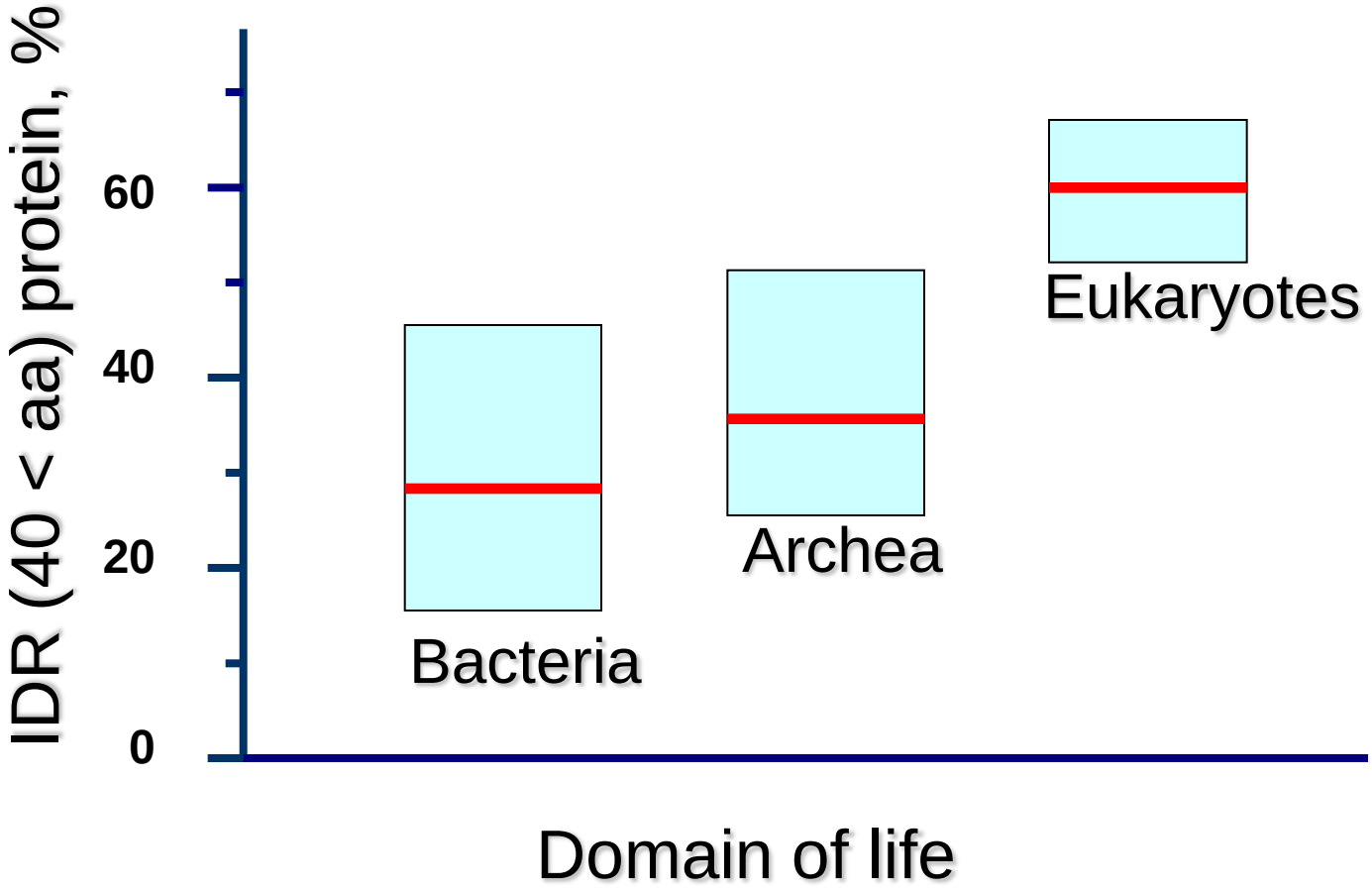
Mechanisms and polymorphism in recognition by intrinsically unstructured proteins

Monika Fuxreiter

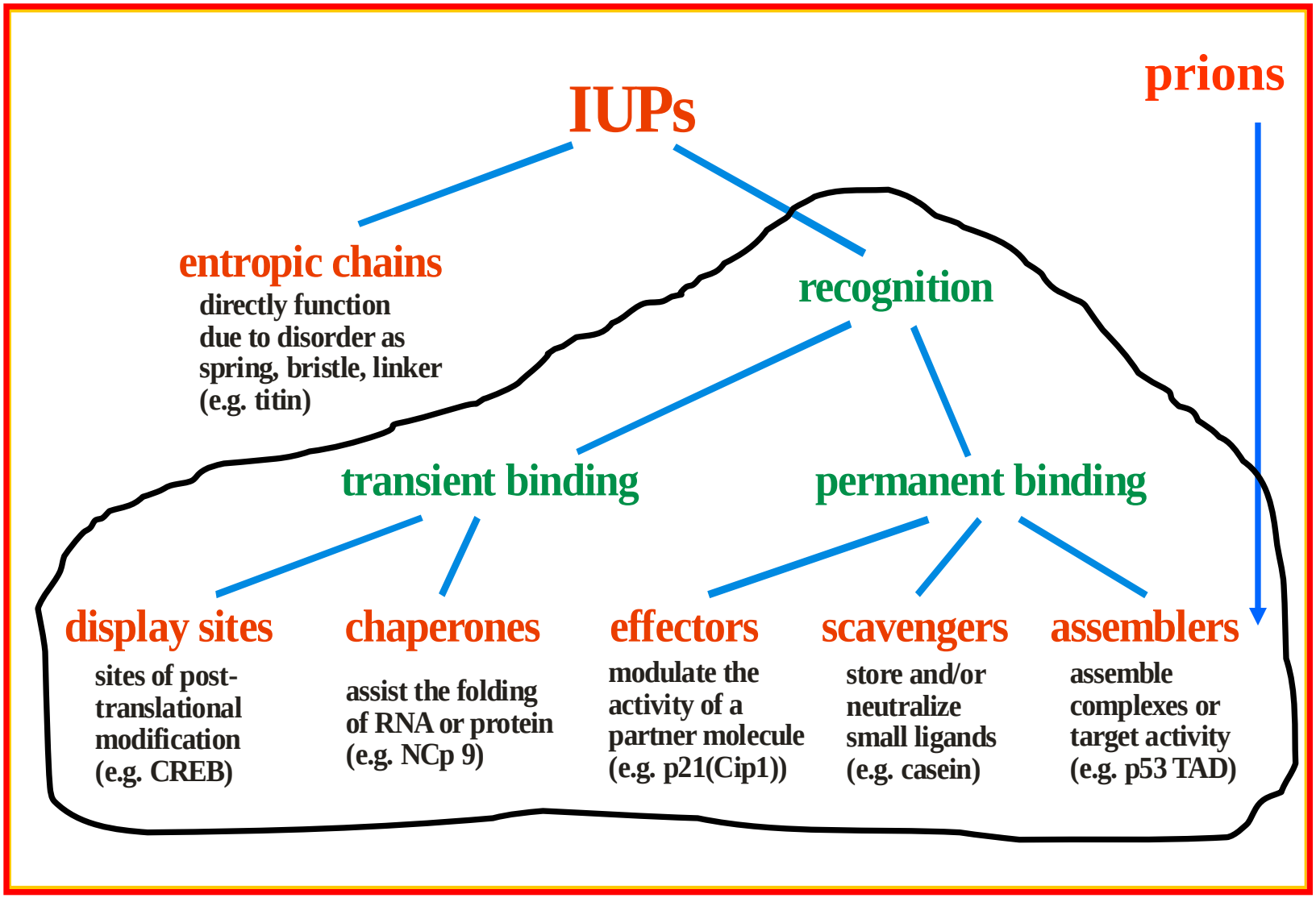
Institute of Enzymology, Budapest

Intrinsically disordered proteins (IDPs)



Vucetic et al. (2002) Proteins 52, 573

IDPs – functional classification



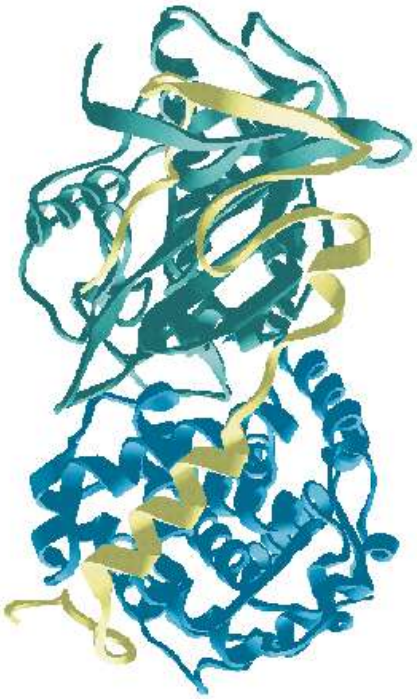
Tompa (2002) TiBS 27, 527; (2005) FEBS Lett 579, 3346

IDP recognition

Benefits

- highly specific
- high k_{on} (fast)
- high k_{off} (reversible)

} facile



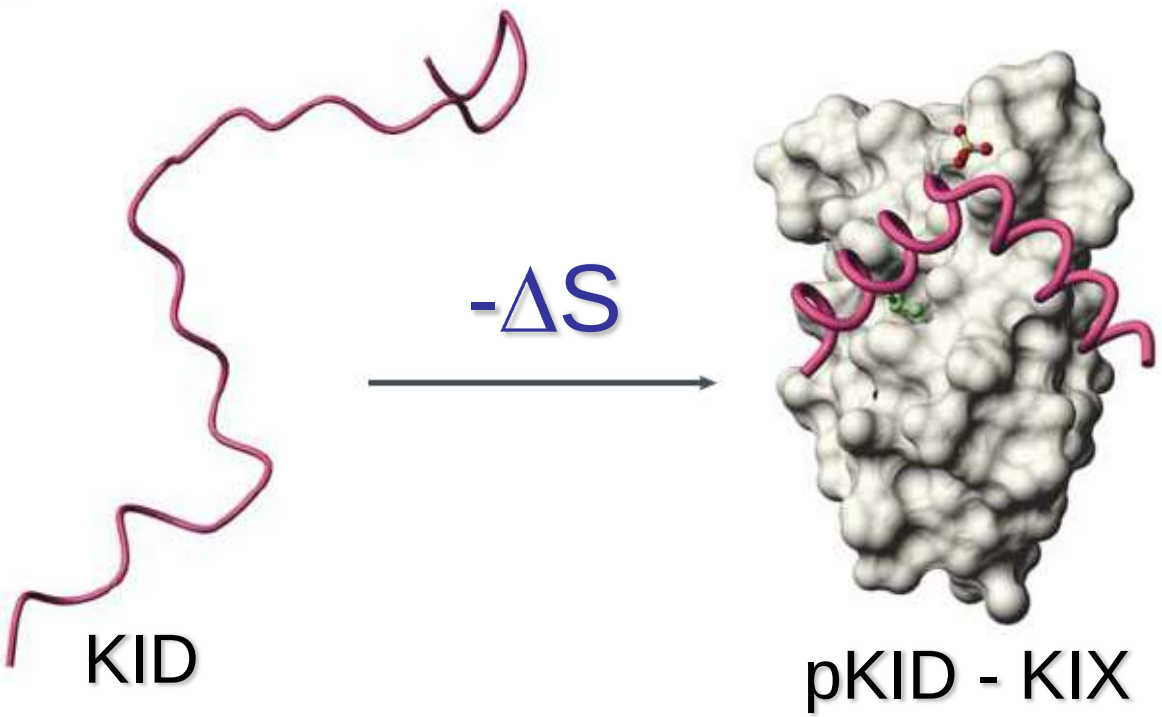
p27^{Kip1}

regulatory functions : signal transduction, regulation of gene expression, cell-cycle

Tompa (2002) TiBS 27, 527; (2005) FEBS Lett 579, 3346

IDP recognition

Problem



large entropic penalty ↔ fast, specific binding

Dyson (2005) Nat Rev Mol Cell Biol 6, 197

IDP recognition models

Thermodynamics

preformed elements

transient secondary structure elements

JMB (2004) 338, 1015

molecular recognition features

motifs that fold upon binding

JMB (2006) 362, 1043

primary contact sites

transiently exposed sites

Biochemistry (2005) 44, 3955

linear motifs

short recognition motifs

Bioinformatics (2007) 23, 950

Kinetics

IDP folding – coupled to binding

mechanism	secondary structure in free and bound form
Induced folding	→ independent
Conformational selection	→ correlation

Preformed elements (PSEs)

IUP	SwissProt code	Number of residues	Partner	Method	PDB code
CREB	P16220	28	CBP KIX	NMR	1kdx
DFF 45	O00273	89	DFF 40	NMR	1ibx
E-cadherin	P09803	57	β-catenin	X-ray	1i7x
FCP1	AAC64549.1*	21	TFIIF/RAP74	NMR	1onv
FnBPA	Q53971	24	Fibronectin	NMR	1o9a
IA_3	P01094	29	Proteinase A	X-ray	1dpj
Killer toxin β chain	P19972	77	Killer toxin α chain	X-ray	1kvd
MAP tau	P10636	13	Pin1 WW	NMR	1i8h
MAX	P25912	86	DNA	X-ray	1an2
OCA-B (Bob-1)	Q16633	22	Oct-1 POU/DNA	X-ray	1cqt
$p27^{Kip1}$	P46527	69	CycA-Cdk2	X-ray	1jsu
p53	P04637	11	MDM 2	X-ray	1ycq
Phe-tRNA synthetase α	P27001	79	Phe-tRNA synthetase β + tRNA	X-ray	1eiy
PKI	P04541	20	PKA	X-ray	1apm
RB3	Q9H169	91	tubulin	X-ray	1ffx
RNA pol II	P04050	17	mRNA capping enzyme	X-ray	1p16
SNAP 25	P13795-2	77	neuronal fusion complex	X-ray	1sfc
SV 40 virus coat protein	P03087	66	assembled coat	X-ray	1sva
$TAF_{II}230$	P51123	67	TBP	NMR	1tba
TBS virus coat chain C	P11795	34	assembled coat	X-ray	2tbv
Tcf3	CAA67686*	41	β-catenin	X-ray	1g3j
Tcf4	Q9NQB0	24	β-catenin	X-ray	1jpw
Troponin I	P19429	17	Troponin C	NMR	1lxf
Vitamin D3 receptor	P11473	89	DNA	X-ray	1kb2

* GenBank code

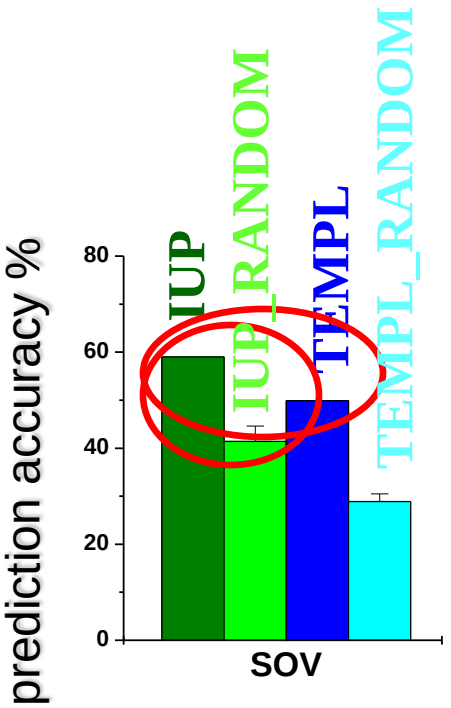
IUP_STR: IUP structures in the bound form
IUP_SEQ: full length sequences of IUPs
IUP_RAN: randomized IUP sequences

TEMPL_STR: template (partner) structures in the complex
TEMPL_SEQ: template sequences
TEMPL_RAN: randomized IUP sequences

GLOB_STR: globular proteins from PDB_select database

Fuxreiter et al. (2004) J Mol Biol 338, 1015-1026

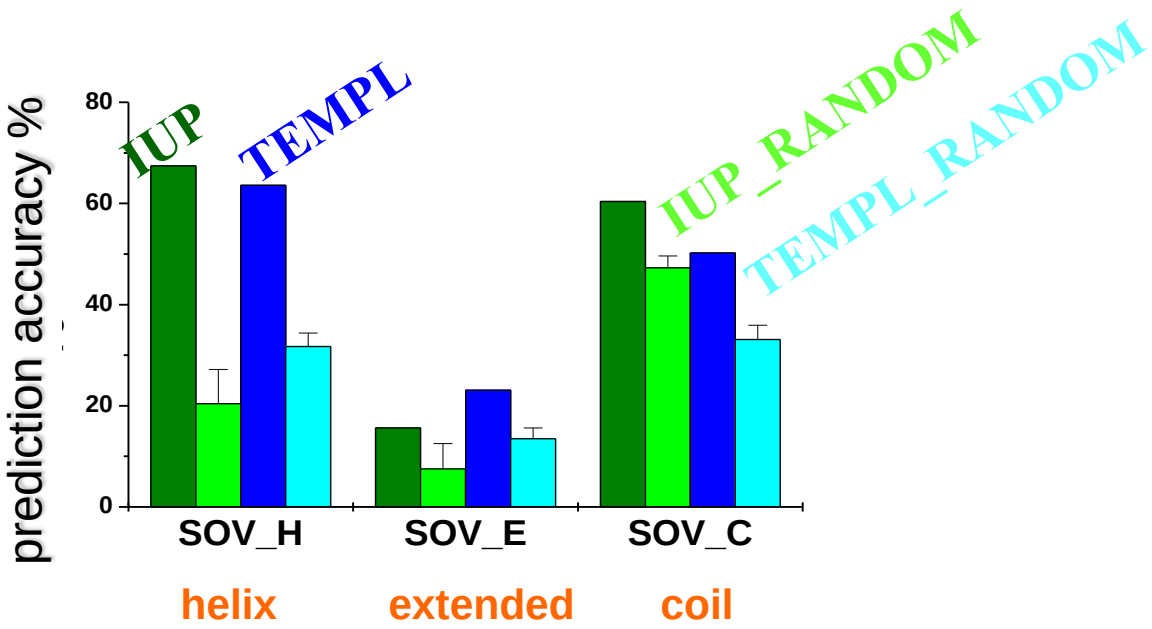
Preformed elements (PSEs)



- IUPs are not random
- presage their final state

Fuxreiter et al. (2004) J Mol Biol 338, 1015-1026

Preformed elements (PSEs)

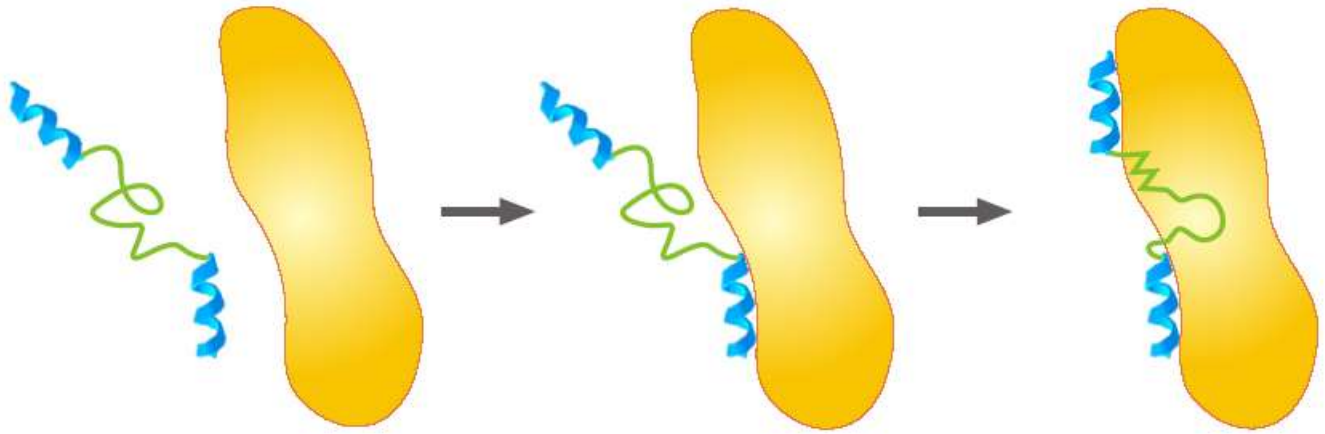


- modular architecture
- highest accuracy for helices

Fuxreiter et al. (2004) J Mol Biol 338, 1015-1026

Preformed elements (PSEs)

preformed elements: transient secondary structures

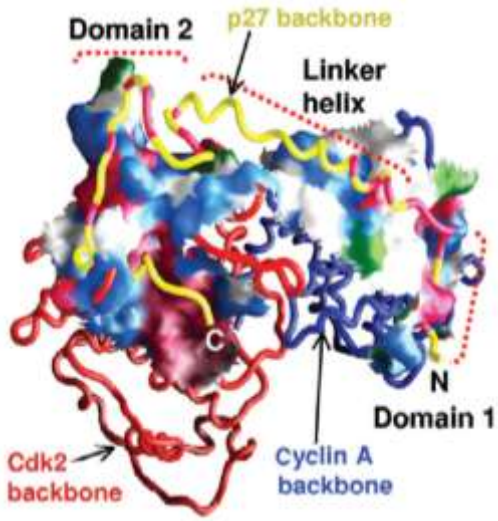


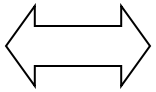
Benefits:

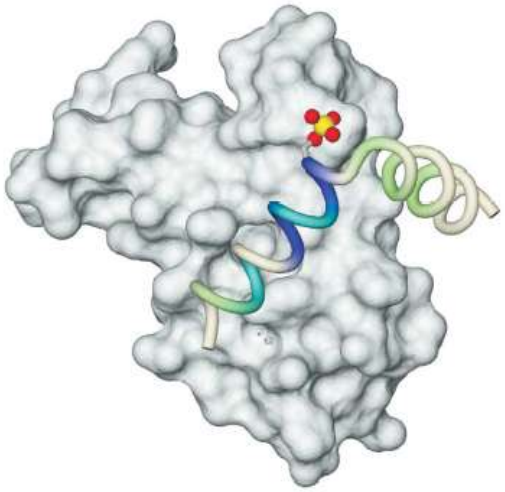
- reduce entropic penalty
- enhance complementarity

Fuxreiter et al. (2004) J Mol Biol 338, 1015-1026

Experimental PSE examples







FlgM, p53, $p27^{Kip1}$, GCN4, CFTR, Measles virus N

pKID (CREB) – KIX (CBP)

Lacy et al. (2005) Nat. Struct. Mol. Biol 11, 358
Sivakolundu et al. (2005) J Mol Biol 353,

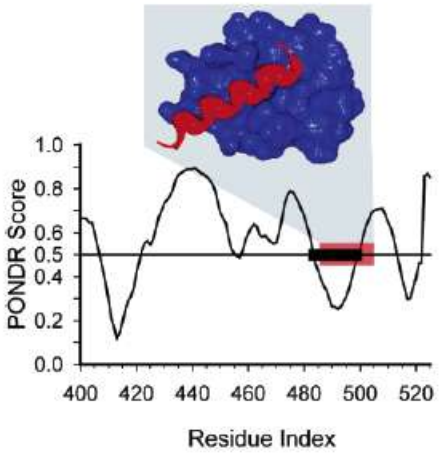
Sugase et al. (2007) Nature 447, 1021
Turjanski et al. (2008) Plos Comp Biol 4, e1000060

Molecular recognition elements (MoREs)

α-MoRE: fold to an α-helix upon binding

Prediction:

- local order within larger disorder
- secondary structure preferences
- physico-chemical features



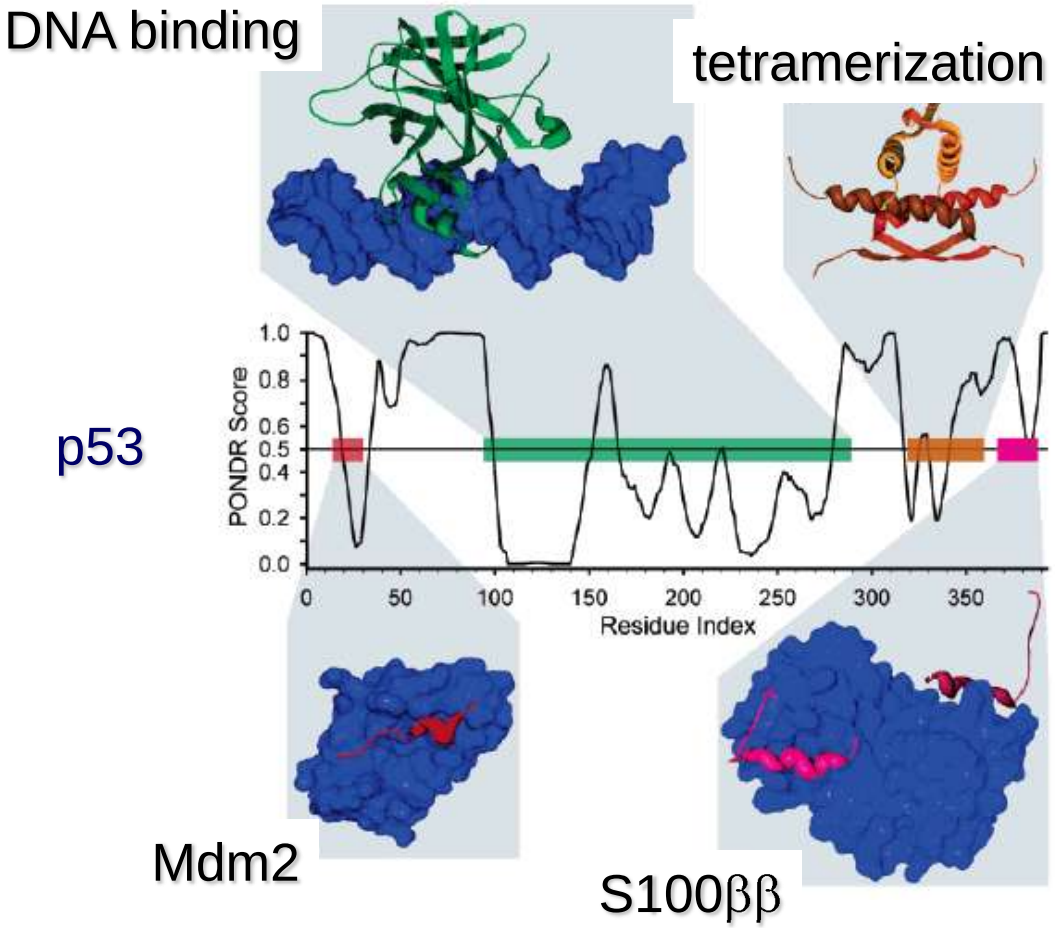
PONDR VL-XT

Measles virus protein N

Oldfield et al. (2005) Biochemistry 44, 12454
Cheng et al. (2007) Biochemistry 46, 13468

Molecular recognition elements (MoREs)

α-MoRE: fold to an α-helix upon binding



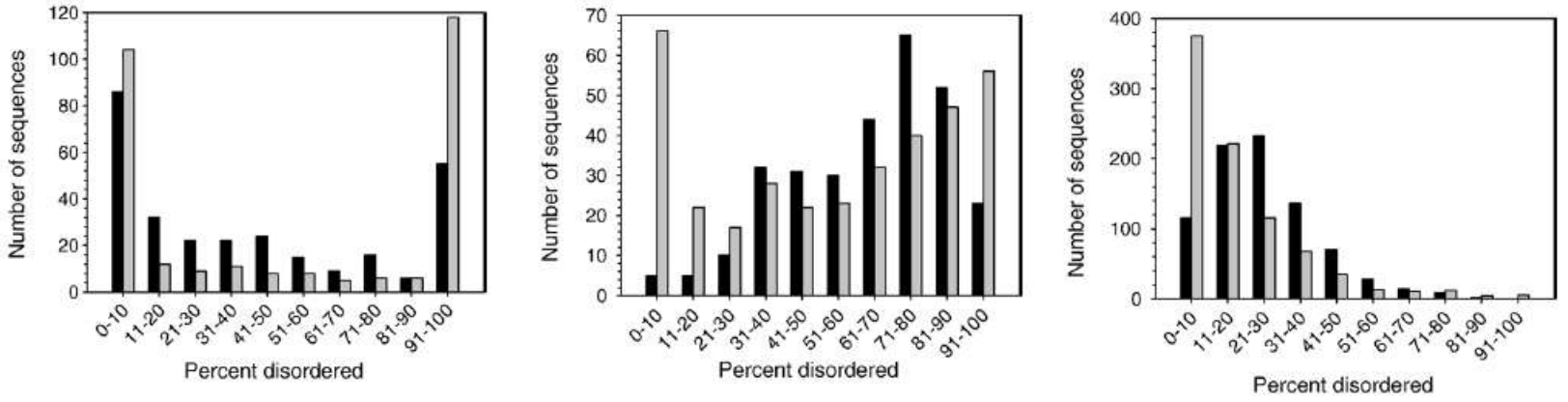
Oldfield et al. (2005) Biochemistry 44, 12454

Molecular recognition features (MoRFs)

MoRF: disordered segment that fold upon binding (α, β, ι)

Database

- 372 segments derived from protein data bank (PDB)
- 10aa <, < 70 aa; partner > 100 aa
- include also S-S bonds and transmembrane segments



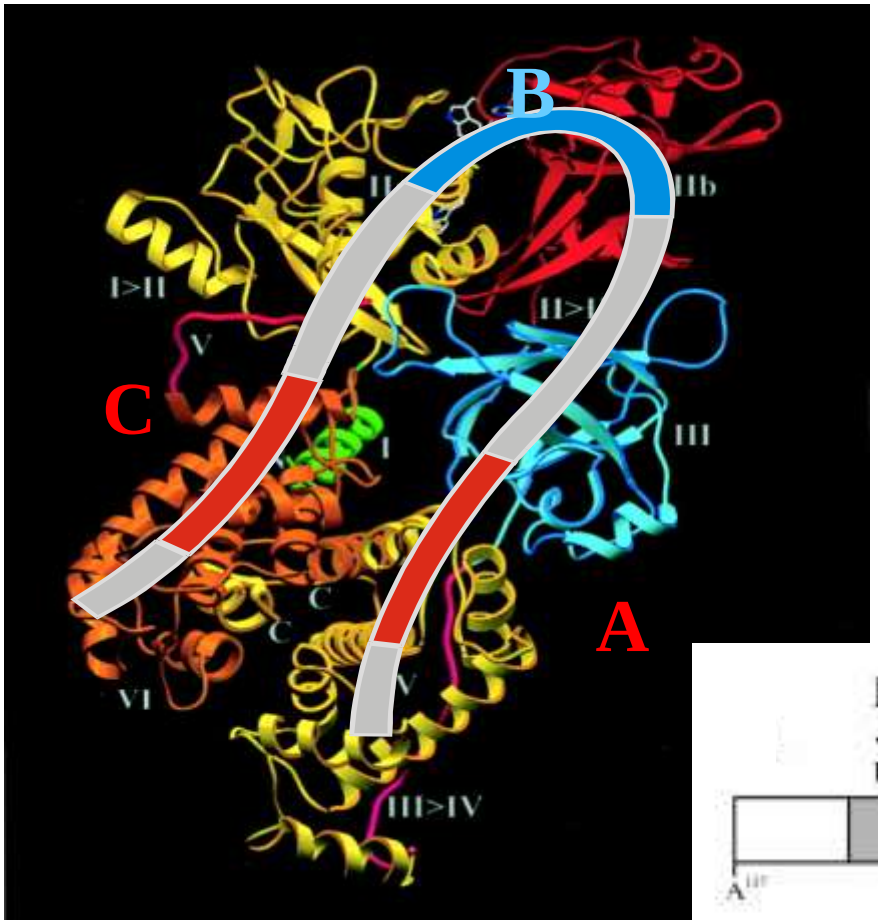
MoRF

MoRF-containing proteins

globular proteins

Mohan et al. (2006) J Mol Biol 362, 1043

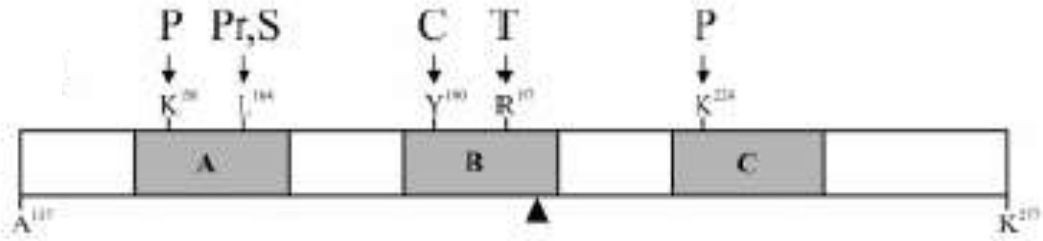
Primary contact sites (PCSs)



Calpastatin (CSD1)

inhibitor of calpain

limited proteolysis

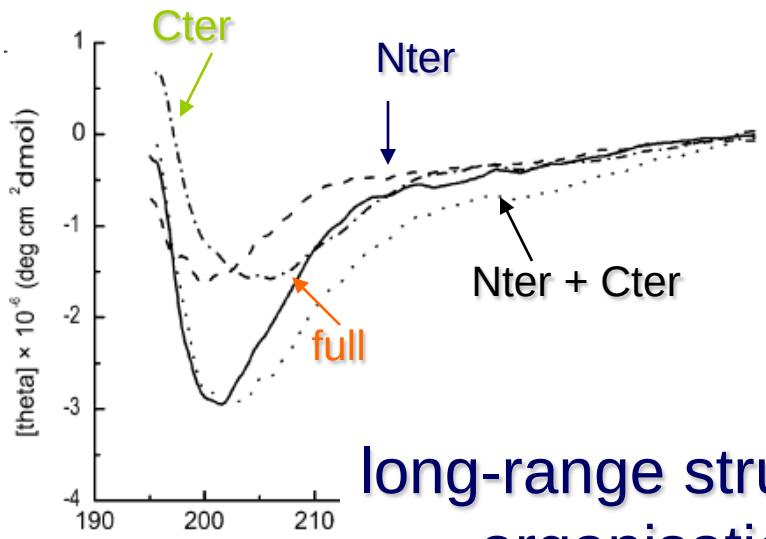


Csizmok et al. (2005) Biochemistry 44, 3955-64

Primary contact sites (PCSs)

Calpastatin (CSD1)

CD spectra

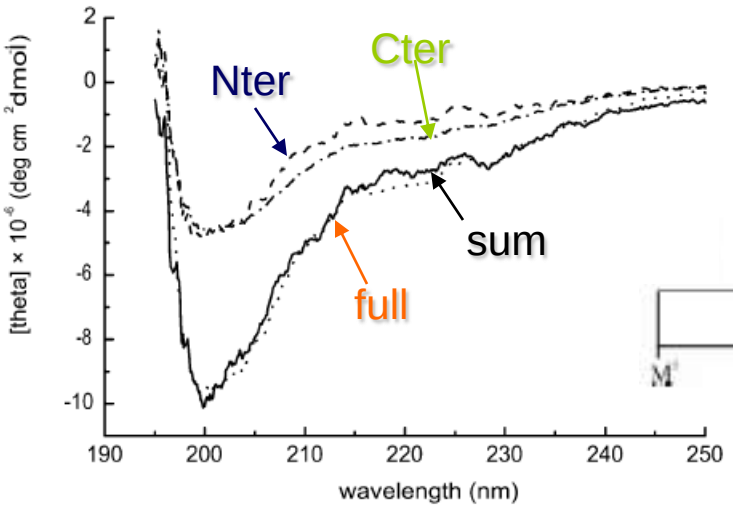


long-range structural organisation

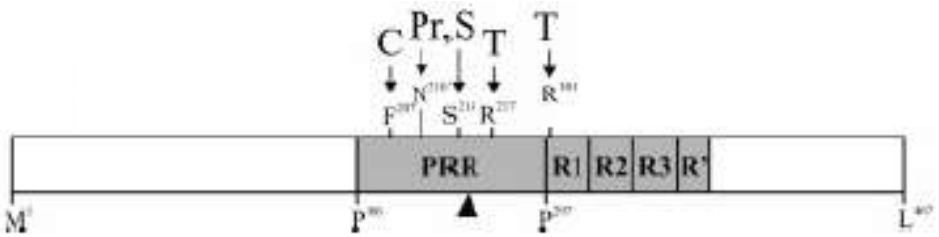
Csizmok et al. (2005) Biochemistry 44, 3955-64

Primary contact sites (PCSs)

MAP2c



limited proteolysis



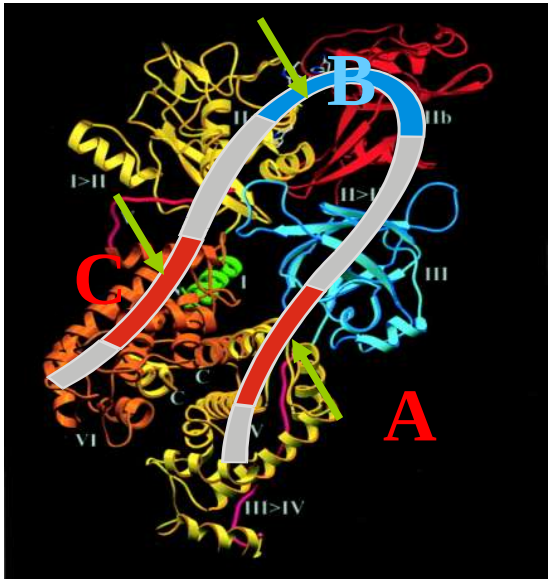
no long-range organisation
only local, PPII conf

Csizmok et al. (2005) Biochemistry 44, 3955-64

Primary contact sites (PCSs)

PCS: transiently exposed site

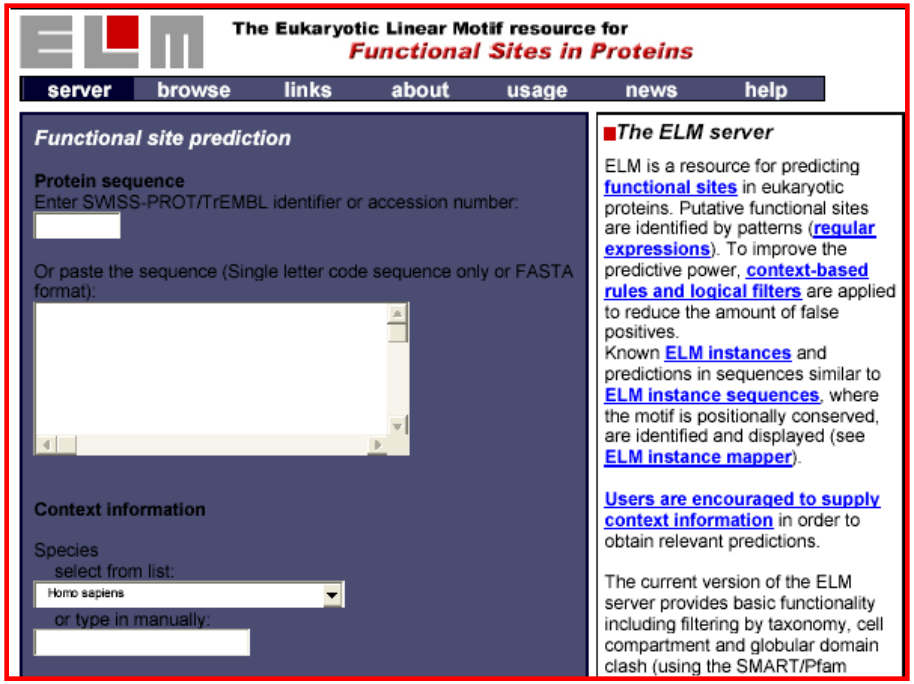
- likely provide first contact point
- no secondary structures needed



Csizmok et al. (2005) Biochemistry 44, 3955-64

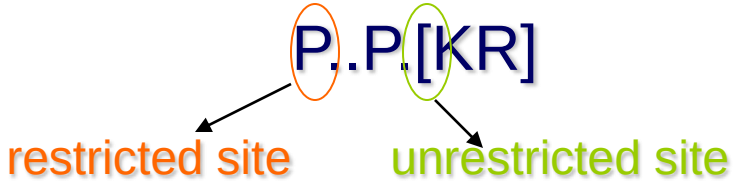
Linear motifs (LMs)

LM: short recognition motifs



(http://elm.eu.org)

- PPI networks
- short, low-complexity motifs
- low-affinity



Puntervoll et al (2003) Nucl Acids Res 31, 3625

Linear motifs (LMs)

- SH3 interaction motifs
- calmodulin binding sites
- 14-3-3 partners
- phosphorylation sites

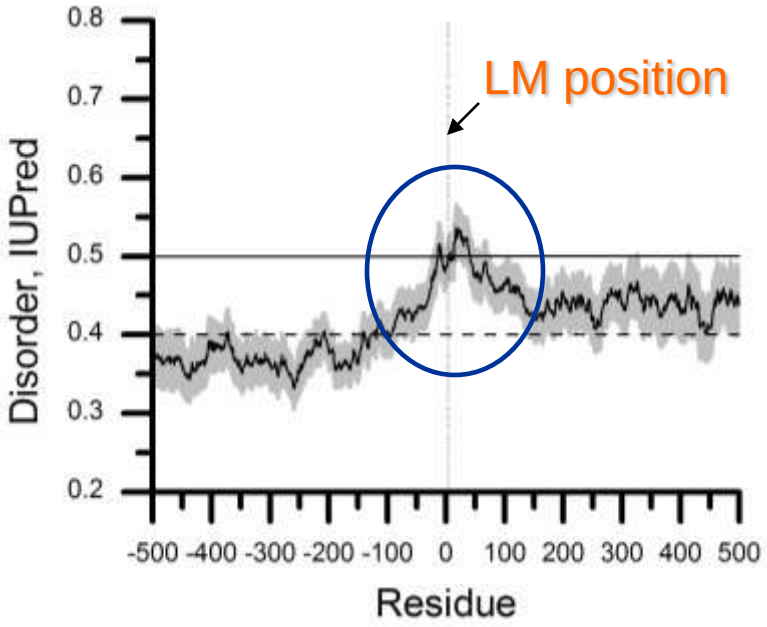
⇨ in disordered regions

GENERAL?

Bustos et al (2006) Proteins 63, 35
Radivojac et. al (2006) Proteins 63, 398-410
Beltrao et al (2005) PLoS Comput Biology 1, 202
Iakucheva et al (2004) Nucl. Acids Res. 32, 1037

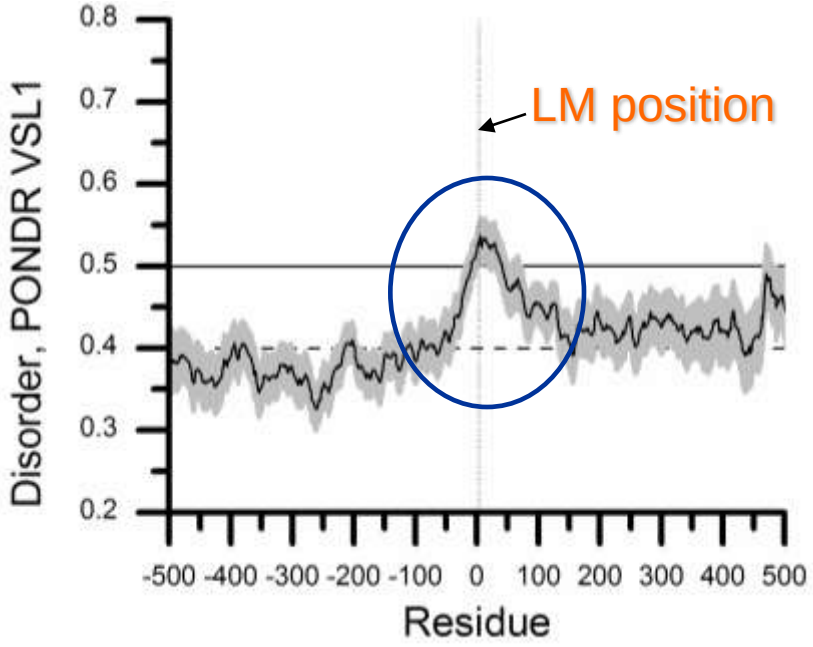
Linear motifs (LMs)

average disorder profiles



IUPred

iupred.enzim.hu

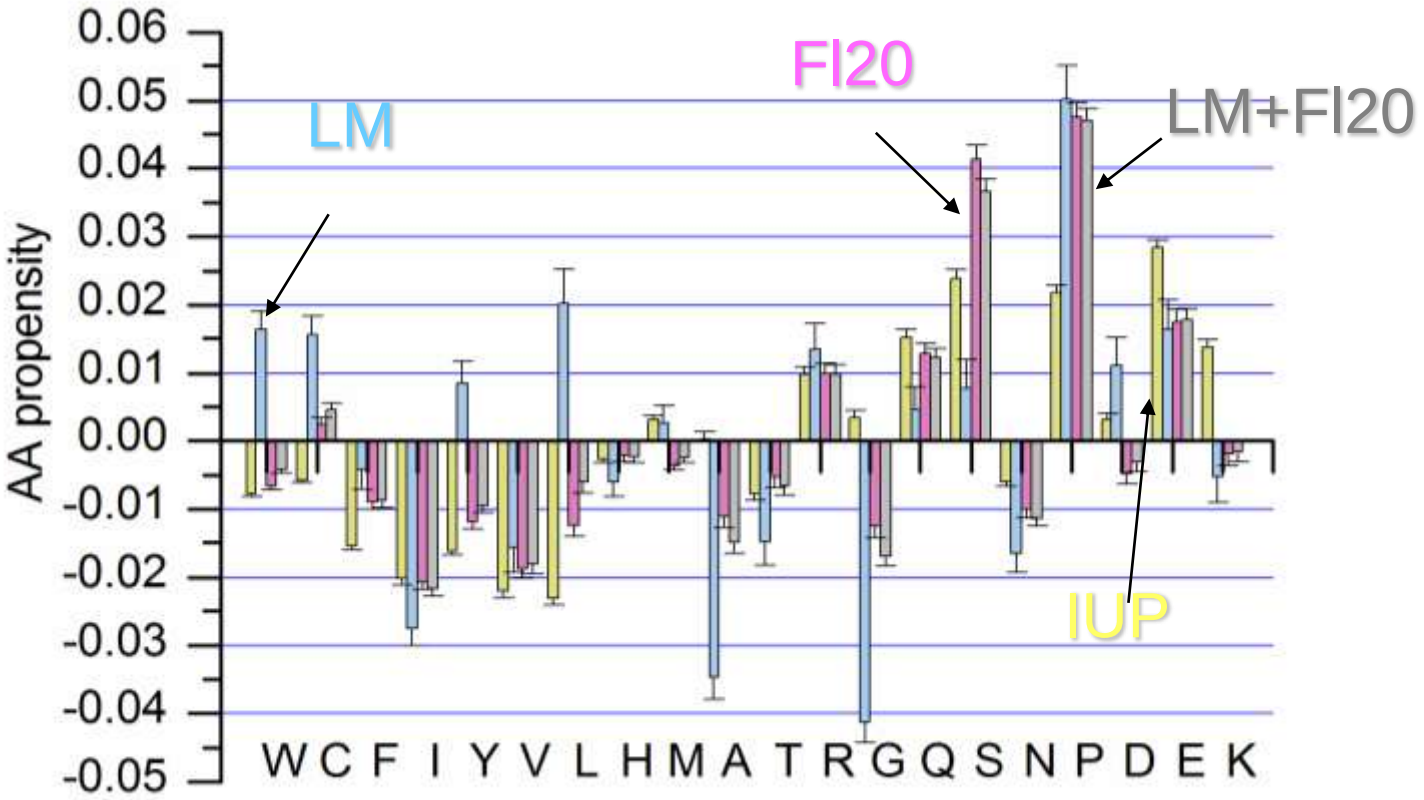


PONDR-VSL1

pondr.com

Fuxreiter et al. (2007) Bioinformatics 23, 950

Linear motifs (LMs)

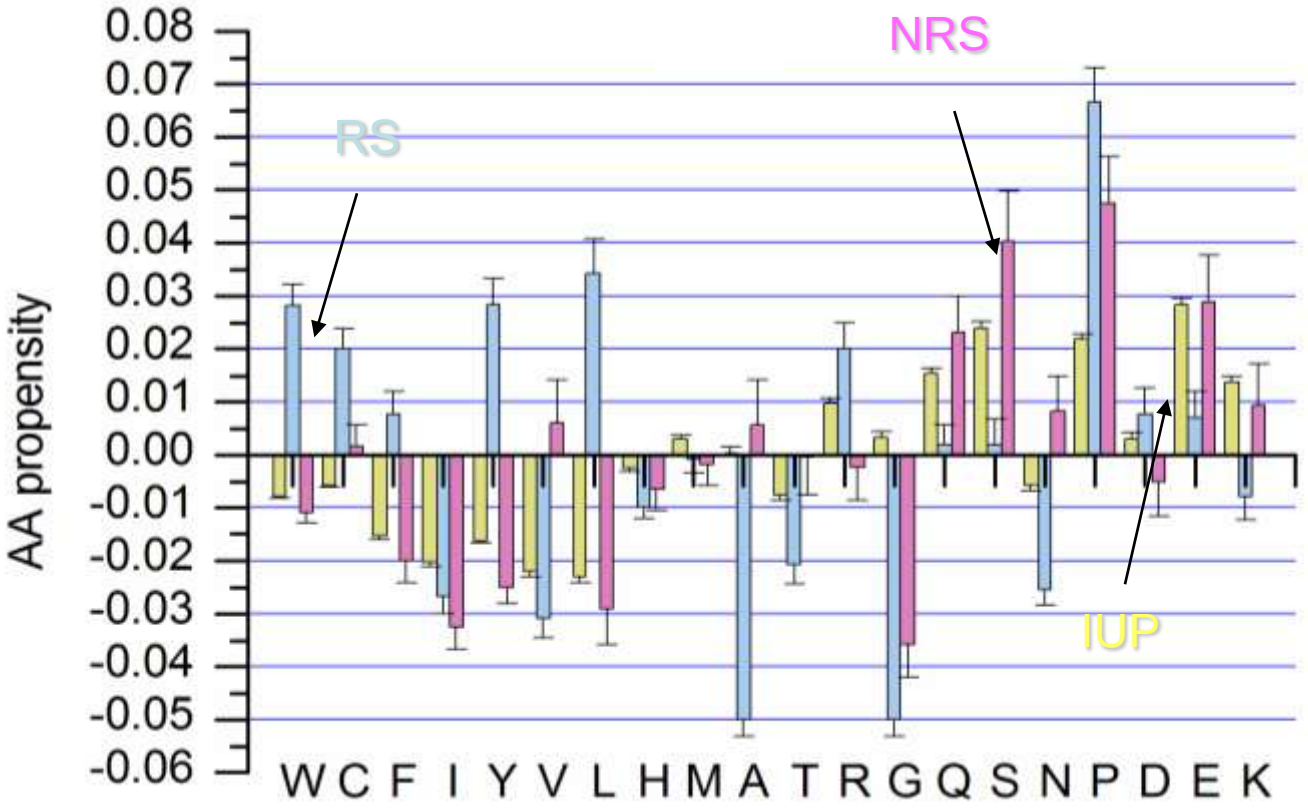


Flanking regions are similar to IUPs

LMs are strange

Fuxreiter et al. (2007) Bioinformatics 23, 950

Linear motifs (LMs)



SH3 binding site

P..P.[KR]

restricted site

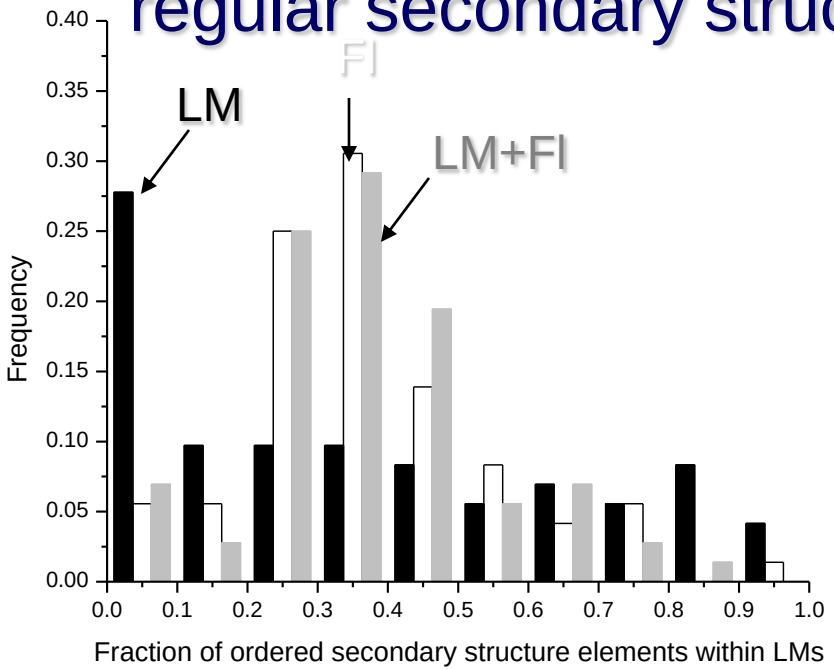
unrestricted site

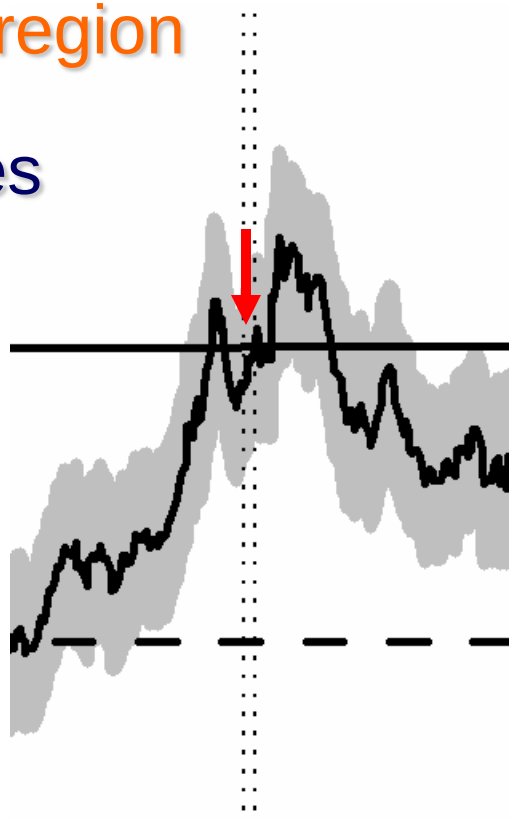
Fuxreiter et al. (2007) Bioinformatics 23, 950

Linear motifs (LMs)

LM: short, locally ordered segment
embedded into a disordered region

regular secondary structures

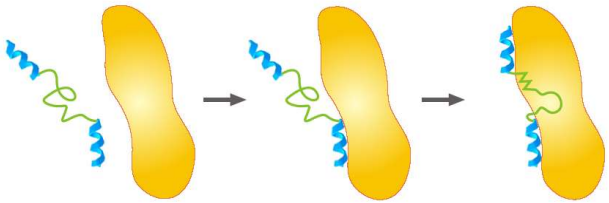




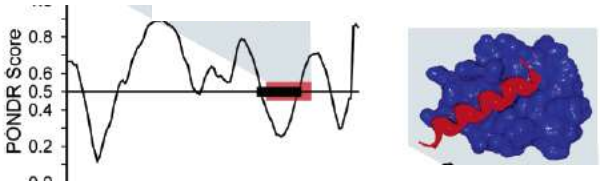
Fuxreiter et al. (2007) Bioinformatics 23, 950

IDP recognition models

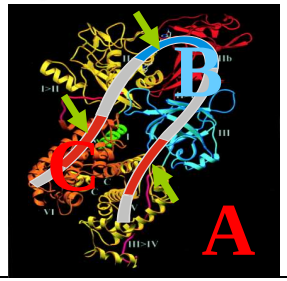
preformed elements



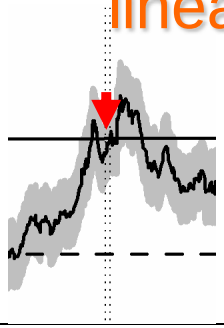
molecular recognition features



primary contact sites

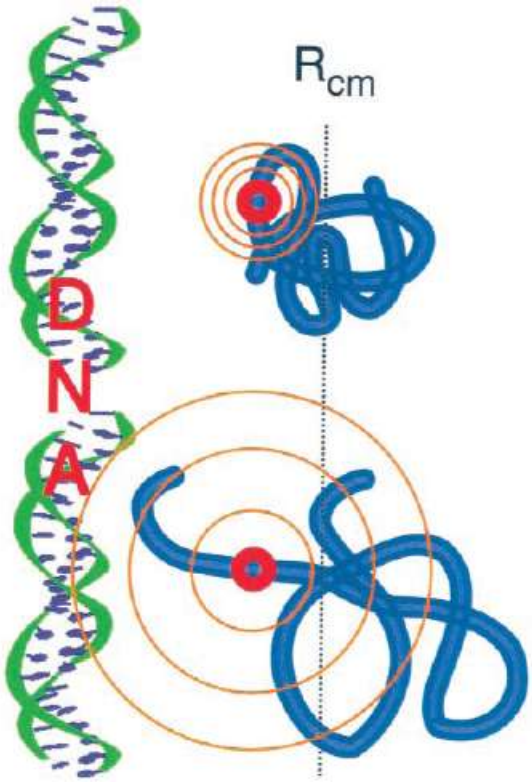


linear motifs



P..P[KR]

IDP recognition – facile binding



fly-casting

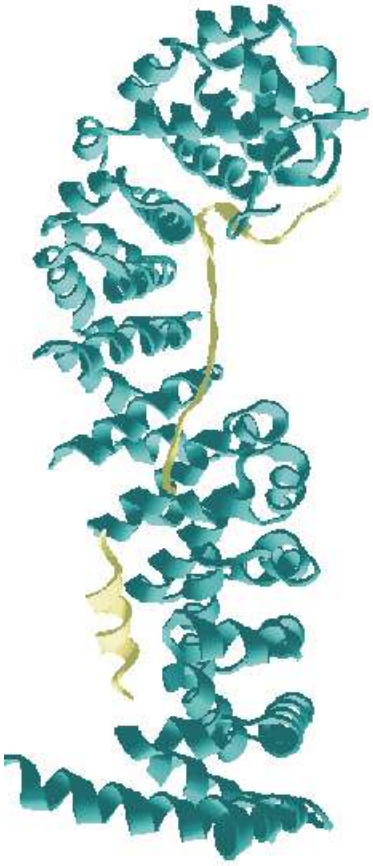
- distinguished short motifs
- secondary structure elements
- large capture radius
- multiple contact sites

⇩

kinetics

Shoemaker et al. (2000) PNAS 97, 8868

IDP recognition – consequences



Tcf3

- weak sequence requirements
- given aa composition

⇩

- less sensitive to sequence
- easy to turn on/off

⇩

structural and functional malleability

IDP recognition – consequences

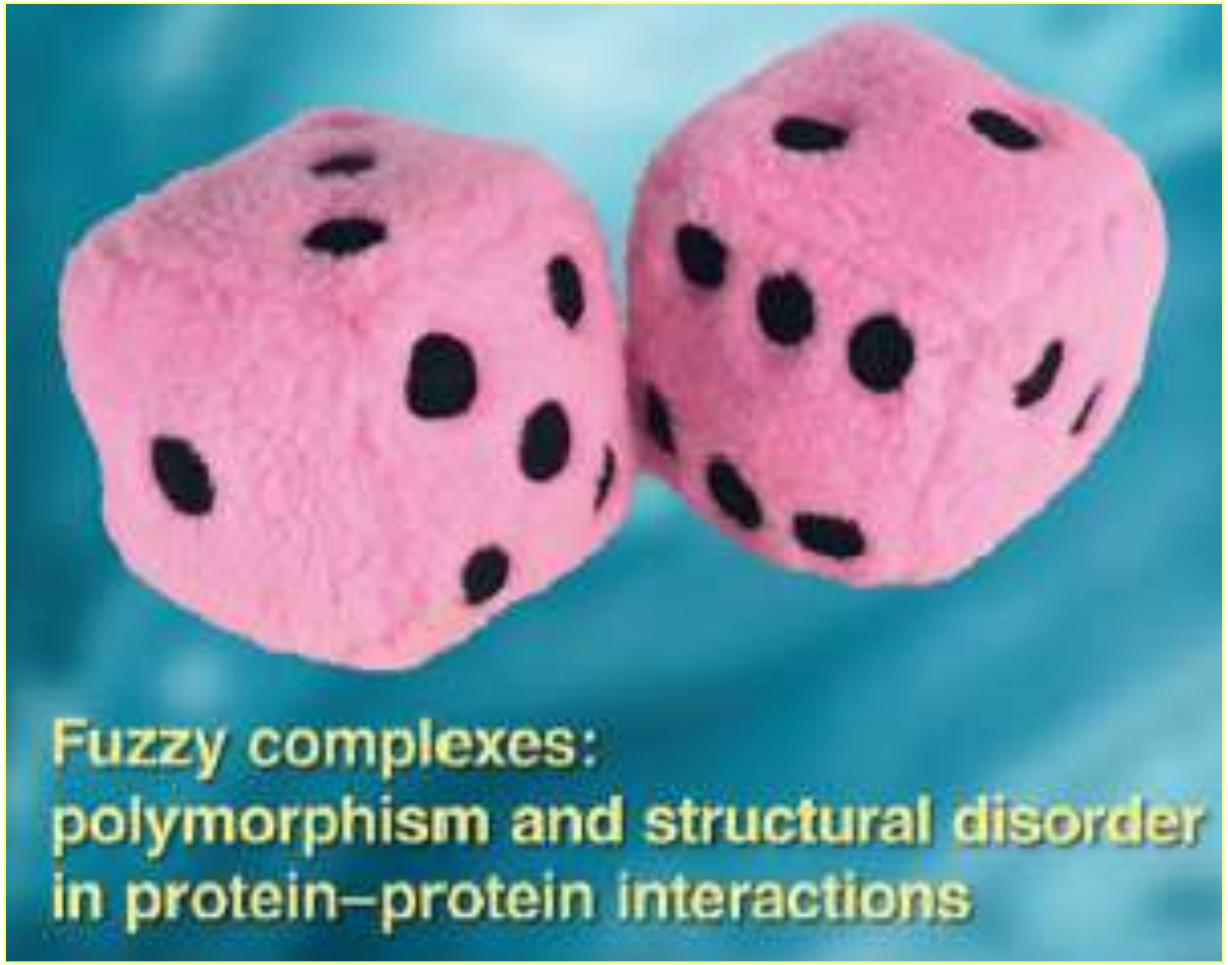
folding coupled binding

restore structure-function paradigm

NO folding coupled to binding – disorder in bound form

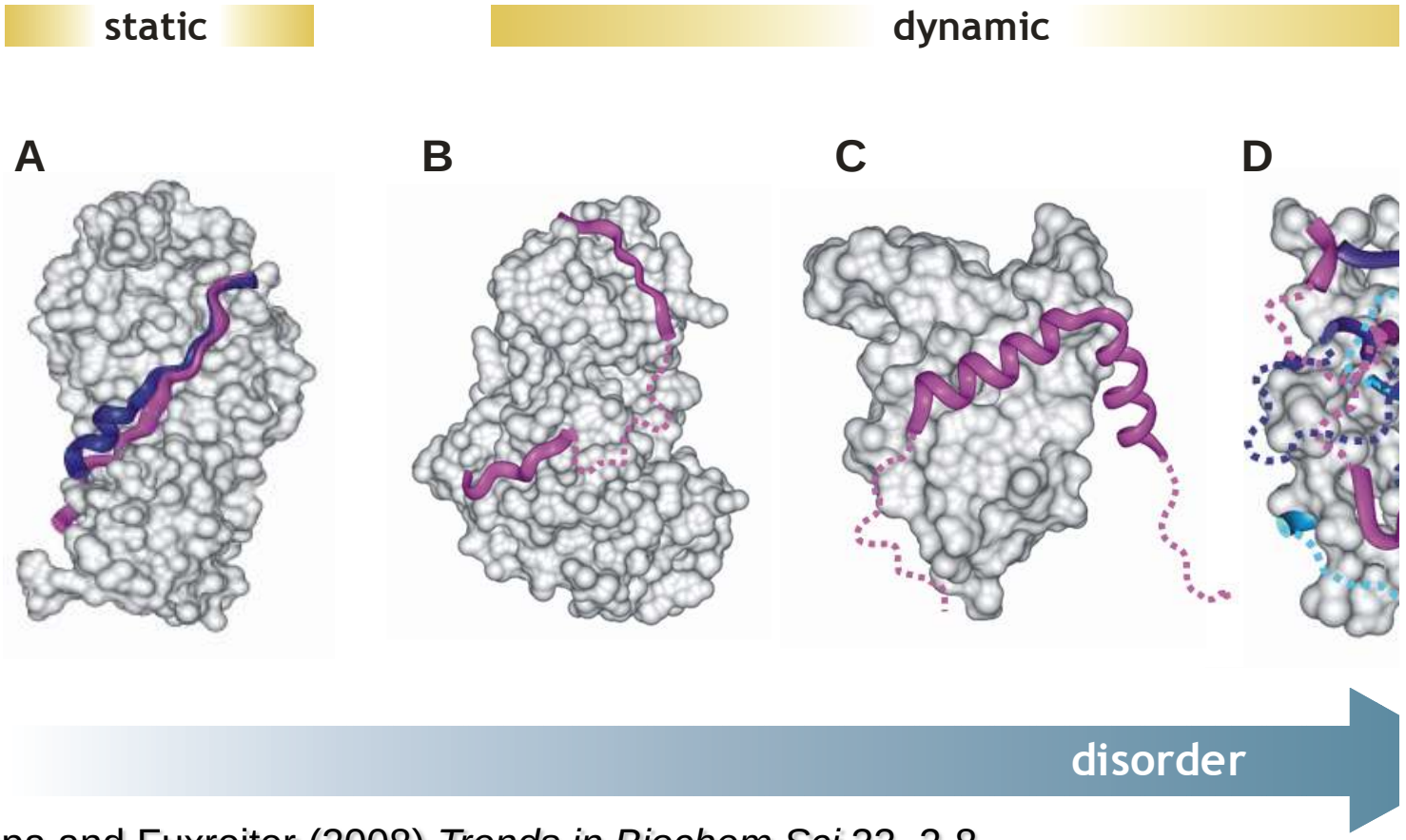
disorder paradigm

IDP recognition – consequences



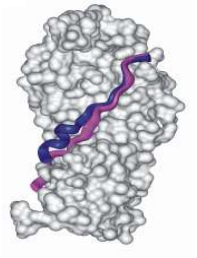
Fuzzy complexes

Fuzziness: structural and functional ambiguity in bound form



Tompa and Fuxreiter (2008) *Trends in Biochem Sci* 33, 2-8

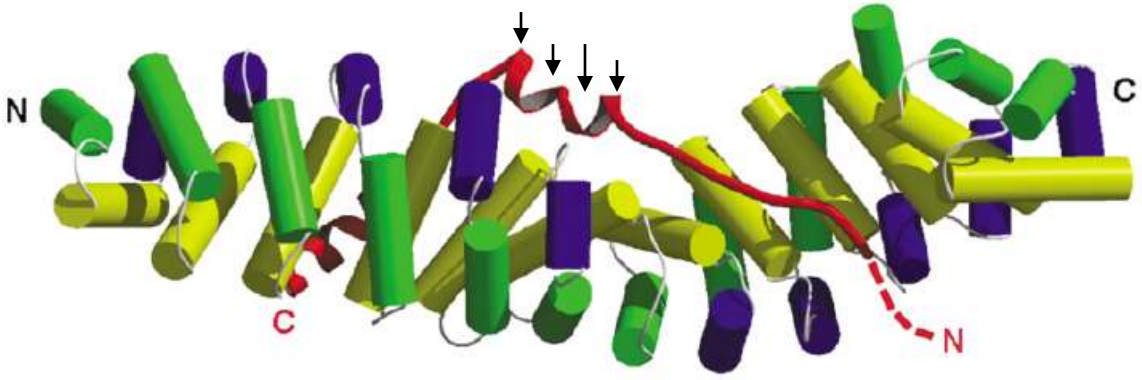
Fuzzy complexes



static different bound structures can be determined

polymorphic

Tcf4 - β-catenin mutagenesis studies

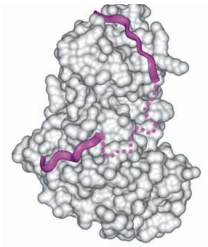


Glu-24, Glu-26,Glu-28,Glu-29
all effectively eliminate binding

alternative binding modes

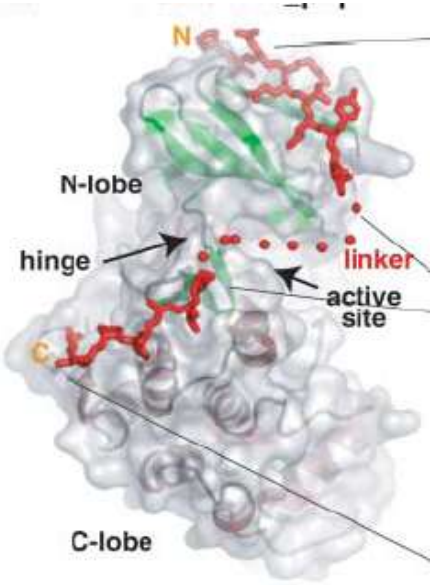
Graham et al. (2001) *Nat Struct. Biol.* 8, 1048

Fuzzy complexes



dynamic interconvert between many structures

clamp



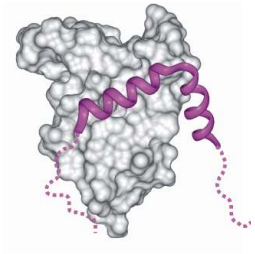
absence of linker – no binding

presence of linker – K_d=4 μM

Ste5 – Fus3

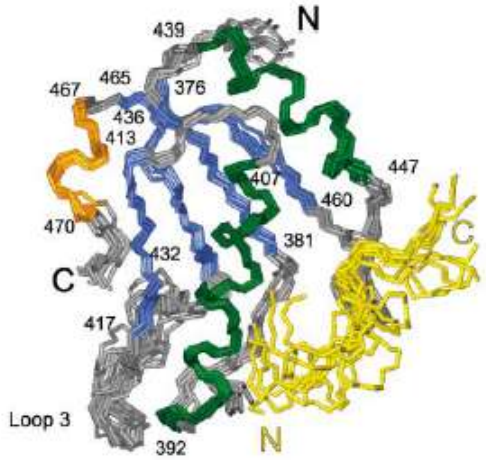
Bhattacharyya et al. (2006) *Science.* 311, 822

Fuzzy complexes



dynamic interconvert between many structures

flanking



SF1 – UA2F[65]

	K_d
full protein	11.8 nM
bound segment	23.8 nM
physically contacting residues	55.6 nM

Selenko et al. (2003) *Mol Cell* 11, 965

Fuzzy complexes

dynamic interconvert between many structures

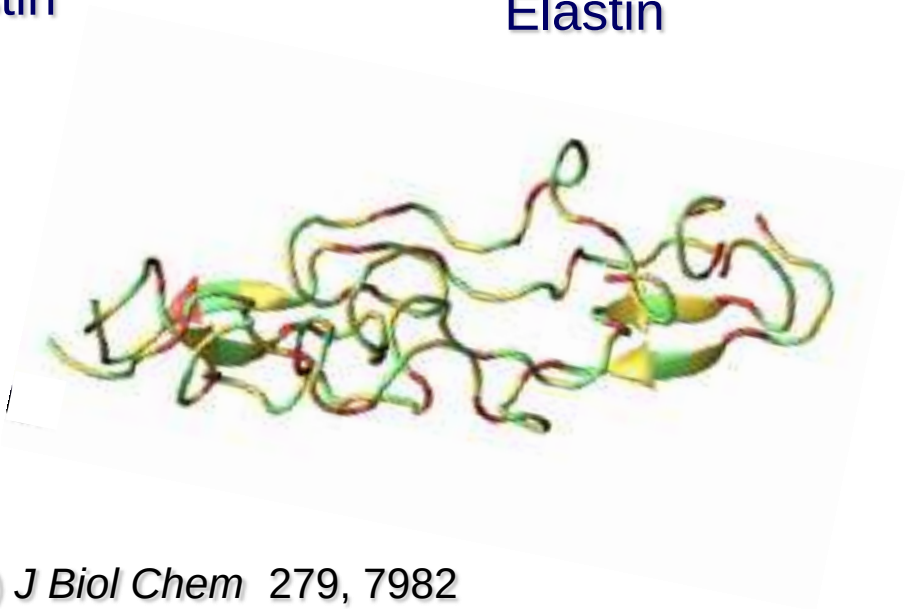
random

IDP

partner

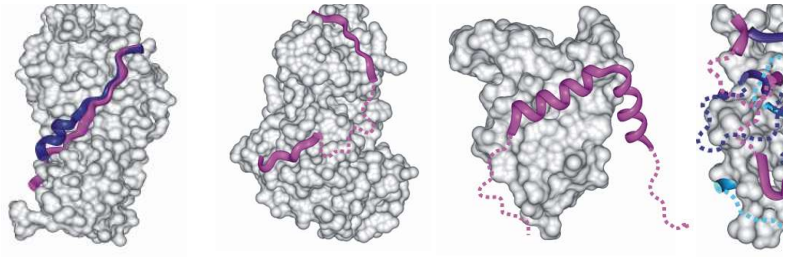
T-cell receptor ζ chain
Elastin

T-cell receptor ζ chain
Elastin

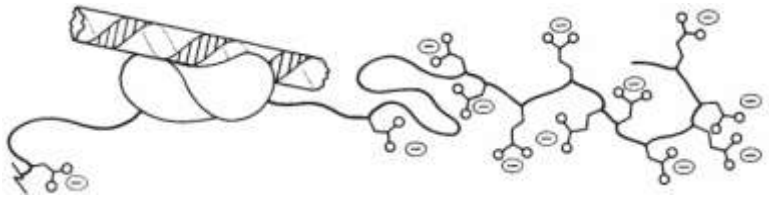


Pometun et al. (2004) *J Biol Chem* 279, 7982

Fuzzy complexes

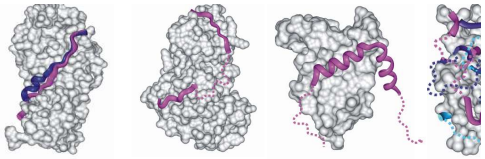


extreme cases: sequence independence



Sigler et al. (1988) *Nature* 333, 210

IDP recognition



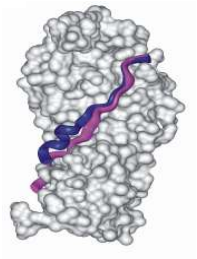
- short recognition motifs
- sequential and structural variability
- multi-functionality
- evolutionary benefits

Thank you

MRTN – CT 2005 019566

Bolyai fellowship

Fuzzy complexes



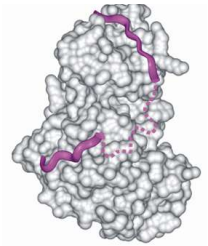
static different bound structures can be determined

polymorphic

IDP	partner
NLS	α-importin
Hsp90 MEEVD	Ppp5 TPR domain
RyR	DHPR
CFTR R domain	CFTR
Inhibitor 1	Protein phosphatase 2
Prion	Prion amyloid

Tompa and Fuxreiter (2008) *Trends in Biochem Sci* 33, 2-8

Fuzzy complexes



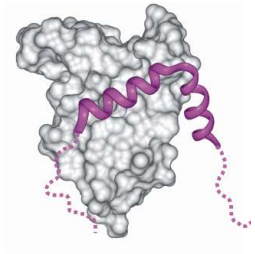
dynamic interconvert between many structures

clamp

IDP	partner
Ste5	Fus3p
Oct-1 TF	Ig-κ promoter
NLS linker	α-importin
Cellulase E	cellulose
Myosine VI	Actin filament

Tompa and Fuxreiter (2008) *Trends in Biochem Sci* 33, 2-8

Fuzzy complexes



dynamic interconvert between many structures

flanking

IDP	partner
Hsp25	α-Acrystalline
RNAP II CTD	m-RNA maturation factors
CREB KID	CBP KIX
Proline rich peptides	SH3 domain
IA_3p	Aspartic acid protease
SF1 splicing factor	$U2AF^{65}$
SP1 TAD	PIC

Tompa and Fuxreiter (2008) *Trends in Biochem Sci* 33, 2-8