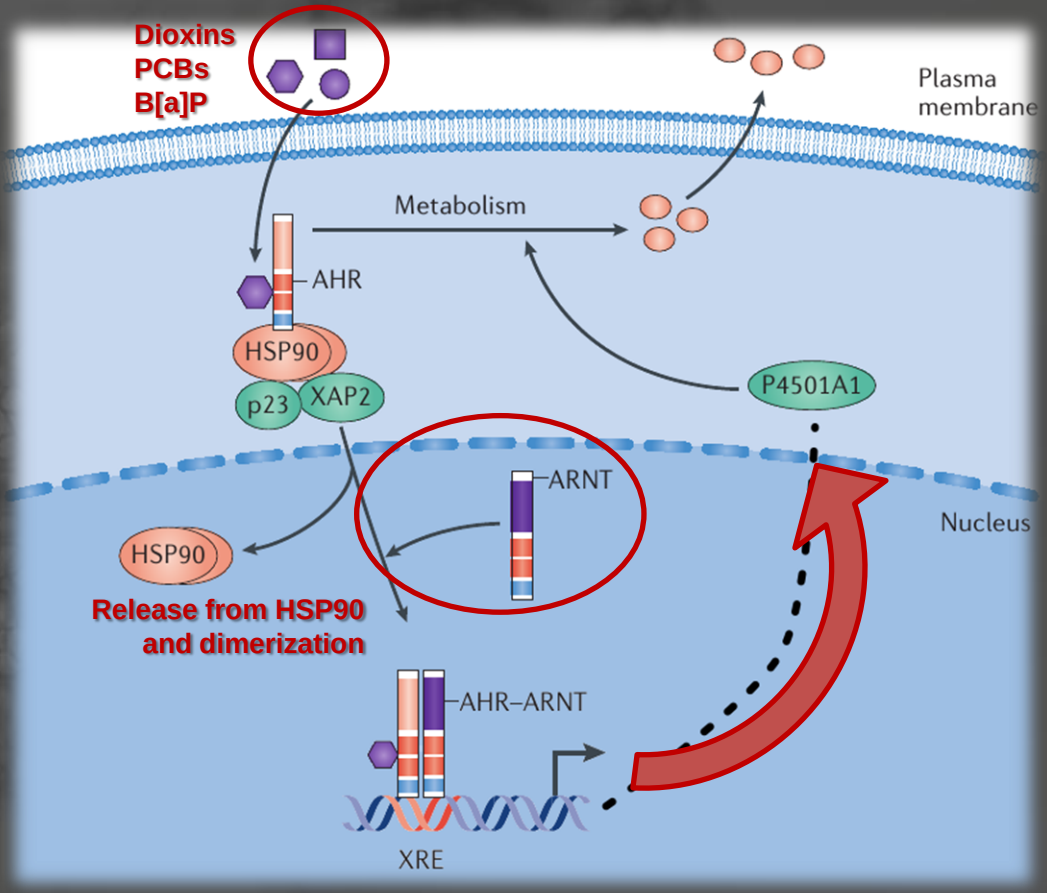
Modelling the Intertwined Network of PPIs along the AhR:ARNT Dimer

Dario CORRADA and Laura BONATI

Department of Earth and Environmental Sciences – University of Milano-Bicocca

Biological Function of Aryl hydrocarbon Receptor (AhR)



Bernsten, D. et al. (2013) *Nat Rev Cancer* 13, 827-41

Classical Mechanism of Action

AhR is a transcription factor maintained in its inactive cytosolic form as part of a larger protein complex

AhR is activated by binding to a wide range of xenobiotics

AhR translocates into the nucleus and dimerizes with the AhR Nuclear Translocator (ARNT)

The AhR:ARNT dimer binds to DNA and promotes the expression of genes involved in metabolic detoxification pathways

The bHLH-PAS Proteins Family

bHLH-PAS *stands for* basic Helix Loop Helix – Per Arnt Sim

bHLH motif
- dimerization
- DNA binding

PAS-B domain
- ligand binding
- environmental sensing

PAS-A domain
- dimerization
- recognition of PAS protein partners

C-terminal region
- regulatory
- transactivation

The bHLH-PAS Protein Family

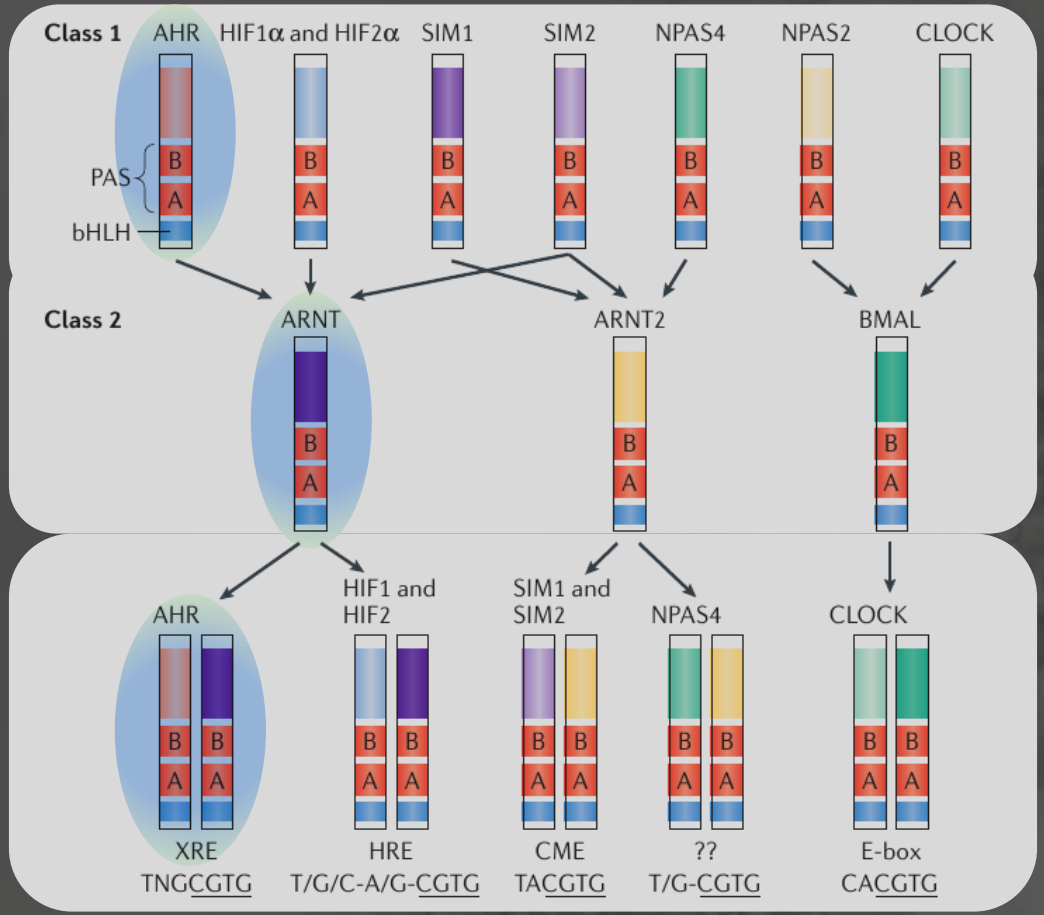
Class I Family Members

Functional roles: sensing of environmental or physiological signals

Class II Family Members

Regulatory roles: dimerization with both class II and class I members

The transcriptionally active complex is always a heterodimer composed of bHLH-PAS proteins



Bernsten, D. et al. (2013) *Nat Rev Cancer* 13, 827-41

Outline – Homology Modelling Strategies

STAGE 1

individual dimers of PAS domains

- Templates Available
- Missing Loops for PAS-A Domains
- Protein-Protein Interactions (PPIs)
- Discriminating Alternative Models
- Experimental Validations

STAGE 2

full-length bHLH-PAS complex

- Templates Available (Brand New)
- Full Length Dimerization Interfaces
- Homology Models Comparison
- Energy Decomposition Analysis
- Known Mutations for AhR and ARNT

incremental availability of structural information (ie brand new templates)

[STAGE 1] Templates Available

Scheuermann, T. et al. (2009) *Proc Natl Acad Sci USA* 106, 450-5

Huang, N. et al. (2012) *Science* 337, 189-94

HIF2α:ARNT
PAS-B
Heterodimer
[3F1P – Res. 1.17Å]

PAS-B domains dimer models

CLOCK:BMAL1
full-length bHLH PAS complex
[4F3L – Res. 2.27Å]

Wu, D. et al. (2013) *Mol Cell Biol* 33, 4346-56

PAS-A domains dimer models

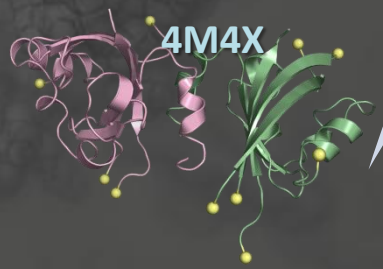
AhR:AhR
PAS-A
Homodimer
[4M4X– Res. 2.55Å]

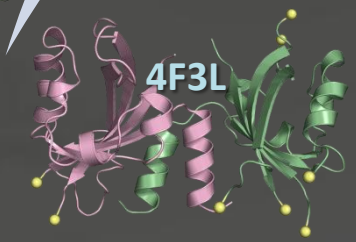
100 putative homology models for every dimerization mode proposed by each template

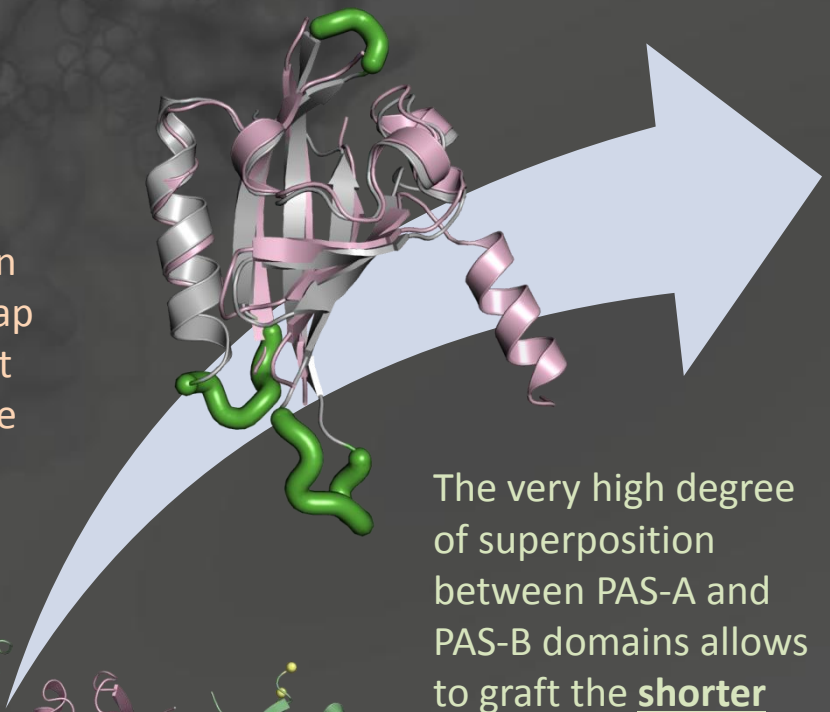
powered by MODELLER v9.13

Missing Loops for PAS-A Domains

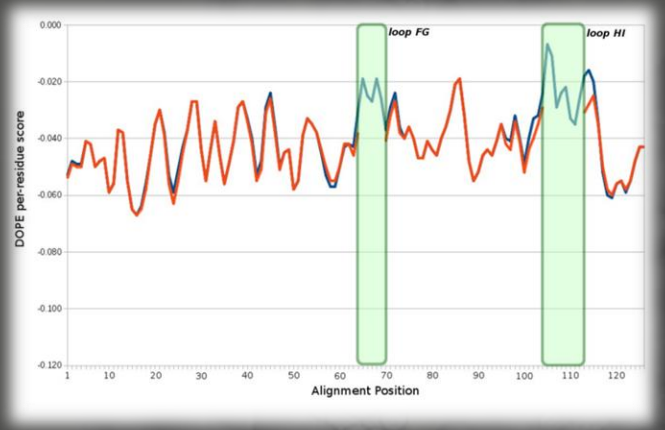
Missing regions in the templates map on loops of about **20 residues** in the target







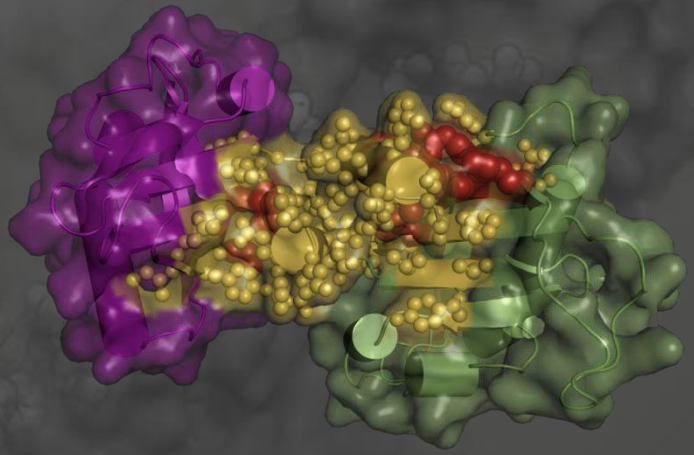
The very high degree of superposition between PAS-A and PAS-B domains allows to graft the **shorter** loops from PAS-B to PAS-A models



The "surgery" outcome seems good: no relevant perturbations emerge by comparison of DOPE energy profiles[1] of PAS-A templates and models

[1] Shen, M. et al. (2006) *Protein Sci* 15, 2507-24

Protein-Protein Interactions (PPIs)



Dimerization interfaces are outlined by variations in Solvent Accessible Surface Area (ΔSASA)[1]

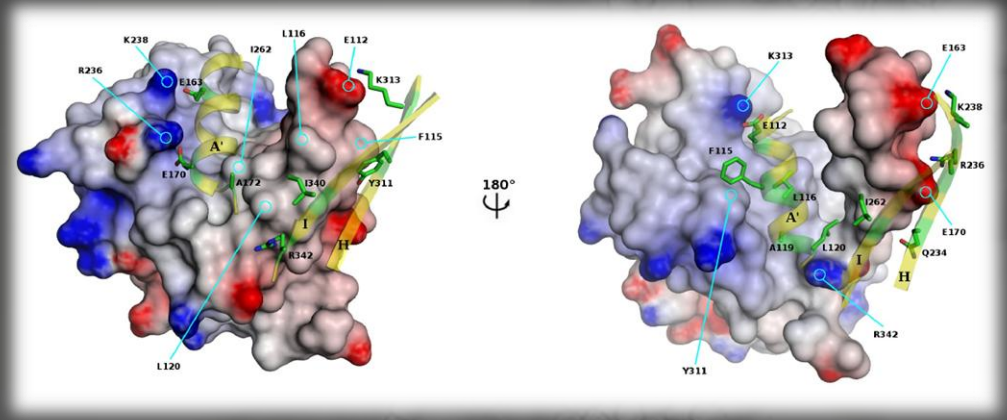
[1] Kleinjung, J. et al. (2005) *Nucleic Acid Res* 33, W342-6

Residues predicted to have critical roles in dimerization are gathered by diverse hot spot prediction algorithms: Robetta[2]; HotPoint[3]; KFC2[4]

[2] Kortemme, T. et al. (2002) *Proc Natl Acad Sci USA* 99, 14116-21
[3] Tuncbag, N. et al. (2010) *Nucleic Acid Res* 38, W402-6
[4] Zhu, X. et al. (2011) *Proteins* 79, 2671-83

Dimerization interfaces exhibit both geometric and electrostatic complementarity, according to the Electrostatic Potential Surfaces (EPS)[5]

[5] Rocchia, W. et al. (2001) *J Phys Chem B* 105, 6507-14

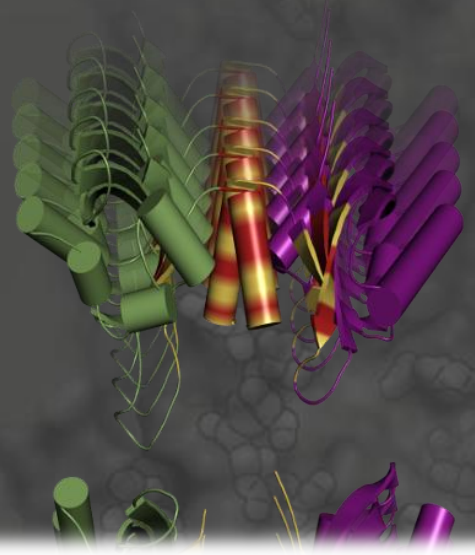


Discriminating Alternative Models

SET A
100 PAS-A(PAS-B) dimers based on CLOCK:BMAL1 complex

VS

SET B
100 PAS-A(PAS-B) dimers based on AhR:AhR(HIF2α:ARNT) dimer



$\Delta G_{binding}$ is calculated by means of MM-GBSA method. Energetic determinants are extracted through Energy (eigen)Decomposition analysis[1]

[1] Corrada, D. et al. (2013) *J Chem Inf Model* 53, 2937-50

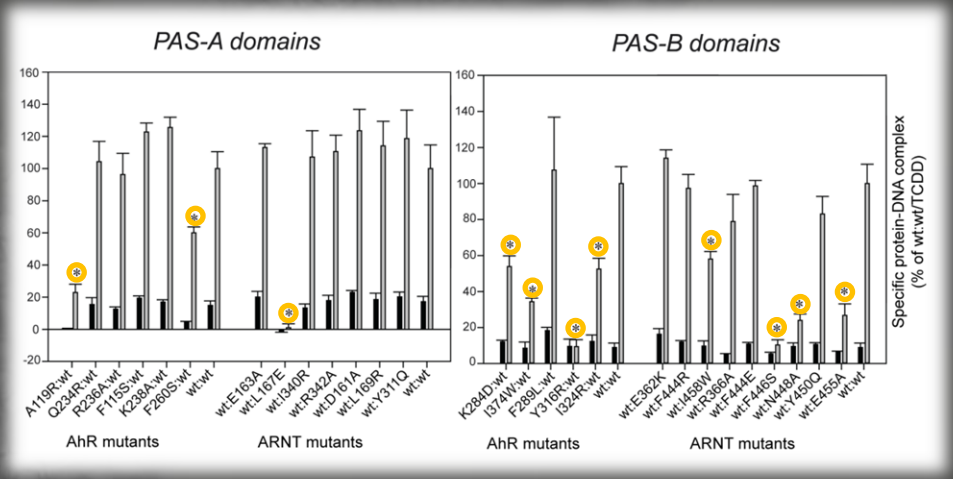
Distinctive patterns of interacting residues are determined by comparing the 100 replicates of each model by the two sample Rank Products algorithm[2]

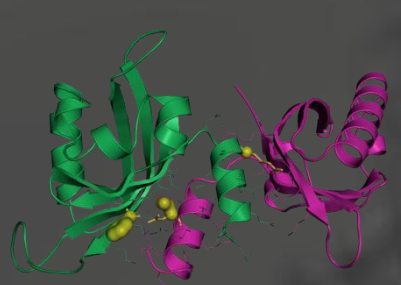
[2] Koziol, J. (2010) *FEBS Lett* 584, 4481-84

Experimental Validations

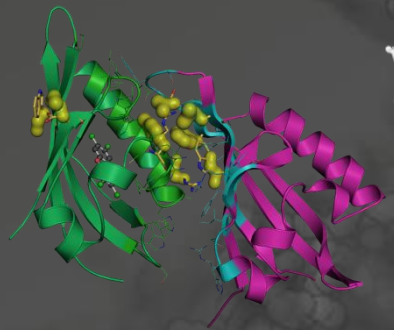
ΔG signatures define pattern of distinct hot spots characterizing each model

mutagenesis experiments for the assessment of the proposed dimerization modes





PAS-A dimer model
based on
CLOCK:BMAL1 complex



PAS-B dimer model
based on
CLOCK:BMAL1 complex

Corrada D, Soshilov AA, Denison MS, Bonati L.

Deciphering Dimerization Modes of PAS Domains: Computational and Experimental Analyses of the AhR:ARNT Complex Reveal New Insights Into the Mechanisms of AhR Transformation.

(2016) PLoS Computational Biology 12(6): e1004981

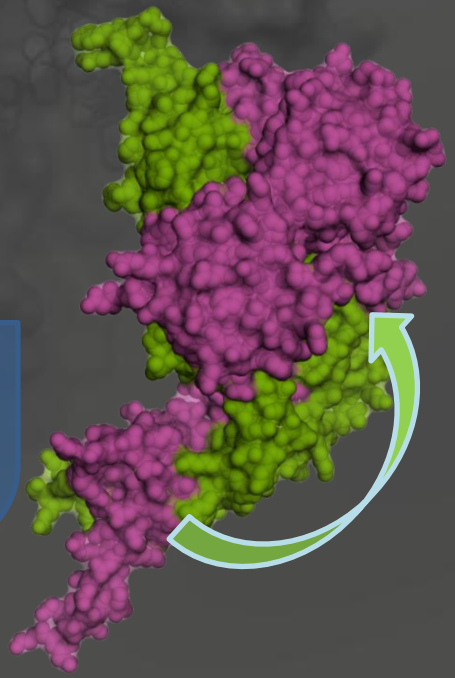
doi:10.1371/ journal.pcbi.1004981

[STAGE 2] Templates Available (Brand New)

Huang, N. et al. (2012) *Science* 337, 189-94

CLOCK:BMAL1
full-length
complex
[4F3L – Res. 2.27Å]

Class I bHLH-PAS (**CLOCK**) wraps around Class II bHLH-PAS (**BMAL1**)



Wu, D. et al. (2015) *Nature* 524, 303-8

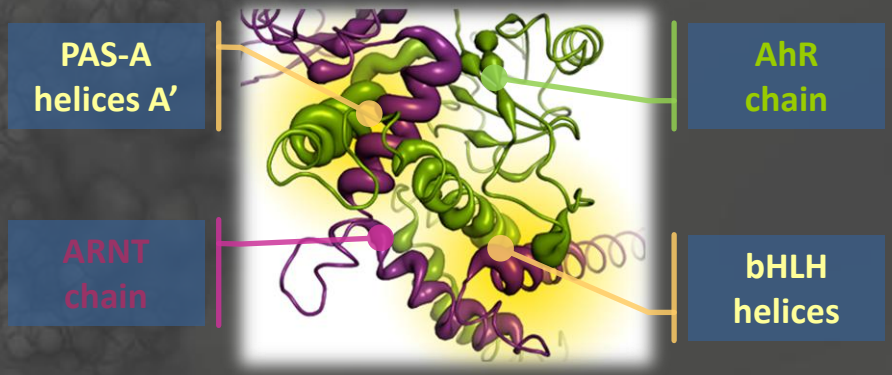
HIF2α:ARNT
full-length
complex
[4ZP4 – Res. 2.36Å]

Class II bHLH-PAS (**ARNT**) wraps around Class I bHLH-PAS (**HIF2α**)

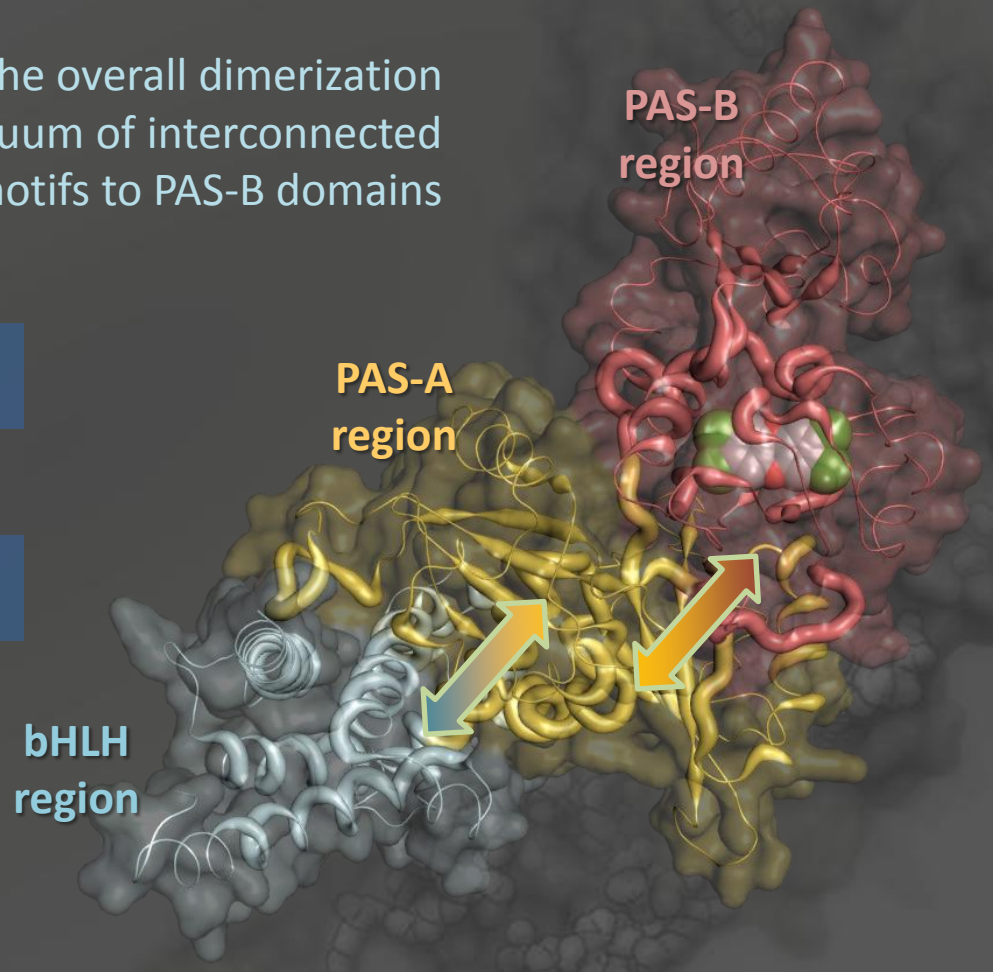
Homomeric PPIs (PAS-A/PAS-A; PAS-B/PAS-B) show very similar interfaces in the two templates and in our dimer models

Full Length dimerization interface

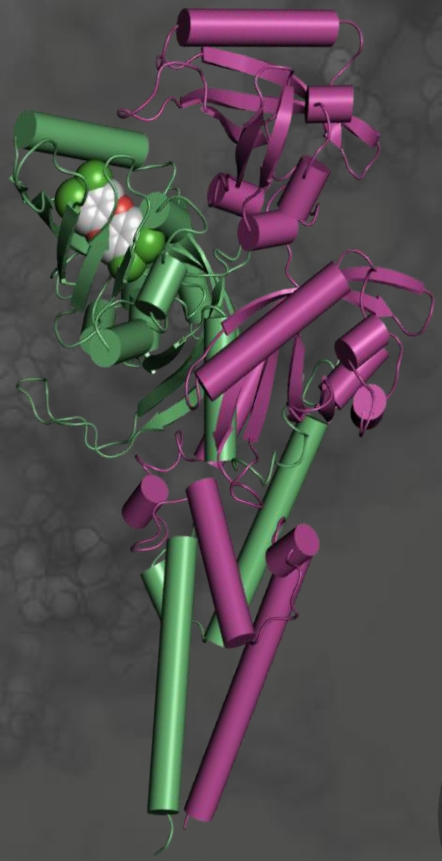
According to ΔSASA analyses, the overall dimerization interface encompass a continuum of interconnected PPIs from bHLH motifs to PAS-B domains



Extended overlaps between homomeric (PAS-A/PAS-A and bHLH/bHLH) and heteromeric (bHLH/PAS-A) interactions

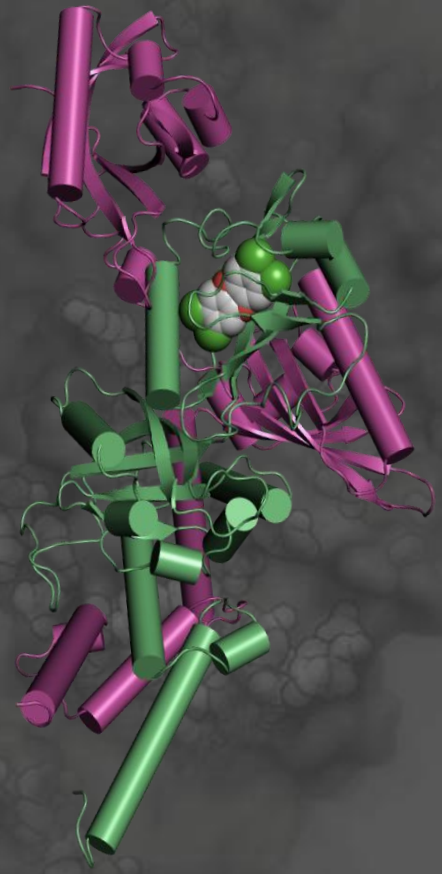


Homology Models Comparison



CLOCK:BMAL1 (4F3L)	*model based on...*	HIF2a:ARNT (4ZP4)
279	*residues*	248
7,046	*ΔSASA (Å²)*	6,536
3.1	*iRMSD (Å)*	1.2
39.2	*indentity (%)*	54.4
54.3	*similarity (%)*	64.6
-323.55 ± 17.67	*$\Delta G_{binding}$ (kcal/mol)*	**-325.70 ± 16.30**

$\Delta G_{binding}$ *values* are nearly identical, despite the overall scaffold of the templates chosen show remarkable differences



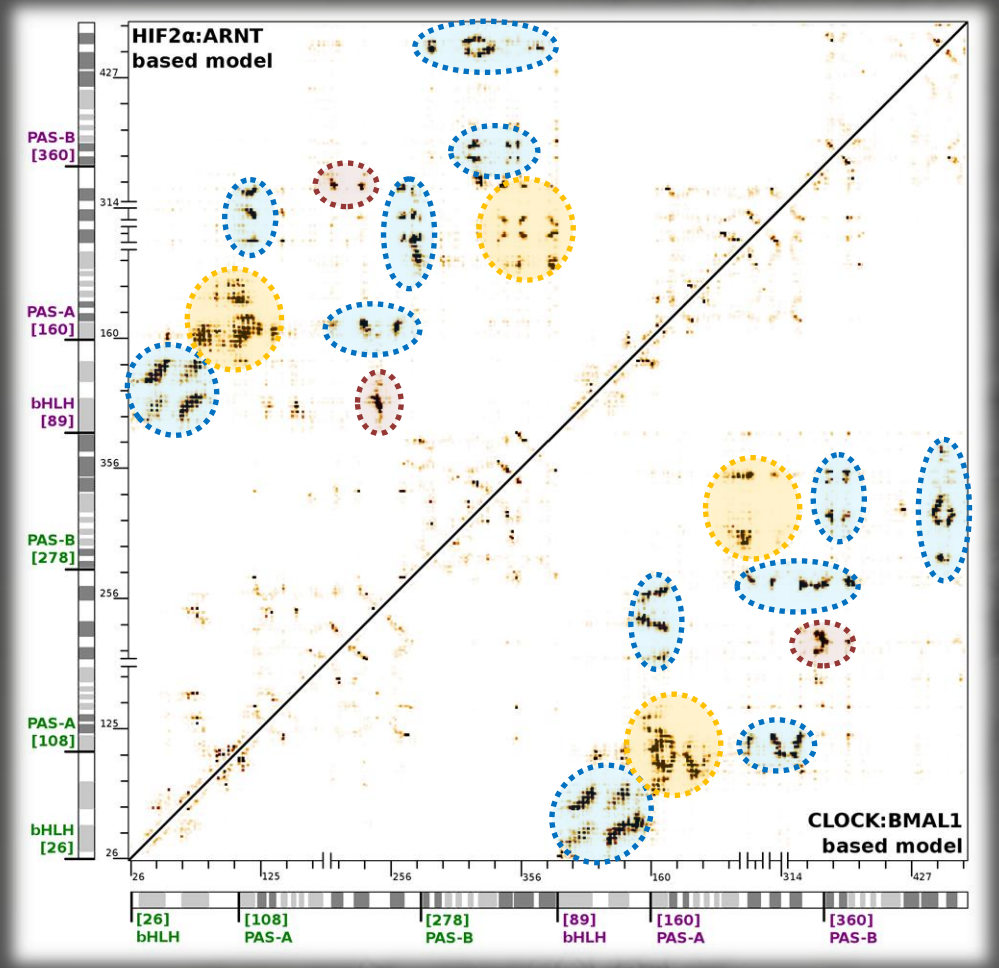
Energy Decomposition Analysis

The distribution of most of the contributions to $\Delta G_{binding}$ greatly overlaps in the two models

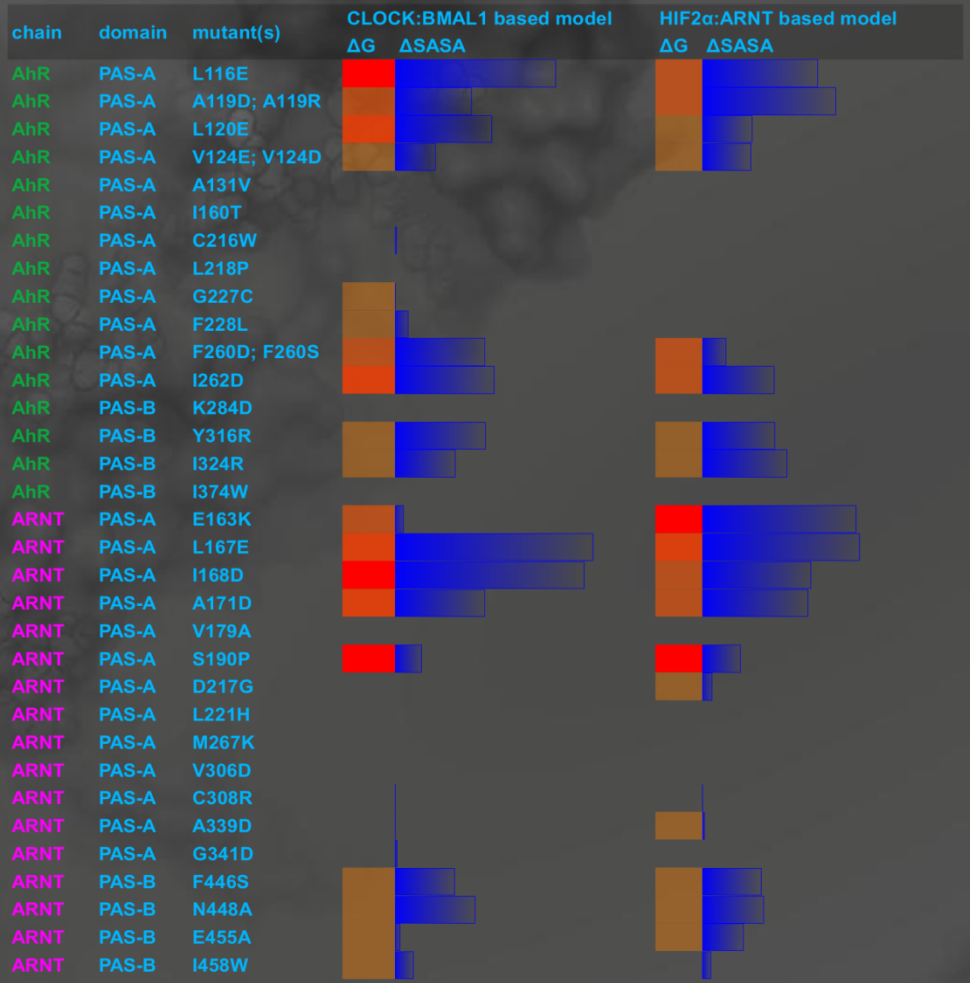
The topology of few energetic couplings slightly varies

Minor stabilizing “islets” can be found, typical of each model

Heteromeric PPIs account for minor contributions to the overall $\Delta G_{binding}$



Known Mutations for AhR and ARNT



The critical importance of several residues along the dimerization interface agree with effective disrupting mutations

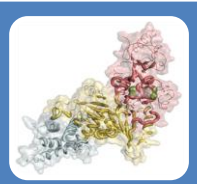
Other known mutations maps outside the dimerization interface, along solvent exposed regions

External partners other than AhR and ARNT involved in the complex stabilization?

Lessons from AhR:ARNT Complex...

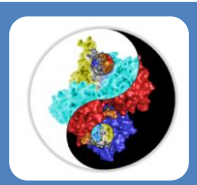
[STAGE 1]

- Our *in silico* workflow strongly suggests reliable dimerization modes among alternative interfaces



[STAGE 2]

- The two full length models share the most critical PPIs, where homomeric interactions overcome the heteromeric ones



The final outcome offers a basic framework of interactions for next studies on the molecular mechanism of transformation and transactivation of the AhR transcription factor

From local definition of PPIs…

…to global characterization of dimerization interface

GRAFTED vs. NATIVE LOOPS
How the native loops could improve the dimerization interface?

DYNAMICAL PERSPECTIVE
May differential behaviours arise from MD simulation?

Thank You for the Attention

Dario CORRADA, PhD
dario.corrada@unimib.it

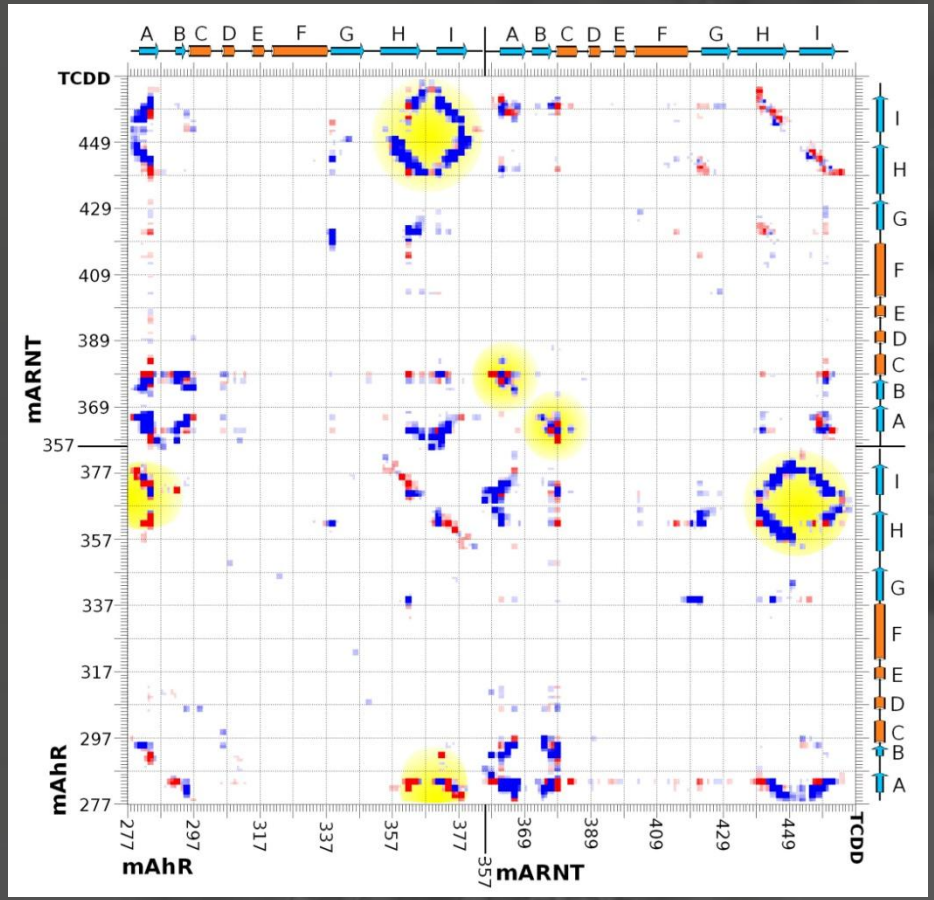
National Institute of Environmental Health Sciences [R01 ES007685]

Energy of the PPI interfaces

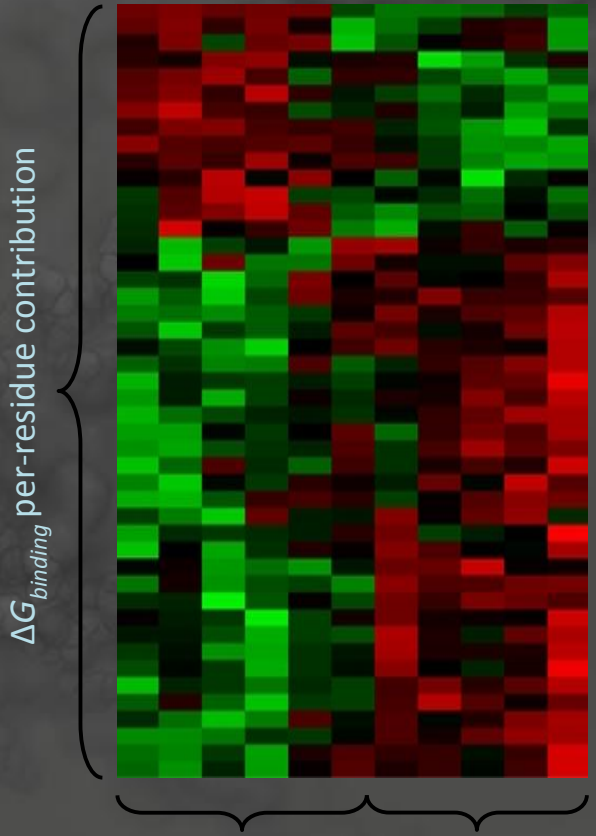
The free energy of binding is calculated by means of the MM-GBSA method, implementing a Single Frame Protocol (SFP) strategy. The choice of receptor and ligand is arbitrary, since both of the protomers (almost) equally contribute to the dimerization process.

The Energy (eigen)Decomposition analysis allows to define a interaction energy matrix in which the most relevant residue pairwises that contribute to dirupting/stabilizing effects are emphasized.

$$A_{ij} \cong \sum_{k=1}^{K} \lambda_k \, w_i^k \, w_j^k$$



Two Sample Rank Products



Rank Products is a positional method to combine ranked lists. In this specific case, the residues that belongs to each model produced are ranked, according to their contributions to the overall free energy of binding of the dimer.

$$R_g = \prod_{i=1}^{K} \frac{r_{i,j}}{N_i}$$

In order to compare the rank products provided from two alternative kind of PPI the two sample variant has been adopted.

$$\log(RP_g) = \log(R_g^A) - \log(R_g^B)$$

Predicting Hot Spots

A hot spot is a residue predicted to have a critical role for the dimerization. The more detailed definition of hot spot varies, according to the different assumptions established by the PPI hot spot prediction methods.

Robetta

Kortemme, T. et al. (2002)
Proc Natl Acad Sci USA 99, 14116-21

the predictions are obtained by performing an *in silico* alanine scanning and calculating the ΔΔG upon mutation with an internal energy function based on rotamers evaluation.

HotPoint

Tuncbag, N. et al. (2010)
Nucleic Acid Res 38, W402-6

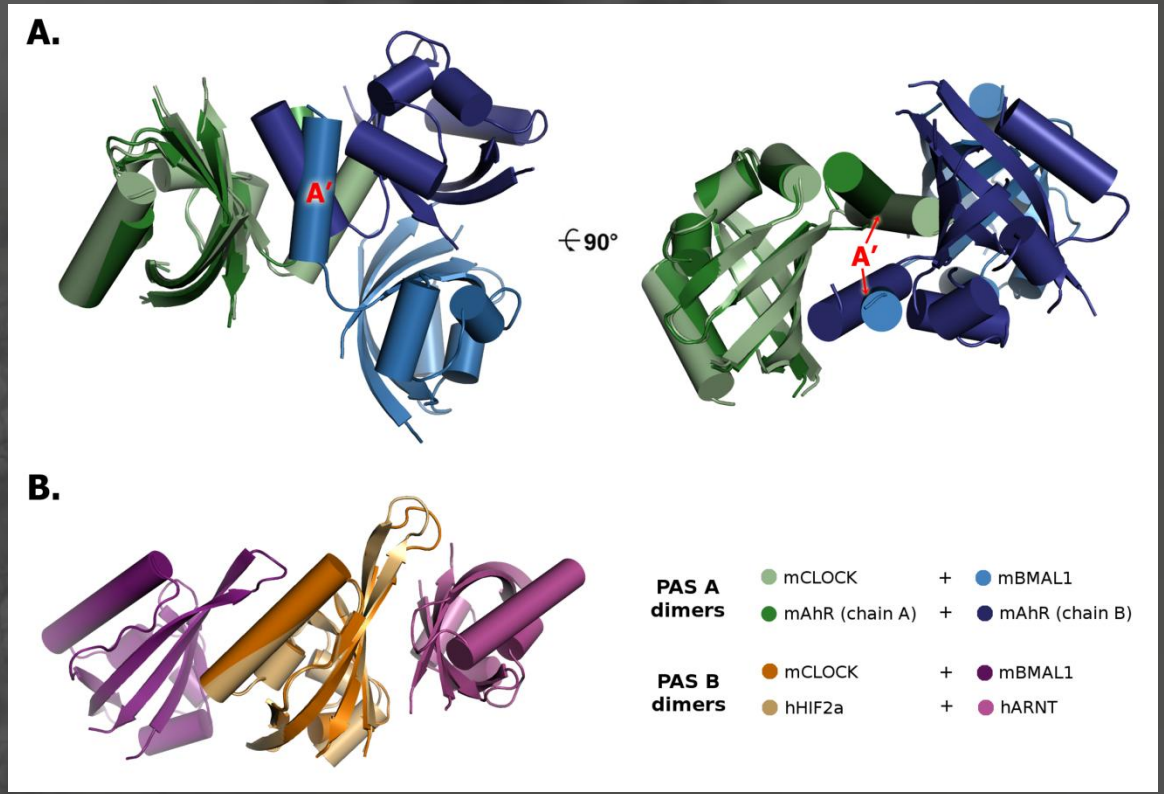
the predictions provided are based both on evaluation of ΔSASA and on a scoring function termed potential contact, based on the number of interacting residues in a shell of 7 Å

KFC2

Zhu, X. et al. (2011)
Proteins 79, 2671-83

the predictions are based on a machine learning method trained on 47 different features derived from solvent accessibility and biochemical properties of the residues (e.g.: hydrophobic profiles, non-bonded interactions and π-stacking interactions)

Homology modelling: single- vs. multi-template

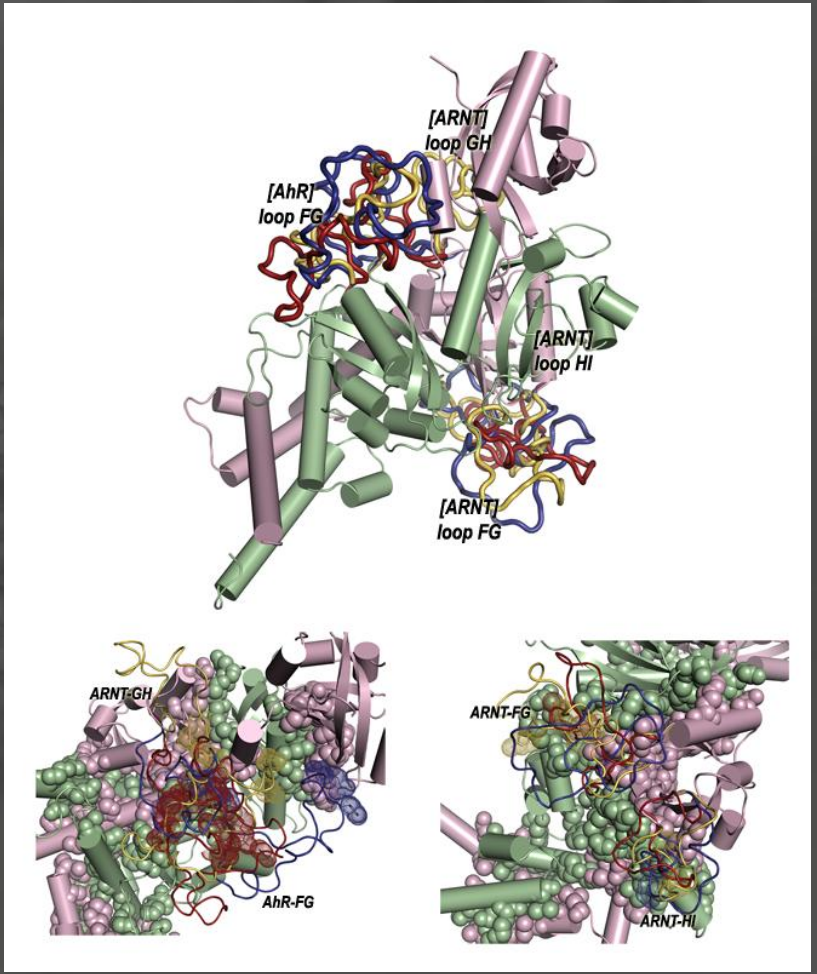


The different dimerization modes observed for the templates span from opposite orientations of the partner (PAS-B dimer) to not negligible spatial arrangement of secondary structural elements.

For such reasons a multi-template approach to homology modelling seems not feasible in this specific case.

NOTE (panel B of the figure): the PAS-B dimerization mode observed for the crystal **4ZP4** (HIF2a:ARNT complex) resemble the dimerization mode of the crystal **4F3L** (CLOCK:BMAL1 complex), in contrast with that showed in the crystal **3F1P** (HIF2a:ARNT PAS-B heterodimer)

Rosetta Native Loops Building



More than 1,000 loop models have been generated, the *next generetion KIC with fragments* method was adopted, according with *Talaris 2014* protocol

Ò Conchùir, S. et al. (2015) *PLoS ONE* 10, e0130433

A limited number of representatives was chosen upon cluster analysis of the putative models

The native loops seems to extend the dimerization interfaces marginally, with minor patches

Main PPIs and Related Energy Couplings