

Ургамлын лейцин баялаг давталт агуулсан уургийн Z_{10} -хеликс – β -эргэлтээс бүрдэх супер хоёрдогч бүтэц ба домен бүтцийн тодорхойлолт

Дашдаваа Батхишиг^{1,2}, Норио Мацushima^{3,4}, Пүрэвжав Энхбаяр^{2,*}

¹ Монгол улсын боловсролын их сургууль, Математик, байгалийн ухааны сургууль, Физикийн тэнхим,

² МУИС, ХШУИС, Мэдээлэл, компьютерын ухааны тэнхим, Биоинформатик ба систем биологийн лаборатори,

³ Тандем давталтын хүрээлэн

⁴ Саппорогийн анагаах ухааны их сургууль

Эукариот, бактери, архей болон вирусээс тодорхойлсон 430,000 гаруй уураг Лейцин Баялаг Давталт (ЛБД) агуулж байгаа нь тогтоогдоод байна. ЛБД-ууд нь 20-30 амин хүчлийн урттай бөгөөд зэрэгцсэн дараалалд тохиолддог. ЛБД агуулсан уураг нь тах, супер хеликс, эсвэл призм хэлбэрийн бүтэцтэй байх ба дотоод хүнхэр гадаргуудаа параллель β -утаслаг, гадаад гүдгэр гадаргуудаа ихэвчлэн хеликс бүтэц агуулна. Тухайн нэг ЛБД нь $LxxLxLxxNxL$ консенсус дараалалтай сайн хадгалагдсан хэсэг (СХХ) болон хувьсах хэсэг (ХХ)-ээс бүрдэнэ. ЛБД-ын найман анги байна. Ургамлын ЛБД-ын хувьсах хэсэг нь Z_{10} -хеликс болон β -эргэлтээр тодорхойлогдох ба консенсус дараалал нь 13 амин хүчлийн урттай ($xGxLPxxLxxLxx$) байна. Ургамлын дархлааны системийн болон сигнал дамжуулах уургуудад ЛБД агуулагддаг. Ургамлын ЛБД агуулсан уургийн бүтцийн онцлогийг ойлгохын тулд бид DSSP-PPII, XTLSSTR ба STRIDE гэсэн гурван программ ашиглан хоёрдогч бүтцийн онолт хийсэн. Уургийн 3D бүтэц дэх альфа-нүүрстөрөгч (Ca) атомын координатыг ашиглан HELFIT программаар хеликсийн (алхам P , радиус r , нэг эргэлтэд оногдох амин хүчлийн тоо n , эргэлтийн чиглэл, тэнхлэгийн нэгж вектор) параметруудийг тодорхойлсон. ЛБД-ын хувьсах хэсэг нь N-терминал талдаа 5-аас 7 амин хүчлийн урттай баруун эргэлттэй Z_{10} -хеликс, C-терминал талдаа I төрлийн бета-эргэлтээс бүрдэж байна. Иймд ургамлын ЛБД-ын хувьсах хэсэг нь Z_{10} -хеликс болон β -эргэлт агуулсан супер хоёрдогч бүтцээр тодорхойлогдож байна. HELFIT-ийн үр дүнгээс харвал I төрлийн β -эргэлт нь баруун (+1) эргэлттэй хеликс байна. HELFIT программаар Z_{10} -хеликс (G), β -эргэлт (B), ЛБД домен (A) тус бүрийн хеликс тэнхлэгийн нэгж векторыг тодорхойлсон. Эдгээр гурван нэгж вектороор тодорхойлогдох Ω_1 , Ω_2 , ба Ω_3 бүтцийн параметр супер хоёрдогч бүтэц болон ЛБД домены бүтцийн онцлогийг тодорхойлж байна.

Түлхүүр үг: ургамлын лейцин баялаг давталт, Z_{10} -хеликс, I төрлийн β -эргэлт, супер хоёрдогч бүтэц, бүтцийн параметр.

I. ОРШИЛ

Лейцин Баялаг Давталт (ЛБД)-ууд нь 20-30 амин хүчлийн урттай бөгөөд зэрэгцсэн дараалалд тохиолддог. ЛБД-ийн тоо 2-оос 97 хүртэл байдаг. InterPro өгөгдлийн санд давхардсан тоогоор 430,000 гаруй ЛБД агуулсан уураг бүртгэгдсэн ба вирусээс эукариот хүртлэх бүх организмуудад илэрсэн [1]. Тухайн ЛБД нь сайн хадгалагдсан хэсэг (СХХ) болон хувьсах хэсэгт (ХХ)-ээс бүрдэнэ. ЛБД-ын найман анги байх ба Мацushima нар саяхан ЛБД-ын 23 төрлийг санал болгосон [2]. Энэ ангилал нь ЛБД-ын хувьсах хэсгийн хоёрдогч бүтцийн ялгаатай байдалд үндэслэгдсэн. Найман анги нь рибонуклеазе ингибитор төст (RI-like), цистейн агуулсан (ЦА), SDS22-төст, IRREKO,

бактерийн, ургамлын, ердийн, трепонема ЛБД юм. СХХ нь 11 эсвэл 12 амин хүчлийн урттай бөгөөд $LxxLxLxx(N/C)(x/-)L$ үүнд “L” нь лейцин, изолейцин, валин, эсвэл фенилаланиныг, “N” нь аспарагин, треонин, серин, эсвэл цистеинийг, “C” нь цистеин, серин эсвэл аспарагиныг, “x” нь ямар нэгэн амин хүчлийг тэмдэглэх ба “-” нь тухайн байрлал дээр амин хүчил арчигдсаныг заана. Богино β -утаслаг бүтэц нь консенсус дарааллын доогуур зураастай 3-5 байрлал дахь амин хүчлүүдээс тогтдог. ЛБД агуулсан уургууд нь ерөнхийдөө морины тах, баруун (+1) эсвэл зүүн (-1) гарын эргэлттэй супер хеликс, эсвэл призм хэлбэртэй байдаг. ЛБД уурагт агуулагдах лейцин, валин, изолейцин, фенилаланин, гэх мэт гидрофоб амин

* Electronic address: enkhbayar.p@seas.num.edu.mn

хүчлүүд уургийн гидрофоб цөмийг үүсгэнэ. ЛБД уургийн N ба C терминал (төгсгөл)-ыг гидрофоб амин хүчлүүд агуулсан дарааллууд тагладаг [3-6].

ЛБД-ын ангиуд нь хувьсах хэсгийн α -хеликс, 3_{10} -хеликс, РРП-хеликс ба дараалсан хоёр эсвэл гурван β -эргэлт гэсэн хоёрдогч бүтцээрээ ялгаатай байдаг [2,7]. Эдгээр ЛБД агуулсан уургийн гүдгэр талын хоёрдогч бүтцүүд нь хүнхэр тал дахь β -ялтас бүтэцтэй хоёр гогцоогоор холбогдоно. Нэг гогцоо нь СХХ-ийн С-терминалыг ХХ-ийн N-терминалтай холбох бөгөөд түүнийг “өгсөх гогцоо” гэдэг. Нөгөө нь ХХ-ийн С-терминалыг дараагийн ЛБД-ийн СХХ-ийн N-терминалтай холбох бөгөөд үүнийг “уруудах гогцоо” гэнэ [2]. ЛБД домеин бүр нь хоорондоо эсрэг гүдгэр ба хүнхэр гадарга, өгсөх ба уруудах гадарга агуулдаг. ЛБД домеин нь уураг, даавар эсвэл лигандтай (нуклеин хүчил, липид, липополисахарид, ургамлын стероид гормон гэх мэт) шууд харилцан үйлчлэлд ордог. ЛБД домеинууд нь бүтцийн хувьд олон төрлийн уургуудтай эсвэл өөр өөр лигандуудтай холбогдох боломжтой байдаг. Учир нь ЛБД агуулсан уургийн нэг амин хүчилд харгалзах гадаргын талбай өндөр буюу глобулар-биш уураг юм [8].

ЛБД агуулсан уургууд нь ургамлын дархлааны хариу үйлдэл болон хөхтөн амьтдын дархлааны төрөлхийн хариу үйлдэл оролцдог. Түүнчлэн тэдгээр нь олон эст организмын эсийн програмчлагдсан үхэл, эсийн доторх биомолекулыг задлах систем, сигнал дамжуулал, мРНХ-ийн тээвэрлэлт, мэдрэлийн эсийн хөгжил гэх мэт олон төрлийн процессод оролцдог. Ургамлын ЛБД агуулсан уургуудын ихэнх нь катализатор буюу фермент, бусад нь рецептор төст уургууд байдаг ба эсийн гэмтлийн дохио өсгөгчийн үүрэг гүйцэтгэхээс гадна амьд организмуудын симбиоз харилцаа болон хөгжлийн процессуудад оролцдог [9-11].

Уургийн дараалал ба бүтцийн хоорондох хамаарал нь уургийн бүтцийн зарчмыг ойлгоход чухал ач холбогдолтой [3,12].

Бид өмнөх ажлаараа Бактерийн ЛБД-ын хувьсах хэсэгт РРП-хеликс ба I төрлийн β -эргэлтээс бүрдэх супер хоёрдогч бүтэц байгааг тогтоосон [9]. Бактерийн ЛБД бүтцийн хувьд РРП-хеликс (**P**), β -эргэлт (**B**) ба ЛБД домеины

үүсгэн супер хеликс (**A**) тэнхлэгүүд болох гурван нэгж векторыг тодорхойлсон бөгөөд эдгээрээр тодорхойлогдох гурван бүтцийн параметрийг дэвшүүлсэн [13].

Энэ судалгааны зорилго нь Ургамлын ЛБД агуулсан бүтцийн онцлогийг тодорхойлох явдал юм. Бид хоёрдогч бүтцийн оноолт гүйцэтгэхэд түгээмэл хэрэглэгддэг DSSP-РРП, XTLSSTR ба STRIDE программ болон HELFIT программыг хослуулан хэрэглэсэн.

Бидний дэвшүүлсэн бүтцийн гурван параметр нь 3_{10} -хеликс болон β -эргэлтийн тэнхлэг хоорондох өнцөг (Ω_1), 3_{10} -хеликс болон ЛБД домеины тэнхлэг хоорондох өнцөг (Ω_2), ЛБД домеин болон $\mathbf{G} \times \mathbf{B}$ векторын хоорондох өнцөг (Ω_3) юм. Эдгээр гурван өнцөг нь супер хоёрдогч бүтэц болон ЛБД домеины бүтцийн онцлогийг тодорхойлж чадаж байна.

II. ӨГӨГДӨЛ БА АРГА ЗҮЙ

A. Ургамлын ЛБД уураг

Уургийн бүтцийн өгөгдлийн сангаас (PDB) ургамлын ЛБД агуулсан бүх уургийн 3-хэмжээст бүтцийн өгөгдлүүдийг судалгаанд хэрэглэсэн. Тэдгээр уургийн өгөгдлүүдэд SIAS онлайн түүл ашиглан амин хүчлийн дарааллын харьцуулалт хийсэн [14]. Дараалал нь 100 хувь адил байх тохиолдолд хамгийн өндөр нарийвчлалтай тодорхойлогдсон бүтцийн өгөгдлийг тооцоололдоо ашигласан. 18 уургийн 3D бүтцийн нарийвчлал ≤ 3.06 Å байсан ба эдгээрийн 19 гинжний дарааллын ялгааны хувь нь 76.96 хувиас бага, дарааллууд дунджаар 93.18 %-ийн ялгаатай байсан.

Эдгээр уургийн амин хүчлийн дараалалд LRRpred алгоритмаар лейцин баялаг давталтыг тогтоож консенсус дараалалд үндэслэн Ургамлын ЛБД-ыг илүү найдвартай таамагласан [15]. Иймд ургамалд PXY z/TDR, RGFR1, AtPSKR, PEPR1, HAESA, BRI1, BRL1, ERL1, DcPSKR1, BIR2, AtSERK1, AtSERK2, AtSERK3, AtPRK3, AtPRK6RR, OsSERK2, PGIP, AtTMK1 гэсэн ЛБД агуулсан уургууд байна (Хүснэгт 1) [16-29].

B. Хоёрдогч бүтцийн оноолт

Хоёрдогч бүтцийн оноолтыг PolyprOnline вэб интерфэйсийн DSSP-РРП ба XTLSSTR [30] болон STRIDE-ийн вэб интерфэйсийг [31] ашиглан гүйцэтгэсэн. α -хеликс, 3_{10} -хеликс, β -

эргэлтийн хоёрдогч бүтцийн оноолтыг DSSP-PPII [32,33], XTLSSTR [34] ба STRIDE [35] програмуудаар хийв.

DSSP-PPII программын оноолт нь устөрөгчийн холбоосын энерги ба давталтыг үндэслэн хоёрдогч бүтцийг тодорхойлдог. STRIDE нь устөрөгчийн холбоосын энерги болон φ, ψ өнцгийн шалгуурыг хослуулан хоёрдогч бүтцийг оноодог [33]. Тус хоёр программын хувьд 3₁₀-хеликсийн хамгийн бага урт нь дараалсан хоёр устөрөгчийн холбоосоор ($i \leftarrow i+3$)

тодорхойлогддог. 3₁₀-хеликс нь хамгийн багадаа 3 дүрмээр 5 амин хүчлийн урттай бөгөөд эхний амин хүчлийн C=O бүлэг болон сүүлийн амин хүчлийн N-H бүлэг устөрөгчийн холбоост оролцдог [36]. XTLSSTR нь хоёр өнцөг ($\text{dison3}=\text{O}(i)\text{N}(i+3)$, $\text{dison4}=\text{O}(i)\text{N}(i+4)$, $\text{discn3}=\text{C}(i)\text{N}(i+3)$) хоёрдогч бүтцийг оноодог [34]. Иймд 3₁₀-хеликс нь хамгийн багадаа хоёр амин хүчлийн урттай байна.

Хүснэгт 1. Ургамлын ЛБД агуулсан уургуудын тодорхойлогдсон буй бүтцүүд.

№	Уураг	PDBid_гинж	Нарийвчлал (Å)	LRR (XX)	XX=13 а.х	Супер хоёрдогч бүтцийн тоо		
						DSSP-PPII	STRIDE	XTLSSTR
1	PXY/TDR	5JFK_A	2.65	22	21	18	18	13
2	RGFR1	5HZ0_B	2.56	25	21	23	23	16
3	AtPSKR	4Z63_A	2.51	21	16	14	14	10
4	PEPR1	5GR8_A	2.59	27	25	20	20	19
5	HAESA	5IXO_A	1.74	22	17	15	14	11
6	BRI1	3RGZ	2.28	25	13	13	13	15
7	BRL1	4J0M_A	2.50	24	12	14	14	13
8	FLS2	4MN8_A	3.06	30	27	24	25	22
9	ERL1	5XJO_A	2.63	20	19	16	19	12
10	DcPSKR1	4Z5W_A	2.20	21	16	13	12	11
11	AtSERK1	6FG8_A	1.25	5	5	4	4	4
12	BIR2	6FG7_A	1.90	5	4	4	4	3
13	AtSERK2	6G3W_A	2.20	5	5	4	4	4
14	AtPRK6PR	5Y9W_A	1.85	6	3	4	4	3
15	AtSERK3	4MN8_B	3.06	5	5	4	4	4
16	AtPRK3	5WLS_A	2.50	7	2	2	3	2
17	PsSERK2	4Q3I_A	2.35	6	5	4	4	7
18	PGIP	1OGQ	1.70	9	6	7	7	5
19	AtTMK1	4HQ1_A	1.55	14	5	7	7	5
				299	227	210	213	179

DSSP-PPII, STRIDE, XTLSSTR гурван программын β-эргэлтийн тодорхойлолт нь хоорондоо ялгаатай байдаг. DSSP нь хамгийн бага урттай 3₁₀-хеликсийн дараалсан хоёр устөрөгчийн холбоосын хоёр дахь устөрөгчийн холбоосыг хасаж дөрвөн амин хүчлийн урттай (i -ээс $i+3$ хүртэл) β-эргэлтийг оноодог [33]. STRIDE нь β-эргэлтийн $i+1$ болон $i+2$ дугаар амин хүчлийн φ, ψ өнцгийн утгаар тодорхойлдог [35]. Харин XTLSSTR нь $\text{O}(i)$ -өөс $\text{N}(i+3)$ хүртэлх зай 3.5 Å-ээс бага бол устөрөгчийн холбоостой β-эргэлт (T), $\text{dison3} < 4.7$ Å бол устөрөгчийн холбоосгүй β-эргэлт (N) гэж онооно [34].

С. Хеликсийн параметр

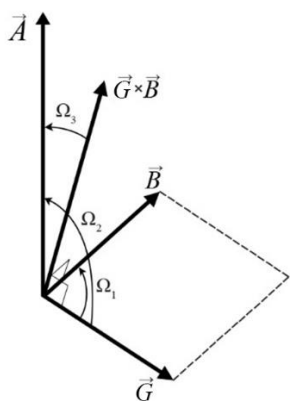
Бид өгөгдлийн цэгүүдэд хамгийн бага квадратын аргаар хеликсийн фиттинг гүйцэтгэдэг HELFIT программаар хеликсийн (ЛБД домеин, 3₁₀-хеликс, β-эргэлт) параметруудийг тодорхойлсон. Үүнд: хеликсийн алхам (P), нэг эргэлтэнд харгалзах амин хүчлийн тоо (n), радиус (R), эргэлтийн чиглэл (Handedness), тэнхлэгийн нэгж вектор (i, j, k), нэг амин хүчилд харгалзах хеликсийн тэнхлэгийн дагуух шилжилт ($\Delta z = P/N$) ба Вороной эзлэхүүнийг $V_c = \pi R^2 \Delta z$ бодно. Түүнчлэн HELFIT нь RMSD бодох бөгөөд d_i нь өгөгдлийн цэгээс хеликсийн мурий хүртэлх хамгийн бага зай.

$$\text{RMSD} = \left[\frac{(\text{the minimum of } \sum d_i)}{n} \right]^{1/2} \quad (1)$$

Энд $p = \text{RMSD}/(n-1)^{1/2}$ өгөгдлийн цэгийн тоо эсвэл хеликсийн уртаас хамааралгүй хеликсийн бүтцийн төлөвийг тогтооно [37]. Зөв буюу алдаа багатай 3₁₀-хеликсийн хувьд шалгуур $p \leq 0.10 \text{ \AA}$ байна [38]. HELFIT нь хамгийн багадаа өгөгдлийн дөрвөн цэг шаардлагатай бөгөөд цэг нь амин хүчил бүрийн α -нүүрстөрөгч атомын (C α) координат байна [39]. ЛБД домеины хеликсийн параметруудийг тодорхойлохдоо СХХ бүрийн β -утаслагийн 4-р байрлал дахь лейцин эсвэл изолейцин амин хүчлийн C α координатыг хэрэглэсэн. Тухайн ЛБД домеины давталтын тоо нь параллель β -ялгаст оролцож буй β -утаслагийн тоогоор тодорхойлно [40]. Энэ тодорхойлолт ёсоор зарим тохиолдолд эхний ЛБД-ад малгай бүтэц орж болно. β -эргэлт нь 4 амин хүчлээс бүрдэх ба түүний амин хүчил бүрийн C α координатаар хеликсийн параметруудийг тооцоолсон.

D. Бүтцийн параметр

Бид ургамлын ЛБД домеин (**A**), 3₁₀-хеликс (**G**) болон β -эргэлт (**B**) тус бүрийн хеликс тэнхлэгийн нэгж векторыг HELFIT-ээр тодорхойлж дараах гурван бүтцийн параметрийг дэвшүүлсэн (Зураг 1).



Зураг 1. Ω_1, Ω_2 , ба Ω_3 өнцгийн тодорхойлолт.

Ω_1 нь 3₁₀-хеликс (**G**) болон β -эргэлтийн (**B**) хеликсийн тэнхлэгийн нэгж векторын хоорондох өнцөг (томъёо 2). Ω_2 нь 3₁₀-хеликс(**G**) ба ЛБД домеины (**A**) хеликсийн нэгж тэнхлэг хоорондох өнцөг (томъёо 3). Ω_3 нь нормаль вектор (**G** × **B**) ба ЛБД домеины хеликсийн тэнхлэгийн нэгж (**A**) векторын хоорондох өнцөг

(томъёо 4). Эдгээр Ω_1, Ω_2 , ба Ω_3 гурван өнцгүүд дараах томъёогоор илэрхийлэгдэнэ.

$$\mathbf{G} \cdot \mathbf{B} = |\mathbf{G}| |\mathbf{B}| \cos \Omega_1 \quad (2)$$

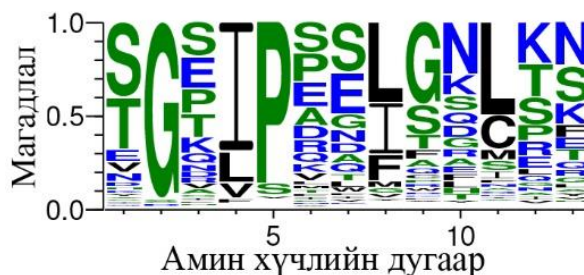
$$\mathbf{G} \cdot \mathbf{A} = |\mathbf{G}| |\mathbf{A}| \cos \Omega_2 \quad (3)$$

$$(\mathbf{G} \times \mathbf{B}) \cdot \mathbf{A} = |\mathbf{G} \times \mathbf{B}| |\mathbf{A}| \cos \Omega_3 \quad (4)$$

III. ҮР ДҮН БА ХЭЛЭЛЦҮҮЛЭГ

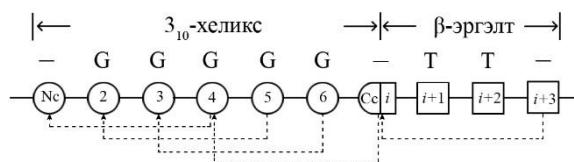
A. Ургамлын ЛБД-ын хувьсах хэсэг

Бид 3D бүтэц нь тодорхойлогдсон 18 уургийн өгөгдлөөс 299 ургамлын ЛБД-ийг тодорхойлсон. Эдгээр давталт нь 24 амин хүчлийн урттай бөгөөд СХХ-ийн урт 11 (LxxLxLxxNxL) бол ХХ-ийн урт нь 13 амин хүчлээс тогтох (xGxLPxxLxxLxx) ЛБД 227 байна (Хүснэгт 1). Ургамлын ЛБД-ын 13 амин хүчлийн урттай ХХ хэсгүүдийн байрлал бүр дэх амин хүчлийн тохиолдох магадлалыг WebLogo онлайн түүл ашиглан тооцоолсон (Зураг 2) [41]. (S/T)Gx(I/L/V)Pxx(L/I/F)(G/S/T/F)x(L/C/M)xx консенсус дарааллын хаалтанд болон тэмдэглэсэн амин хүчлүүдийн тохиолдох магадлал 0.7-оос дээш байна.



Зураг 2. Ургамлын ЛБД-ын ХХ-ийн консенсус дарааллын тухайн байрлалд амин хүчил тохиолдох магадлал.

Хоёрдогч бүтцийн оноолтоос харвал нийт ургамлын ЛБД-ын 202-т нь төвийн мужид долоон амин хүчлийн урттай 3₁₀-хеликс, С-терминал талд β -эргэлт, зарим тохиолдолд 3₁₀-хеликс нь тав эсвэл зургаан амин хүчлийн урттай байна. Иймд ургамлын ЛБД-ын ХХ хэсгүүд нь 3₁₀-хеликс болон β -эргэлтээс бүрдэх супер хоёрдогч бүтцээр тодорхойлогдсон (Зураг 3).

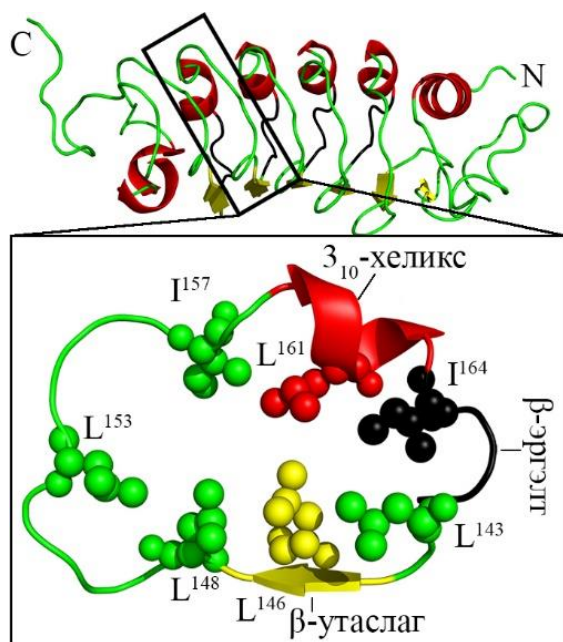


Зураг 3. Ургамлын ЛБД-д тодорхойлогдсон 3₁₀-хеликс ба β -эргэлтээс бүрдсэн супер хоёрдогч бүтцийн бүдүүвч. 3₁₀-

хеликс (G) нь долоон амин хүчлийн урттай ба β-эргэлтийн (T) 4 амин хүчил ($i, i+1, i+2, i+3$). Устөрөгчийн холбоосыг тасархай шулуунаар үзүүлэв.

В. XX-ийн хоёрдогч бүтцийн оноолт

Уургийн хоёрдогч бүтцийн сегментийн захын амин хүчлийг хоёрдмол утгагүй шийдвэрлэхэд бэрхшээлтэй байдаг [42]. Бид β-эргэлийн сонгодог тодорхойлолтыг үндэслэн түүний i дүгээр амин хүчил нь DSSP-ППИ-ийн “ТТ” оноолтын өмнөх N-терминалд ба STRIDE болон XTLSSTR-ын “ТТТ” эсвэл “NNN” оноолтын N-терминалд байгааг тогтоосон [43]. XX-ийн C-терминалын дараалсан хоёр β-эргэлтийг DSSP-ППИ “-ТТ-ТТ-” гэж оноосон бол STRIDE болон XTLSSTR-ийн оноолт ижилхэн “ТТТТТТ” байна.



Зураг 4. AtSERK1 (6FG8) уургийн А гинжний 4 дэх ЛБД. Тус давталтын XX-т 3₁₀-хеликс болон устөрөгчийн холбоосгүй I төрлийн β-эргэлтээс бүрдэх супер хоёрдогч бүтэц. Сферээр лейцин болон изолейцин амин хүчлийн хажуугийн гинжний (радикалын бүлэг) атомууд.

Жишээ нь, бидний өгөгдлөөс хамгийн өндөр нарийвчлалтай 3D бүтэц нь тодорхойлогдсон AtSERK1 (6FG8) уургийн А гинжний 4 дүгээр ЛБД-ын XX-ийн “TGSIPMSLTNIIT” хоёрдогч бүтцийн оноолт “EPPPPGGGGG-TT” (DSSP-ППИ), “CCCCCGGGGGTTT” (STRIDE), ба “r-PPpGGGGgNNN” (XTLSSTR) байна. Үүнд PMSLTNI дараалал нь 7 амин хүчлийн урттай 3₁₀-хеликс бол C-терминалын ITTL нь β-эргэлт болно. 3₁₀-хеликсийн C-малгай амин хүчил болон β-эргэлтийн i дүгээр амин хүчил нь

изолейцин I бөгөөд түүний сүүлийн лейцин “L” амин хүчил нь дараагийн ЛБД-ын CXX-ийн эхний амин хүчил юм (Зураг 4). Ургамлын ЛБД дахь XX-ийн N-терминалын эхний 5 амин хүчлийн хоёрдогч бүтцийн оноолт гурван программ бүр хоорондоо ялгаатай байна.

С. 3₁₀-хеликс ба β-эргэлтийн хеликсийн параметр

Бид ургамлын ЛБД агуулсан уургуудын 19 гинжийг DSSP-ППИ, XTLSSTR ба STRIDE програмуудаар хоёрдогч бүтцийн оноолт хийсэн (Хүснэгт 1). DSSP-ППИ программаар 5-аас 11 амин хүчлийн урттай 210, XTLSSTR-ээр 4-өөс 10 урттай 179 ба STRIDE-аар 5-аас 11 амин хүчлийн урттай 213 3₁₀-хеликсийг илрүүлсэн. Гурван программ тус бүрийн оноосон 3₁₀-хеликс ба β-эргэлтийн хеликсийн параметруудийг HELFIT программаар тооцоолж олсон (Хүснэгт 2). Харин хүснэгтэнд оруулаагүй хеликсийн эргэлтийн чиглэл параметрийн утга β-эргэлт болон 3₁₀-хеликс бүгд +1 буюу баруун байсан. 1987 онд Барлов, Торнтон нар 57 глобулар уургийн кристалл бүтцээс нийт 71 3₁₀-хеликсийг илрүүлсэн [44] бол Перутз (1951), Паулинг нар (1951) каноник 3₁₀-хеликсийн параметруудийг тодорхойлсон байдаг [45,46]. Мөн Энхбаяр нар 3₁₀-хеликсүүдийн параметрийг HELFIT программаар тодорхойлсон [38,47]. Харин ургамлын ЛБД-ын XX-т бидний илрүүлсэн 3₁₀-хеликсийн бүх параметрууд дээрх 3₁₀-хеликсүүдийнхээс их гарсан.

ЛБД агуулсан уураг нь глобулар бус бөгөөд түүний молекулын масст харгалзах усан орчинтой харилцан үйлчлэлцэх гадаргын талбай, холбоосын хэсгийн гадаргын талбай, цэнгийн гадарга нягт зэрэг хэмжигдэхүүний дундаж утга глобулар уургуудаас их байдаг [48,49].

Д. Супер хоёрдогч бүтэц ба ЛБД домеины бүтцийн параметр

Бид өмнөх ажлаараа бактерийн ЛБД-д уургийн XX-ийн ППИ-хеликс ба β-эргэлт агуулсан супер хоёрдогч бүтэц ба ЛБД домеин бүтцийн параметруудийг тодорхойлсон (Зураг 1). Энэ удаад DSSP-ППИ, XTLSSTR ба STRIDE программаар тодорхойлсон 3₁₀-хеликс ба β-эргэлтээс бүрдэх супер хоёрдогч бүтэц болон

ЛБД домеины гурван бүтцийн параметрийг тодорхойлсон утгыг 2-р хүснэгтэд үзүүлэв.

DSSP-PPII программын оноолтыг үндэслэн тодорхойлсон супер хоёрдогч бүтцийн Ω_1

өнцгийн утга 38.80° -аас 110.99° хооронд байхад дундаж утга 75.61° байна. Ургамлын ЛБД домеин бүтцийн Ω_2 болон Ω_3 өнцгийн дундаж 139.39° ба 85.34° утгатай байна (Хүснэгт 3).

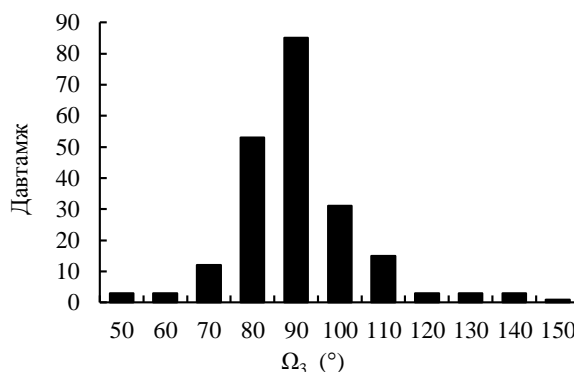
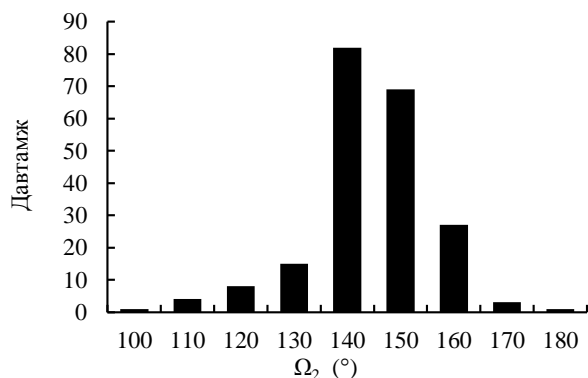
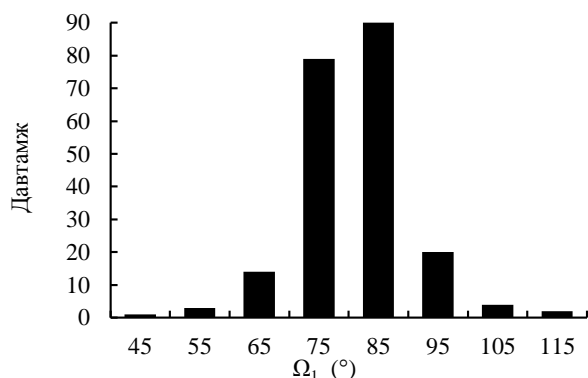
Хүснэгт 2. 3_{10} -хеликс ба β -эргэлийн хеликсийн параметруудийн дундаж. Хаалтанд стандарт хазайлт.

Программ	3_{10} -хеликс						
	N	$P, \text{ \AA}$	n	$R, \text{ \AA}$	$p, \text{ \AA}$	$\Delta z, \text{ \AA}$	$V_c, \text{ \AA}^3$
DSSP-PPII	6.74(0.79)	6.08(0.16)	3.24(0.10)	2.02(0.07)	0.09(0.04)	1.88(0.07)	24.01(0.99)
XTLSSTR	5.24(0.77)	6.00(0.22)	3.20(0.16)	1.99(0.10)	0.06(0.04)	1.88(0.12)	23.33(0.90)
STRIDE	6.57(0.84)	6.05(0.15)	3.24(0.09)	2.02(0.06)	0.08(0.03)	1.87(0.07)	23.96(0.83)
Программ	β -эргэлт						
	N	$P, \text{ \AA}$	n	$R, \text{ \AA}$	$p, \text{ \AA}$	$\Delta z, \text{ \AA}$	$V_c, \text{ \AA}^3$
DSSP-PPII	4	6.09(0.75)	3.74(0.39)	2.29(0.30)	0.03(0.04)	1.66(0.36)	26.45(3.35)
XTLSSTR	4	6.08(0.77)	3.78(0.29)	2.32(0.22)	0.03(0.04)	1.63(0.29)	26.88(2.70)
STRIDE	4	6.09(0.78)	3.76(0.36)	2.30(0.27)	0.03(0.04)	1.65(0.33)	26.66(3.01)

Хүснэгт 3. Ургамлын ЛБД агуулсан уургийн бүтцийн параметр.

Программ	Дундаж ($^\circ$)	MAX ($^\circ$)	MIN ($^\circ$)
DSSP-PPII	75.61 ± 9.20	110.99	38.80
Ω_1 XTLSSTR	75.39 ± 10.85	123.04	37.39
STRIDE	74.43 ± 11.89	121.39	29.42
DSSP-PPII	139.39 ± 11.18	174.99	99.48
Ω_2 XTLSSTR	138.72 ± 10.96	172.95	102.40
STRIDE	139.48 ± 10.64	174.99	101.60
DSSP-PPII	85.34 ± 14.83	149.01	46.93
Ω_3 XTLSSTR	83.97 ± 16.72	147.95	41.62
STRIDE	85.90 ± 16.73	154.82	46.50

DSSP-PPII программаар тус гурван өнцгийн давтамжийн түгэлтийг 5-р зурагт үзүүлэв.



Зураг 5. Ω_1 , Ω_2 , ба Ω_3 өнцгийн давтамжийн түгэлт (DSSP-PPII).

IV. ДҮГНЭЛТ

Бид уургийн хоёрдогч бүтцийн оноолтын гурван программаар ургамлын ЛБД-ын ХХ нь 3_{10} -хеликс болон β -эргэлтээс бүрдэх супер хоёрдогч бүтэцтэй болохыг тогтоосон. Мөн түүний консенсус дараалал 13 амин хүчлийн урттай (xGxLPxxLxxLxx) ба “L” нь лейцин, изолейцин, валин, фенилаланин, цистеин, метионин зэрэг гидрофобик амин хүчлүүд байна. HELFIT программаар 3_{10} -хеликс (**G**), β -эргэлт (**B**), ЛБД домеин (**A**) тус бүрийн хеликсийн тэнхлэгийн нэгж векторыг тодорхойлсон. Эдгээр нэгж вектороор тодорхойлогдох Ω_1 , Ω_2 , ба Ω_3 бүтцийн параметр нь супер хоёрдогч бүтэц ба ЛБД домеин бүтцийн онцлогийг тодорхойлж байна.

АШИГЛАСАН МАТЕРИАЛ

[1] R. D. Finn, T. K. Attwood, P. C. Babbitt, A. Bateman, P. Bork, Z. Dosztányi, S. El-Gebali,

- and M. Fraser, *Nucleic acids research* **45**, D190 (2016).
- [2] N. Matsushima and R. Kretsinger, Saarbrücken: LAMBERT Academic Publishing, 1 (2016).
- [3] B. Kobe and A. V. Kajava, *Current opinion in structural biology* **11**, 725 (2001).
- [4] N. Matsushima, H. Miyashita, T. Mikami, and Y. Kuroki, *BMC microbiology* **10**, 235 (2010).
- [5] N. Matsushima, S. Takatsuka, H. Miyashita, and R. H. Kretsinger, *Protein and peptide letters* **26**, 108 (2019).
- [6] I. Miras, F. Saul, M. Nowakowski, P. Weber, A. Haouz, W. Shepard, and M. Picardeau, *Acta Crystallographica Section D: Biological Crystallography* **71**, 1351 (2015).
- [7] A. A. Adzhubei, M. J. Sternberg, and A. A. Makarov, *Journal of molecular biology* **425**, 2100 (2013).
- [8] B. Kobe and J. Deisenhofer, *Current opinion in structural biology* **5**, 409 (1995).
- [9] I. P. Calil and E. P. B. Fontes, *Annals of botany* **119**, 711 (2017).
- [10] K. Toth and G. Stacey, *Frontiers in plant science* **6**, 401 (2015).
- [11] L. McHale, X. Tan, P. Koehl, and R. W. Michels, *Genome biology* **7**, 212 (2006).
- [12] C. Bystroff, Y. Shao, and X. Yuan, *Applied bioinformatics* **3**, 97 (2004).
- [13] D. Batkhishig, K. Bilguun, P. Enkhbayar, H. Miyashita, R. H. Kretsinger, and N. Matsushima, *The protein journal* **37**, 223 (2018).
- [14] D. Kanduc, *Journal of peptide science: an official publication of the European Peptide Society* **18**, 487 (2012).
- [15] N. Matsushima and H. Miyashita, *Biomolecules* **2**, 288 (2012).
- [16] Z. Li, S. Chakraborty, and G. Xu, *PLoS One* **12**, e0175317 (2017).
- [17] W. Song, L. Liu, J. Wang, Z. Wu, W. Li, Z. Han, H. Guo, and J. Chai, *Cell Res* **26**, 674 (2016).
- [18] J. Wang, H. Li, Z. Han, H. Zhang, T. Wang, G. Lin, J. Chang, W. Yang, and J. Chai, *Nature* **525**, 265 (2015).
- [19] J. Tang, Z. Han, Y. Sun, H. Zhang, X. Gong, and J. Chai, *Cell research* **25**, 110 (2015).
- [20] J. Santiago, B. Brandt, M. Wildhagen, U. Hohmann, L. A. Hothorn, M. A. Butenko, and M. Hothorn, *eLife* **5**, e15075 (2016).
- [21] J. She, Z. Han, T.-W. Kim, J. Wang, W. Cheng, J. Chang, and J. Chai, *Nature* **474**, 472 (2011).
- [22] J. She, Z. Han, B. Zhou, and J. Chai, *Protein & cell* **4**, 475 (2013).
- [23] Y. Sun, L. Li, A. P. Macho, Z. Han, Z. Hu, C. Zipfel, J. M. Zhou, and J. Chai, *Science* **342**, 624 (2013).
- [24] U. Hohmann, J. Nicolet, A. Moretti, L. A. Hothorn, and M. Hothorn, *Nature plants* **4**, 345 (2018).
- [25] X. Zhang, W. Liu, T. T. Nagae, H. Takeuchi, H. Zhang, Z. Han, T. Higashiyama, and J. Chai, *Nature communications* **8**, 1331 (2017).
- [26] R. McAndrew, R. N. Pruitt, S. G. Kamita, J. H. Pereira, D. Majumdar, B. D. Hammock, P. D. Adams, and P. C. Ronald, *Acta crystallographica. Section D, Biological crystallography* **70**, 3080 (2014).
- [27] A. Di Matteo, L. Federici, B. Mattei, G. Salvi, K. A. Johnson, C. Savino, G. De Lorenzo, D. Tsernoglou, and F. Cervone, *Proceedings of the National Academy of Sciences of the United States of America* **100**, 10124 (2003).
- [28] P. Liu, Z. Hu, B. Zhou, S. Liu, and J. Chai, *Cell Res* **23**, 303 (2013).
- [29] G. Lin, L. Zhang, Z. Han, X. Yang, W. Liu, E. Li, J. Chang, and J. Chai, *Genes & development* **31**, 927 (2017).
- [30] R. Chebrek, S. Leonard, A. G. de Brevern, and J. C. Gelly, *Database : the journal of biological databases and curation* **2014** (2014).
- [31] M. Heinig and D. Frishman, *Nucleic acids research* **32**, W500 (2004).
- [32] Y. Mansiaux, A. P. Joseph, J. C. Gelly, and A. G. de Brevern, *PLoS One* **6**, e18401 (2011).
- [33] W. Kabsch and C. Sander, *Biopolymers* **22**, 2577 (1983).
- [34] S. M. King and W. C. Johnson, *Proteins* **35**, 313 (1999).
- [35] D. Frishman and P. Argos, *Proteins* **23**, 566 (1995).
- [36] *The Biochemical journal* **121**, 577 (1971).
- [37] P. Enkhbayar, S. Damdinsuren, M. Osaki, and N. Matsushima, *Computational biology and chemistry* **32**, 307 (2008).

- [38] P. Enkhbayar, K. Hikichi, M. Osaki, R. H. Kretsinger, and N. Matsushima, *Proteins* **64**, 691 (2006).
- [39] P. Enkhbayar, S. Damdinsuren, M. Osaki, and N. Matsushima, *Computational biology and chemistry* **32**, 307 (2008).
- [40] N. Matsushima, H. Miyashita, P. Enkhbayar, and R. H. Kretsinger, *Biomolecules* **5**, 1955 (2015).
- [41] G. E. Crooks, G. Hon, J. M. Chandonia, and S. E. Brenner, *Genome research* **14**, 1188 (2004).
- [42] J. Martin, G. Letellier, A. Marin, J. F. Taly, A. G. de Brevern, and J. F. Gibrat, *BMC structural biology* **5**, 17 (2005).
- [43] E. G. Hutchinson and J. M. Thornton, *Protein Science* **5**, 212 (1996).
- [44] D. J. Barlow and J. M. Thornton, *Journal of molecular biology* **201**, 601 (1988).
- [45] L. Pauling, R. B. Corey, and H. R. Branson, *Proceedings of the National Academy of Sciences* **37**, 205 (1951).
- [46] M. Perutz, *Nature* **167**, 1053 (1951).
- [47] P. Enkhbayar and N. Matsushima, *AIP Conference Proceedings* **1479**, 1942 (2012).
- [48] Q. R. Fan and W. A. Hendrickson, *Proteins* **72**, 393 (2008).
- [49] U. Hohmann, J. Santiago, J. Nicolet, V. Olsson, F. M. Spiga, L. A. Hothorn, M. A. Butenko, and M. Hothorn, *Proceedings of the National Academy of Sciences of the United States of America* **115**, 3488 (2018).