

11. Inteiny

Přednáška o tom, jak příroda sama vyvinula a používá nativní ligaci.

Chemical
Science



T. Muir,
Princeton

PERSPECTIVE

[View Article Online](#)
[View Journal](#) | [View Issue](#)

Inteins: nature's gift to protein chemists

Cite this: *Chem. Sci.*, 2014, 5, 446

Neel H. Shah and Tom W. Muir*

Inteins are auto-processing domains found in organisms from all domains of life. These proteins carry out a process known as protein splicing, which is a multi-step biochemical reaction comprised of both the cleavage and formation of peptide bonds. While the endogenous substrates of protein splicing are specific essential proteins found in intein-containing host organisms, inteins are also functional in exogenous contexts and can be used to chemically manipulate virtually any polypeptide backbone. Given this, protein chemists have exploited various facets of intein reactivity to modify proteins in myriad ways for both basic biological research as well as potential therapeutic applications. Here, we review the intein field, first focusing on the biological context and phylogenetic diversity of inteins, followed by a description of intein structure and biochemical function. Finally, we discuss prevalent intein-based technologies, focusing on their applications in chemical biology, followed by persistent caveats of intein chemistry and approaches to alleviate these shortcomings. The findings summarized herein describe two and a half decades of research, leading from a biochemical curiosity to the development of powerful protein engineering tools.

Received 24th October 2013
Accepted 27th November 2013

DOI: 10.1039/c3sc52951g

www.rsc.org/chemicalscience



P. Dawson,
Scripps

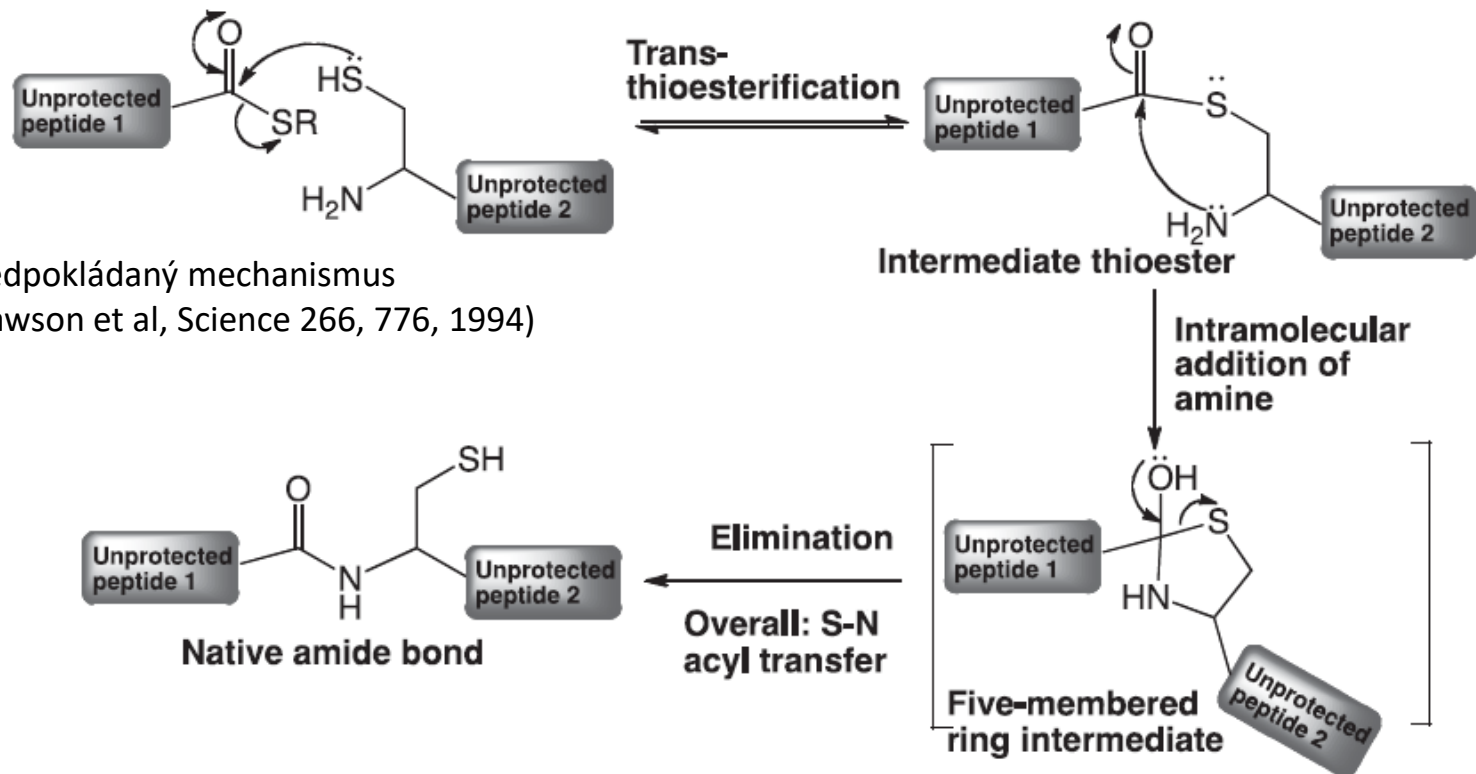


Stephen Kent
(Scripps, Univ. Chicago)

Nativní chemické ligace

1953 Theodor Wieland: Val-CO-S-R a Cys reagovaly na Val-Cys (reakce aktivního esteru s aminem)

1990s, Stephen Kent ve The Scripps Research Institute vyvinul "Native Chemical Ligation (NCL)" na stejném principu a aplikoval jí na ligace nechráněných peptidových fragmentů a syntézu proteinů



Předpokládaný mechanismus
(Dawson et al, Science 266, 776, 1994)

V jádru se jedná o chemoselektivní a regioselektivní transthioesterifikaci následovanou S-to-N acyltransferem, která probíhá ve vodě. Většinou se přidává nějaký SH katalyzátor. Reakce je ale dost pomalá a má i vedlejší reakce.

Nativní chemické ligace jsou obecně limitovány chemickou syntézou C-koncových thioesterů.

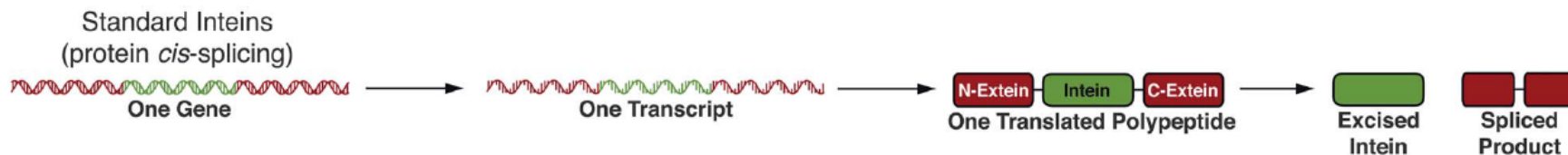
Intein, z *intervening protein*

Jsou to proteiny s auto-procesující aktivitou (tzv. splicing).

Samy se vyštěpují ze sekvence a zároveň spojí zbylé fragmenty, **exteiny** (**external protein**).

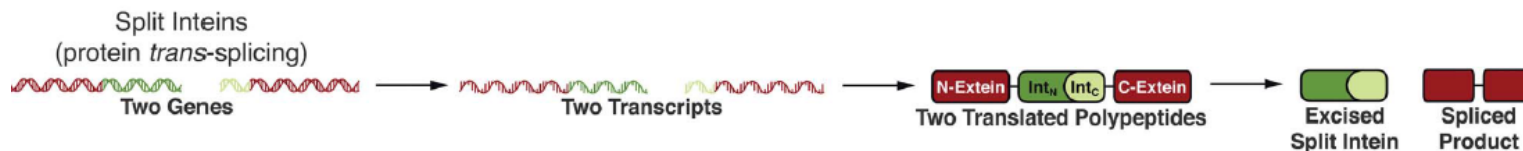
To znamená, že tyto proteiny štěpí a opět spojují peptidové vazby.

Klasický nejčastější typ inteinů (95%)



Sbalením do správné 3D konformace
nabývá intein své aktivity

Minoritní typ inteinů (tzv. „split inteins“)

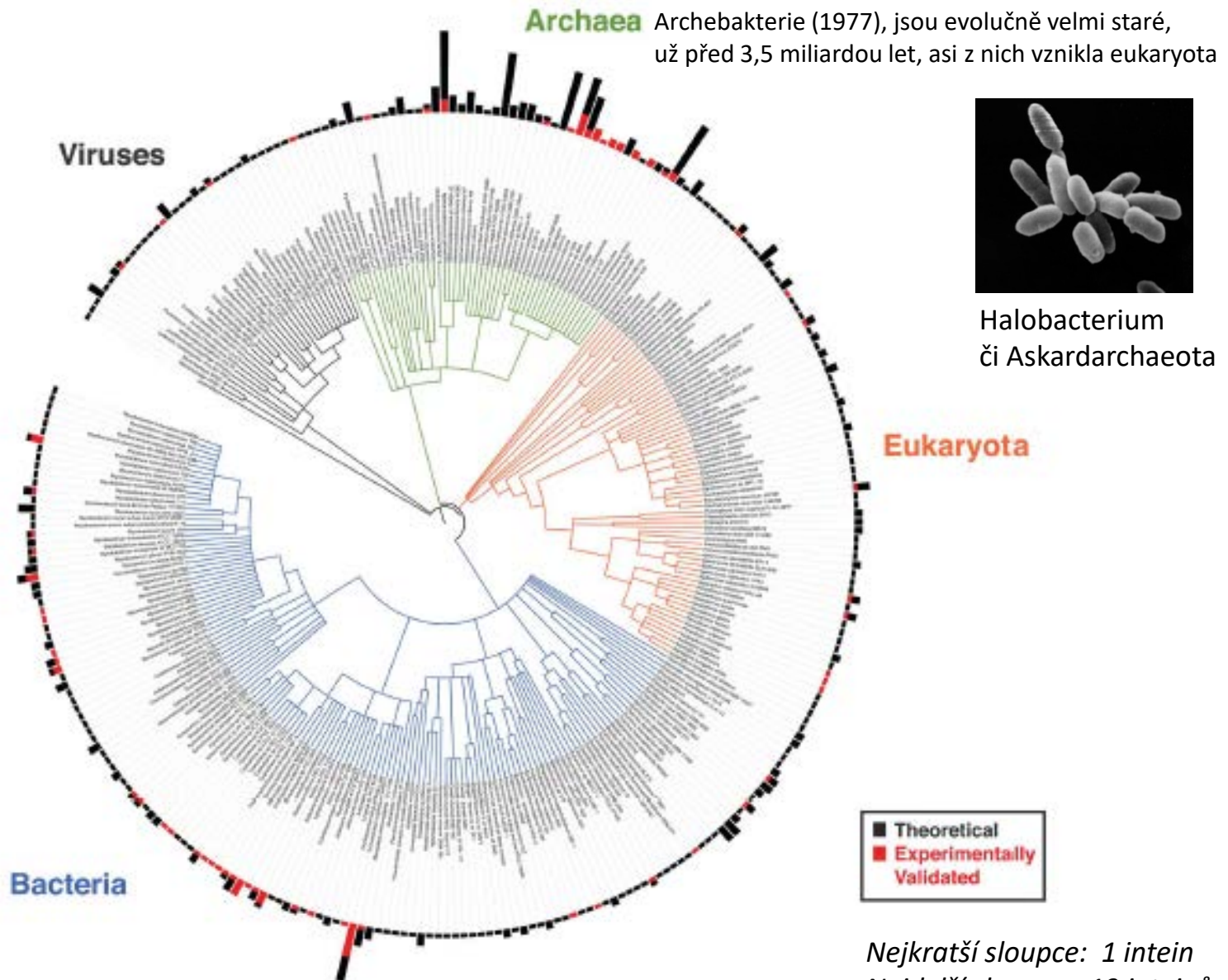


Po translaci se dva fragmenty nekovalentně spojí
a nabývají své aktivity

$k_{on} 10^7-10^5 \text{ M}^{-1} \text{ s}^{-1}$, $K_d 2-30 \text{ nM}$

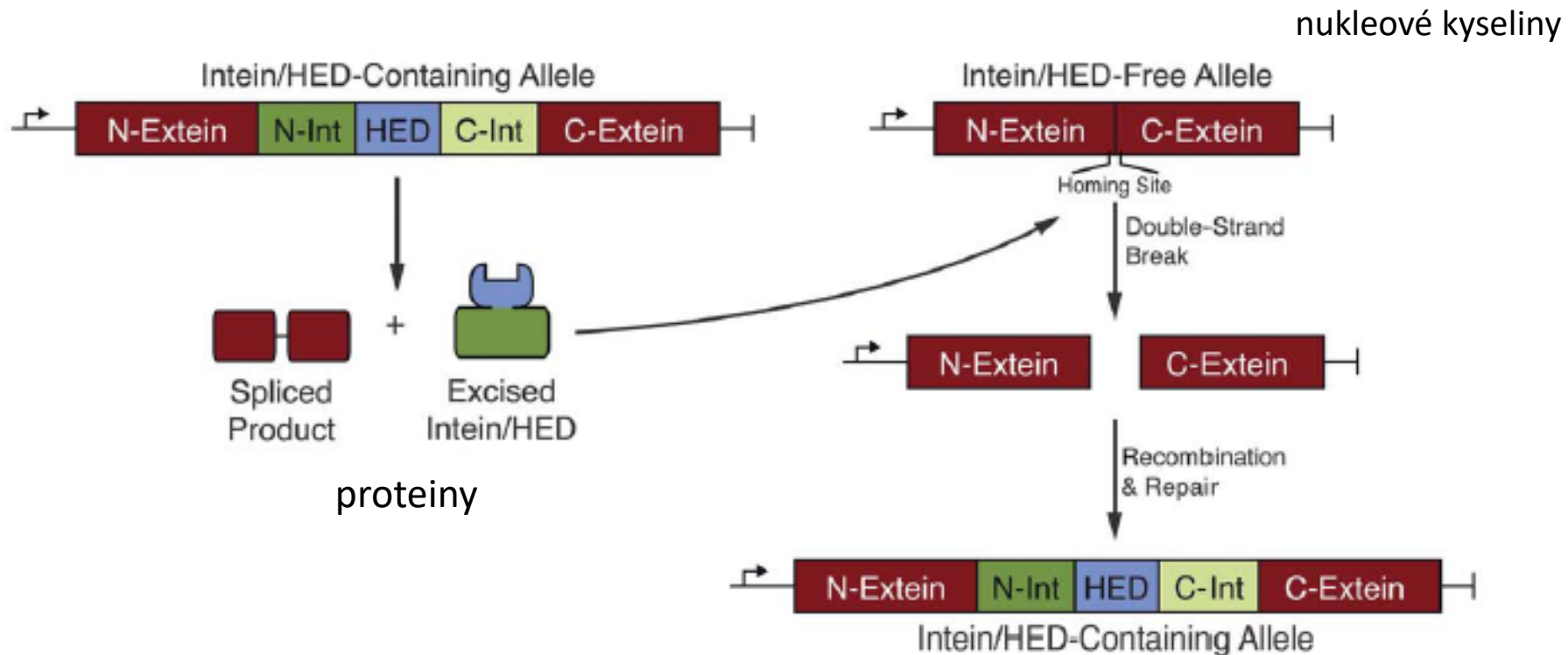
Fylogenetika inteinů

První intein byl objeven v roce 1990. Od té doby jich bylo predikováno asi 600 napříč organismy.



*Mechnismus rozšiřování **inteinů** v rámci organismu(ů)*

Některé inteiny mají do své sekvence vmezeženou tzv. **HED doménu** (**homing endonuclease domain**), která je po vyštěpení z původního proteinu schopná znovu začlenit gen inteinu do genu hostitele.



Inteiny bývají obvykle „zabudovány“ do sekvencí esenciálních proteinů

zapojených do metabolismu DNA jako jsou DNA a RNA polymerázy, DNA helikázy a gyrázy.

V těchto hostujících proteinech jsou obvykle umístěny poblíž sekvencí, které jsou krucální pro fungování hostitele, např. aktivní místo enzymu, či vazba ligandu.

V tomto ohledu se zdá jisté, že vyštěpení inteinů je nutné pro funkci hostitele, ale doposud nebyla prokázána žádná jejich regulační role či nějaká výhoda pro hostitele způsobená jejich existencí.

Proč tedy inteiny vůbec existují? Zdá se, že nemají žádnou biologickou roli.

Jsou to možná jen artefakty evoluce, „selfish genes“, sobecké geny, starají se jen o sebe...

Dawkins, Richard (1976). The Selfish Gene. Oxford University Press.

According to Dawkins' gene centered view of evolution, all genes are selfish, even those that cooperate with other genes to build the organisms that carry the genes into the next generation.

*Podle teorie sobeckého genu je **gen** hlavní jednotkou přirozeného výběru, ne jednotlivec, nebo dokonce skupina, jak předpokládal výběr skupinový. Jednotlivé alely mezi sebou soupeří o zastoupení v dalších generacích a organismy jsou jejich pouhými výtvoři, nosiči či vehikly, které si postavily, aby se mohly účinněji prosazovat ve světě.*

Mem je termín pro kulturní obdobu genu – replikující se jednotku kulturní informace. Evoluce nemusí být nutně postavena jen na replikaci DNA, ale jejím principem je šíření jakékoli informace, která podléhá změnám (mutacím) a selekčnímu tlaku.

Tento termín poprvé použil v roce 1976 ve své knize Sobecký gen Richard Dawkins.

Clinton Richard Dawkins, [FRS FRS](#) (born 26 March 1941) is an English [ethologist](#), [evolutionary biologist](#), and [author](#). He is an [emeritus fellow](#) of [New College, Oxford](#), and was the [University of Oxford's Professor for Public Understanding of Science](#) from 1995 until 2008.

Dawkins first came to prominence with his 1976 book [The Selfish Gene](#), which popularised the [gene-centred view of evolution](#) and introduced the term [meme](#). With his book [The Extended Phenotype](#) (1982), he introduced into [evolutionary biology](#) the influential concept that the [phenotypic](#) effects of a gene are not necessarily limited to an organism's body, but can stretch far into the environment. In 2006, he founded the [Richard Dawkins Foundation for Reason and Science](#).

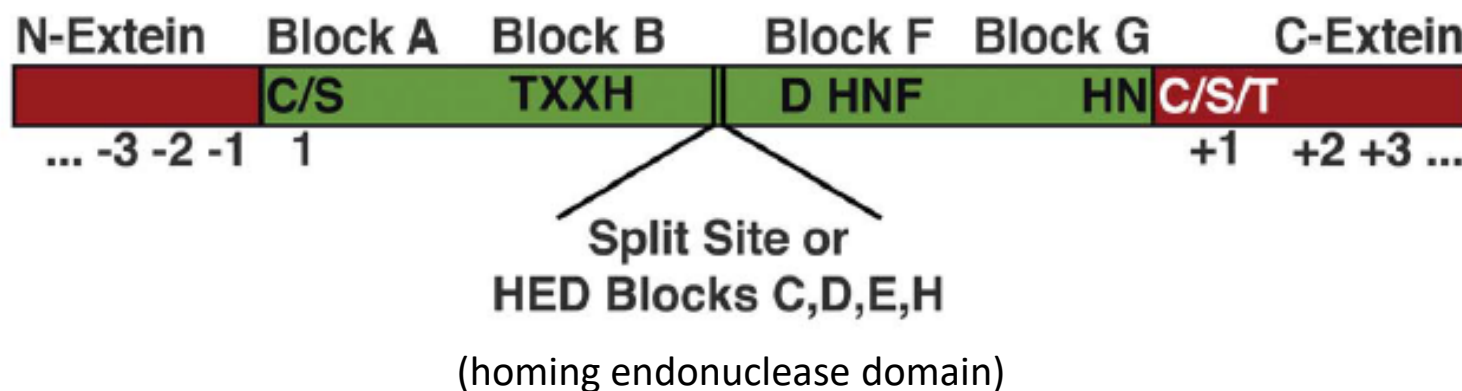
Dawkins is known as an outspoken [atheist](#). He is well known for his criticism of [creationism](#) and [intelligent design](#). In [The Blind Watchmaker](#) (1986), he argues against the [watchmaker analogy](#), an argument for the existence of a [supernatural creator](#) based upon the complexity of living organisms. Instead, he describes evolutionary processes as analogous to a *blind* watchmaker, in that reproduction, mutation, and selection are unguided by any designer. In [The God Delusion](#) (2006), Dawkins contends that a supernatural creator almost certainly does not exist and that religious faith is a [delusion](#).

Dawkins has been awarded many prestigious academic and writing awards, and he makes regular television, radio, and Internet appearances, predominantly discussing his books, his atheism, and his ideas and opinions as a [public intellectual](#).

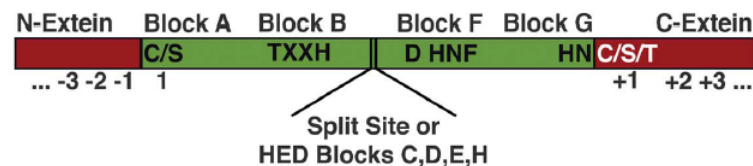
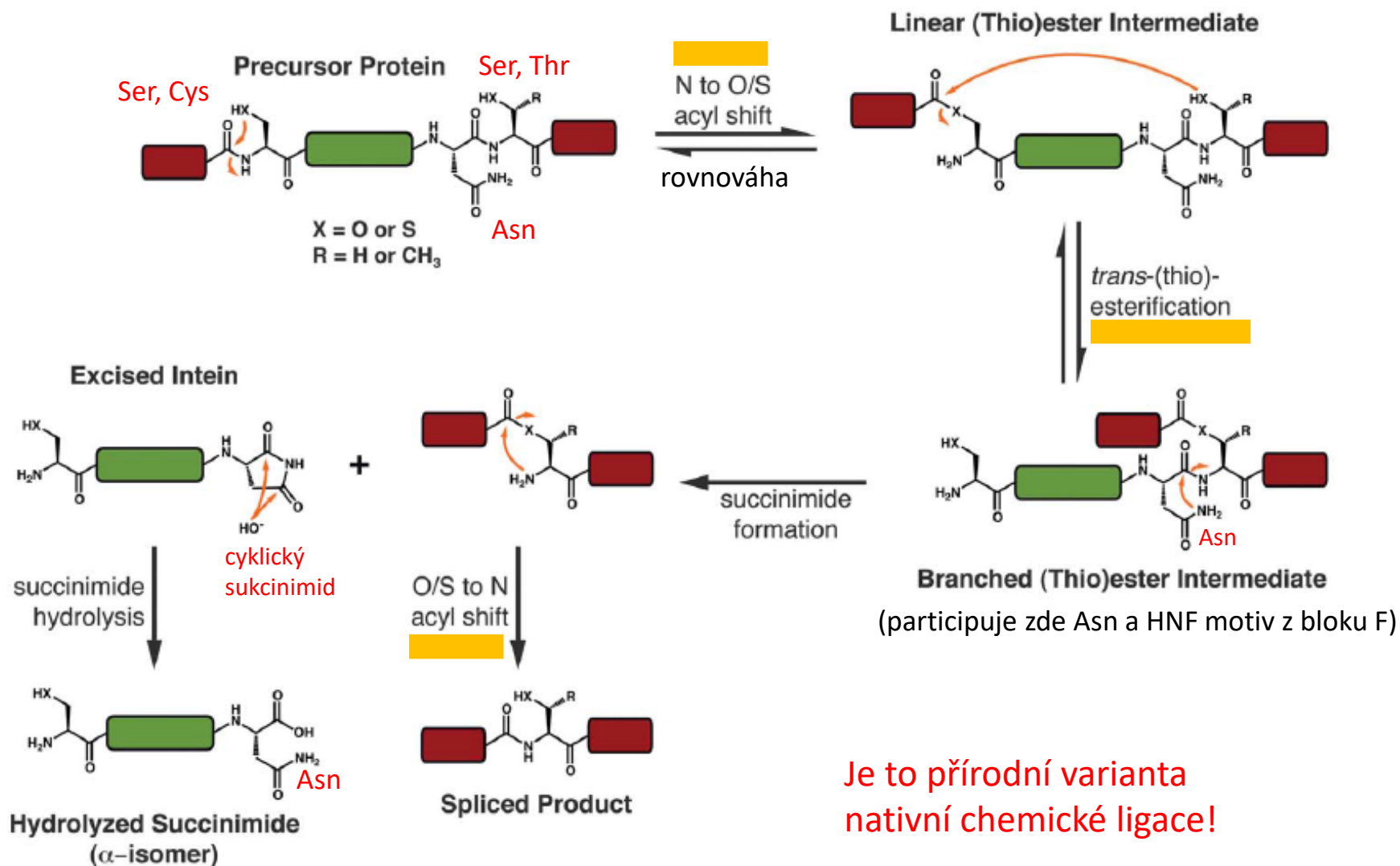


Inteiny mají mezidruhově spíše malou homologii sekvencí (< 40%),
ale sdílejí stejný reakční mechanismus díky několika **konzervovaným sekvenčním motivům**

C je Cys, S je Ser, T je Thr, H je His, N je Asn, F je Phe

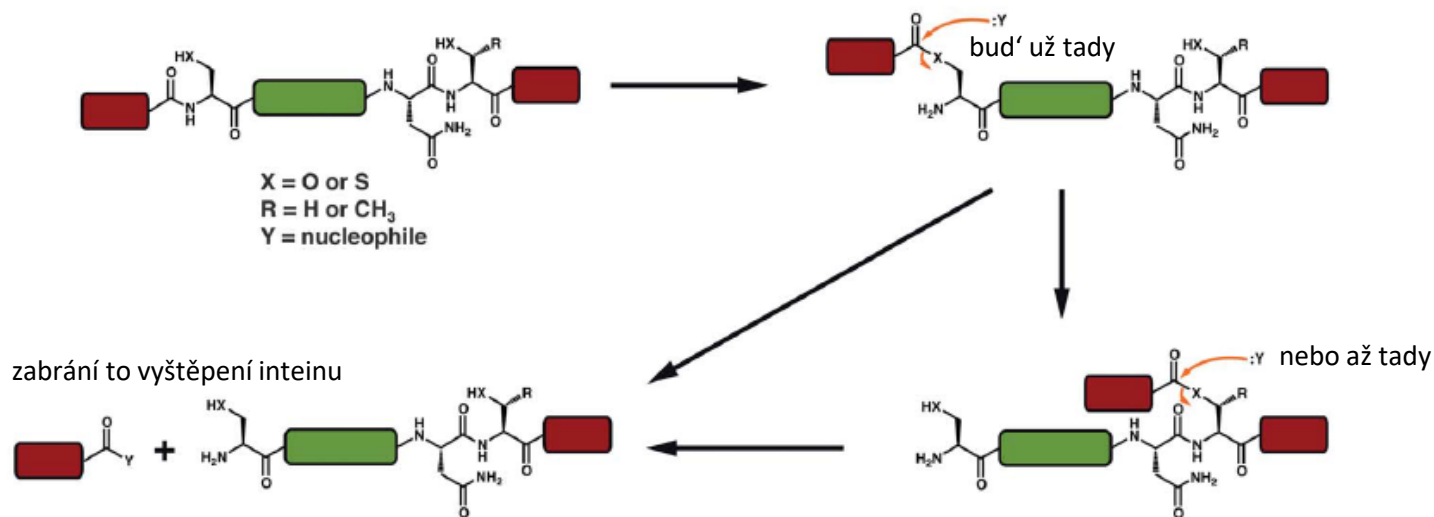


Chemicky se jedná o sled reakcí přenášejících karboxylovou skupinu (acyl-transfer)

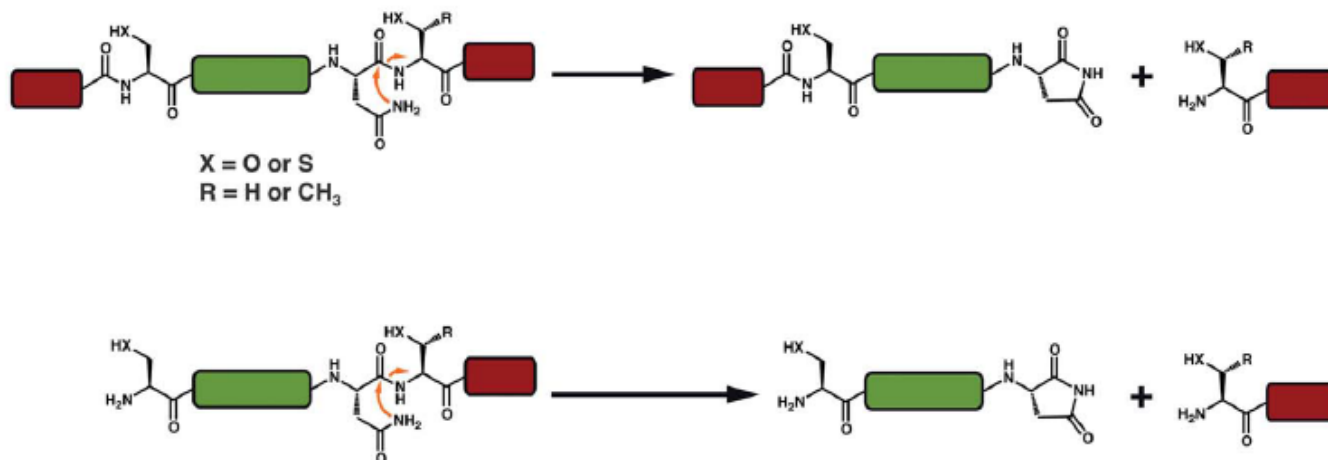


Vedlejší reakce při štěpení inteiny

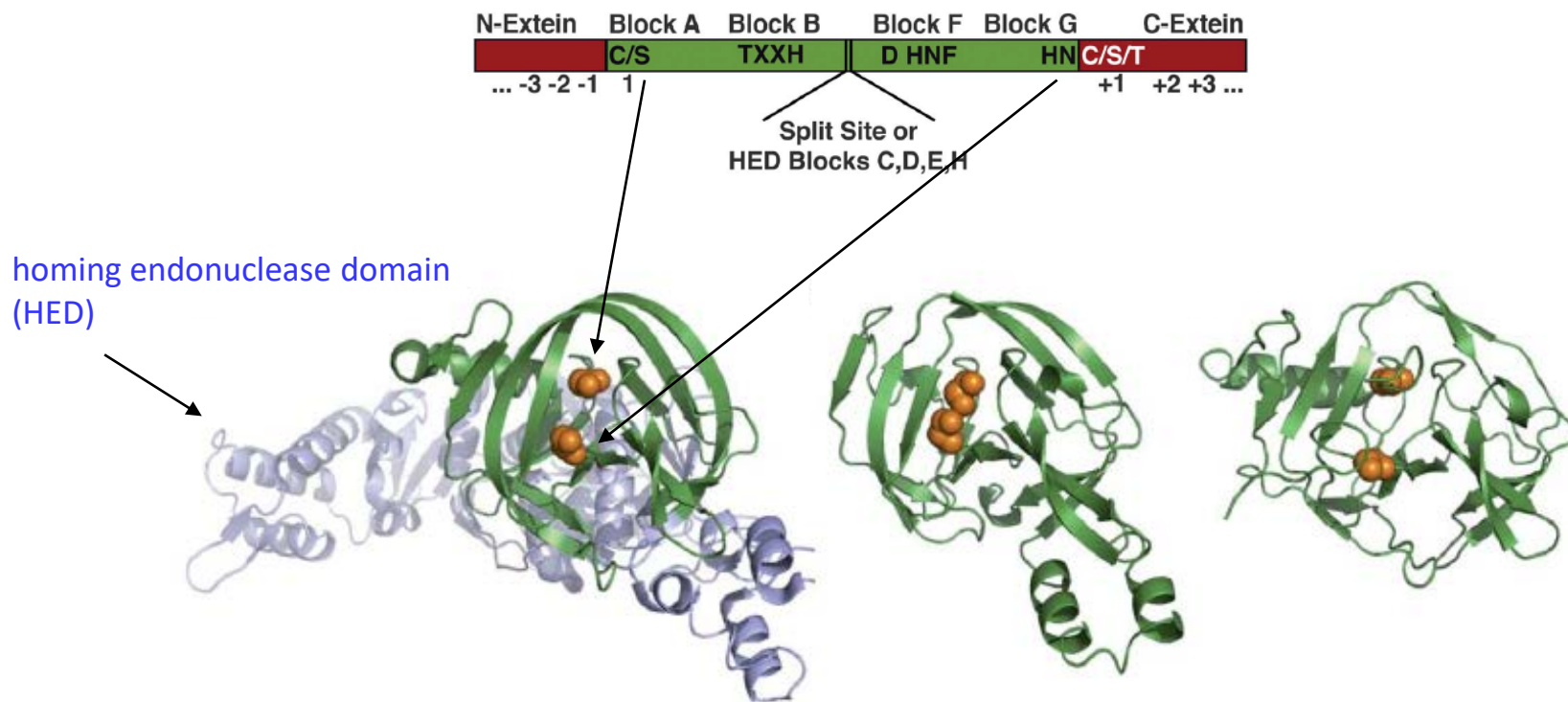
Přítomný nukleofil (:Y), alkohol či thiol, může atakovat ester



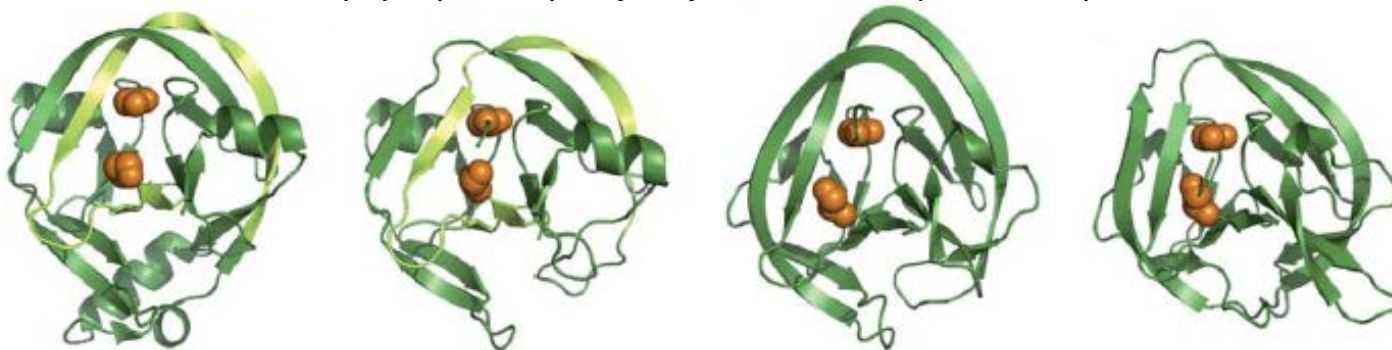
V případě, že **Asn** cyklizuje již ve struktuře původního proteinu, může docházet k C-terminálnímu štěpení ještě před tvorbou thioesteru na N-konci, **výsledkem je, že se exteiny nespojí**



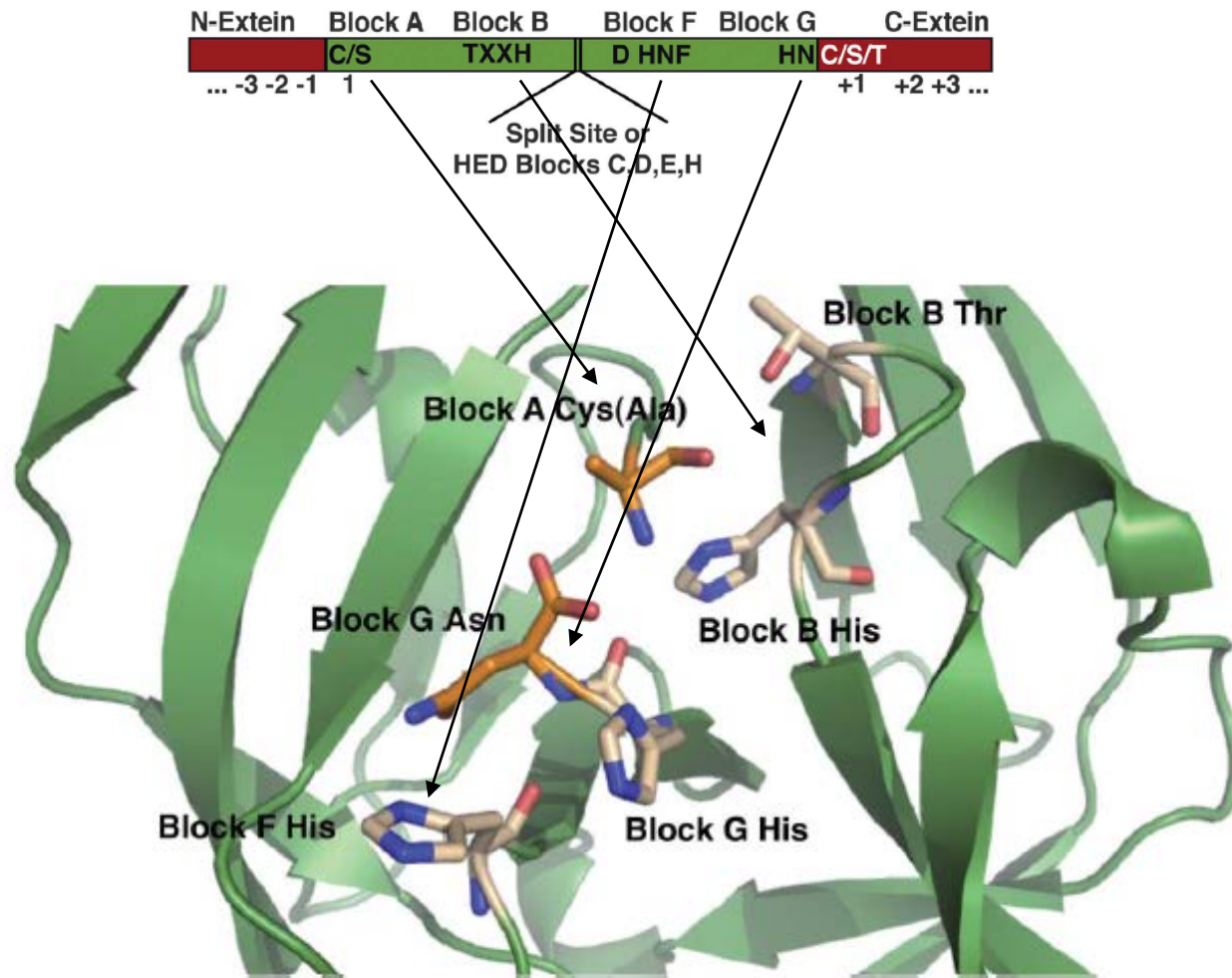
Inteiny patří do větší skupiny proteinů nazývaných **Hedgehog/intein (Hint) domény**
 a patří mezi ně *inteiny, bakteriální intein-like domény a Hedgehog auto-processing domains*



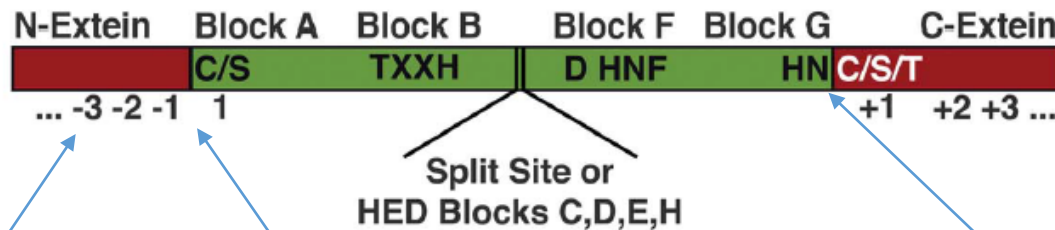
Všechny tyto proteiny mají stejnou strukturu podobnou podkově



Detail aktivního místa jednoho z inteinů



Aplikace inteinů pro purifikaci proteinů bez Tagu (His atd.)

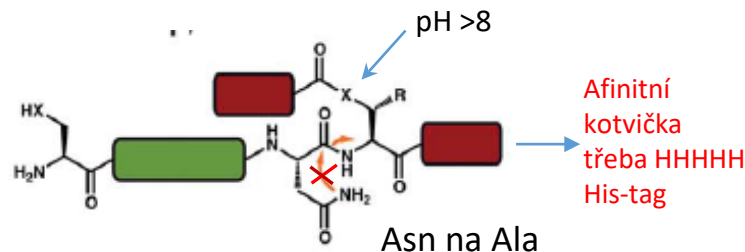


Na C-konec vložíme afinitní kotvičku

Požadovaný protein je na místě N-exteinu, kde dochází k rovnováze mezi amidem a thioesterem

zmutuji-li Asn za Ala, nedochází k C-terminálnímu štěpení

Takže po purifikaci konstruktu pomocí C-terminálního Tagu zvýšíme pH do bazické oblasti a tím dojde k hydrolýze thioesteru a odštěpení požadovaného proteinu (N-exteinu) bez tagu.

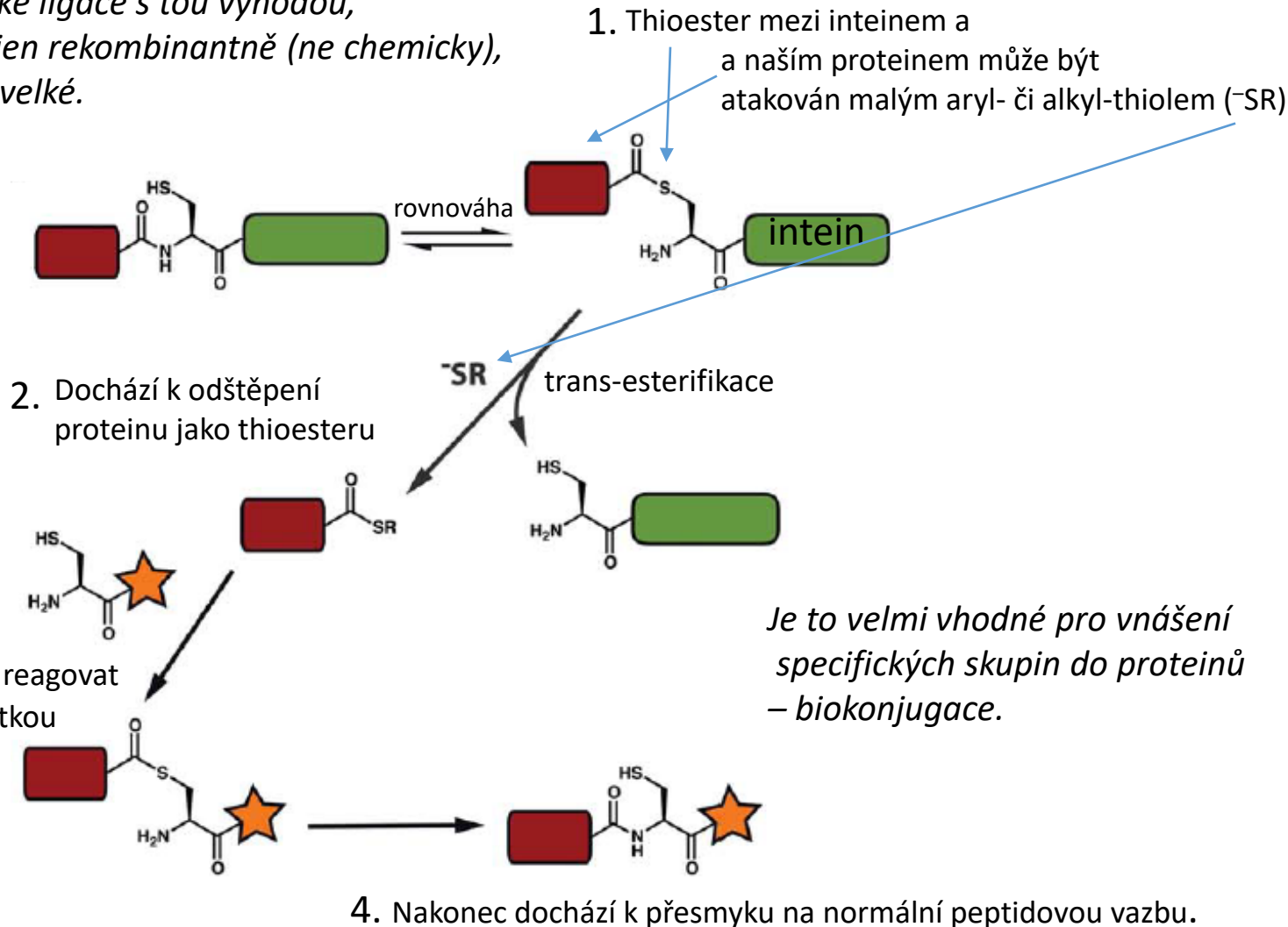


Afinitní kotvička třeba HHHHH His-tag

Aplikace klasických **inteinů** pro semisyntézu proteinů

„Expressed Protein Ligation (EPL)“

Obdoba Nativní chemické ligace s tou výhodou, že můžeme vše vyrobit jen rekombinantně (ne chemicky), čili proteiny mohou být velké.

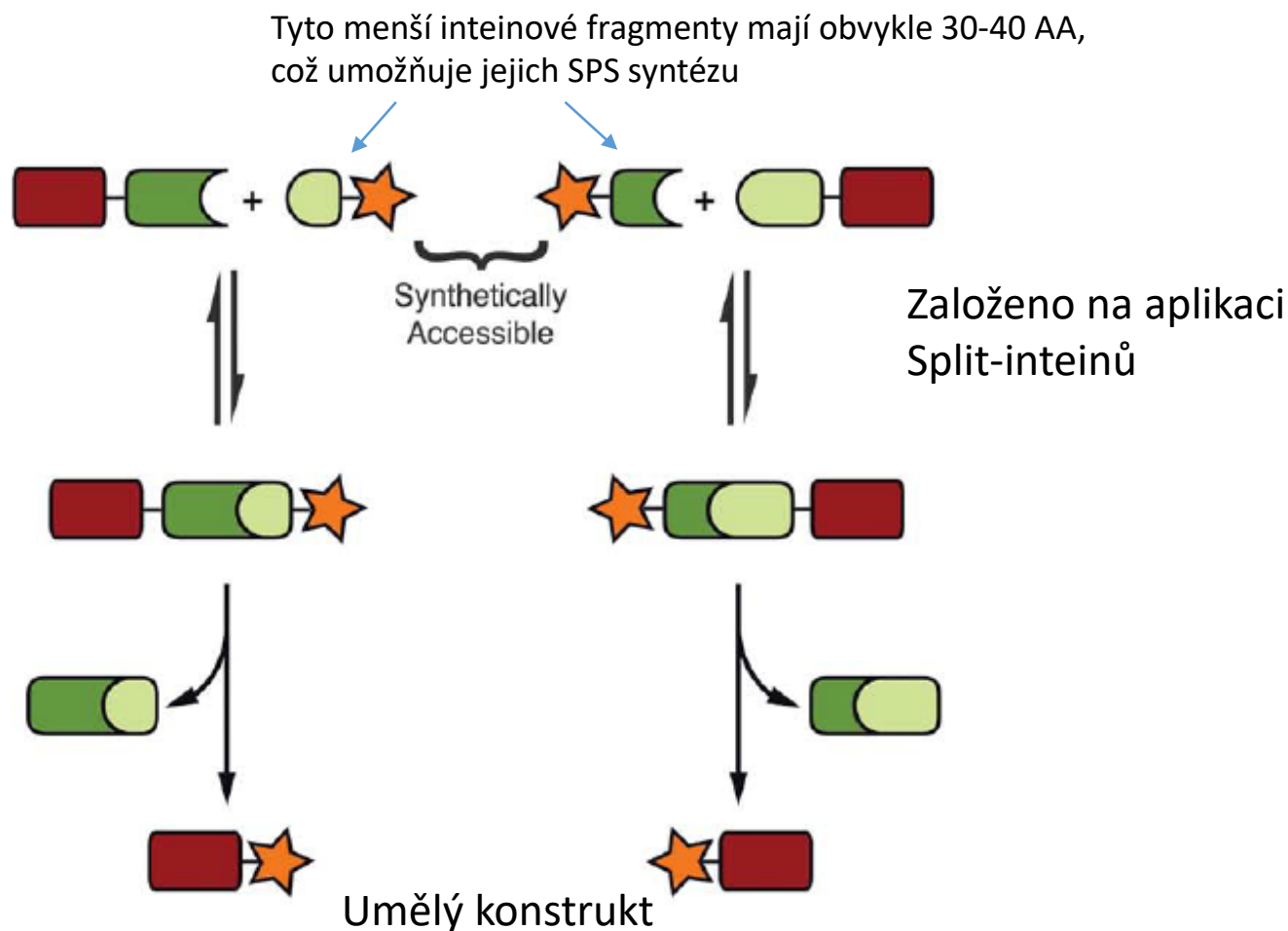


Většinou za denaturujících podmínek, protože jsou potřebné vysoké koncentrace proteinů, pro efektivní srážky peptidů.

Aplikace *split* inteinů pro *semisyntézu* proteinů

„Protein Trans-Splicing (PTS)“

Slouží to ke spojování peptidů a proteinů v nedenaturujícím prostředí i za malých koncentracích, neboť je to řízeno vysoce afinitní interakcí inteinů.

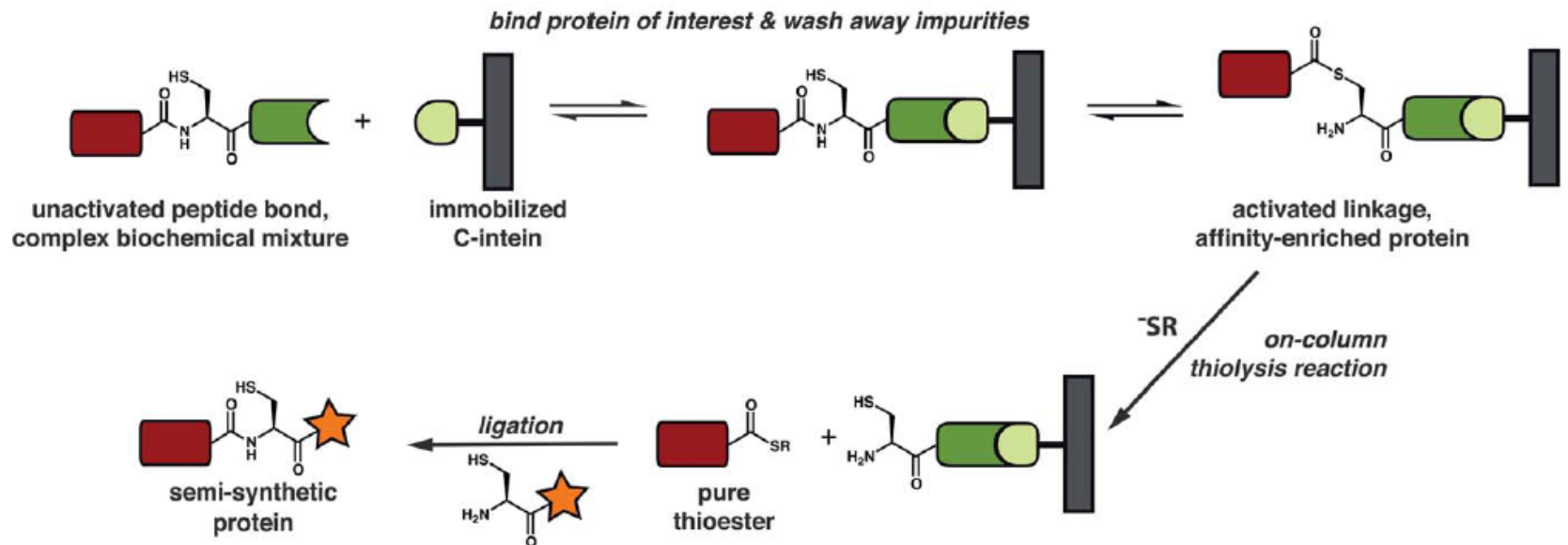


Aplikace split inteinů pro *purifikaci rekombinantních proteinů* a jejich následnou *semisyntézu*

DnaE intein se neštěpí
dokud není kompletní

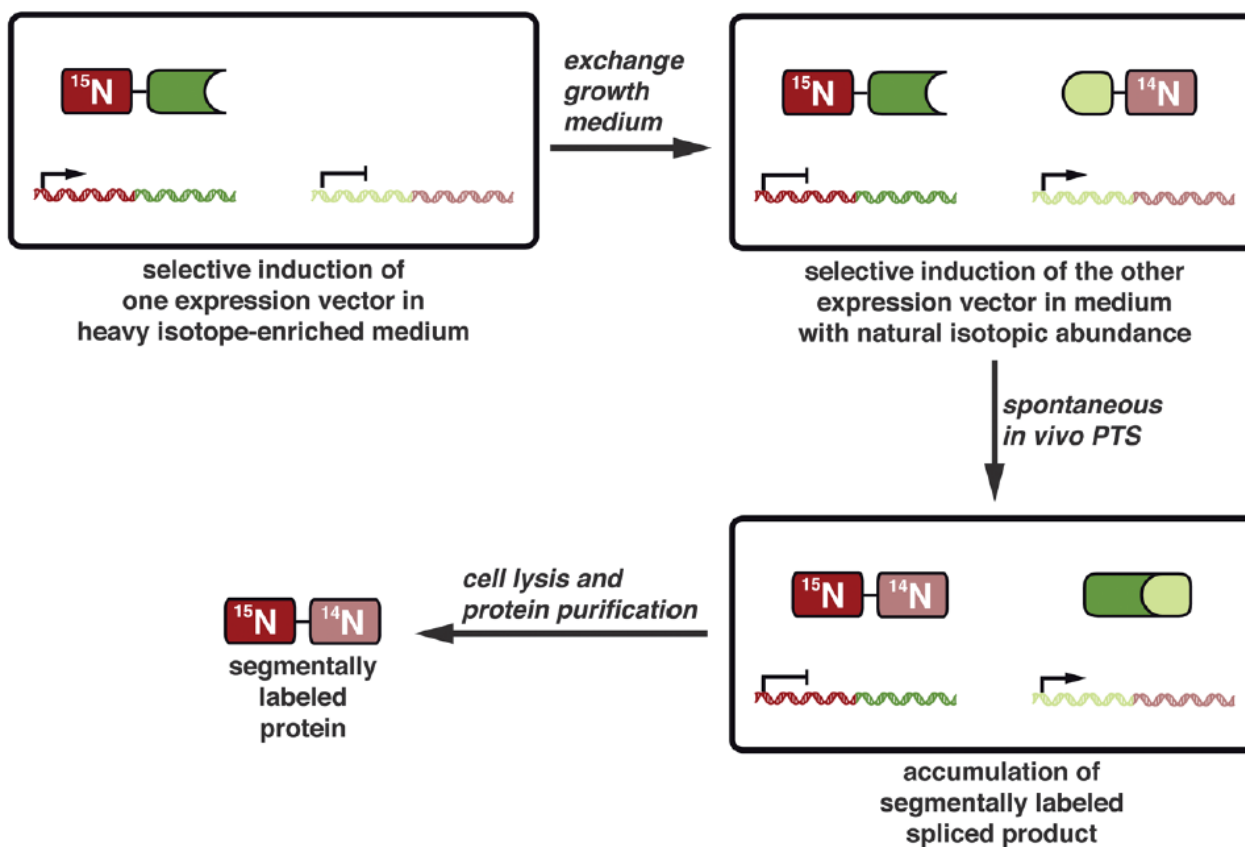
Založeno na aplikaci
Split-inteinů

Po fúzi s C-terminální částí
intein přesmykne na thioester

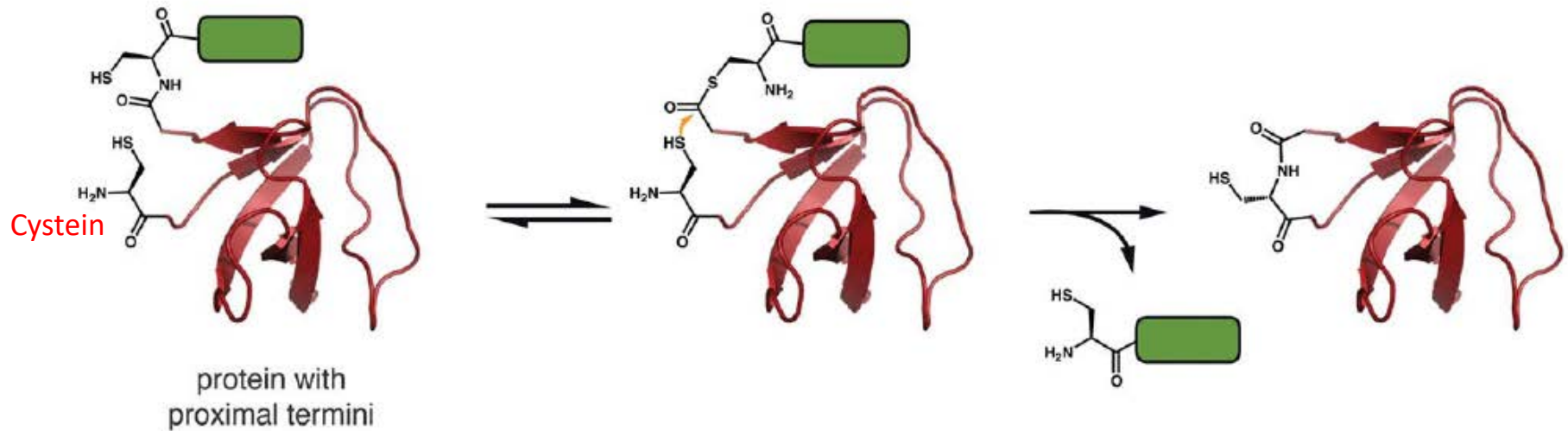


Aplikace *split inteinů pro segmentové izotopické značení* např. pro NMR měření s izotopy ^{15}N , ^{13}C

Příliš velké proteiny (150 AA a víc, či proteiny s degenerovanými sekvencemi)
mají NMR spektra příliš komplikovaná a to i přes izotopické ^{15}N a ^{13}C značení proteinů.
Proto je výhodné značení jen některých úseků.



Aplikace *inteinů* pro *cyklizace* proteinů a peptidů

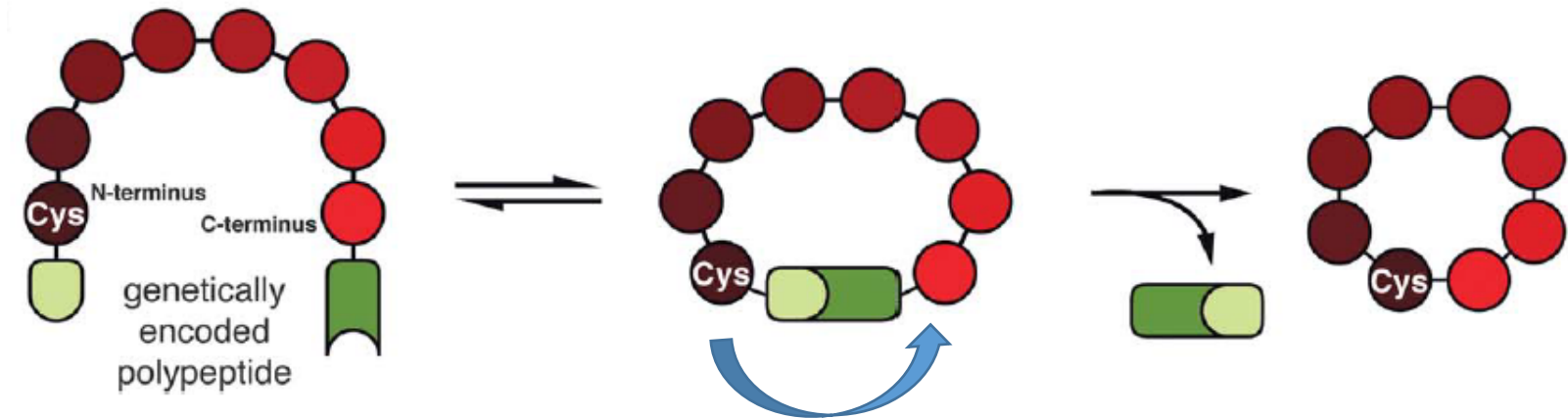


Úspěch zde ale velmi záleží na přístupnosti a blízkosti obou reagujících jednotek

Výsledný cyklický protein může mít vyšší termodynamickou stabilitu a rozdílné vlastnosti

Aplikace **inteinů** pro **cyklizace** proteinů a peptidů

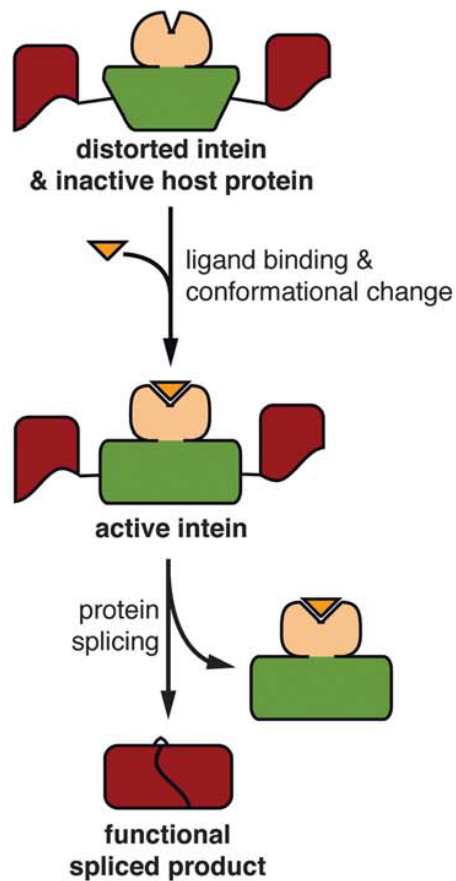
Split Intein-mediated Circular Ligation Of Peptides and Proteins (SICLOPPS)



Touto metodou byly vytvořeny knihovny cyklických peptidů a objeveny nové inhibitory biologicky relevantních enzymů.

Inteiny s *podmíněnou auto-procesující aktivitou* „Conditional Protein Splicing (CPS)“

Jedná se o aktivaci či inhibici procesování proteinu vnějším modulátorem. Umožní to zapínání funkce proteinu dokonce i *in vivo*.



V tomto případě je kontinuální intein zřůzován s doménou vážící ligand. Po navázání malé molekuly dojde **ke konformační změně** v konstruktu a ke splicingu.

Regulační
jednotka

Ligand-binding domain of the human
thyroid hormone receptor β (TR):



Open form



Closed form

Allosteric fusion prototype
(Δ I-SM^{TR} intein):



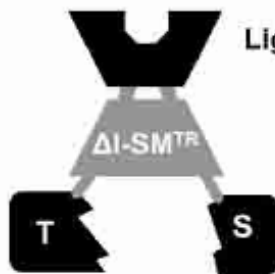
Intein in inactive state

Mycobacterium tuberculosis RecA
artificial mini-intein (Δ I-SM intein):



intein

Inactive intein chimera



Abolished TS activity

Ligand-induced splicing



Spliced intein

+



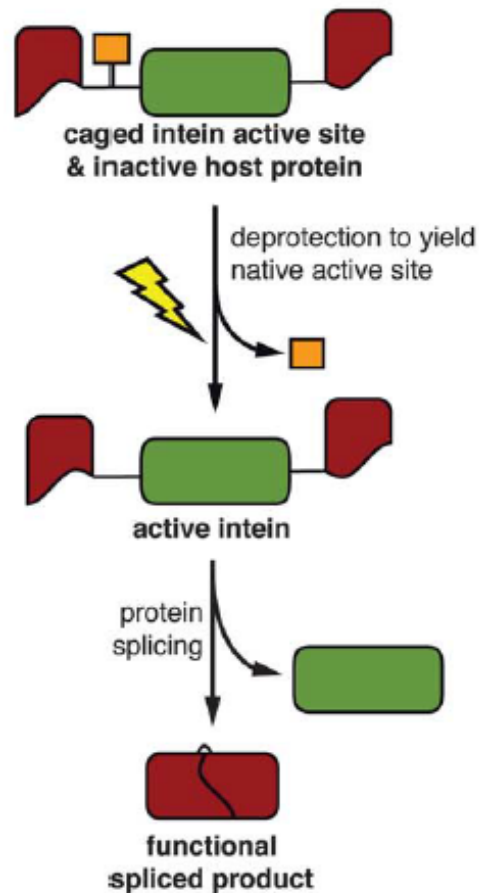
Functional TS

fungovalo to v bakteriích

extein

T4 thymidylát synthasa, kterou tak lze vypínat a zapínat pomocí ligandu

V tomto případě je funkce **inteinu** zablokována pomocí specifické skupiny a je odblokována proteázou či světlem.

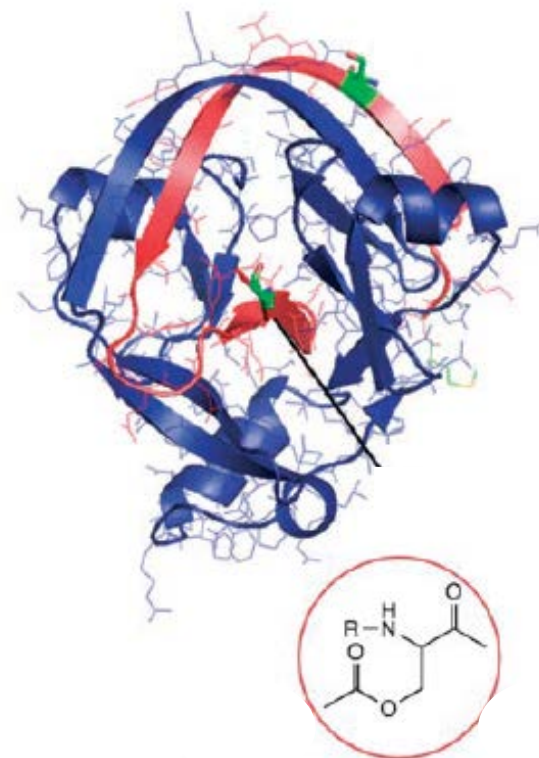
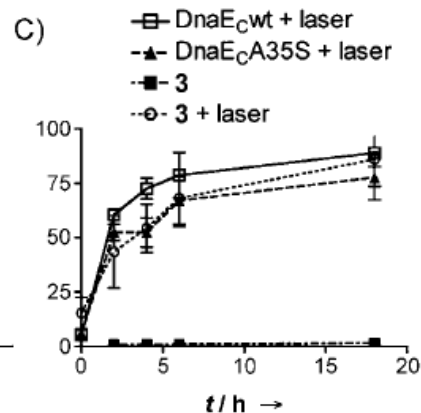
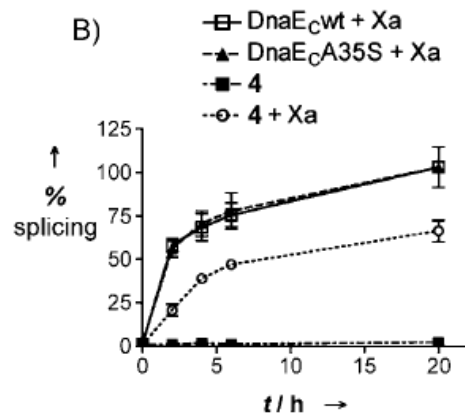
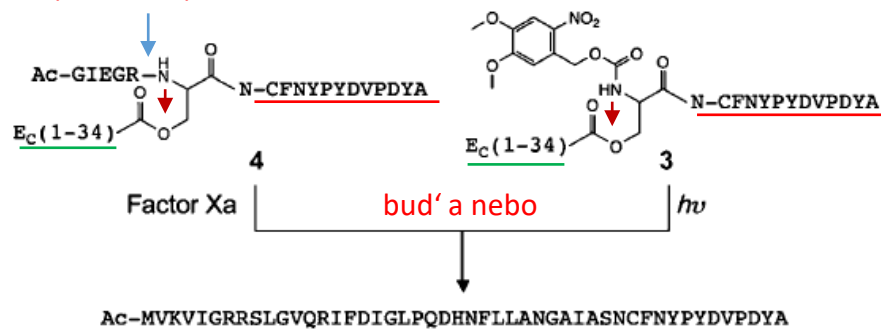


Activation of Protein Splicing by Protease- or Light-Triggered O to N Acyl Migration**

Angew. Chem. Int. Ed. 2008, 47, 7764–7767

Proteázou (Factor Xa)
štěpitelná skupina

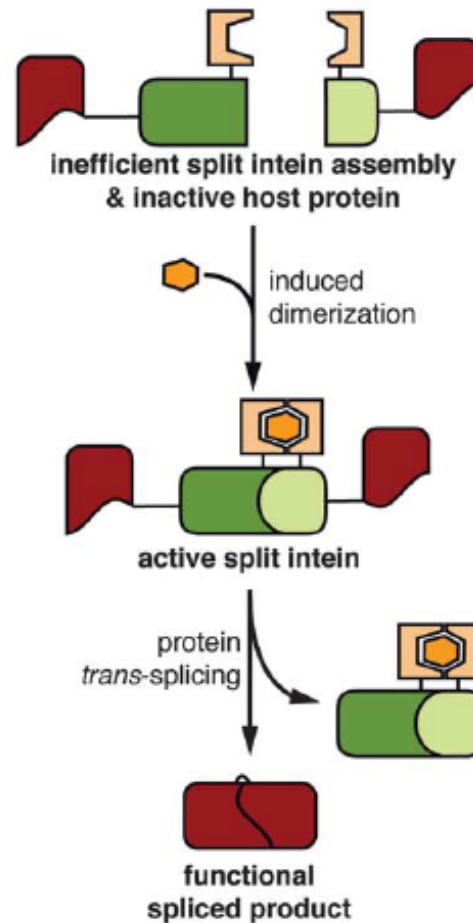
fotoštěpitelná skupina



Volný amin po odchránění napadá ester – O to N acyl shift

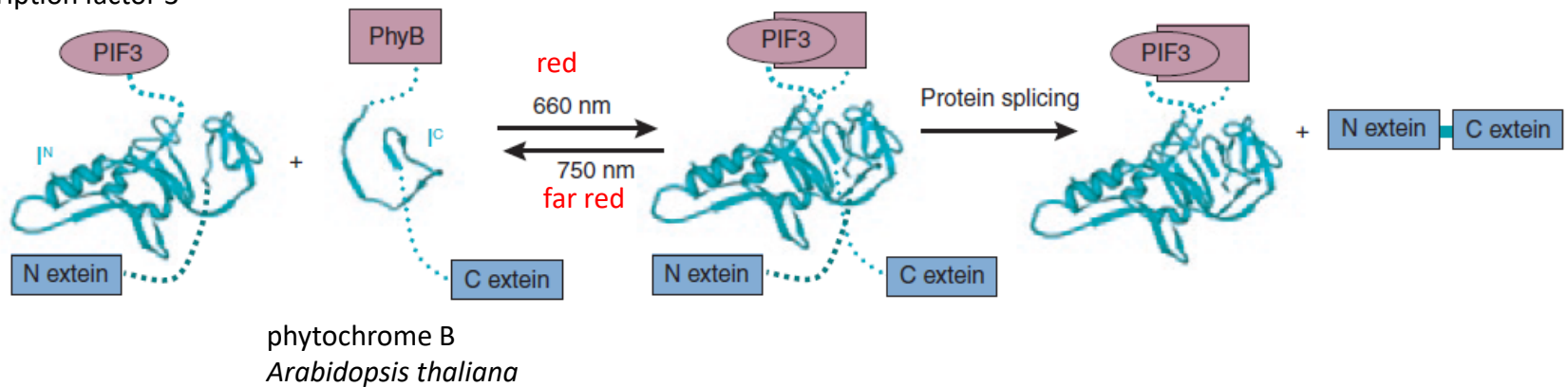
Produkovali tak řízeně (kdy chtěli) antimikrobiální peptid

Rozdělené **inteiny**, neschopné samy zfúzovat jsou modifikovány proteinem, který asociaci iniciuje v odpověď na vazbu ligandu či světlo.

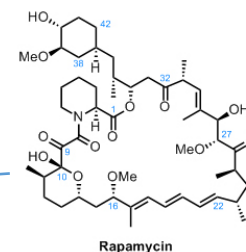


Rostlinné chromoproteiny (phytochrome proteins) asociují či disociují v odpověď na červené světlo.

Phytochrome interacting
transcription factor 3

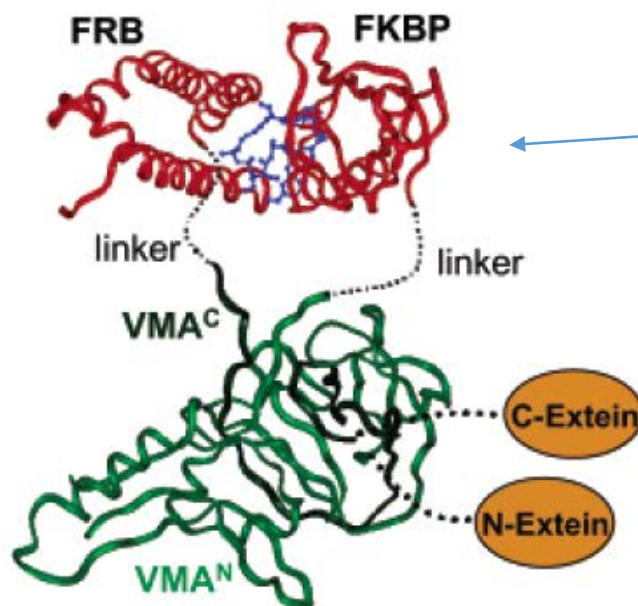


J. AM. CHEM. SOC. 2003, 125, 10561–10569 ■ 10561



dimerizuje proteiny
FRB a FKBP

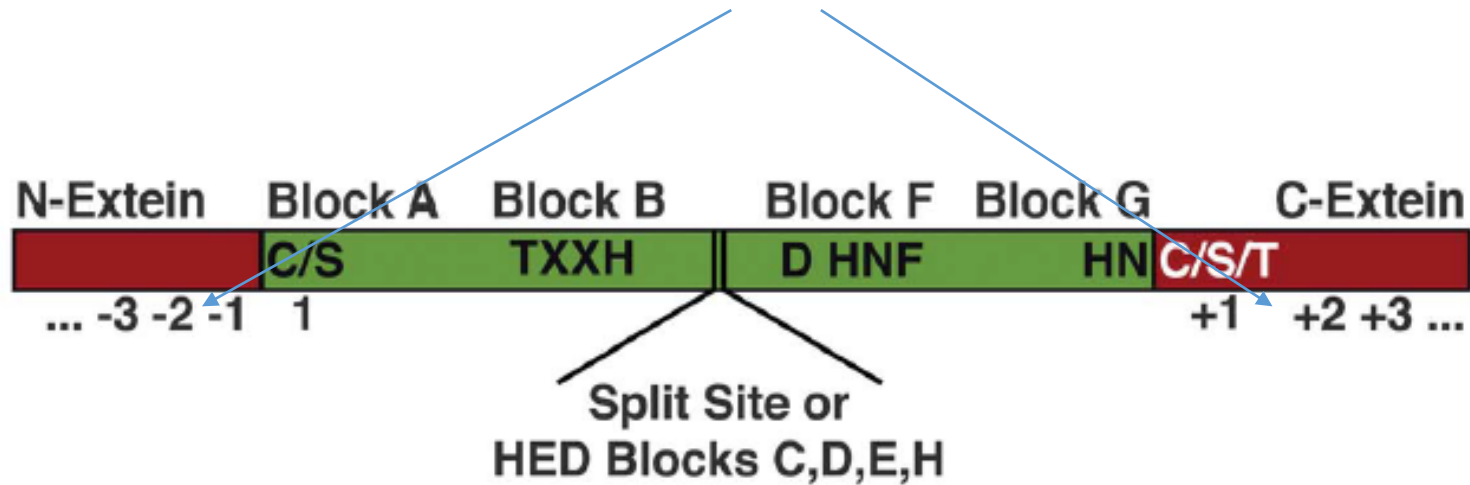
Inteinovou část
rozdělili na dvě půlky
VMA^C a VMA^N



VMA intein z *S. cerevisiae* byl rozdělen na dvě poloviny
VMA^C a VMA^N

Nedostatky inteinů

1. Závislost na sekvenci lokálních aminokyselin v exteinu



Záleží na pozicích -1, -2, 1, 2 atd. speciálně je citlivá pozice -1.

2. Pomalá kinetika splicingu

Většina dříve používaných inteinů provádí reakci spíše pomalu, v rozmezí minut až hodin. Nové studie ale odhalily řadu inteinů s kinetikou v řádu sekund.

Table 1 *In vitro* splicing kinetics and yields for several inteins^a

Intein name	Type	Temperature (°C)	k_{splice} (s ⁻¹)	$t_{1/2}$ ^b	Splicing yield	Reference
<i>MxeGyrA</i>	Contiguous	25	1.9×10^{-5}	10 h	>90% ^c	26
<i>PabPolIII</i>	Contiguous	70	1.6×10^{-5}	12 h	74% ^d	101
<i>MjaKlbA</i>	Contiguous	42	2.2×10^{-3}	5 min	64% ^d	102
<i>SspDnaB</i>	Artificially split	25	9.9×10^{-4}	12 min	32–56% ^e	103
<i>SceVMA</i>	Artificially split	25	1.2×10^{-3}	10 min	67–73% ^e	103
<i>SspDnaE</i>	Naturally split	37	1.5×10^{-4}	76 min	<50% ^c	70
<i>NpuDnaE</i>	Naturally split	37	3.7×10^{-2}	19 s	>90% ^f	70
<i>AvaDnaE</i>	Naturally split	37	3.1×10^{-2}	23 s	>90% ^f	70
<i>CraDnaE</i>	Naturally split	37	1.2×10^{-2}	58 s	>90% ^f	70
<i>CspDnaE</i>	Naturally split	37	1.8×10^{-2}	39 s	>90% ^f	70
<i>CwaDnaE</i>	Naturally split	37	5.0×10^{-3}	140 s	80–90% ^f	70
<i>MchtdNaE</i>	Naturally split	37	2.4×10^{-2}	29 s	>90% ^f	70
<i>OliDnaE</i>	Naturally split	37	1.6×10^{-2}	44 s	>90% ^f	70
<i>TerDnaE</i>	Naturally split	37	8.5×10^{-3}	82 s	>90% ^f	70
gp41-1	Naturally split	45	1.8×10^{-1}	4 s	85–95% ^d	100
gp41-8	Naturally split	37	4.5×10^{-2}	15 s	85–95% ^d	100
NrdJ-1	Naturally split	37	9.8×10^{-2}	7 s	85–95% ^d	100
IMPDH-1	Naturally split	37	8.7×10^{-2}	8 s	90–95% ^d	100