. I., Koonin E. V. (2007). Evolution of the genetic code: partial optimization of a random code for robustness to translation error in a rugged fitness landscape. *Biology Direct* 2: 24.
- Ohtsuki T., Kimoto M., Ishikawa M., Mitsui T., Hirao I., Yokoyama S. (2001). Unnatural base pairs for specific transcription. *Proceedings of the National Academy of Sciences* 98: 4922–4925.
- Pelc S. R. (1965). Correlation between coding-triplets and amino-acids. *Nature* 207: 597–599.
- Polyansky A. A., Zagrovic B. (2013). Evidence of direct complementary interactions between messenger RNAs and their cognate proteins. *Nucleic Acids Research* 41: 8434–8443.
- Robertson M. P., Joyce G. F. (2010). The origins of the RNA world. *Cold Spring Harbor Perspectives in Biology* 4: a003608.
- Sengupta S., Higgs P. G. (2015). Pathways of genetic code evolution in ancient and modern organisms. *Journal of Molecular Evolution* 80: 229–243.
- Shandell M. A., Tan Z., Cornish V. W. (2021). Genetic code expansion: a brief history and perspective. *Biochemistry*[online]. Dostupno na: <https://doi.org/10.1021/acs.biochem.1c00286> (pristupljeno 28. 8. 2021.)
- Trifonov E. N. (2004). The triplet code from first principles. *Journal of Biomolecular Structure and Dynamics* 22: 1–11.
- Vetsigian K., Woese C., Goldenfeld N. (2006). Collective evolution and the genetic code. *Proceedings of the National Academy of Sciences* 103: 10696–10701.
- Vnučec D., Kutnar A., Goršek A. (2016). Soy-based adhesives for wood-bonding – a review. *Journal of Adhesion Science and Technology* 31: 910–931.
- Wang A., Nairn N., Marelli M. (2012). Protein engineering with non-natural amino acids. U: Pravin Kaumaya (ur.) *Protein engineering*. Rijeka, InTech, str. 253–290.
- Winkler W. C., Cohen-Chalamish S., Breaker R. R. (2002). An mRNA structure that controls gene expression by binding FMN. *Proceedings of the National Academy of Sciences* 99: 15908–15913.
- Woese C. R. (1965). On the evolution of the genetic code. *Proceedings of the National Academy of Sciences* 54: 1546–1552.



Woese C. R. (1968). The fundamental nature of the genetic code: prebiotic interactions between polynucleotides and polyamino acids or their derivatives. *Proceedings of the National Academy of Sciences* 59: 110–117.

Woese C. R., Dugre D. H., Saxinger W. C., Dugre S. A. (1966). The molecular basis for the genetic code. *Proceedings of the National Academy of Sciences* 55: 966–974.

Woese C. R., Hinegardner R. T., Engelberg J. (1964). Universality in the genetic code. *Science* 144: 1030–1031.

Wong J. T. F. (1975). A co-evolution theory of the genetic code. *Proceedings of the National Academy of Sciences* 72: 1909–1912.

Yarus M. (2021). Evolution of the standard genetic code. *Journal of Molecular Evolution* 89: 19–44.

Yarus M., Caporaso J. G., Knight R. (2005). Origins of the genetic code: the escaped triplet theory. *Annual Review of Biochemistry* 74: 179–198.

Yarus M., Widmann J. J., Knight R. (2009). RNA–amino acid binding: a stereochemical era for the genetic code. *Journal of Molecular Evolution* 69: 406–429.

Yuan J., O'Donoghue P., Ambrogelly A., Gundllapalli S., Sherrer R. L., Palioura S., Simonović M., Söll D. (2010). Distinct genetic code expansion strategies for selenocysteine and pyrrolysine are reflected in different aminoacyl-tRNA formation systems. *FEBS Letters* 584: 342–349.

Zaia D. A. M., Zaia C. T. B. V., De Santana H. (2008). Which amino acids should be used in prebiotic chemistry studies? *Origins of Life and Evolution of Biospheres* 38: 469–488.

Zhang Y., Baranov P. V., Atkins J. F., Gladyshev V. N. (2005). Pyrrolysine and selenocysteine use dissimilar decoding strategies. *Journal of Biological Chemistry* 280: 20740–20751.

[www.pediaa.com/difference-between-sense-and-antisense-strand/](http://www.pediaa.com/difference-between-sense-and-antisense-strand/) (pristupljeno 30. 7. 2021.)

[www.ddbj.nig.ac.jp/ddbj/code-e.html](http://www.ddbj.nig.ac.jp/ddbj/code-e.html) (pristupljeno 30. 7. 2021.)

[www.pubchem.ncbi.nlm.nih.gov/compound/Pyrrolysine#section=2D-Structure](http://www.pubchem.ncbi.nlm.nih.gov/compound/Pyrrolysine#section=2D-Structure) (pristupljeno 1. 8. 2021.)

[www.pubchem.ncbi.nlm.nih.gov](http://www.pubchem.ncbi.nlm.nih.gov) (pristupljeno 17. 8. 2021.)

## 7. SAŽETAK

Standardni genetski kod – informacija u obliku 64 tripleta dušičnih baza (kodona) na mRNA koji kodiraju 20 proteinogenih aminokiselina, gotovo je univerzalan, uz ograničene iznimke, u svim organizmima na Zemlji i otporan na štetne učinke mutacija u sekvenci kodona. Pokušavajući objasniti evoluciju navedenih svojstava razvijene su 4 hipoteze. Hipoteza zamrznute slučajnosti tvrdi da su aminokiseline nasumično povezane s pripadajućim kodonima te da je univerzalnost koda rezultat štetnih učinaka potencijalnih promjena. Stereokemijska hipoteza objašnjava povezanost aminokiselina s pripadajućim kodonima ili antikodonima putem direktnih stereokemijskih interakcija. Koevolucijska hipoteza tvrdi da su rane aminokiseline uklopljene u kod iz abiotskih prekursora, a kasne usporedno s evolucijom njihovih biosintetskih puteva. Hipoteza umanjivanja grešaka predlaže da je glavni selekcijski pritisak evolucije koda bio prema većoj otpornosti na učinke mutacija, no novija istraživanja potvrđuju da je otpornost koda najvjerojatnije nusprodukt drugih selekcijskih pritisaka. Eksperimenti koji povezuju ideje četiriju hipoteza s evolucijom translacijske mašinerije nude najbolje mogućnosti rekonstrukcije evolucijskog puta genetskog koda, s obzirom da je svaka od hipoteza pojedinačno u tome neuspješna. Genetski kod se danas nalazi u novoj, *in vitro* evolucijskoj fazi zahvaljujući razvoju nestandardnih aminokiselina i sintetskih baza. Razumijevanje prošlosti genetskog koda, odnosno principa njegove evolucije, važno je za ostvarivanje napretka u novoj evolucijskoj fazi.

## 8. SUMMARY

The standard genetic code – the information stored in 64 triplets of nitrogenous bases (codons) in mRNA, which encode 20 proteinogenic amino acids, is nearly universal, with limited exceptions, in all life forms on Earth and it exhibits robustness to harmful effects of mutations in codon sequence. Four hypotheses were put forward in an effort to explain those properties. The frozen accident hypothesis posits that amino acids were coincidentally connected with cognate codons and that the code's universality is the result of harmful effects of potential alterations. The stereochemical hypothesis explains the connection of amino acids and cognate codons or anticodons by direct stereochemical interactions. The coevolution hypothesis claims that early amino acids were incorporated into the code from abiotic precursors, and the incorporation of late amino acids was concurrent with the evolution of their respective biosynthetic pathways. Error minimization hypothesis proposes that the main selective pressure of code evolution was to obtain greater robustness to effects of mutations, but more recent experiments confirmed that the code's robustness is most likely a by-product of other selective pressures. Experiments integrating the ideas of the 4 hypotheses with the evolution of the translational apparatus put forward the best possibility of reconstructing the code's evolutionary path, since none of the hypotheses individually is able to do so. The genetic code today is in a new, *in vitro* evolutionary stage owing to the development of nonstandard amino acids and synthetic bases. Understanding the code's past, that is the principles of its evolution, is important in order to accomplish progress in its new evolutionary stage.

## 9. ŽIVOTOPIS

### OBRAZOVANJE

- Od 2018. Prirodoslovno-matematički fakultet Sveučilišta u Zagrebu, Biološki odsjek, preddiplomski studij Molekularne biologije
- 2014. – 2018. V. gimnazija, Zagreb

### PUBLIKACIJE

- Vuković, B.B. (2017). Non-coding RNAs: molecules behind cellular development. Program and abstracts 10th ISABS Conference on Forensic and Anthropologic Genetics and Mayo Clinic Lectures in Individualized Medicine. International Society for Applied Biological Science, Zagreb.

### NAGRADE I PRIZNANJA

- 2018./19. STEM stipendija Ministarstva znanosti i obrazovanja Republike Hrvatske
- 2017. The ISABS Future Scientist Award
- 2017. priznanje za najbolje prezentirani istraživački projekt na državnom Natjecanju iz biologije
- 2017. priznanje za najbolji plakat prema glasovima učenika na državnom Natjecanju iz biologije

### AKTIVNOSTI

- 2020./21. sudjelovanje u radu Organizacijskog odbora Šestog Simpozija studenata bioloških usmjerenja (SiSB6) u okviru Tima za planiranje
- 2019. sudjelovanje na Mutimiru, godišnjem okupljanju mladih hrvatskih istraživača iz znanosti i industrije, u organizaciji udruge Penkala, u Stubičkim Toplicama
- 2019. sudjelovanje na Petom Simpoziju studenata bioloških usmjerenja (SiSB5) održanom na Kemijskom odsjeku PMF-a u Zagrebu
- 2019. održano predavanje u sklopu Noći biologije na PMF-u u Zagrebu
- 2018. sudjelovanje u radu Vijeća učenika V. gimnazije u Zagrebu kao predstavnik razreda, sudjelovanje u organizaciji Božićnog sajma V. gimnazije i Otvorenog dana V. gimnazije

- 2017. sudjelovanje te prezentacija plakata na 10. konferenciji Međunarodnog društva primjenjenih bioloških znanosti (ISABS) u Dubrovniku
- 2017. sudjelovanje na državnom Natjecanju iz biologije u kategoriji istraživački rad
- 2015. sudjelovanje na državnom Natjecanju iz biologije u kategoriji znanje
- 2014. sudjelovanje na državnom Natjecanju iz povijesti

#### STRANI JEZICI

- engleski jezik – Cambridge English Level 2 Certificate in ESOL International – Level C1 (Grade A)
- poznavanje razgovornog njemačkog jezika

#### OSTALE INFORMACIJE

- poznavanje osnova programiranja u Pythonu
- iskustvo u obradi RNA-seq podataka u programskom jeziku R