



Τμήμα Χημείας

Πανεπιστήμιο Κρήτης

Department of Chemistry

University of Crete

Εικονική κλωνοποίηση γονιδίου και δημιουργία βιβλιοθηκών μεταλλαγμάτων

2nd Seminar

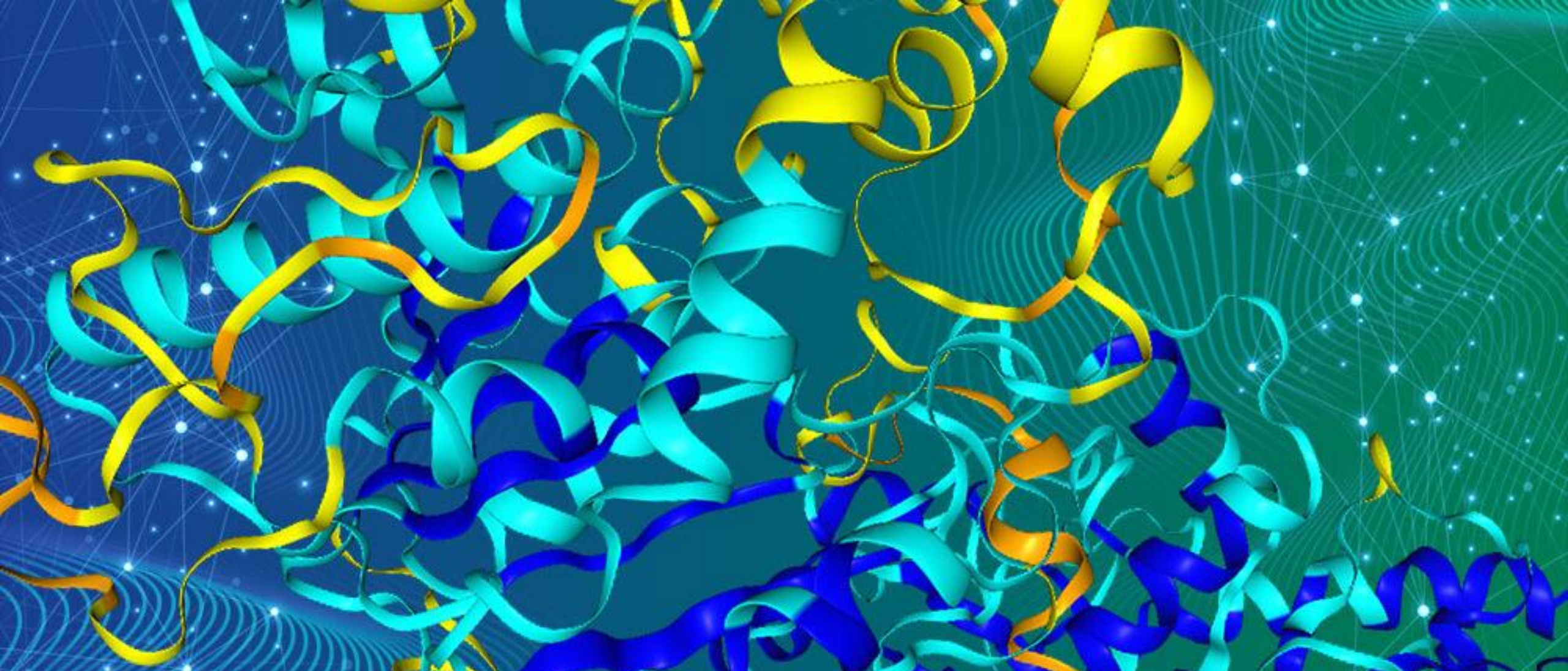
Πρωτεϊνική Μηχανική

Διδάσκων Καθηγητής: Ιωάννης Παυλίδης

[Τσέτσερη Εύα \(AM:1264\)](#)

[Ντόκο Αριστέα \(AM:1321\)](#)

Μάιος 2026, Ηράκλειο



ΜΕΡΟΣ 1: ΕΙΣΑΓΩΓΗ

Εισαγωγή: α/β υδρολάσες

Κατηγορία ενζύμων: **EC3**

Καταλύουν τη διάσπαση δεσμών μέσω αντίδρασης με H_2O

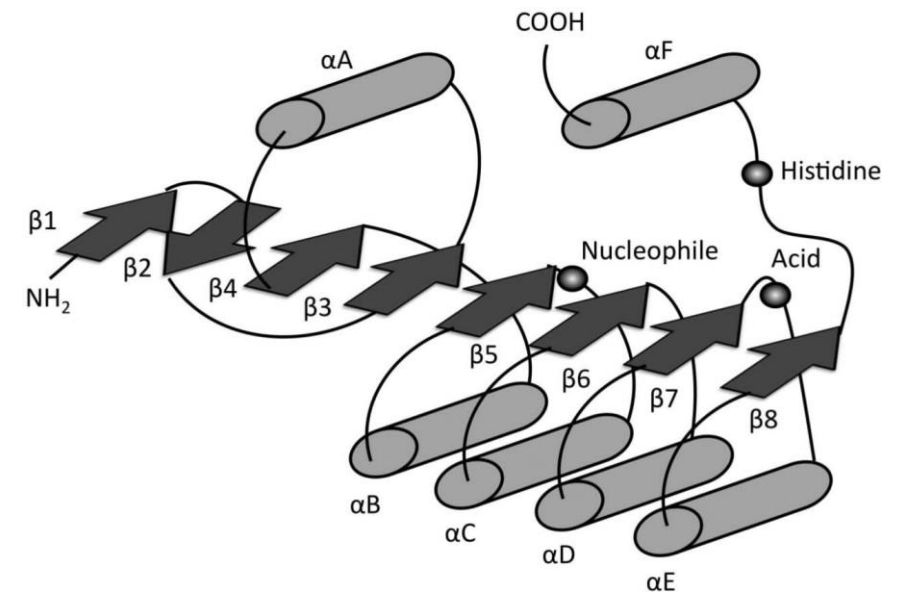
Συμμετέχουν σημαντικά στον μεταβολισμό και την αποικοδόμηση βιομορίων

Κοινός τρόπος αναδίπλωσης (α/β fold) και παρόμοιο ενεργό κέντρο

Καταλυτική τριάδα: νουκλεόφιλο-ιστιδίνη-όξινο αμινοξύ

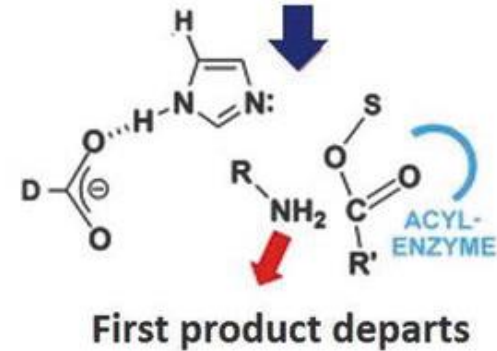
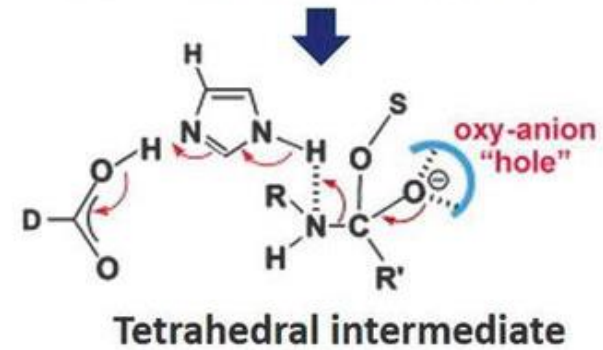
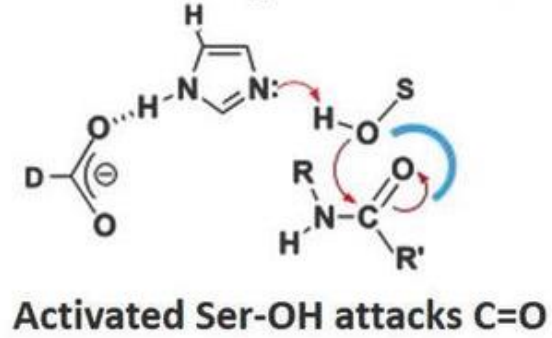
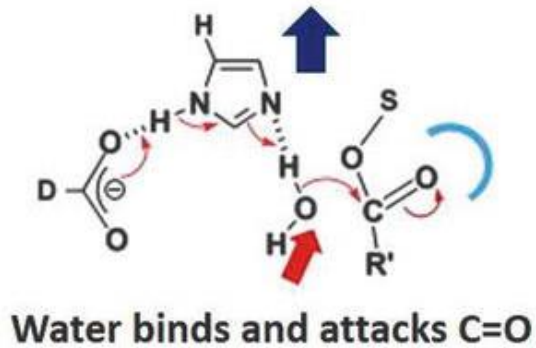
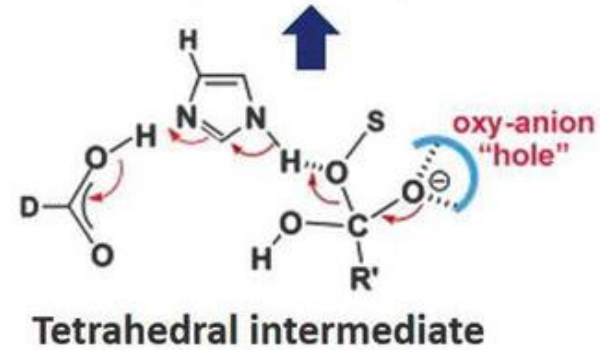
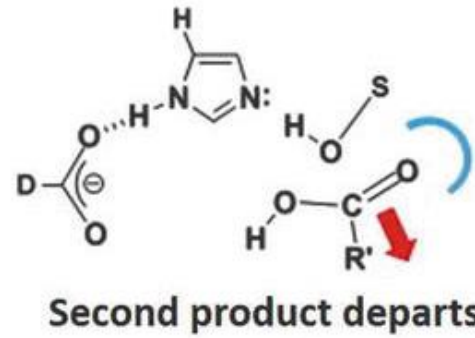
Γύρω από το ενεργό κέντρο σχηματίζεται η «οπή οξυανιόντος» (oxyanion hole) για την σταθεροποίηση της μεταβατικής κατάστασης (δομή Gly-X-Ser-Gly)

Ποικιλία αλληλουχιών και λειτουργιών (λιπάσες, εστεράσες, πρωτεάσες, αμιδάσες κ.ά)



α/β hydrolase fold (core domain)

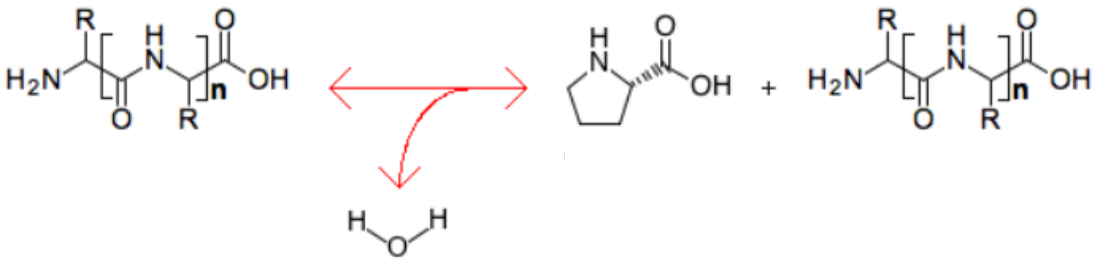
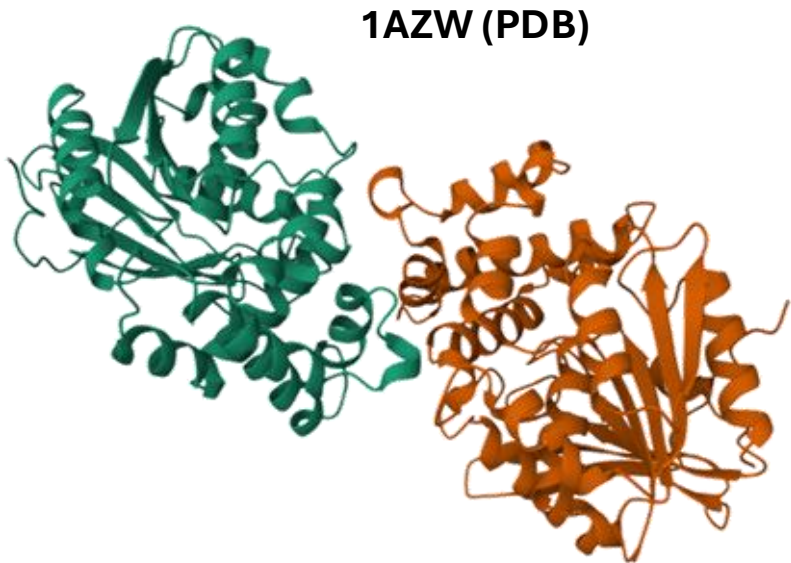
Εισαγωγή: α/β υδρολάσες – Μηχανισμός κατάλυσης



Εισαγωγή: α/β υδρολάσες – G2LE64_CHLTF

Function	Proline Iminopeptidase (Pip)
3DM Subfamily	1AZWA (73% core identity to subfamily)
Organism	Chloracidobacterium thermophilum (strain B)
EC Number	3.4.11.5
Active Residues	S113, D267, H295
Active Residues (3DM)	32, 64, 87
α/β fold	38-299

Accession Numbers	
NCBI	WP_014099022.1
Uniprot ID	G2LE64
KEGG Entry	K01259
3DM	G2LE64_CHLTF
GenBank	AEP11284



R00135 reaction

Εισαγωγή: α/β υδρολάσες – G2LE64

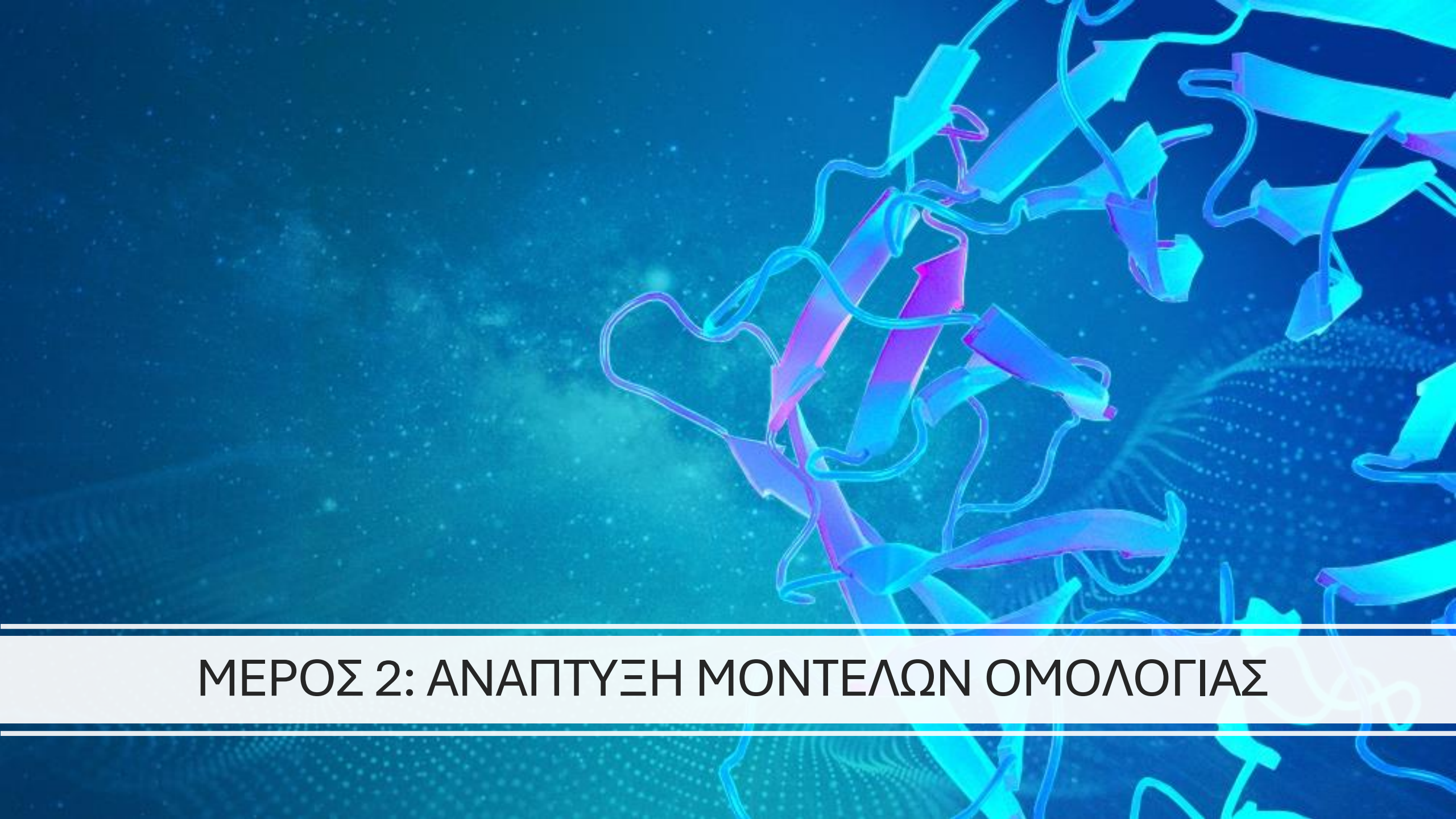
Protein Parameters	
Molecular Weight (kDa)	36.4 (36436.28 Da)
pI (theoretical)	5.63
Number of Amino Acids	318
Extinction coefficient (M ⁻¹ cm ⁻¹)	66600
Extinction coefficient (M ⁻¹ cm ⁻¹) (Cys-Reduced)	66350
Estimated half-life	30 hours (mammalian reticulocytes, in vitro)
	>20 hours (yeast, in vivo)
	>10 hours (Escherichia coli, in vivo)
Instability index	41.98 (unstable)
Aliphatic index	82.86
Grand average of hydropathicity (GRAVY)	-0.356

Amino Acid Sequence

MTDARTLYPPIEPYETGMLPVSERHTLYEVSNGNPQGKPVVFLHGGPGGGTS PDHRRYFD
PERYRIVLFDQRGAGKSTPYACLEENTTWDLVADIERLRQHLDIRKWVVFGG^SWGSTLAL
AYAETHPDRVRALVLRGIFLCRKKEIDWIFYQEGANAIFPDWEPYRDLIPPDERHDMVAA
YYRRLTSDDEQVRLAAARAWSVWEGSTSKLLPDPELIKDFDERALAIARIECHYFINRIF
MESEHYLLEHVHRIRHIPTVIVQGRY^DVVCPVMTAWELHKAFPEADFQIIPNAG^HSASEP
GTTAALVAATDRFALLPD

Nucleotide Sequence

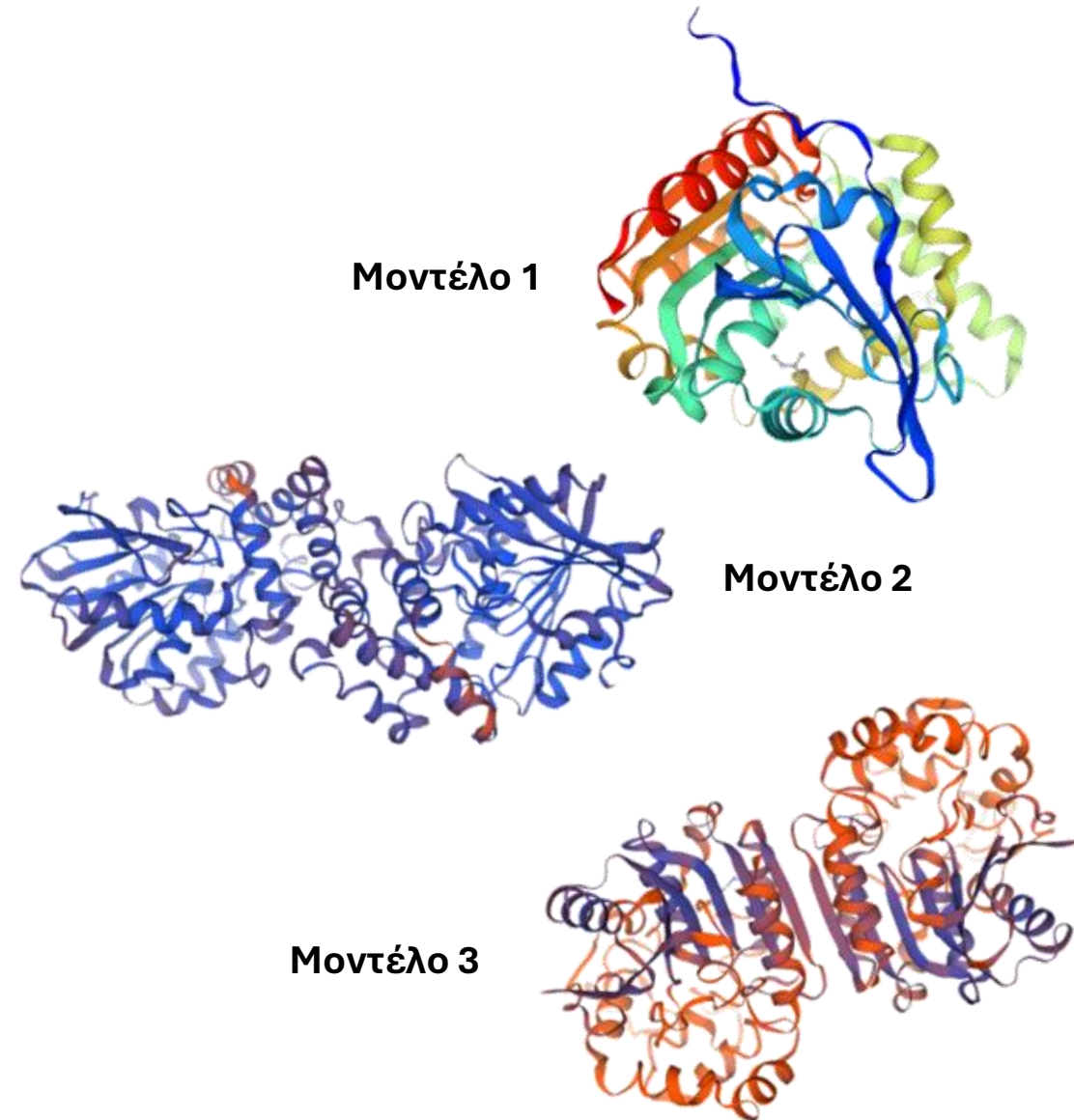
^{ATG}ACCGACGCGCGCACGCTCTATCCGCCCATCGAACCCCTATGAAACCGGGATGCTGCCG
GTTTCCGAACGGCACACGCTGTACTACGAAGTCAGCGGCAATCCCCAGGGGAAGCCGGTG
GTCTTTCTCCACGGCGGGCCCGGCGGCGGCACCTCCCCTGACCACCGGCGGTACTTCGAC
CCGGAGCGTTACCGCATCGTCCTGTTTGACCAGCGCGGTGCCGGGAAAAGCACGCCCTAC
GCCTGTCTGGAAGAGAACACGACGTGGGATTTGGTTGCCGACATCGAGCGTTTGCGCCAA
CACCTGGACATCCGAAAGTGGGTCGTGTTTGGCGGCTCCTGGGGCAGTACCCTGGCGCTG
GCTTACGCAGAAACCCATCCCGACCGCGTCCGGGCGCTGGTGCTGCGGGGAATTTTTCTG
TGCCGCAAAAAGGAGATTGACTGGTTTTACCAGGAAGGGGCGAATGCCATTTTCCCGGAC
GCCTGGGAGCCGTACCGCGATCTCATTCCCCCGGATGAACGCCACGACATGGTGGCCGCT
TACTACCGCCGGTTGACGAGCGACGATGAGCAAGTGCGGCTGGCGGCTGCCCGCGCCTGG
AGTGTCTGGGAAGGCAGTACCTCAAACTGCTCCCGGACCCGGAACATCAAGGATTTTC
GACGAGCGTGCGCTGGCCATTGCCCGCATTGAGTGCCACTACTTCATCAACCGGATTTTC
ATGGAGTCCGAGCATTACCTGCTCGAACACGTCCATCGCATTCGCCACATTCCGACGGTC
ATCGTGCAAGGGCGTTACGATGTGGTGTGTCCGGTGATGACCGCCTGGGAGCTGCACAAG
GCGTTTCCCGAAGCCGATTTTCAAATCATTCCGAACGCCGGGCATTCCGCCTCGGAGCCG
GGGACGACGGCGGCGCTGGTTGCGGCGACGGACCGTTTCGCCCTGCTGCCCGAC^{TAG}



ΜΕΡΟΣ 2: ΑΝΑΠΤΥΞΗ ΜΟΝΤΕΛΩΝ ΟΜΟΛΟΓΙΑΣ

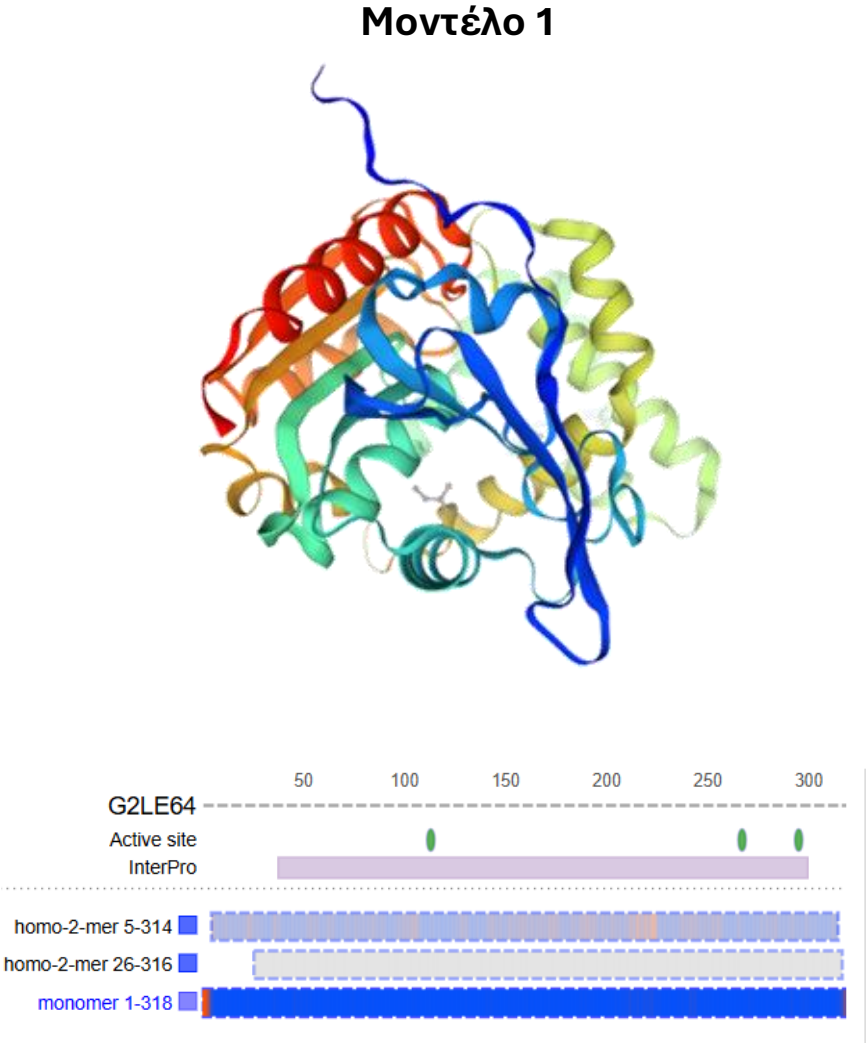
SWISS-MODEL

	MODEL 1	MODEL 2	MODEL 3
Sequence Identity	100%	64.84%	20.85%
Sequence Similarity	0.63	0.50	0.29
Template	G2LE64.1.A	1azw.1.A	1u2e.1.A
Oligo-State	Monomer	Homo-dimer (matching prediction)	Homo-dimer (matching prediction)
GMQE Score	0.98	0.87	0.45
QSQE	–	0.48	0.20
Range	1-318	5-314	26-316
Coverage	1.00	0.97	0.81
Method	AlphaFold v2	X-ray, 2.70 Å	X-ray, 2.10 Å
Ligands	–	–	–



SWISS-MODEL - BEST

	MODEL 1	MODEL 2	MODEL 3
Sequence Identity	100%	64.84%	20.85%
Sequence Similarity	0.63	0.50	0.29
Template	G2LE64.1.A	1azw.1.A	1u2e.1.A
Oligo-State	Monomer	Homo-dimer (matching prediction)	Homo-dimer (matching prediction)
GMQE Score	0.98	0.87	0.45
QSQE	–	0.48	0.20
Range	1-318	5-314	26-316
Coverage	1.00	0.97	0.81
Method	AlphaFold v2	X-ray, 2.70 Å	X-ray, 2.10 Å
Ligands	–	–	–



D-I-TASSER

Η αξιοπιστία του μοντέλου προσδιορίζεται μέσω του TM-score

	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5
TM Score	0.80	0.64	0.53	0.53	0.53

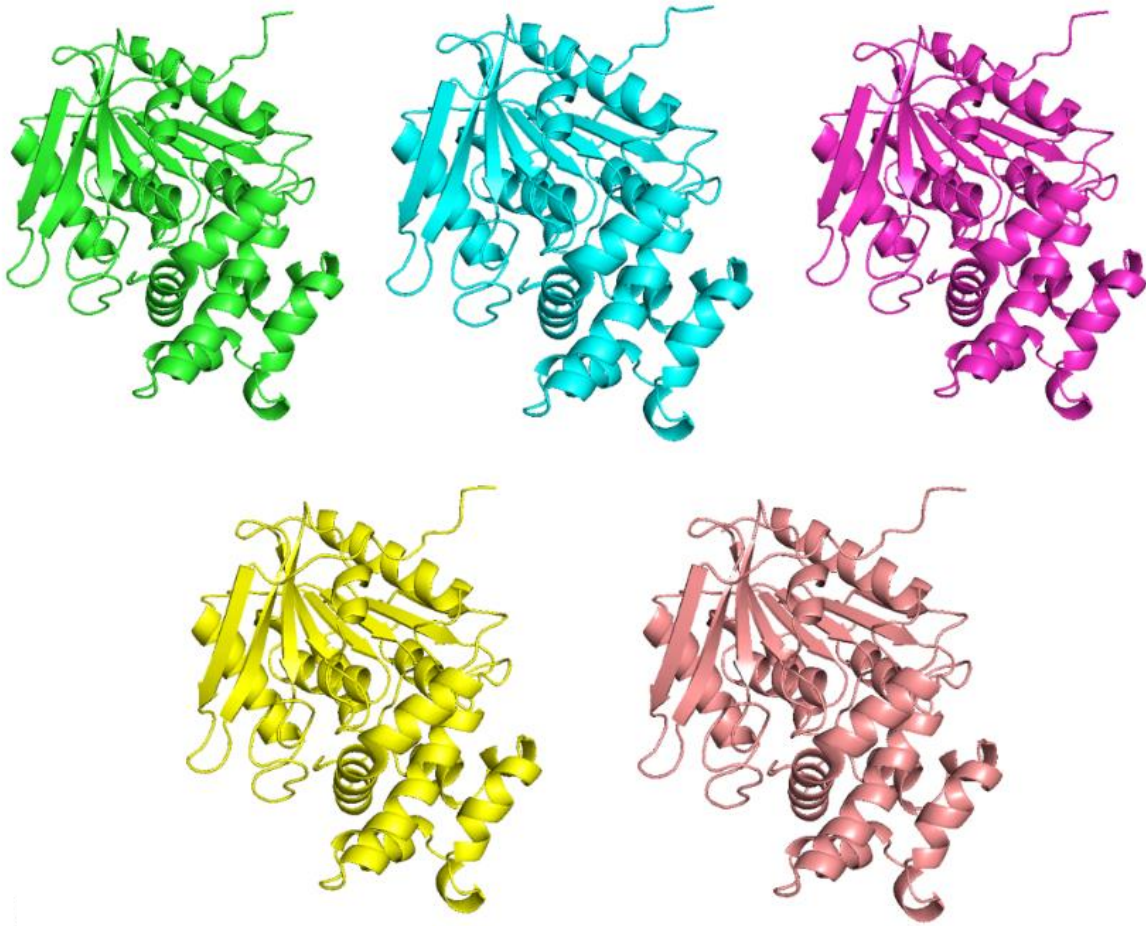


Περισσότερο αξιόπιστο μοντέλο: Model 1

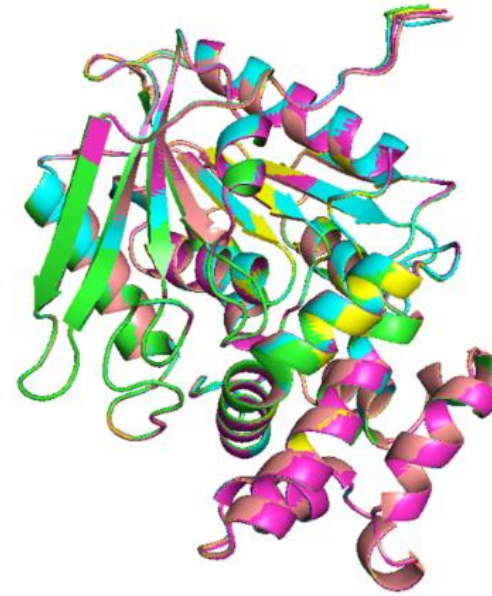
Top 10 threading templates used by D-I-TASSER

Rank	PDB hit	ID1	ID2	Cov	Norm. Zscore	Download alignment
1	5yhpA	0.49	0.46	13.42	1.33	DEthreader
2	5yhpA	0.52	0.42	12.28	1.17	DEthreader
3	1azwA	0.65	0.63	18.04	1.94	SPARKS-K
4	1azwA	0.65	0.63	18.04	0.30	CEthreader
5	1azwA	0.65	0.63	18.04	1.71	MUSTER
6	1qtrA	0.58	0.57	16.27	1.03	HHsearch
7	1qtrA	0.58	0.57	16.26	2.80	FFAS-3D
8	1azwA	0.65	0.63	18.04	2.41	CNFpred
9	1qtrA	0.50	0.47	13.68	1.33	DEthreader
10	1qtrA	0.51	0.42	12.03	1.17	DEthreader

Alpha-Fold v.3



5 generated 3D structures

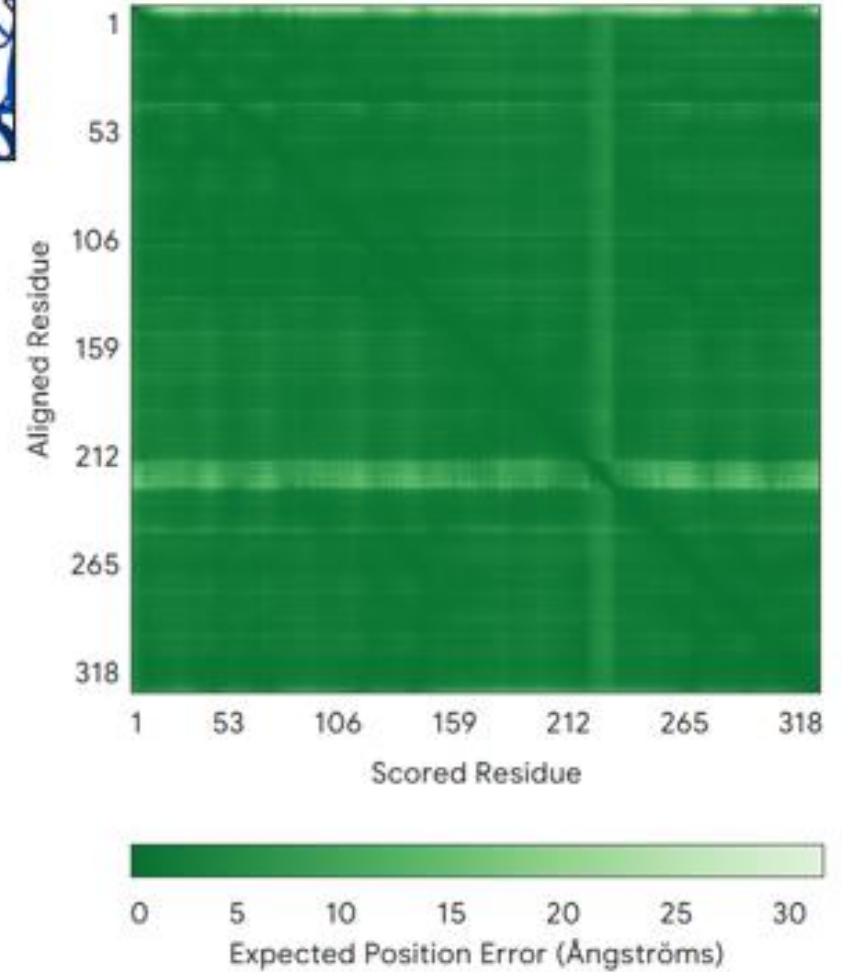
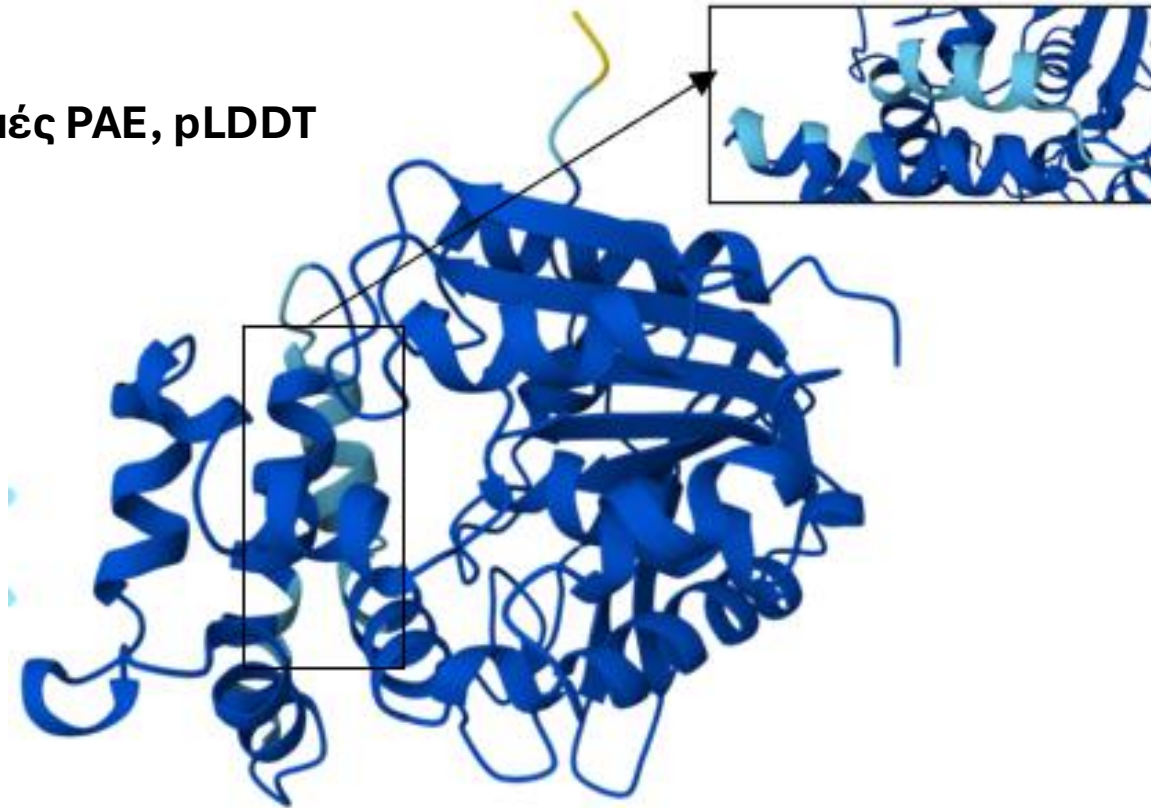


Structural Alignment of
predicted models

- Μέθοδος βαθιάς μάθησης (*deep learning*)
- Δεν χρησιμοποιήθηκε template για την πρόβλεψη της 3D δομής
- Δημιουργήθηκαν 5 μοντέλα
- Όλα τα αμινοξέα μοντελοποιήθηκαν

Alpha-Fold v.3 – Best Model

Υψηλές τιμές PAE, pLDDT
pTM = 0.95



PVFLVLLH	GYRVVAPD	RVALLVGHHS	GGALLALLALRL	ERVVKGLVLLSP	PTLIHGEDDP	PDE
QPVVLIH	GYRVITYD	DAVLVGHHS	GAGVALRYASR	DRVAKAVLLAS	PVLLILHGTDDR	PAT
QPIVFHH	GYRVIAHD	NAVHIGHS	GGGEVARYVAR	GRVAKAVLVSA	PTLVLHGGDDDDQ	ADS
RPVVFVH	GYRVIAHD	DVTLVGHHS	GGGEVARYIGR	GRVRSLLVLVGA	PTLVVHGGDDDDQ	DAT
QPVVFSH	GYRVIAHD	DAVLVGHHS	AGQYIPYIA DA	ARVAKAVLIGA	PTLVIHGGDDDDQ	DAS
PLVIWLN	LANLLLFVD	DIYIAGES	GGASVGMHLLS	INLKGLLLIGNG	RVLLLFNGDLDL	NGT
PVMVWIY	DVILVSMN	RVTIFFGES	GGASVHLQTL	GLFHRAILQSG	PILLGVNKKDEG	ICP
PVMVWIH	NVIVVTIN	NITLFFGES	GGASVHLQTL	GLFKRAISQSG	PYIAGYNTMDG	LVP
ACVIWLH	DTRFVFPQ	RIFLAGFS	GGAMALHTALR	EPLAGVIALSG	PVLLAHGQQDP	LAM
KPVVFLH	KYRIVLFD	KWVVFGGGS	GSTLALAYAQ	ERVVTGLVLRGI	PGVIVQGGRYDV	LDS
PPVLFLLH	RHRCIAPD	RVVLLVGH D	GGAILGLRLAAR	ERVVRGLVLLNT	PTLLVWGTTRDP	PAF
PPVLLLLH	RFRVIAPD	KAHVVGHS	GGAVALELAAE	ERVVDSVVLVGS	PVLLLLWGRQDR	LET
PVLFWIF	PIIFVAVN	KVTIWGES	GAISVFDQLLA	PLFRGAIMQSG	PIIIGDQEDEG	TLP
PVVLWLN	NANVIFLD	DFHIAGES	AGHYIPAFASE	INLKGVLLIGNG	RVLIYAGDADL	QAW
PPVLLCH	GYRVIAPD	QAVLIGH D	GGAVVWRMALF	ERVRAVASLNT	PALMVTAEKDP	PEM
DPIVFLH	LVRLIAPD	RVVLLVH D	GSALGF D WARR	ERVQGIAYME A	PKLFINAEPGA	GRI
PGIVVIQ	GYVVLAPD	KVG VVG YC	GGKLAYLAAAR	TDV DAAVSYYG	PLLLHFGEQDP	REA
PAVYLL D	GVNVVMPV	GNAIVGLS	GGTAALILAAAR	GQFKYAAASLSG	RLYVSCGNGT	LRT
PALLYGY	GGVLAVAN	KLAINGGS	GGLLVAAACINQ	DLFGCAIAQVG	ATLLLTADHDD	PLH
PPLLLLLH	SYRVIAPD	KAHVVGH D	GGVVALLLLA D K	ERVKKLTIIAS	PVLLLIWGEHDP	KAH
PALVYFH	GAVVVSVD	RIAVGG D S	GGNLA AVVALL	QKL AGVVALSG	PALIA TAGFDP	RDE
ATVIFLH	HVKFICPH	RIVLGGFS	GGALS LYTALT	DYFKYFMPLSG	PILQC HGDADP	YEF
NVLYLMH	PLIVVTPT	HRAFGGFS	GSVTTWYVFIN	SQLRGFVPVAP	YIFAATGSDDI	KPQ
FSVLLLLH	GYRAVAID	PPVLLVSPS	SGMYSLPFLMA	ERVRRLLVLMGS	PTLIVYGDQDP	QTS
PPVLLLIH	GYRVIAPD	KAHLVGHHS	GGAVALLALLALR	INLKGLVLLNS	PTLI IWGEEDR	LEH
PLVLWLN	VANVLFLE	DFYITGES	AGHYVPQLAQL	ELVKGLVLLNS	RVLVYSGD TDS	RYW
PPLVLLH	HYRVIAPD	PVVLVGNSS	GGAILALQLAAL	PSLKAALPLTF	PTLI IWGEEDDP	VEV
GAVAIS P	GFVVITID	RLAVMGHS	GGGGGLEAAKB		PTLIIGAEENBT	ATH
PVLYLLH	PMIVVMPD	GRAIAGLS	GGYOALVIGLN	DLFGYVGSFSG	LF FIDCGDKDF	YKD
PVLIYFH	GAVVVSVD	KIGVVGHS	GGVVALVIGLN	PTLIATAEFDP	PTLISVGE LDP	RDE
PVVVYIH	GAVVVSVD	RLVVGHS	GGVVALVIGLN	PTFISVGE LDP	PVALY YQGNW	RDE
PVVL LQW	GYDVWLGN	KLWYVGHHS	GGVVALVIGLN	PTFISVGE LDP		PEB
PVLLICIP	RHSVYALP	PFVLSGYYS	GGVVAYALASH	RNPLGVVLI D T	PVLLVRAKDR I	RIN
PAIVKVH	GYATLEIG	RIGVTGGGS	GGALSIAA AAL	DRVKA AVA DYP	PVLM SIGLND E	PTT
PVLLERT	GYAVVRQD	KVGM MGIS	GGFTQLQA AAL	PALKAI V PWEG	PVLAVGGWY D G	LNG
PVIMTAS	GEAVVYVS	KVAMTGVSS	IGTISNAVATT	DG IKTIIPEAA	SVVFEVHGINDW	PKH

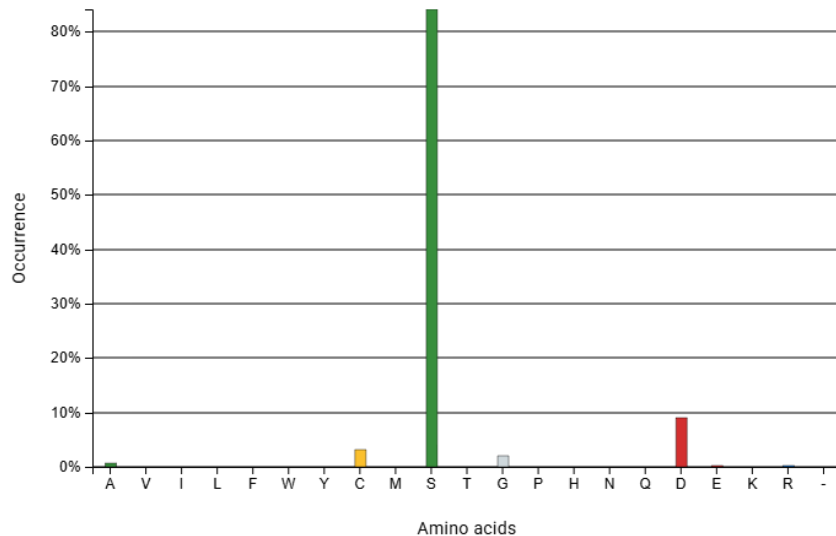
ΜΕΡΟΣ 3: Ανάλυση 3DM

3DM – Active Residues

ACTIVE SITE RESIDUES		
Residue	No. Sequency	No. 3DM
Serine (Ser, S) <i>nucleophile</i>	113	32
Aspartic Acid (Asp, D)	267	64
Histidine (His, H) <i>protein donor</i>	295	87

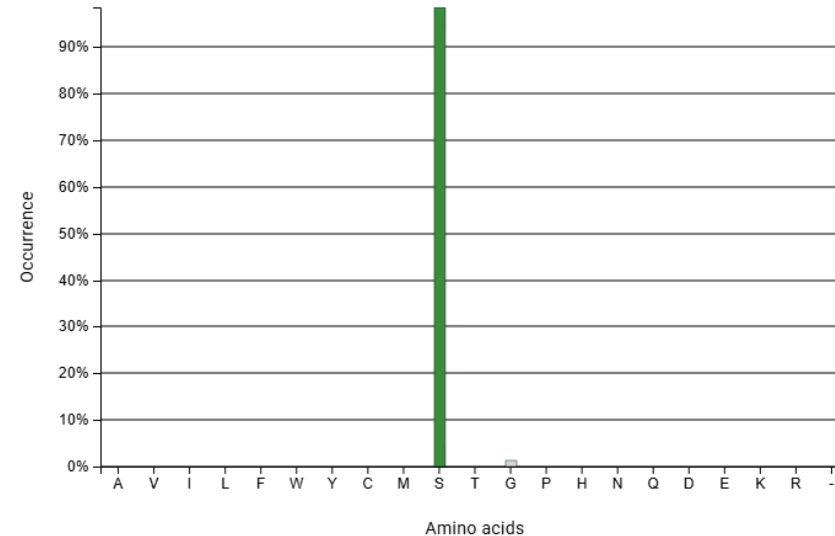
Τα κατάλοιπα της καταλυτικής τριάδας είναι συντηρημένα σε ποσοστό μεγαλύτερο του 75%

Θέση 32 - υπεροικογένεια



Ser: 84.15%
Asp: 9.02%
Cys: 3.22%
Gly: 1.99%

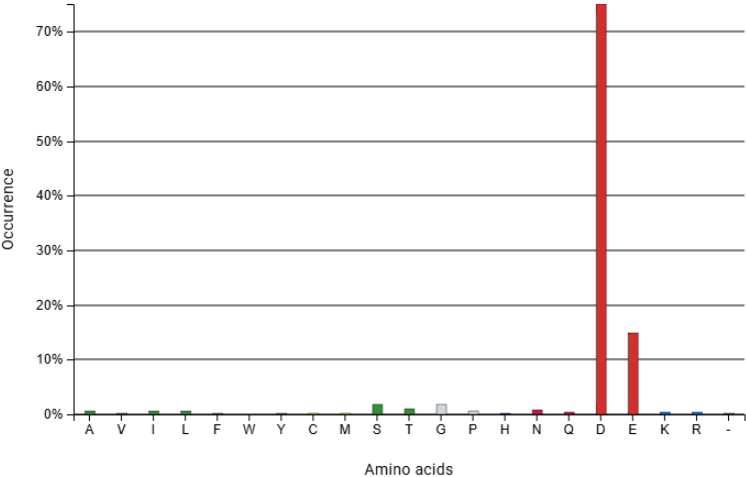
Θέση 32 – υποοικογένεια (1AZWA)



Ser: 98.59%
Gly: 1.41%

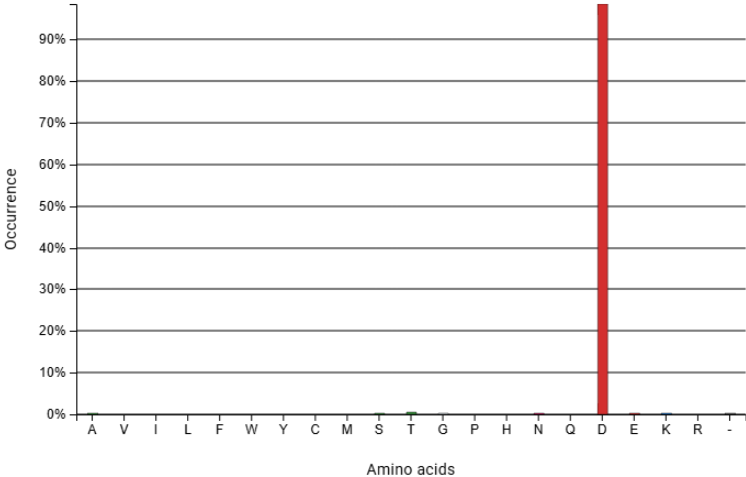
3DM-Active Residues

Θέση 64 - υπεροικογένεια



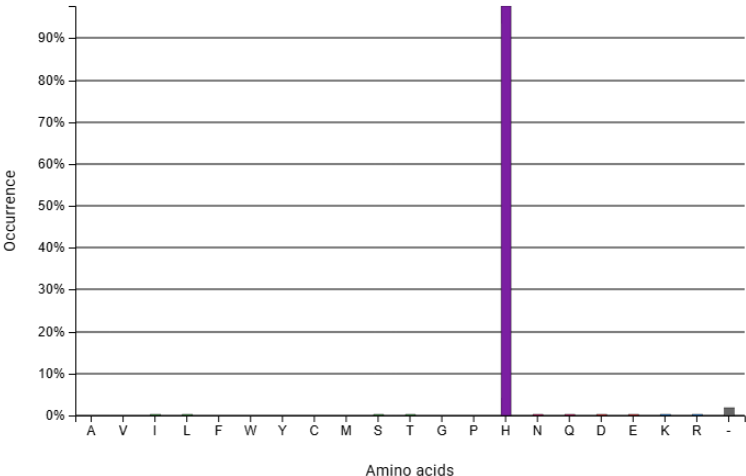
Asp: 75.13%
Glu: 14.98%

Θέση 64 – υποοικογένεια (1AZWA)



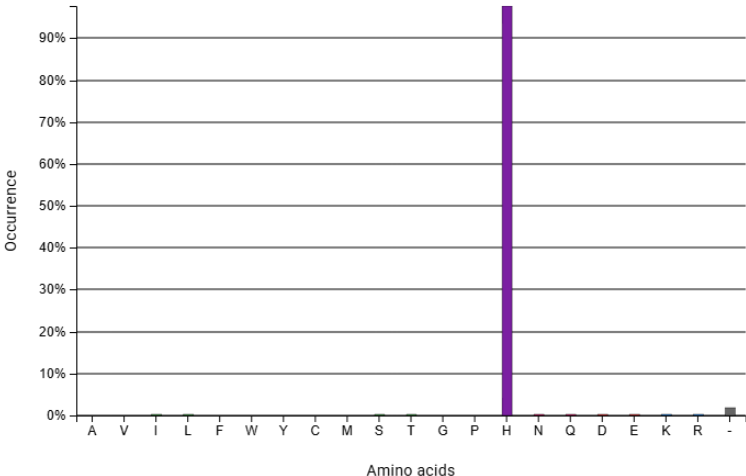
Asp: 98.58%
Others: <1%

Θέση 87 - υπεροικογένεια



His: 97.71%

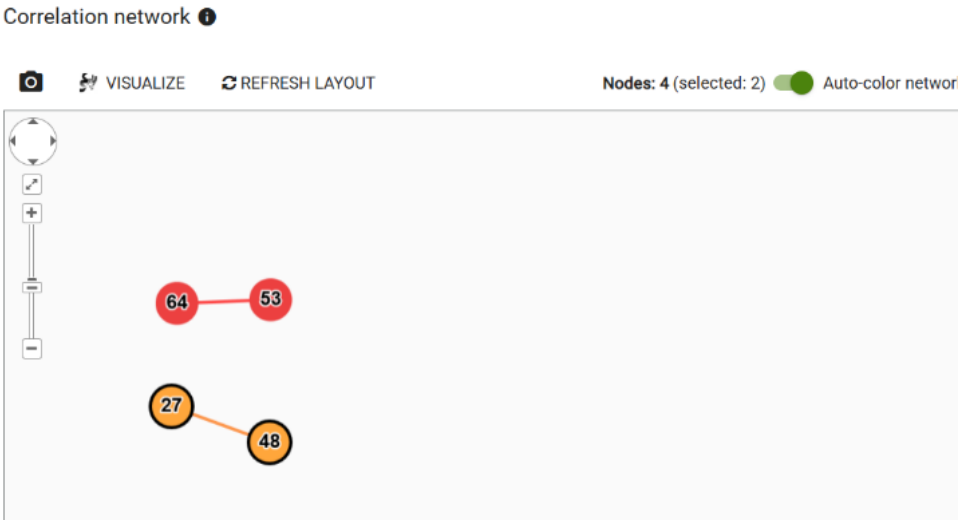
Θέση 87 – υποοικογένεια (1AZWA)



His: 99.52%

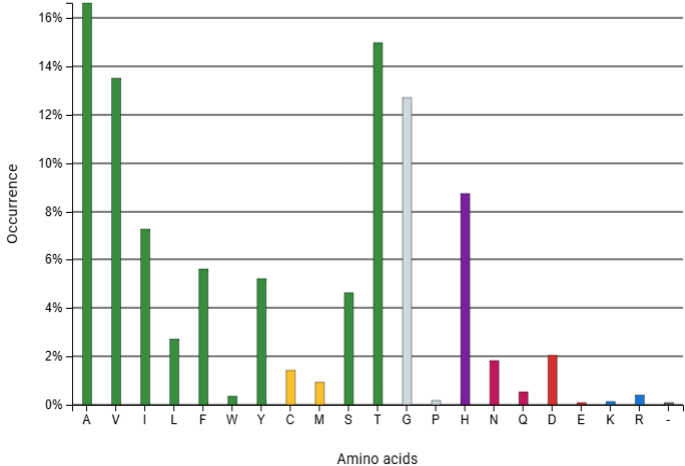
3DM – Correlation Network

Number	27	48	Proteins	Percentage	Number	53	64	Proteins	Percentage
1.	G	A	3,597	6.25%	1.	A	D	7,867	13.67%
2.	T	R	3,582	6.23%	2.	S	E	7,431	12.92%
3.	V	G	3,259	5.66%	3.	S	D	7,172	12.47%
4.	H	S	3,174	5.52%	4.	N	D	5,765	10.02%
5.	A	A	3,088	5.37%	5.	Y	D	5,464	9.50%
6.	V	A	2,229	3.87%	6.	G	D	5,124	8.91%
7.	Y	G	2,215	3.85%	7.	V	D	3,597	6.25%
8.	F	G	2,138	3.72%	8.	D	D	2,425	4.21%
9.	T	K	1,832	3.18%	9.	C	D	1,540	2.68%
10.	I	G	1,367	2.38%	10.	H	D	684	1.19%



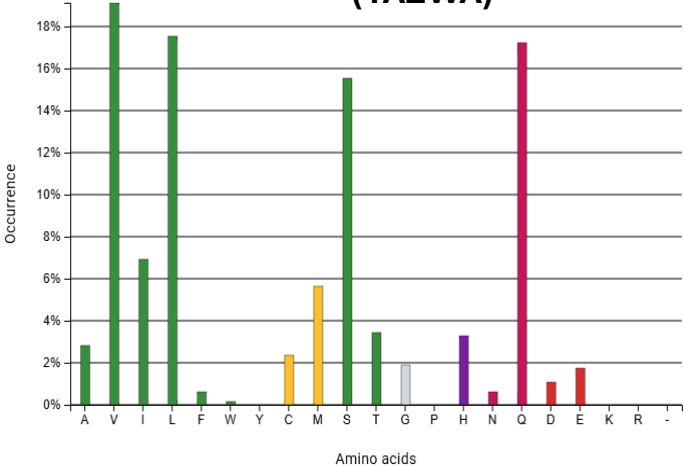
Ζεύγη συνεξέλιξης: **27-48** και **53-64** (κατά 3DM)
(cutoff 0.8)

Υπεροικογένεια



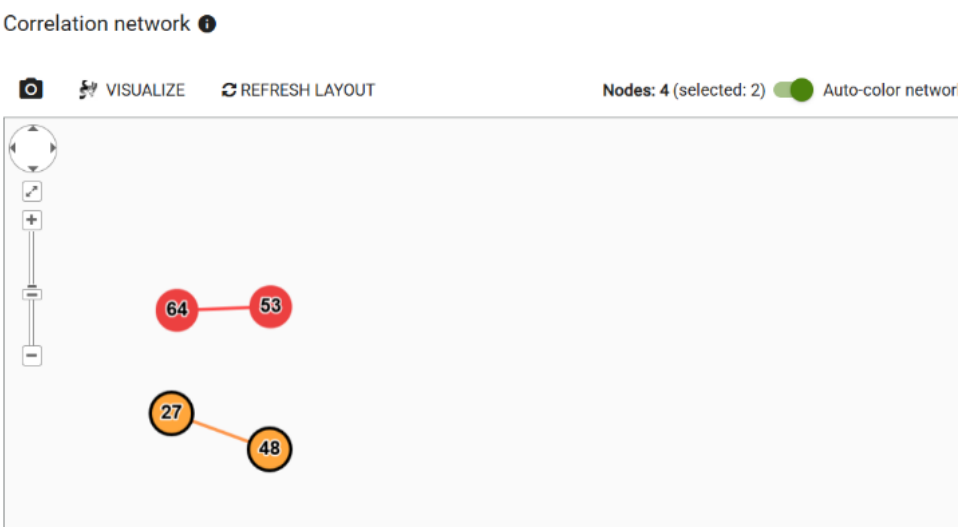
ΘΕΣΗ 27

Υποοικογένεια
(1A2WA)



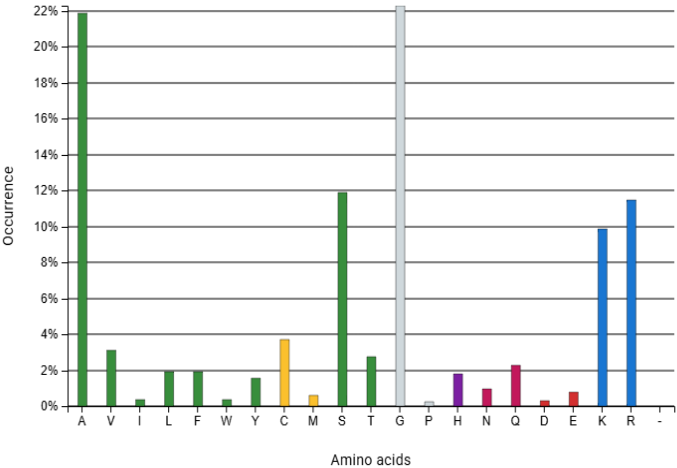
3DM – Correlation Network

Number	27	48	Proteins	Percentage	Number	53	64	Proteins	Percentage
1.	G	A	3,597	6.25%	1.	A	D	7,867	13.67%
2.	T	R	3,582	6.23%	2.	S	E	7,431	12.92%
3.	V	G	3,259	5.66%	3.	S	D	7,172	12.47%
4.	H	S	3,174	5.52%	4.	N	D	5,765	10.02%
5.	A	A	3,088	5.37%	5.	Y	D	5,464	9.50%
6.	V	A	2,229	3.87%	6.	G	D	5,124	8.91%
7.	Y	G	2,215	3.85%	7.	V	D	3,597	6.25%
8.	F	G	2,138	3.72%	8.	D	D	2,425	4.21%
9.	T	K	1,832	3.18%	9.	C	D	1,540	2.68%
10.	I	G	1,367	2.38%	10.	H	D	684	1.19%



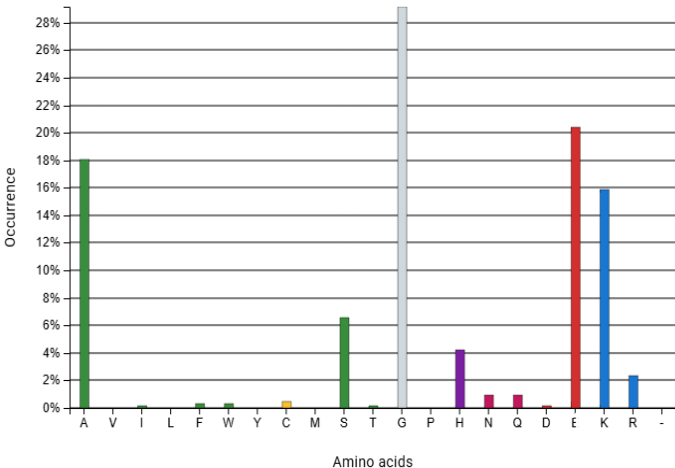
Ζεύγη συνεξέλιξης: **27-48** και **53-64** (κατά 3DM)
(cutoff 0.8)

Υπεροικογένεια



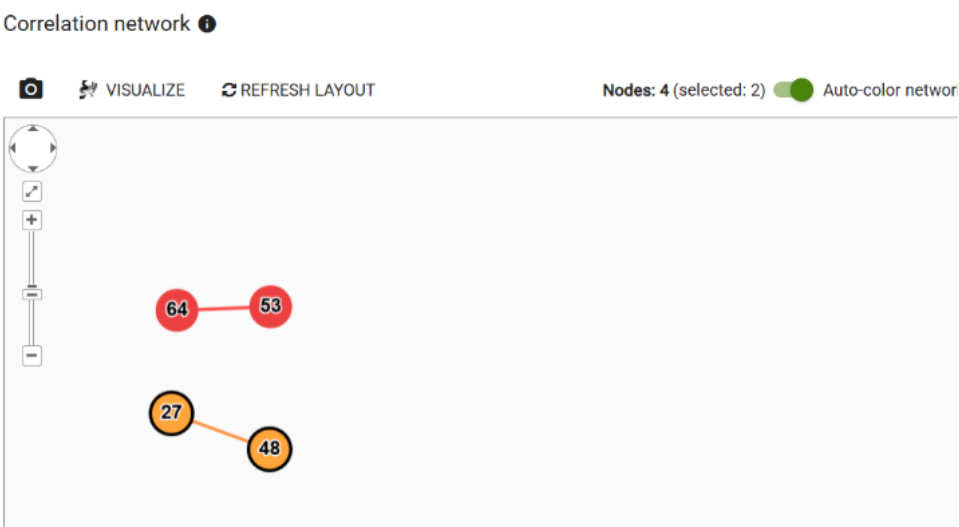
ΘΕΣΗ 48

Υποοικογένεια
(1A2WA)



3DM – Correlation Network

Number	27	48	Proteins	Percentage	Number	53	64	Proteins	Percentage
1.	G	A	3,597	6.25%	1.	A	D	7,867	13.67%
2.	T	R	3,582	6.23%	2.	S	E	7,431	12.92%
3.	V	G	3,259	5.66%	3.	S	D	7,172	12.47%
4.	H	S	3,174	5.52%	4.	N	D	5,765	10.02%
5.	A	A	3,088	5.37%	5.	Y	D	5,464	9.50%
6.	V	A	2,229	3.87%	6.	G	D	5,124	8.91%
7.	Y	G	2,215	3.85%	7.	V	D	3,597	6.25%
8.	F	G	2,138	3.72%	8.	D	D	2,425	4.21%
9.	T	K	1,832	3.18%	9.	C	D	1,540	2.68%
10.	I	G	1,367	2.38%	10.	H	D	684	1.19%

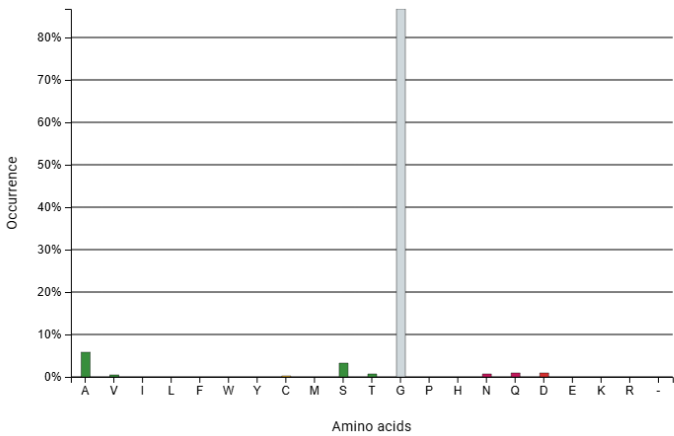
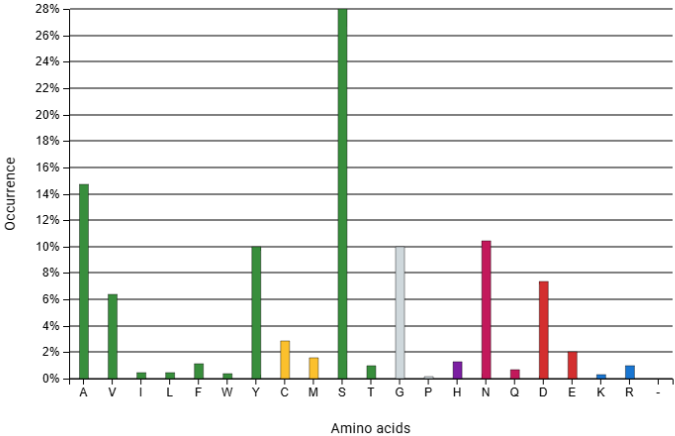


Ζεύγη συνεξέλιξης: **27-48** και **53-64** (κατά 3DM)
(cutoff 0.8)

ΘΕΣΗ 53

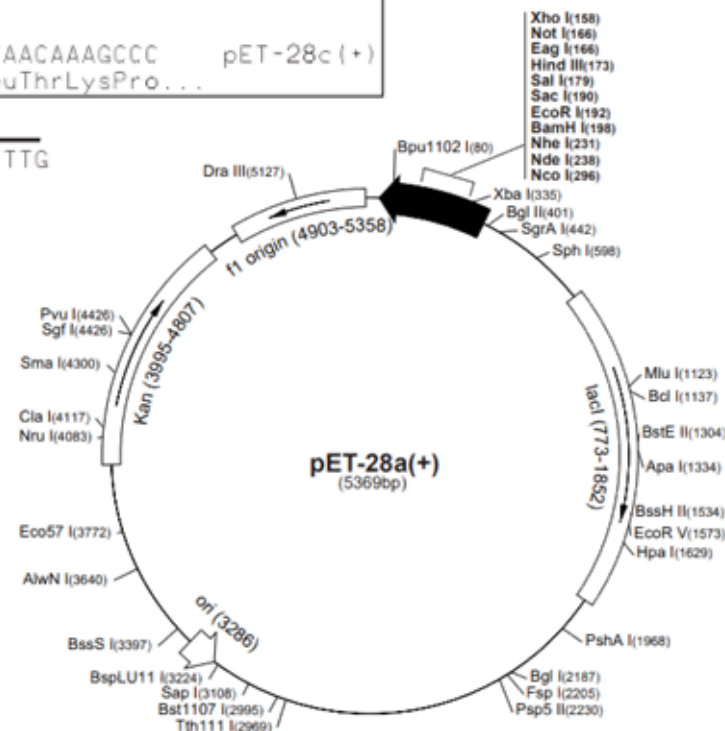
Υποοικογένεια
(1AZWA)

Υπεροικογένεια





ΜΕΡΟΣ 4: Εικονική Κλωνοποίηση



Εικονική Κλωνοποίηση

Name	Cat #	Temp °C	Supplied Buffer	Add SAM	% Activity in NEBuffer™			
					r1.1	r2.1	r3.1	rCutSmart
NdeI	R0111	37	rCutSmart Buffer	No	75	100	100	100
SacI- HF	R3156	37	rCutSmart Buffer	No	10	50	10	100

Name	Time- Saver™	Heat Inactivation (°C)	Methylation Sensitivity
NdeI	Yes	65	None
SacI- HF	Yes	65	CpG:Blocked by Some Combinations of Overlapping



- Βέλτιστη Θερμοκρασία δράσης : **37 °C**
- Απενεργοποίηση: **65 °C**
- 100% ενεργότητα στο ρυθμιστικό διάλυμα rCutSmart Buffer και στα δύο ένζυμα
- Δεν χρειάζεται κάποιο συμπαράγοντα

Βελτιστοποίηση Γονιδίου

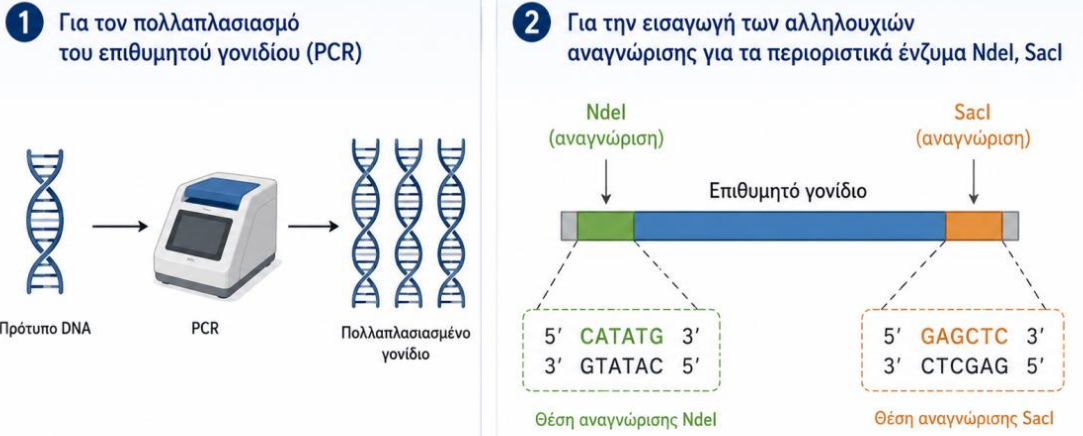
- Βακτήριο Escherichia Coli str. K-12 substr. MG1655
- Αποκλείστηκαν οι αλληλουχίες αναγνώρισης των δύο περιοριστικών ενζύμων

Βελτιστοποιημένη Αλληλουχία: GC= **57.05%** , CAI = **0.95**

```
ATGACCGACGCGCGCACTCTGTATCCGCCGATTGAACCGTATGAAACCGGCATGCTGCCGGT
AAGCGAACGCCACACCCCTGTATTATGAAGTGAGCGGCAACCCGCAGGGCAAACCGGTTGTGT
TTTTACACGGCGGCCCGGGCGGCGGCACCAGCCCGGATCATCGCCGCTATTTTGATCCGGAA
CGTTATCGCATTGTGCTGTTTGATCAGCGTGGCGCCGGCAAAAGCACCCCGTATGCCTGTCT
GGAAGAAAATACCACCTGGGATCTGGTTGCGGATATTGAACGCCTTCGCCAGCACCTGGACA
TTCGTAAATGGGTGGTGTGTTGGTGGCAGCTGGGGCAGCACCCCTGGCGCTGGCGTATGCCGAA
ACCCATCCGGATCGCGTTCGCGCGCTGGTGCTGCGCGGCATTTTTCTGTGCCGTAAAAAAGA
AATTGATTGGTTTTTATCAGGAAGGCGCCAACGCGATTTTTCCGGATGCGTGGGAACCGTACC
GTGATCTGATTCCGCCGGATGAACGTCATGATATGGTGGCAGCCTACTACCGCCGCCTGACC
AGCGATGATGAACAGGTGCGTCTGGCGGGCGGCGCGCGCTGGAGCGTGTGGGAAGGCAGCAC
CAGCAAACCTGCTGCCGGATCCGGAACCTGATCAAAGATTTTCGATGAACGCGCACTGGCCATTG
CCCGTATTGAATGCCACTATTTTCATTAACCGCATTTTCATGGAAAGCGAACATTATCTGCTG
GAACATGTGCATCGTATTCGTACATTCCGACCGTGATTGTGCAGGGCCGCTATGATGTGGT
GTGCCCCGGTGATGACCGCCTGGGAACTGCATAAAGCCTTCCCGGAAGCGGATTTTCAGATTA
TTCCGAATGCCGGTCATAGCGCGAGCGAACCGGGCACCAACCGCAGCGCTGGTGGCGGCGACC
GATCGCTTTGCACTGCTGCCGGATTAA
```

Σχεδιασμός Εκκινητών

Σκοπός:



- ✓ 10–15 βάσεις από το γονίδιο
- ✓ να έχουν και οι δύο παρόμοια θερμοκρασία τήξης (πάνω από 50 °C)
- ✓ την θέση αναγνώρισης του περιοριστικού ενζύμου
- ✓ στο 3' άκρο να έχουν ισχυρά νουκλεοτίδια (1 ή 2 C/G), τα οποία σχηματίζουν τρεις δεσμούς υδρογόνου κατά τον υβριδισμό
- ✓ 4–5 βάσεις προς το 5' άκρο ώστε η θέση περιοριστικής πέψης να μην είναι στην άκρη

Oligonucleotide Properties Calculator:

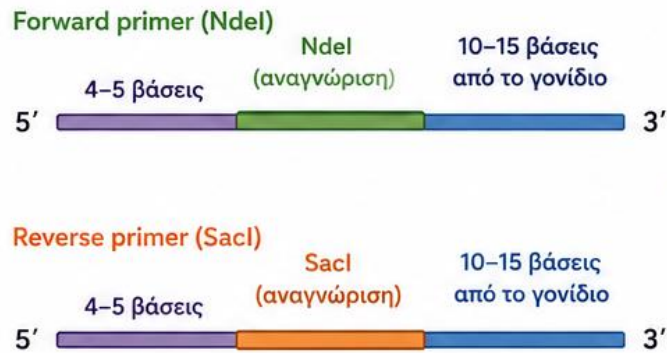
Εκκινητές	Αλληλουχία Εκκινητών	Μήκος	Μοριακό βάρος	Περιεχόμενο GC%	T _M (Salt Adjusted)	Αυτοσυμπληρωματικότητα
1 ^{ος} Primer	5´ GCGCCATATGACCGACG 3´	17	5180.4	65	57	None
2 ^{ος} Reverse Primer	5´ GCGCGAGCTCTTAATCCG 3´	18	5475.6	61	58	None

Εικονική Κλωνοποίηση - PCR - Πέψη - Κλωνοποίηση

Ροή εργασίας: Από τον σχεδιασμό εκκινητών μέχρι το ανασυνδυασμένο πλασμίδιο

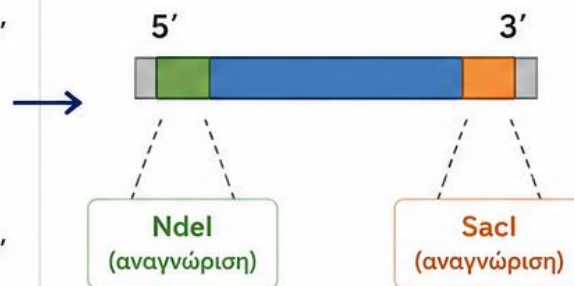
1 Σχεδιασμός εκκινητών

Οι εκκινητές περιέχουν θέσεις αναγνώρισης για τα ένζυμα NdeI και SacI.



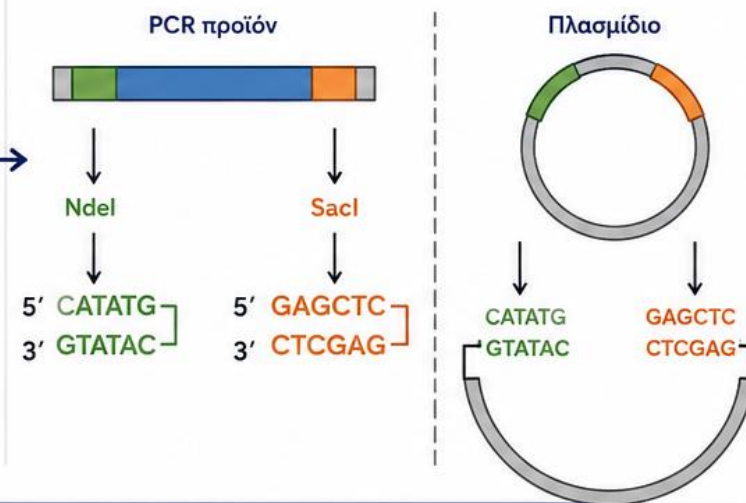
2 PCR

Μετά την PCR προκύπτει το προϊόν που φέρει τους εκκινητές στα 5' και 3' άκρα.



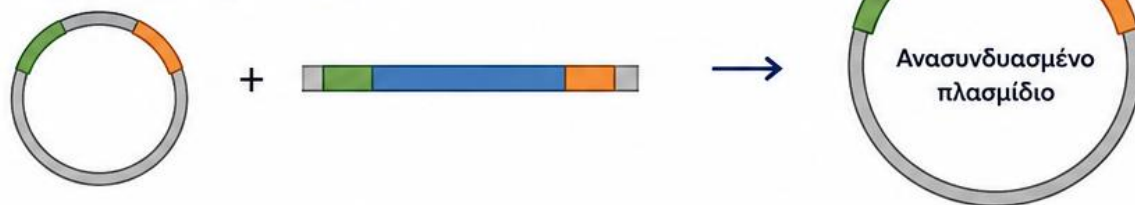
3 Πέψη με περιοριστικά ένζυμα

Πέψη τόσο του γονιδίου (PCR προϊόντος) όσο και του πλασμιδίου με τα ίδια ένζυμα (NdeI και SacI) για τη δημιουργία συμπληρωματικών άκρων.



4 Ligation

Τα συμπληρωματικά άκρα υβριδίζονται και η DNA λιγάση συνδέει το ένθετο με το πλασμίδιο.



5 Εργαλείο που χρησιμοποιήθηκε

Για τη διαδικασία αυτή χρησιμοποιήθηκε το εργαλείο Geneious.

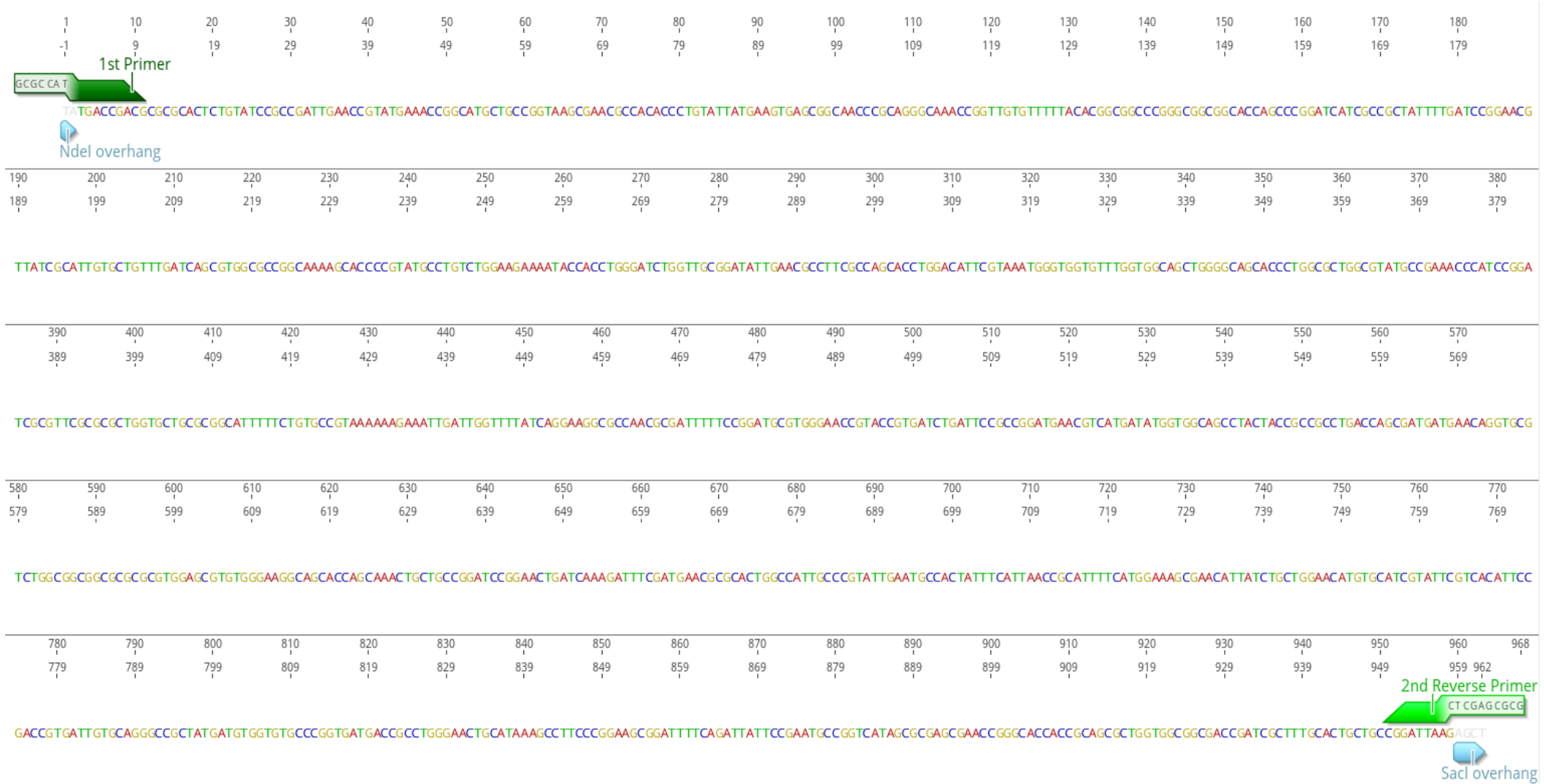


PCR Product



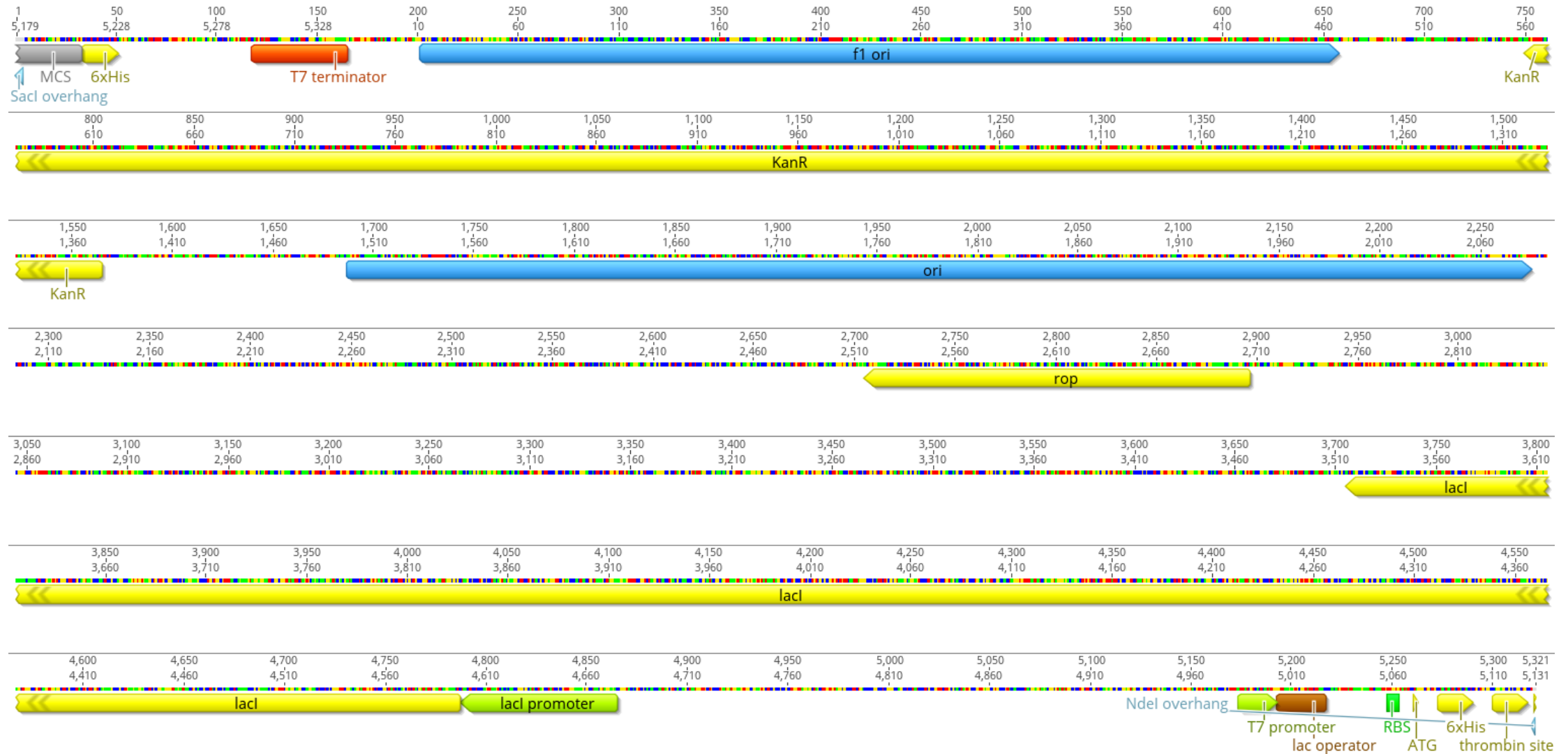
Εικονική Κλωνοποίηση

Προϊόν Πέψης (PCR product με NdeI , SacI)



Εικονική Κλωνοποίηση

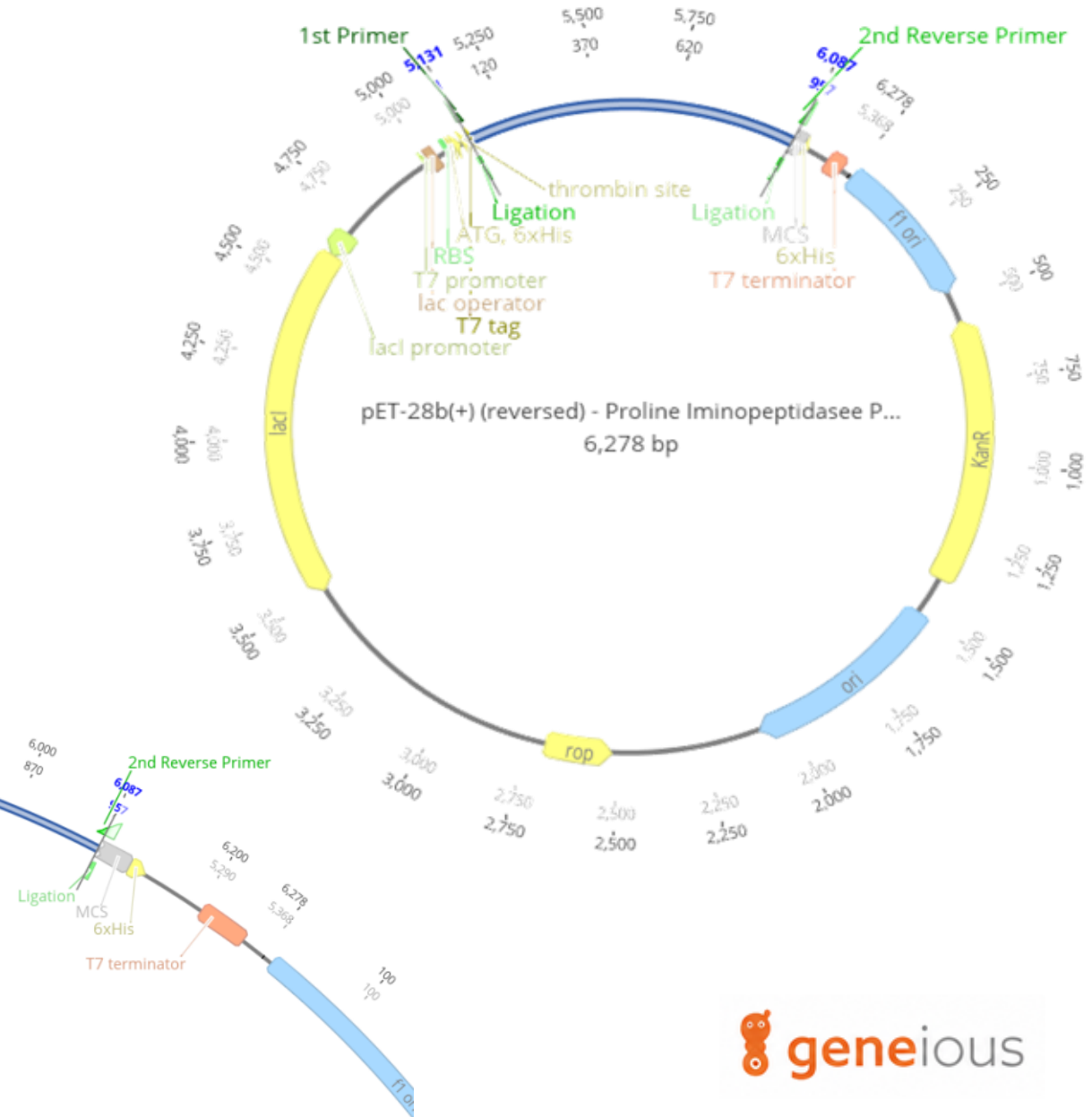
Προϊόν Πέψης (pET28b με NdeI, SacI)

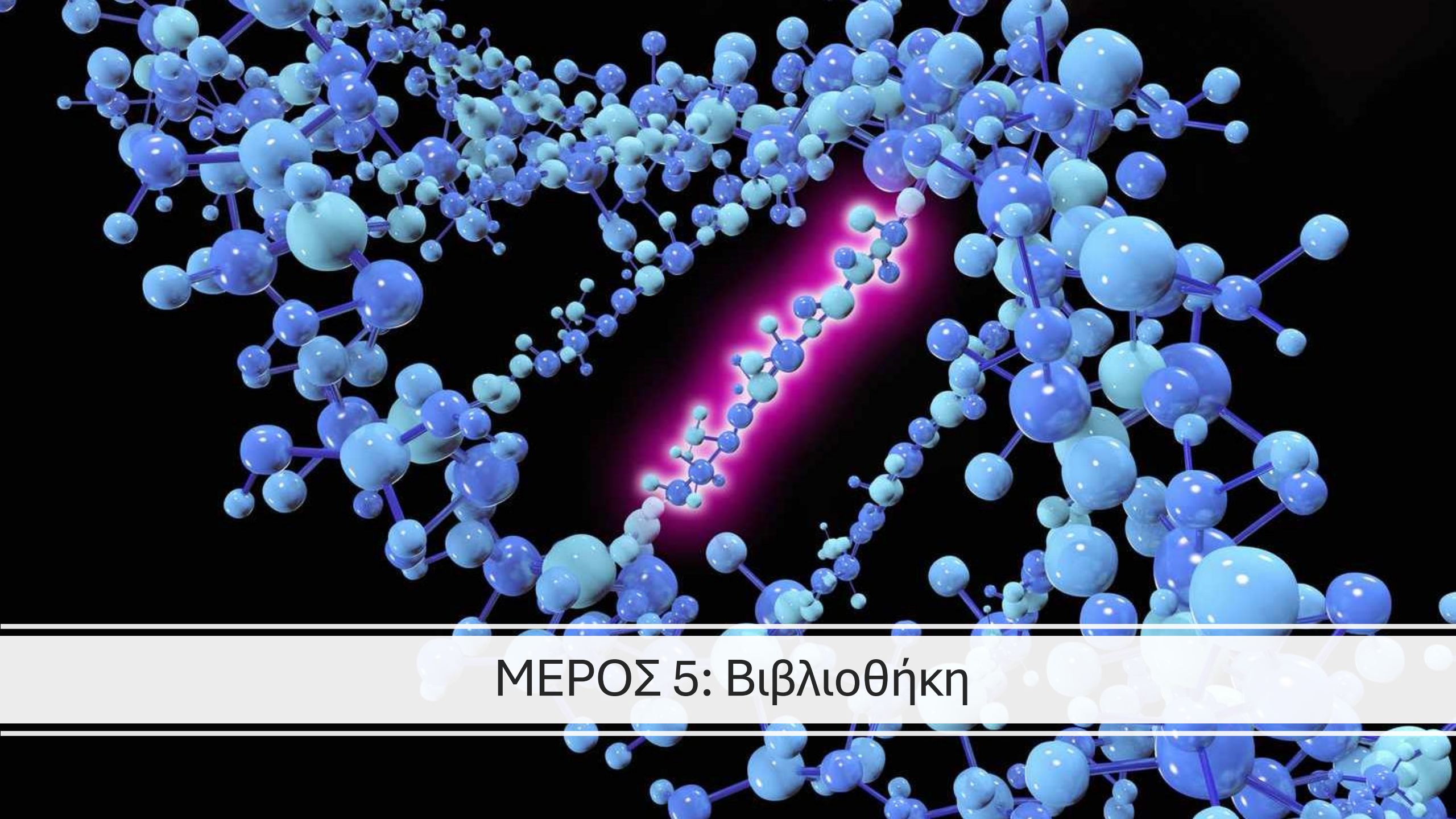


Εικονική Κλωνοποίηση

Ligation

✓ Ελέγχοντας την εισαχθείσα αλληλουχία του γονιδίου στο ανασυνδυασμένο πλασμίδιο διαπιστώθηκε ότι έχει 100% ομοιότητα με την αρχική. Η διαπίστωση αυτή έγινε με το sequence alignment tool του

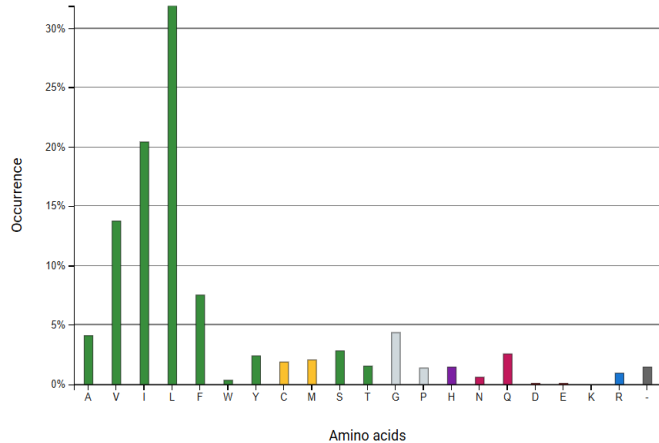
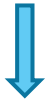




ΜΕΡΟΣ 5: Βιβλιοθήκη

Επιλέχθηκαν από την 3DM τρεις τυχαίες συντηρημένες θέσεις από το 1 έως το 88:

Θέση
15



Επιτρεπτά Αμινοξέα:

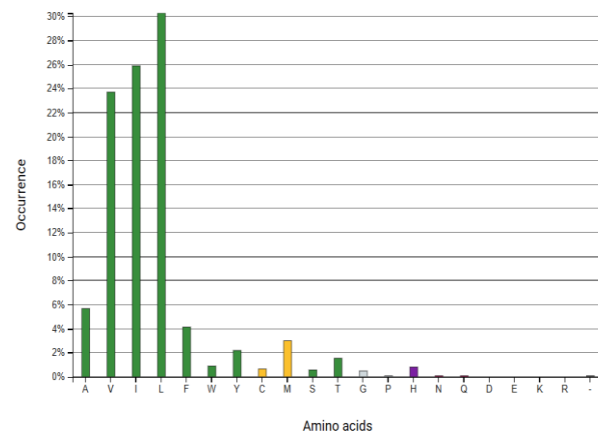
- Βαλίνη (13.72%)
- Ισολευκίνη (20.44%)
- Λευκίνη (31.84%)
- Φαινυλαλανίνη (7.52%)



DTM

Degenerate Κωδικόνια

Θέση
28



Επιτρεπτά Αμινοξέα:

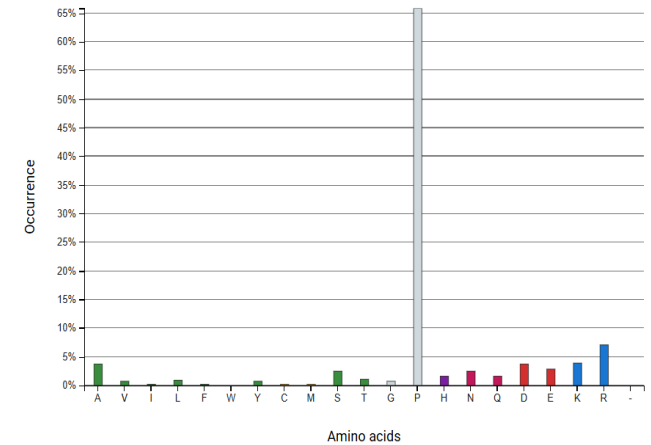
- Αλανίνη (5.64%)
- Βαλίνη (23.71%)
- Ισολευκίνη (25.90%)
- Λευκίνη (30.27%)



DYA

Degenerate Κωδικόνια

Θέση
55



Επιτρεπτά Αμινοξέα:

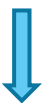
- Προλίνη (65.84%)
- Αργινίνη (7.11%)



CSA

Προβλεφθείσες βιβλιοθήκες για κάθε θέση

Θέση
15

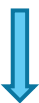


IV. Genetic Code - Codon Distribution																	
SECOND POSITION OF THE CODON																	
		T			C			A			G						
F	T	TTT	Phe [F]	TCT	Ser [S]	TAT	Tyr [Y]	TGT	Cys [C]	T	T						
I	T	TTC	Phe [F]	TCC	Ser [S]	TAC	Tyr [Y]	TGC	Cys [C]	C	H						
R	S	TTA	Leu [L]	TCA	Ser [S]	TAA	Ter [end]	TGA	Ter [end]	A	I						
S		TTG	Leu [L]	TCG	Ser [S]	TAG	Ter [end]	TGG	Trp [W]	G	R						
T		CTT	Leu [L]	CCT	Pro [P]	CAT	His [H]	CGT	Arg [R]	T	D						
C		CTC	Leu [L]	CCC	Pro [P]	CAC	His [H]	CGC	Arg [R]	C							
P	O	CTA	Leu [L]	CCA	Pro [P]	CAA	Gln [Q]	CGA	Arg [R]	A	P						
S		CTG	Leu [L]	CCG	Pro [P]	CAG	Gln [Q]	CGG	Arg [R]	G	O						
I	A	ATT	Ile [I]	ACT	Thr [T]	AAT	Asn [N]	AGT	Ser [S]	T	S						
T		ATC	Ile [I]	ACC	Thr [T]	AAC	Asn [N]	AGC	Ser [S]	C	I						
I		ATA	Ile [I]	ACA	Thr [T]	AAA	Lys [K]	AGA	Arg [R]	A	T						
O	N	ATG	Met [M]	ACG	Thr [T]	AAG	Lys [K]	AGG	Arg [R]	G	I						
		GTT	Val [V]	GCT	Ala [A]	GAT	Asp [D]	GGT	Gly [G]	T	O						
		GTC	Val [V]	GCC	Ala [A]	GAC	Asp [D]	GGC	Gly [G]	C	N						
		GTA	Val [V]	GCA	Ala [A]	GAA	Glu [E]	GGA	Gly [G]	A							
		GTG	Val [V]	GCG	Ala [A]	GAG	Glu [E]	GGG	Gly [G]	G							

DTM

Degenerate Κωδικόνια

Θέση
28



IV. Genetic Code - Codon Distribution																	
SECOND POSITION OF THE CODON																	
		T			C			A			G						
F	T	TTT	Phe [F]	TCT	Ser [S]	TAT	Tyr [Y]	TGT	Cys [C]	T	T						
I	T	TTC	Phe [F]	TCC	Ser [S]	TAC	Tyr [Y]	TGC	Cys [C]	C	H						
R	S	TTA	Leu [L]	TCA	Ser [S]	TAA	Ter [end]	TGA	Ter [end]	A	I						
S		TTG	Leu [L]	TCG	Ser [S]	TAG	Ter [end]	TGG	Trp [W]	G	R						
T		CTT	Leu [L]	CCT	Pro [P]	CAT	His [H]	CGT	Arg [R]	T	D						
C		CTC	Leu [L]	CCC	Pro [P]	CAC	His [H]	CGC	Arg [R]	C							
P	O	CTA	Leu [L]	CCA	Pro [P]	CAA	Gln [Q]	CGA	Arg [R]	A	P						
S		CTG	Leu [L]	CCG	Pro [P]	CAG	Gln [Q]	CGG	Arg [R]	G	O						
I	A	ATT	Ile [I]	ACT	Thr [T]	AAT	Asn [N]	AGT	Ser [S]	T	S						
T		ATC	Ile [I]	ACC	Thr [T]	AAC	Asn [N]	AGC	Ser [S]	C	I						
I		ATA	Ile [I]	ACA	Thr [T]	AAA	Lys [K]	AGA	Arg [R]	A	T						
O	N	ATG	Met [M]	ACG	Thr [T]	AAG	Lys [K]	AGG	Arg [R]	G	I						
		GTT	Val [V]	GCT	Ala [A]	GAT	Asp [D]	GGT	Gly [G]	T	O						
		GTC	Val [V]	GCC	Ala [A]	GAC	Asp [D]	GGC	Gly [G]	C	N						
		GTA	Val [V]	GCA	Ala [A]	GAA	Glu [E]	GGA	Gly [G]	A							
		GTG	Val [V]	GCG	Ala [A]	GAG	Glu [E]	GGG	Gly [G]	G							

DYA

Degenerate Κωδικόνια

Θέση
55



IV. Genetic Code - Codon Distribution																	
SECOND POSITION OF THE CODON																	
		T			C			A			G						
F	T	TTT	Phe [F]	TCT	Ser [S]	TAT	Tyr [Y]	TGT	Cys [C]	T	T						
I	T	TTC	Phe [F]	TCC	Ser [S]	TAC	Tyr [Y]	TGC	Cys [C]	C	H						
R	S	TTA	Leu [L]	TCA	Ser [S]	TAA	Ter [end]	TGA	Ter [end]	A	I						
S		TTG	Leu [L]	TCG	Ser [S]	TAG	Ter [end]	TGG	Trp [W]	G	R						
T		CTT	Leu [L]	CCT	Pro [P]	CAT	His [H]	CGT	Arg [R]	T	D						
C		CTC	Leu [L]	CCC	Pro [P]	CAC	His [H]	CGC	Arg [R]	C							
P	O	CTA	Leu [L]	CCA	Pro [P]	CAA	Gln [Q]	CGA	Arg [R]	A	P						
S		CTG	Leu [L]	CCG	Pro [P]	CAG	Gln [Q]	CGG	Arg [R]	G	O						
I	A	ATT	Ile [I]	ACT	Thr [T]	AAT	Asn [N]	AGT	Ser [S]	T	S						
T		ATC	Ile [I]	ACC	Thr [T]	AAC	Asn [N]	AGC	Ser [S]	C	I						
I		ATA	Ile [I]	ACA	Thr [T]	AAA	Lys [K]	AGA	Arg [R]	A	T						
O	N	ATG	Met [M]	ACG	Thr [T]	AAG	Lys [K]	AGG	Arg [R]	G	I						
		GTT	Val [V]	GCT	Ala [A]	GAT	Asp [D]	GGT	Gly [G]	T	O						
		GTC	Val [V]	GCC	Ala [A]	GAC	Asp [D]	GGC	Gly [G]	C	N						
		GTA	Val [V]	GCA	Ala [A]	GAA	Glu [E]	GGA	Gly [G]	A							
		GTG	Val [V]	GCG	Ala [A]	GAG	Glu [E]	GGG	Gly [G]	G							

CSA

Βιβλιοθήκη - CASTER

Amino acids [AA] encoded in each position					
	1	2	3	4	5
Ala [A]	0	1	0	0	0
Arg [R]	0	0	1	0	0
Asn [N]	0	0	0	0	0
Asp [D]	0	0	0	0	0
Cys [C]	0	0	0	0	0
Gln [Q]	0	0	0	0	0
Glu [E]	0	0	0	0	0
Gly [G]	0	0	0	0	0
His [H]	0	0	0	0	0
Ile [I]	2	1	0	0	0
Leu [L]	1	1	0	0	0
Lys [K]	0	0	0	0	0
Met [M]	0	0	0	0	0
Phe [F]	1	0	0	0	0
Pro [P]	0	0	1	0	0
Ser [S]	0	1	0	0	0
Thr [T]	0	1	0	0	0
Trp [W]	0	0	0	0	0
Tyr [Y]	0	0	0	0	0
Val [V]	2	1	0	0	0
Stop	0	0	0	0	0
Codons	6	6	2	0	0
AA	4	6	2	0	0

Amino acids [AA] encoded in each position					
	1	2	3	4	5
Ala [A]	2	2	2	0	0
Arg [R]	3	3	3	0	0
Asn [N]	1	1	1	0	0
Asp [D]	1	1	1	0	0
Cys [C]	1	1	1	0	0
Gln [Q]	1	1	1	0	0
Glu [E]	1	1	1	0	0
Gly [G]	2	2	2	0	0
His [H]	1	1	1	0	0
Ile [I]	1	1	1	0	0
Leu [L]	3	3	3	0	0
Lys [K]	1	1	1	0	0
Met [M]	1	1	1	0	0
Phe [F]	1	1	1	0	0
Pro [P]	2	2	2	0	0
Ser [S]	3	3	3	0	0
Thr [T]	2	2	2	0	0
Trp [W]	1	1	1	0	0
Tyr [Y]	1	1	1	0	0
Val [V]	2	2	2	0	0
Stop	1	1	1	0	0
Codons	32	32	32	0	0
AA	20	20	20	0	0

Amino acids [AA] encoded in each position					
	1	2	3	4	5
Ala [A]	0	0	0	0	0
Arg [R]	1	1	1	0	0
Asn [N]	1	1	1	0	0
Asp [D]	1	1	1	0	0
Cys [C]	1	1	1	0	0
Gln [Q]	0	0	0	0	0
Glu [E]	0	0	0	0	0
Gly [G]	1	1	1	0	0
His [H]	1	1	1	0	0
Ile [I]	1	1	1	0	0
Leu [L]	1	1	1	0	0
Lys [K]	0	0	0	0	0
Met [M]	0	0	0	0	0
Phe [F]	1	1	1	0	0
Pro [P]	0	0	0	0	0
Ser [S]	1	1	1	0	0
Thr [T]	0	0	0	0	0
Trp [W]	0	0	0	0	0
Tyr [Y]	1	1	1	0	0
Val [V]	1	1	1	0	0
Stop	0	0	0	0	0
Codons	12	12	12	0	0
AA	12	12	12	0	0

Συνδυαστική Βιβλιοθήκη

Degenerate κωδικόνια :

DTM, DYA, CSA

Codons: **72**

Colonies: **214**

NNK

Codons: **32768**

Colonies: **98163**

NDT

Codons: **1728**

Colonies: **5175**

Βιβλιοθήκη – Συνδυαστική Βιβλιοθήκη vs NNK

Συνδυαστική Βιβλιοθήκη

- Σημαντικά μικρότερη και πιο εύκολα διαχειρίσιμη πειραματικά
- Μέγεθος: **214 αποικίες**
- Δεν περιλαμβάνει κωδικόνιο έναρξης ή λήξης
- Αποφεύγονται ανεπιθύμητα αποτελέσματα, όπως πρόωρος τερματισμός μετάφρασης ή λανθασμένη έκφραση γονιδίου
- Περιλαμβάνει κυρίως τα «επιτρεπτά» αμινοξέα με μικρό ποσοστό «μη επιτρεπτών»
- Πιο στοχευμένος και αποδοτικός σχεδιασμός για το σύστημα

Βιβλιοθήκη NNK

- Πολύ μεγαλύτερη και πιο απαιτητική στη διαχείριση
- Μέγεθος: **98.163 αποικίες**
- Περιλαμβάνει κωδικόνιο έναρξης και κωδικόνιο λήξης
- Πιθανότητα ανεπιθύμητων αποτελεσμάτων λόγω παρουσίας start/stop κωδικονίων
- Καλύπτει και τα 20 φυσικά αμινοξέα, αλλά με αυξημένο αριθμό «μη επιτρεπτών» για το συγκεκριμένο σύστημα
- Μεγάλη ποικιλότητα, αλλά μικρότερη εξειδίκευση

Βιβλιοθήκη – Συνδυαστική Βιβλιοθήκη vs NDT

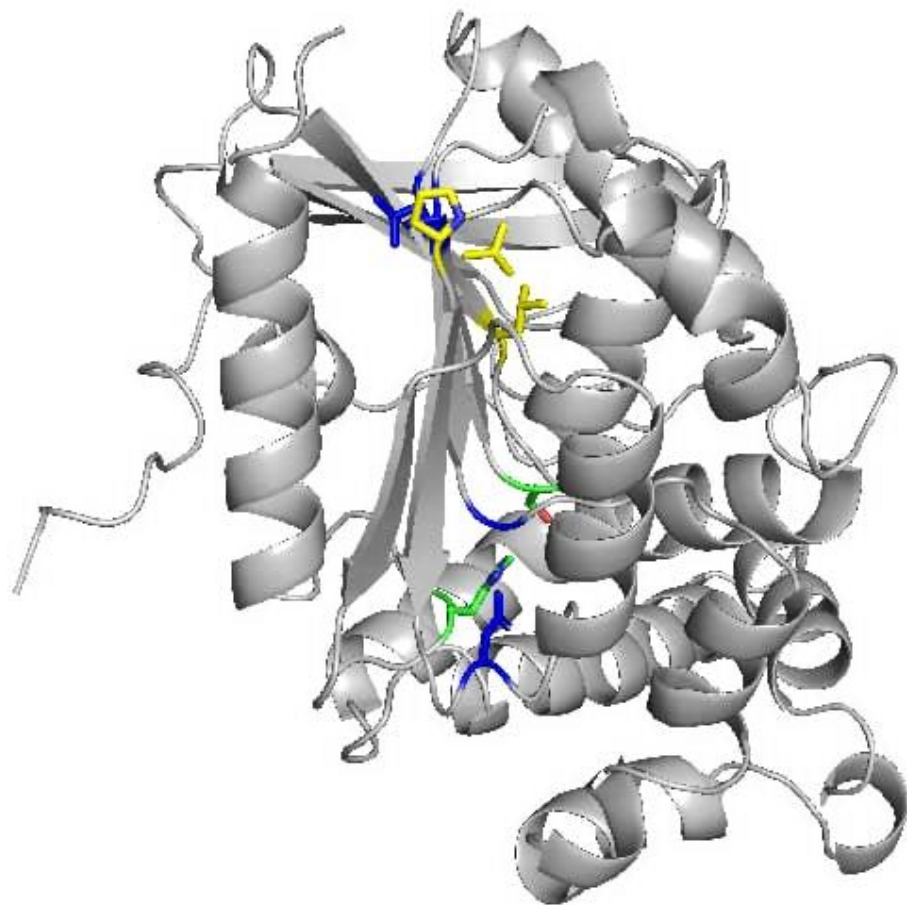
Συνδυαστική Βιβλιοθήκη

- Σημαντικά μικρότερη και πιο εύκολα διαχειρίσιμη πειραματικά
- Μέγεθος: **214 αποικίες**
- Δεν περιλαμβάνει κωδικόνιο έναρξης ή λήξης
- Εξασφαλίζει καλύτερα αποτελέσματα και μειωμένο κίνδυνο ανεπιθύμητων εκφράσεων
- Περιλαμβάνει κυρίως όλα τα «επιτρεπτά» αμινοξέα με μικρό ποσοστό «μη επιτρεπτών»
- Πιο στοχευμένος και αποδοτικός σχεδιασμός για το συγκεκριμένο σύστημα

Βιβλιοθήκη NDT

- Μεγαλύτερη από τη συνδυαστική, αλλά σαφώς μικρότερη από τη NNK
- Μέγεθος: **5.175 αποικίες**
- Δεν περιλαμβάνει κωδικόνιο έναρξης ή λήξης
- Επίσης εξασφαλίζει καλύτερα αποτελέσματα χωρίς start/stop κωδικόνια
- Καλύπτει **12 αμινοξέα**, αλλά περιλαμβάνει περισσότερα «μη επιτρεπτά» σε σχέση με τη συνδυαστική
- Δεν περιλαμβάνει τα επιτρεπτά αλανίνη (Ala) και προλίνη (Pro)
- Καλή ισορροπία ποικιλότητας και μεγέθους, αλλά λιγότερο εξειδικευμένη από τη συνδυαστική

Τελικό μοντέλο ομολογίας



Model 1 από SWISS-MODEL με την εξής παρουσίαση:

cartoon, σε γκρι χρώμα, με μόνο τις πλευρικές ομάδες των εξής αμινοξέων εμφανισμένες ως *sticks*.

Χρώματα : καταλυτικά αμινοξέα πράσινα, τα αμινοξέα του *correlation network* με μπλε , τα αμινοξέα της βιβλιοθήκης ως κίτρινα. Δεν υπάρχουν *ramachandran outliers*. Το αμινοξύ στην θέση 64 (κατά 3DM) και 267 (κατά αμινοξική αλληλουχία) αποτελεί και αμινοξύ της καταλυτικής τριάδας και αμινοξύ στα *correlation networks* και έχει χρωματιστεί μόνο με μπλέ χρώμα.

[1] Paul D. Carr and David L. Ollis. *α/β Hydrolase Fold: An Update*. Protein & peptide letters/Protein and peptide letters 2009, 16 , 1137–1148.
<https://doi.org/10.2174/092986609789071298>

[2] Tabea L. Bauer ,Patrick C. F. Buchholz and Jurgen Pleiss . *The modular structure of a/b -hydrolases*, 2019. <https://doi.org/10.1111/febs.15071>

[3] Mats Holmquist. *Alpha/Beta-Hydrolase Fold Enzymes: Structures, Functions and Mechanisms*, Current Protein and Peptide Science, 2000, 1, 209-235.
<https://doi.org/10.2174/1389203003381405>

[4] [Fatih Ozhelvaci](#) 1, [Kamil Steczkiewicz](#). *α/β Hydrolases: Toward Unraveling Entangled Classification*, 2024. <https://doi.org/10.1002/prot.26776>

Βιβλιογραφία - Links

[5] 3DM Bio- Prodict , https://app3dm.bio-prodict.nl/system/a-b-hydro_newprot_2013/alignment/1/subset/1/numberingscheme/-1/proteins/protein/G2LE64/info

[6] NCBI, National Center for Biotechnology Information , <https://www.ncbi.nlm.nih.gov/>.

[7] UniProt , <https://www.uniprot.org/uniprotkb/G2LE64/entry>

[8] KEGG , https://www.genome.jp/dbget-bin/www_bget?ctm:Cabther_A0526

[9] SWISS-MODEL , <https://swissmodel.expasy.org/>

[10] D-i-TASSER , <https://www.aideepmed.com/D-I-TASSER/>

[11] AlphaFold 3 , <https://alphafoldserver.com/welcome>

[12] MolProbity , <http://molprobity.biochem.duke.edu/>

[13] VectorBuilder , Codon Optimization , <https://en.vectorbuilder.com/tool/sequence-alignment/e6ae77cd-37a6-4b14-861a-3c6588691594.html>

[14] Oligonucleotide Properties Calculator , <https://www.biosyn.com/Gizmo/Tools/Oligo/Oligonucleotide%20Properties%20Calculator.htm>

[15] NEBcloner - version 1.17.0 , <https://nebcloner.neb.com/#!/redigest>

[16] Geneious Platform

[17] Εργαλείο CASTER



Thank you for your attention!