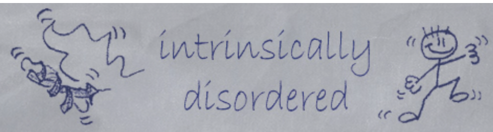
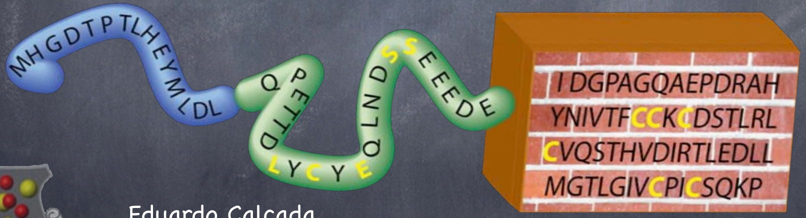
IDPbyNMR



Discovering the complete construct of intrinsically disordered oncoprotein E7 from human papilloma virus type 16



Eduardo Calçada
PhD student, CERM, Italy

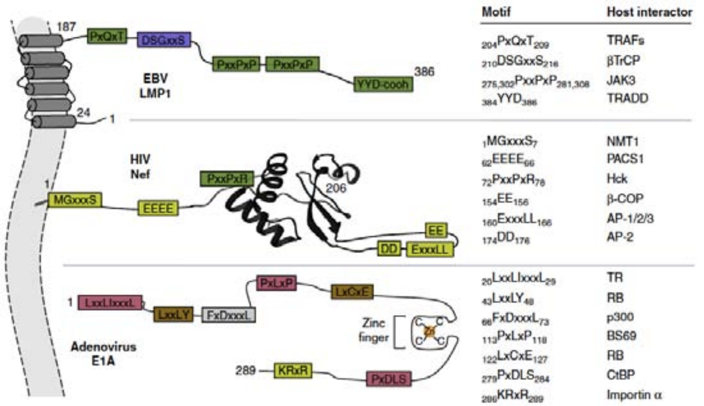
Outline

1. Introducing Intrinsically Disordered Proteins (IDP)
    2.1. E7 from Human Papilloma virus
    2.2. E1A from Adenovirus 2/5
2. Results and Goals
3. Perspectives

Intrinsically disordered proteins

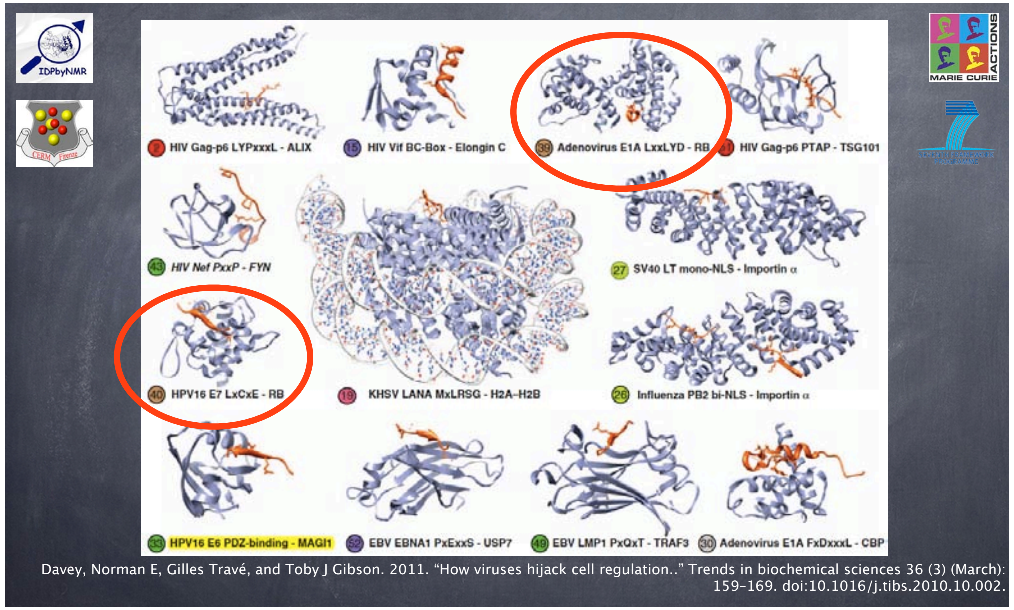
- lack stable 3D structure
- Flexibility means advantages:
    - plasticity
    - protein interactions
- IDPs: regulatory and signaling functions
- Viral proteins can hijack and manipulate regulatory proteins

How viruses hijack cell regulation

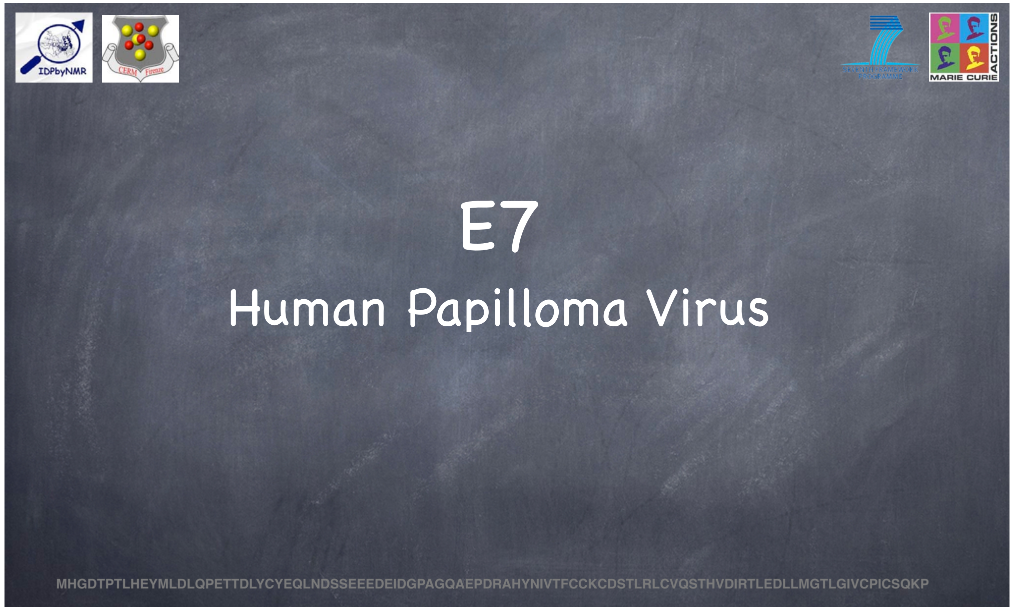


Motif	Host interactor
$_{204}$PxQxT$_{209}$	TRAFs
$_{210}$DSGxxS$_{216}$	βTrCP
$_{275,302}$PxxPxP$_{281,308}$	JAK3
$_{384}$YYD$_{386}$	TRADD
$_{1}$MGxxxS$_{7}$	NMT1
$_{62}$EEEE$_{66}$	PACS1
$_{72}$PxxPxR$_{78}$	Hck
$_{154}$EE$_{156}$	β-COP
$_{160}$ExxxLL$_{166}$	AP-1/2/3
$_{174}$DD$_{176}$	AP-2
$_{20}$LxxLIxxxL$_{29}$	TR
$_{43}$LxxLY$_{48}$	RB
$_{66}$FxDxxxL$_{73}$	p300
$_{113}$PxLxP$_{118}$	BS69
$_{122}$LxCxE$_{127}$	RB
$_{279}$PxDLS$_{284}$	CtBP
$_{286}$KRxR$_{289}$	Importin α

Davey, Norman E, Gilles Travé, and Toby J Gibson. 2011. “How viruses hijack cell regulation..” Trends in biochemical sciences 36 (3) (March): 159–169. doi:10.1016/j.tibs.2010.10.002.



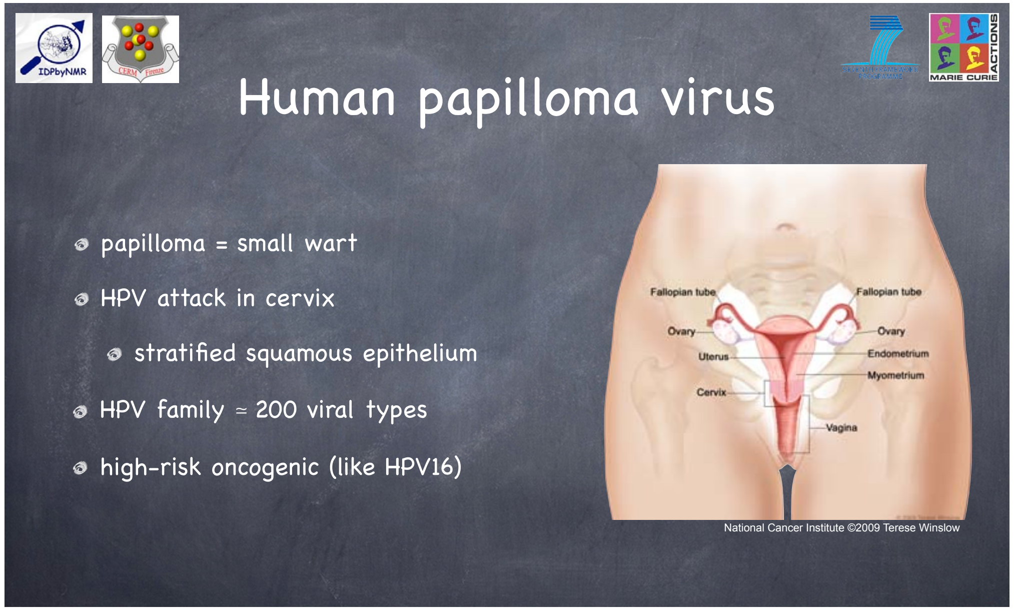
Davey, Norman E, Gilles Travé, and Toby J Gibson. 2011. "How viruses hijack cell regulation.." Trends in biochemical sciences 36 (3) (March): 159–169. doi:10.1016/j.tibs.2010.10.002.



E7

Human Papilloma Virus

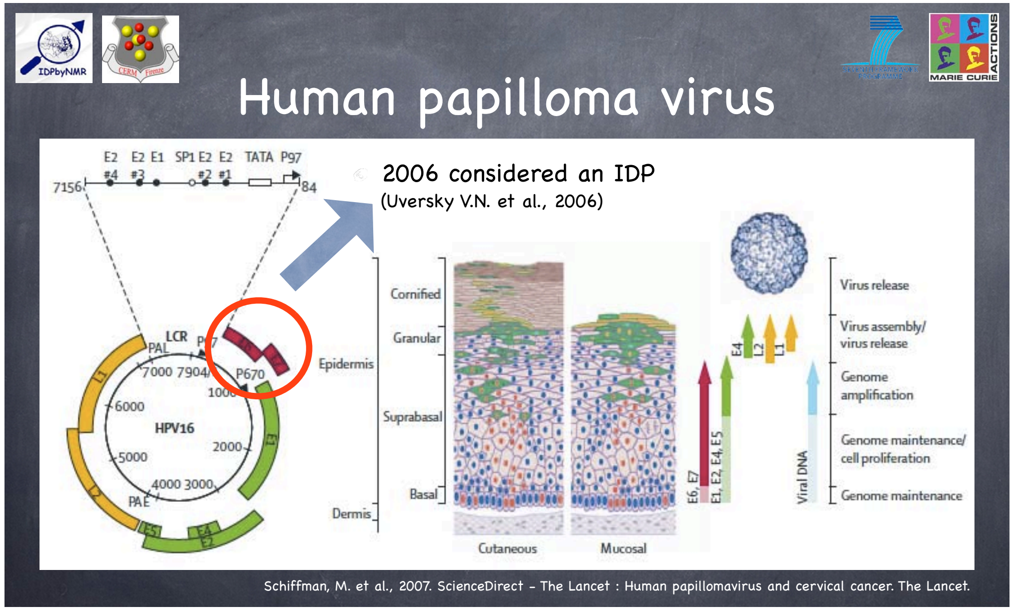
MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP



Human papilloma virus

- papilloma = small wart
- HPV attack in cervix
  - stratified squamous epithelium
- HPV family ≃ 200 viral types
- high-risk oncogenic (like HPV16)

National Cancer Institute ©2009 Terese Winslow

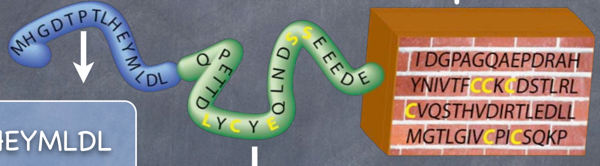


Human papilloma virus

2006 considered an IDP
(Uversky V.N. et al., 2006)

Schiffman, M. et al., 2007. ScienceDirect - The Lancet : Human papillomavirus and cervical cancer. The Lancet.

HPV16 E7 entire sequence



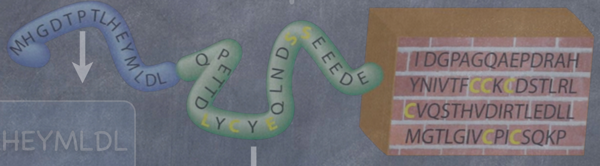
MHGDTPTLHEYMLDL

QPETTDLYCYEQLNDSSEEEDE

IDGPAGQAEPDRAHYNIVTFCCKCDSTLRLC
VQSTHVDIRTLEDLLMGTLGIVCPICSQKP

MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP

E7 sequence



MHGDTPTPTLHEYMLDL

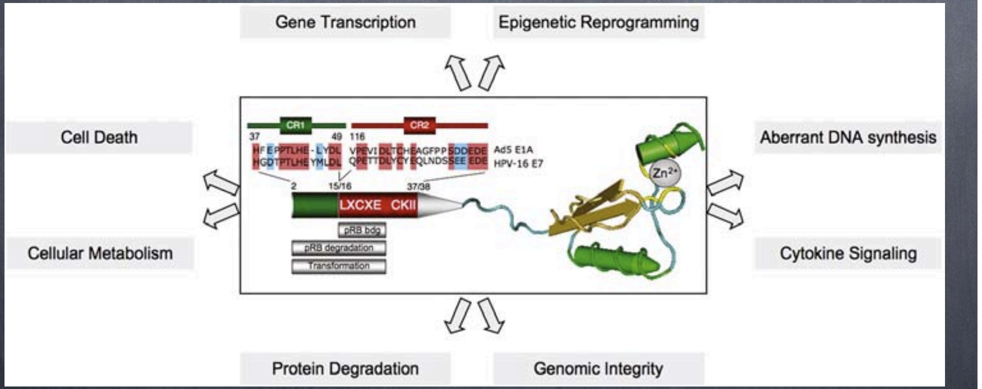
QPETTDLYCYEQLNDSSEEEDE

- CR3 contains two CxxC zinc-binding motifs
- separated by 29 residues

IDGPAGQAEPDRAHYNIVTFCCKCDSTLRLC
VQSTHVDIRTLEDLLMGTLGIVCPICSQKP

Balog, E.R.M. et al., 2011. Crystal structure of the unliganded retinoblastoma protein pocket domain. Proteins, 79(6), pp.2010–2014.

Cellular processes affected by E7



McLaughlin-Drubin, M.E. & Münger, K., 2009. The human papillomavirus E7 oncoprotein. Virology, 384(2), pp.335–344.

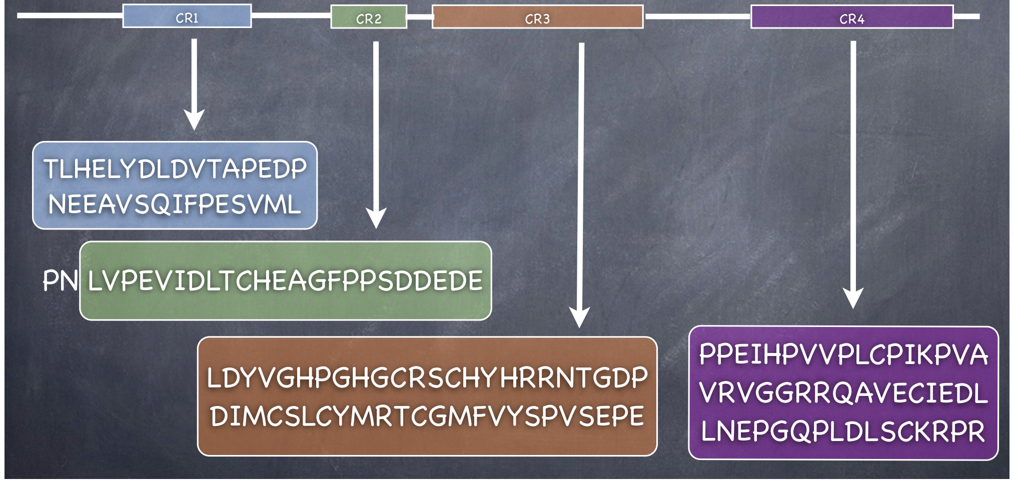
E1A

Adenovirus 5

E1A protein

- key role many cell regulatory processes
- it contains much more linear motifs
- 289 aa

Ad5 E1A entire sequence



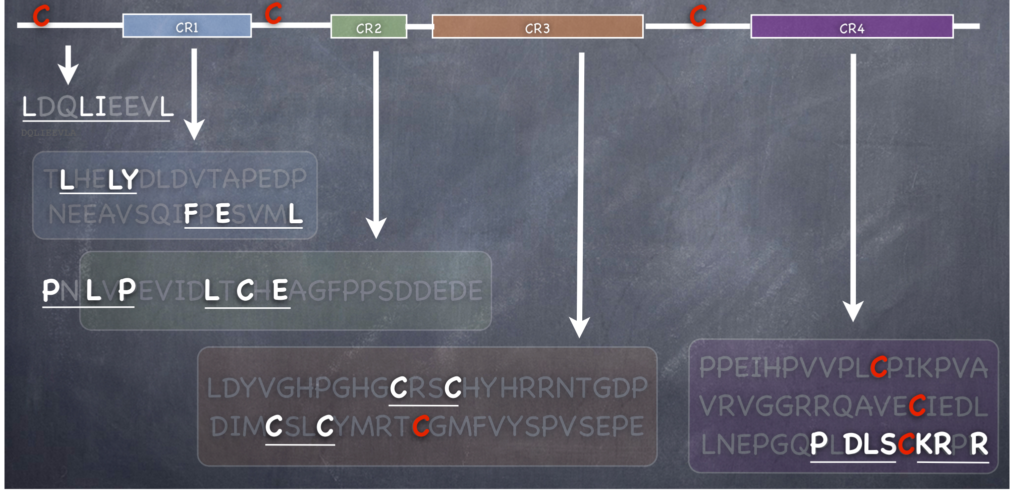
TLHELYDLDVTAPEDP
NEEAVSQIFPESVML

PN LVPEVIDLTCHEAGFPPSDDEDE

LDYVGHPGHGCRSCHYHRRNTGDP
DIMCSLCYMRTCGMFVYSPVSEPE

PPEIHPVVPLCPIKPVA
VRVGGRRQAVECIEDL
LNEPGQPLDLSCKRPR

Ad5 E1A entire sequence



LDQLIEEVL

DQLIEEVLA

TLHELYDLDVTAPEDP
NEEAVSQIFPESVML

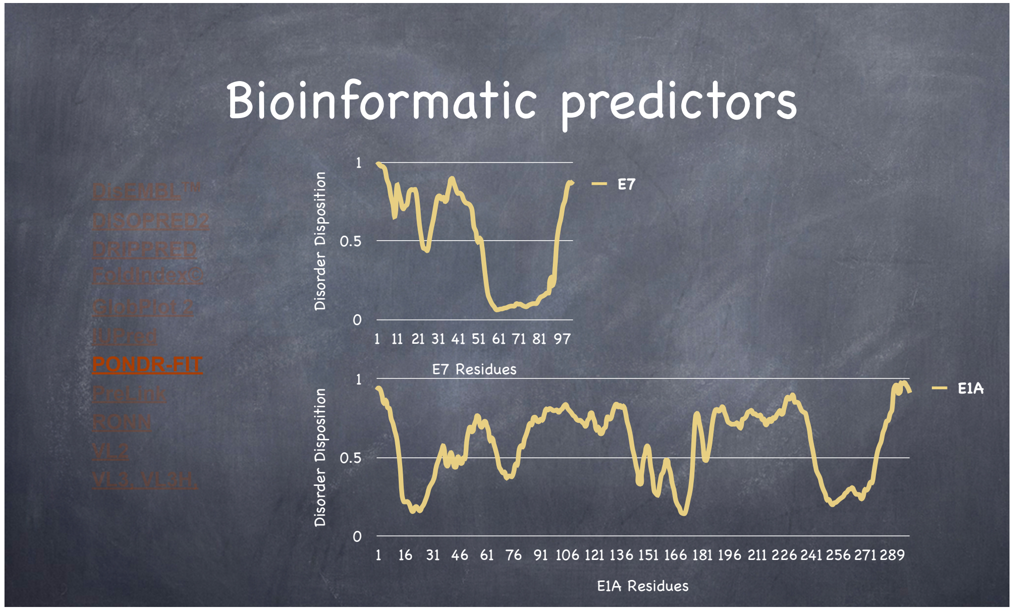
PNLVPEVIDLTCHEAGFPPSDDEDE

LDYVGHPGHGCRSCHYHRRNTGDP
DIMCSLCYMRTCGMFVYSPVSEPE

PPEIHPVVPLCPIKPVA
VRVGGRRQAVECIEDL
LNEPGQPLDLSCKRPR

Bioinformatic predictors

DisEMBL™	Intrinsic Protein Disorder Prediction
DISOPRED2	Disorder Prediction Server
DRIPPRED	Web based predictor for disordered regions in proteins
FoldIndex©	Estimate the fold probability of a protein
GlobPlot 2	Intrinsic Protein Disorder, Domain & Globularity Prediction
IUPred	Prediction of Intrinsically Unstructured Proteins
PONDR-FIT	Predictors of Natural Disordered Regions
PreLink	Prediction of unfolded segments in a protein sequence based on
RONN	Regional Order Neural Network
VL2	DisProt Predictor of Intrinsically Disordered Regions
VL3, VL3H,	DisProt Predictor of Intrinsically Disordered Regions



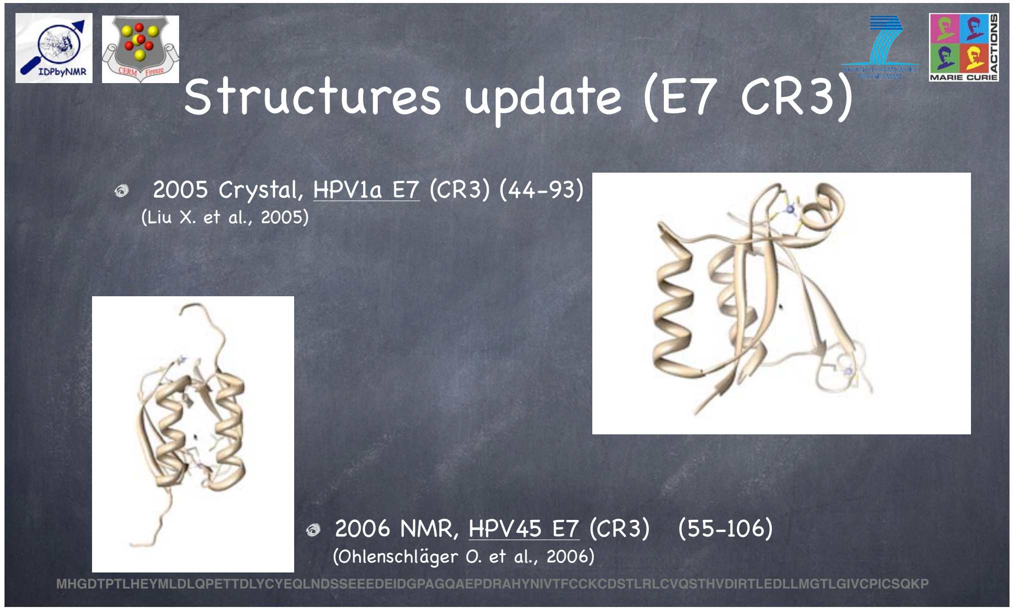
Bioinformatic predictors

- DisEMBL™
- DISOPRED2
- DRIPPRED
- FoldIndex©
- GlobPlot 2
- IUPred
- **PONDR-FIT**
- PreLink
- RONN
- VL2
- VL3, VL3H,

E7

Human Papilloma Virus

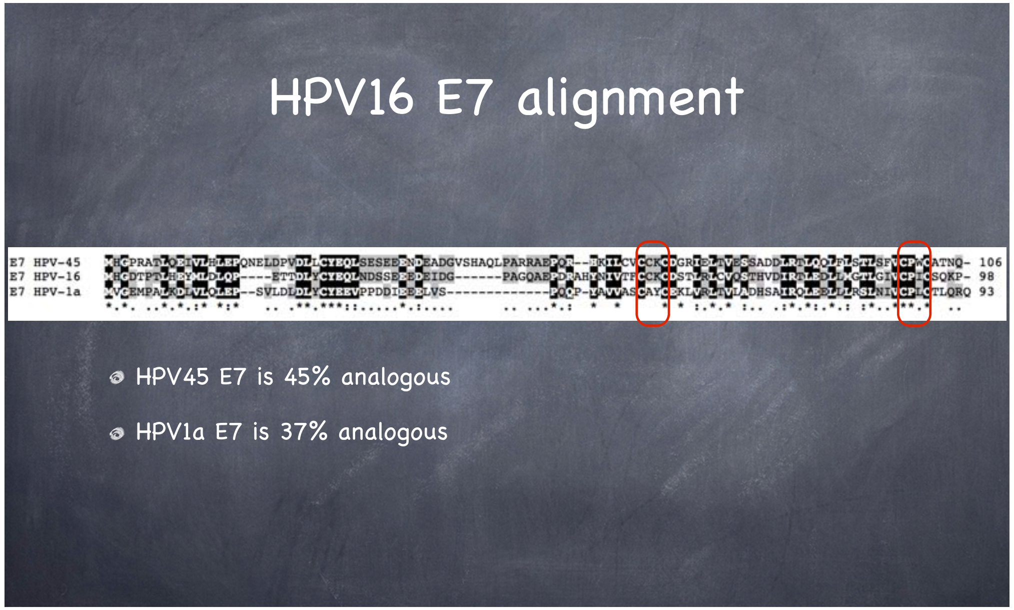
MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP



Structures update (E7 CR3)

- 2005 Crystal, HPV1a E7 (CR3) (44-93)
  (Liu X. et al., 2005)
- 2006 NMR, HPV45 E7 (CR3)   (55-106)
  (Ohlenschläger O. et al., 2006)

MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP

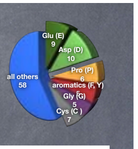


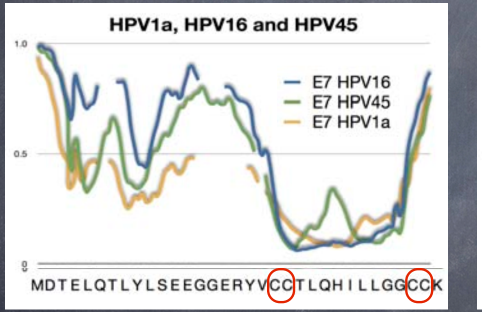
HPV16 E7 alignment

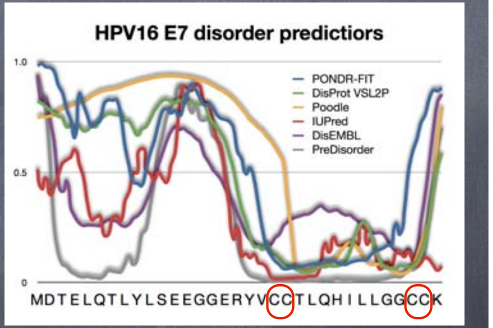
```
E7 HPV-45    MHGPRATLQEIVLHLEPQNELDPVDLLCYEQLSESEEENDEADGVSHAQLPARRAEPQR--HKILCVCCKCDGRIELTVESSADDLRTLQQLFLSTLSFVCPWCATNQ- 106
E7 HPV-16    MHGDTPTLHEYMLDLQP----ETTDLYCYEQLNDSSEEEDEIDG------PAGQAEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP- 98
E7 HPV-1a    MVGEMPALKDLVLQLEP--SVLDLDLYCYEEVPPDDIEEELVS-------------PQQP-YAVVASCAYCEKLVRLTVLADHSAIRQLEELLLRSLNIVCPLCTLQRQ 93
             *.*. ..*.*.:* *:*   .. .**.***::......*.:....        .. ...*.:  *  *    * * :.*.* :.. .:*.*:.*.: :*..***.*  ..
```

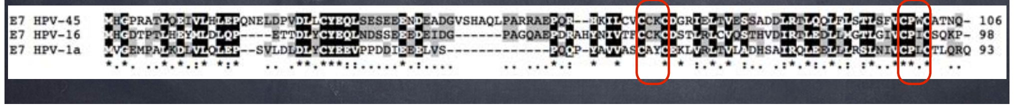
- HPV45 E7 is 45% analogous
- HPV1a E7 is 37% analogous

HPV16 E7 prediction



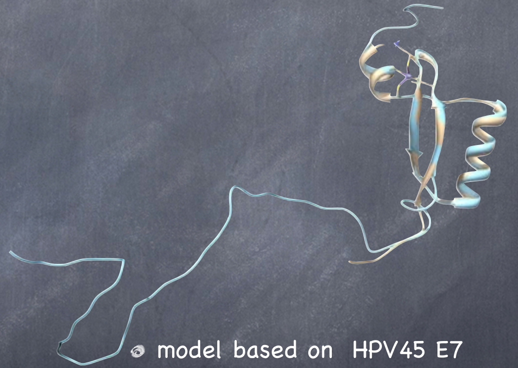






```
E7 HPV-45   MHGPRATLQEIVLHLEPQNELDPVDLLCYEQLSESEEENDEADGVSHAQLPARRAEPQR--HKILCVCCKCDGRIELTVESSADDLRTLQQLFLSTLSFVCPWCATNQ- 106
E7 HPV-16   MHGDTPTLHEYMLDLQP----ETTDLYCYEQLNDSSEEEDEIDG------PAGQAEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP- 98
E7 HPV-1a   MVGEMPALKDLVLQLEP--SVLDLDLYCYEEVPPDDIEEELVS-------------PQQP-YAVVASCAYCEKLVRLTVLADHSAIRQLEELLLRSLNIVCPLCTLQRQ 93
```

HPV16 E7 modeling

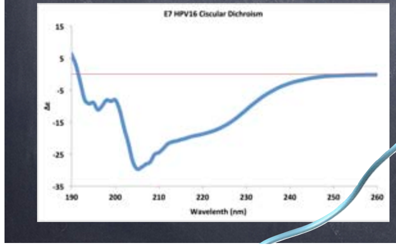


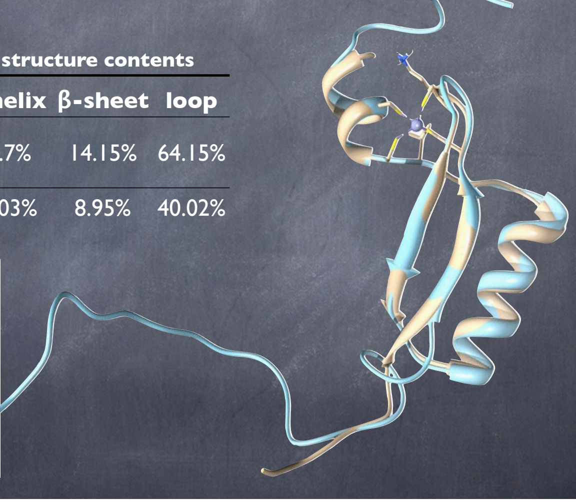
- model based on HPV1a E7
  - 33 atoms, RMSD = 1.165 Å
  - HPV1a (37% homologous)
- model based on HPV45 E7
  - 52 atoms, RMSD = 0.496 Å
  - HPV45 E7 (45% homologous)

HPV16 E7 modeled with HPV45 E7

HPV16 E7 secondary structure contents

E7 protein	α-helix	β-sheet	loop
E7 HPV16 modeled with E7 HPV45 CR3	21.7%	14.15%	64.15%
E7 HPV16 CD data	51.03%	8.95%	40.02%





viral proteins expression

E7 protein

- E. coli BL21(DE3)pLys
- pET20b(+) plasmid
- 2 silent mutations
- 2 purification steps
- E7 yield: 20mg/L

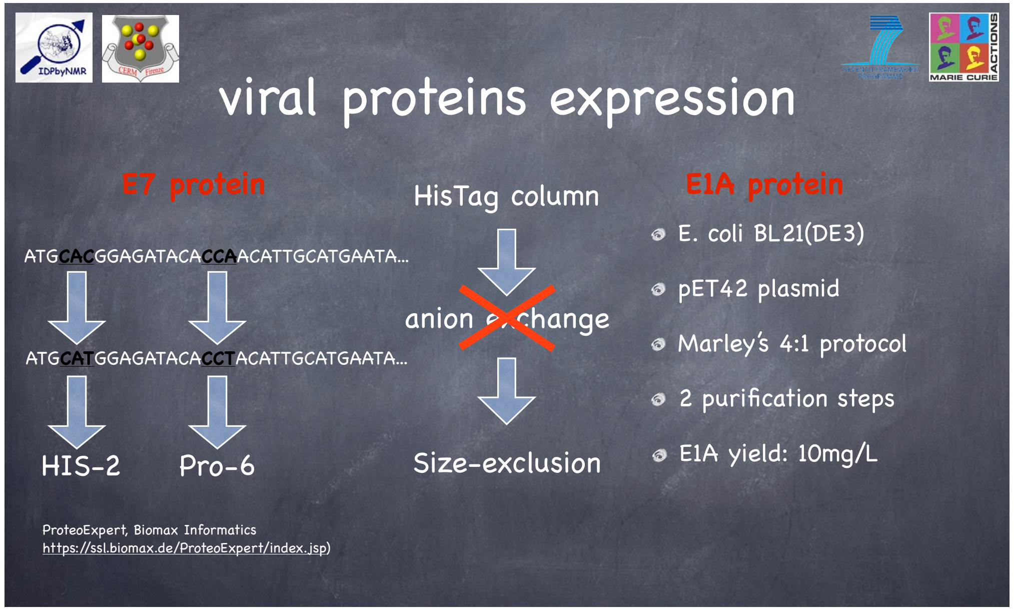
HisTag column

~~anion exchange~~

Size-exclusion

E1A protein

- E. coli BL21(DE3)
- pET42 plasmid
- Marley's 4:1 protocol
- 2 purification steps
- E1A yield: 10mg/L



viral proteins expression

E7 protein

ATGCACGGAGATACACCAACATTGCATGAATA...

ATGCATGGAGATACACCTACATTGCATGAATA...

HIS-2 Pro-6

HisTag column

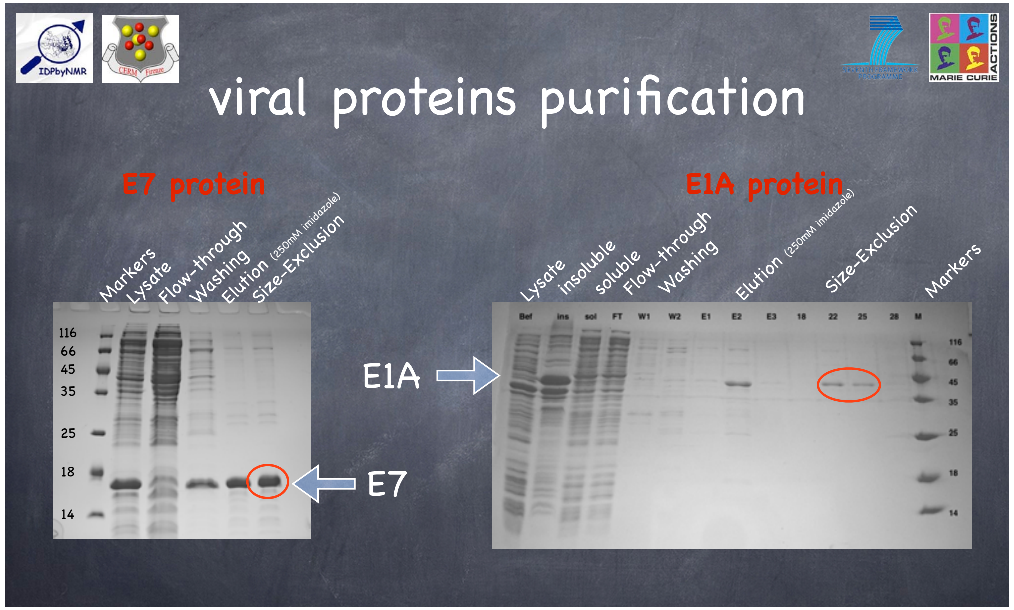
~~anion exchange~~

Size-exclusion

E1A protein

- E. coli BL21(DE3)
- pET42 plasmid
- Marley's 4:1 protocol
- 2 purification steps
- E1A yield: 10mg/L

ProteoExpert, Biomax Informatics
https://ssl.biomax.de/ProteoExpert/index.jsp)



viral proteins purification

E7 protein

E1A protein

Optimization

reductant agent: TCEP, DTT, ß-mercaptoethanol	DTT
pH 5.5 - 8.5	pH 7.5
buffers: PBS, HEPES	HEPES
temperature 278 - 308 K	298 K
anaerobic vs aerobic purification steps	☹
denaturing conditions till 8M urea	☹

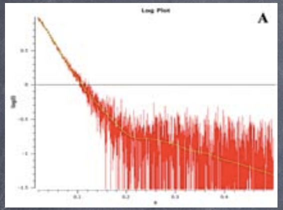
MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP

Biophysical techniques

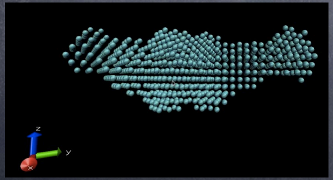
- **ESI and MALDI mass spectrometry**
- **SAXS**
- **Dynamic Light Scattering**
- **NMR**

HPV16 E7 protein molecular weight	
Theoretic - ExPASy	12152 Da
ESI	12149 Da
MALDI	12219 Da
DLS	24 ± 6 kDa
SAXS	30 ± 4 kDa

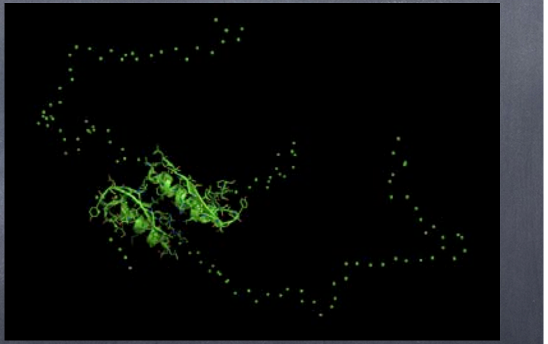
SAXS on E7 protein



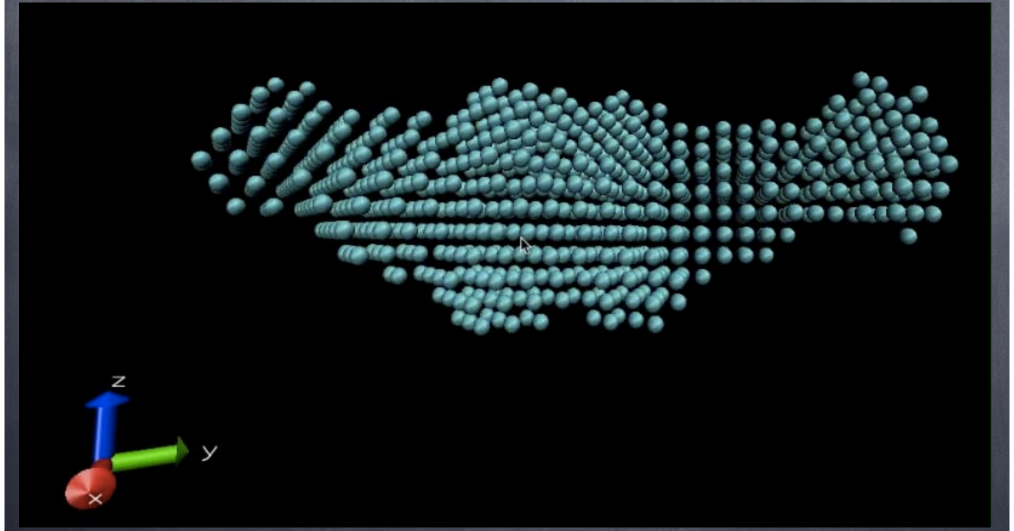
Experimental X-ray scattering



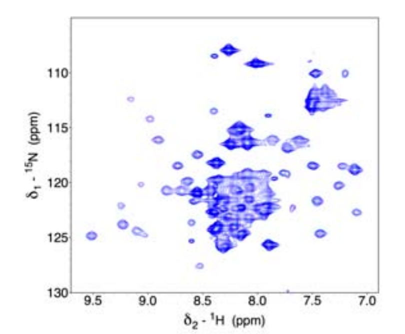
DAMMIF analysis



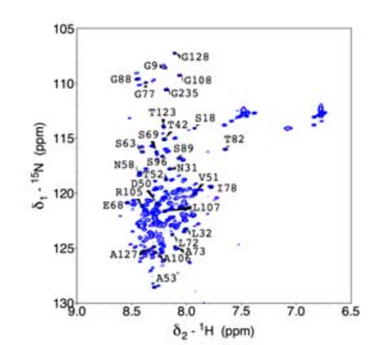
EOM models using a homologous HPV1 structured part



HPV16 E7



Ad2/5 E1A



Future goals

- E7 HPV16 protein
    - mutate specific Cysteines
    - phosphorylation studies of Serine residues
    - interaction with partner - tumor suppressor pRB
    - in-cell NMR studies using human cells (electroporation)
- E1A Adenovirus 2/5
    - analyze the 12S E1A protein with 243 residues (no CR3)

MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP

Acknowledgments

- Professor Roberta Pierattelli
- Professor Isabella Felli
- Tomáš Hošek
- CERM
- IDPbyNMR

Thank you