iScience, Volume *25*

Supplemental information

Trogocytosis in innate immunity to cancer is an intimate relationship with unexpected outcomes

Fabrizio Mattei, Sara Andreone, Francesca Spadaro, Francesco Noto, Antonella Tinari, Mario Falchi, Silvia Piconese, Claudia Afferni, and Giovanna Schiavoni

Data S1. Protein docking analysis and generation of 3D *in silico* protein models. Related to Figures 7 and 8.

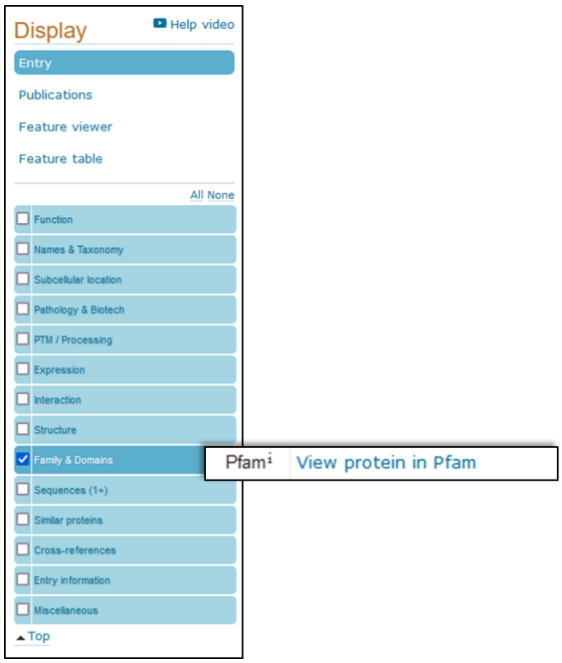
UniProt to RCSB-PDB protein code conversion by highest similarity score procedure.

Whenever required, protein structure database codes from The Universal Protein Resource (UniProt, https://www.uniprot.org) were transformed into the equivalent for the US Data Center for the Global Protein Data Bank (RCSB-PDB, https://www.rcsb.org).

This procedure allows to convert a protein code and 3D structure from UniProt to the exact RCSB-PDB structural equivalent. For the conversion of the human talin-2 protein starting from UniProt, the protein code must be inserted in the Uniprot search box (Q9Y4G6, Suppl. Figure 1).

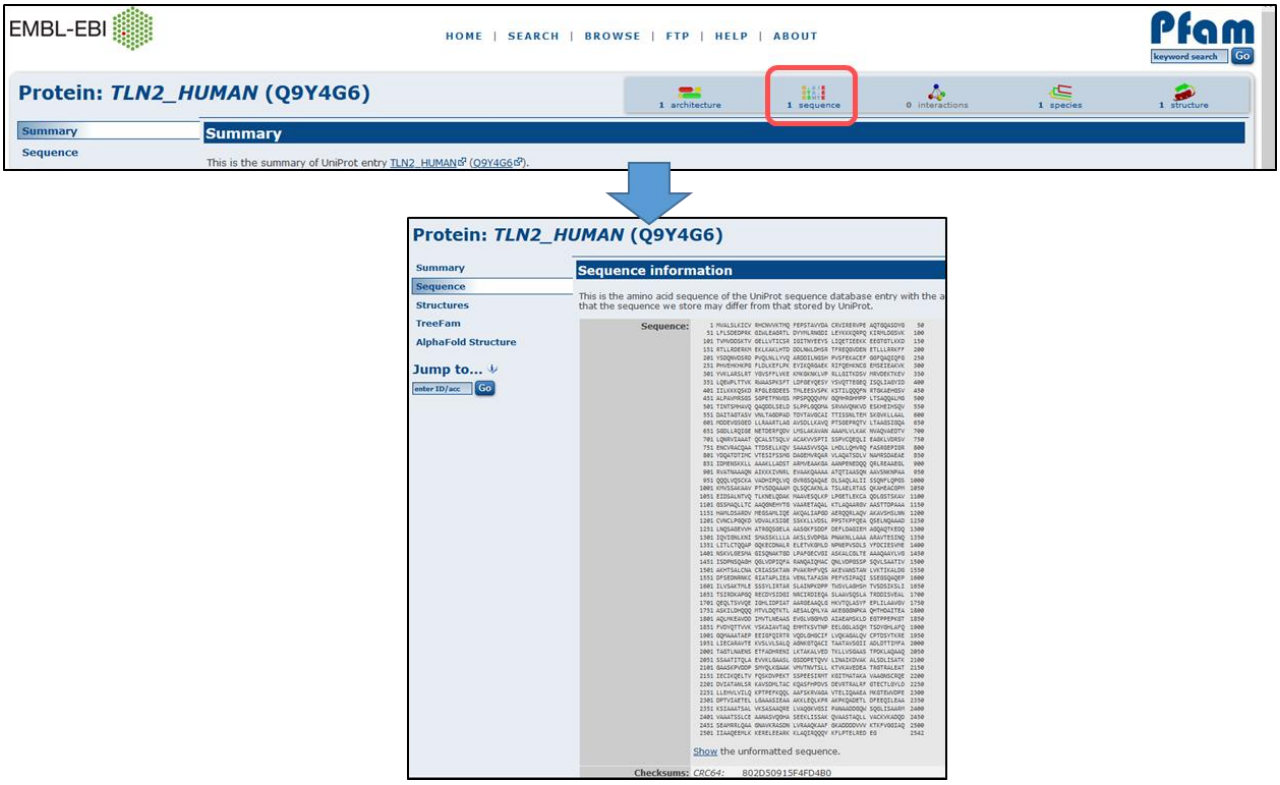
Supplementary Figure 1. Finding a 3D protein structure in the UniProt database (human talin-2, Q9Y4G6)

In the resulting search page, the link "Family domains" at the left-sided menu of the page (Suppl. Figure 2) was selected. A set of options were activated, and page has been scrolled to click a link allowing to view our desired protein in the Protein Families Database (PFAM, http://pfam.xfam.org). This link is named as "View protein in Pfam". PFAM does list a wide collection of protein families, identified by their aminoacidic sequence or by alignments.



Supplementary Figure 2. Checking the category to be viewed ("Family domains") and scrolling the page to select the link "View protein in Pfam" inside the category "Family and domains" databases.

These sets of actions allowed to migrate our job inside the PFAM website, where there was possible to select and copy in the clipboard the aminoacidic protein sequence of human talin-2.

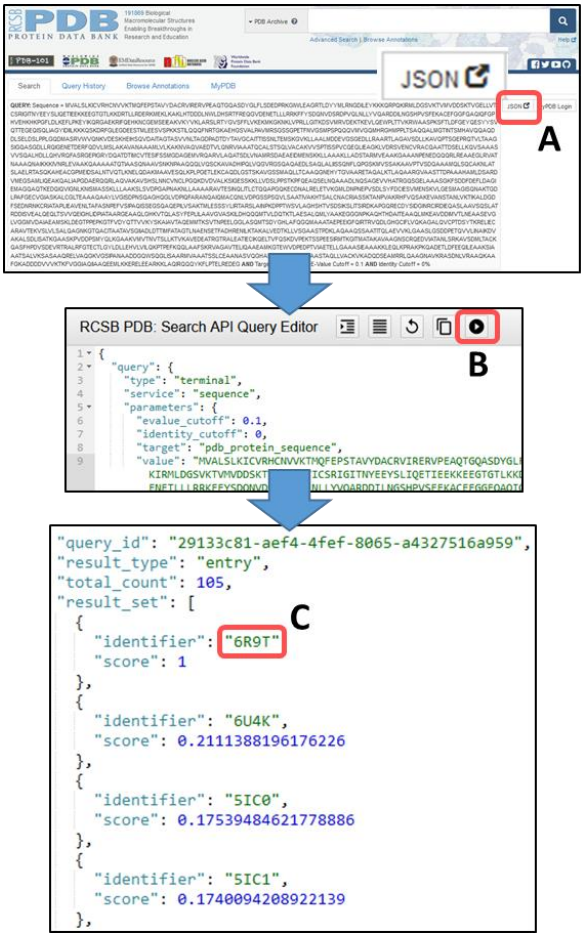


Supplementary Figure 3. Selecting the talin-2 whole primary sequence in the PFAM website. Red box depicts the link to be clicked to view the aminoacidic sequence.

This was performed by clicking on the link “Sequence”, and the primary talin-2 sequence does appear (Suppl. Figure 3). There was now the possibility to copy that sequence (in the unformatted version, without side numbers) and paste it into the RCSB-PDB webpage.

Supplementary Figure 4. Transferring the talin-2 sequence into the RCSB-PDB website. Navigate into the displayed webpage, click to “Advanced search” (A, red box) and then choose the tab “Sequence” (B, red box). Then do a copy&paste action to paste the talin-2 protein sequence (C) in this box and finally click on the “Search button at the right bottom of the page (D, red box).

By browsing into the RCSB-PDB webpage, a copy and paste operation for the talin-2 protein sequence has been done to search for the corresponding sequence in this database (Suppl. Figure 4). At this point, a search tab is activated in this page, which contains the pasted sequence.



Supplementary Figure 5. Recovery of the talin-2 code into the RCSB-PDB database. To start searching the RCSB-PDB database for the talin-2 sequence, JSON) script, the appropriate JSON tab must be selected (A, red box). The script will be displayed in the newly generated page, where it must be executed by selecting the appropriate button (B, red box). At this point a list of RCSB-PDB codes will appear, classified through a similarity-based score. The code whose score is equal to the value 1 is the real RCSB-PDB code for the talin-2 protein (C, red box).

To start the search, a Javascript Object Notation (JSON) script has been created by selecting the correspondent link into the page (Suppl. Figure 5). This provided the final code used for the RCSB-PDB database for human talin-2, namely the code 6R9T (red box C in Suppl. Figure 5). Therefore, the specific protein information (sequence and structures) can be retrieved by using the two codes Q9Y4G6 and 6R9T (for Uniprot and RCSB-PDB databases, respectively).

In order to proceed with in silico 3D protein modelling, this procedure has been applied for the other Uniprot-derived proteins displayed in Figures 7B and 8C-D of the paper. This is due to the fact that ClusPro, our preferred protein-protein docking webtool, only accepts RCSB-PDB coded proteins, or User-defined PDB files. This has been explained in detail in the section below.

Supplementary Table I evidences the Uniprot proteins and equivalent RCSB-PDB codes obtained by applying our aforementioned highest similarity score procedure.

Supplememtary Table 1. UniProt proteins (human) and equivalent RCSB-PDB codes selected by the highest similarity score.

Protein name	Referring manuscript figure	UniProt code	Returned RCSB-PDB code
Integrin alpha-M	Figure 3A, 3B	P11215	7P2D
Talin-2	Figure 3A, 3B	Q9Y4G6	6R9T
Kindlin-3	Figure 3A, 3B	Q86UX7	7C3M

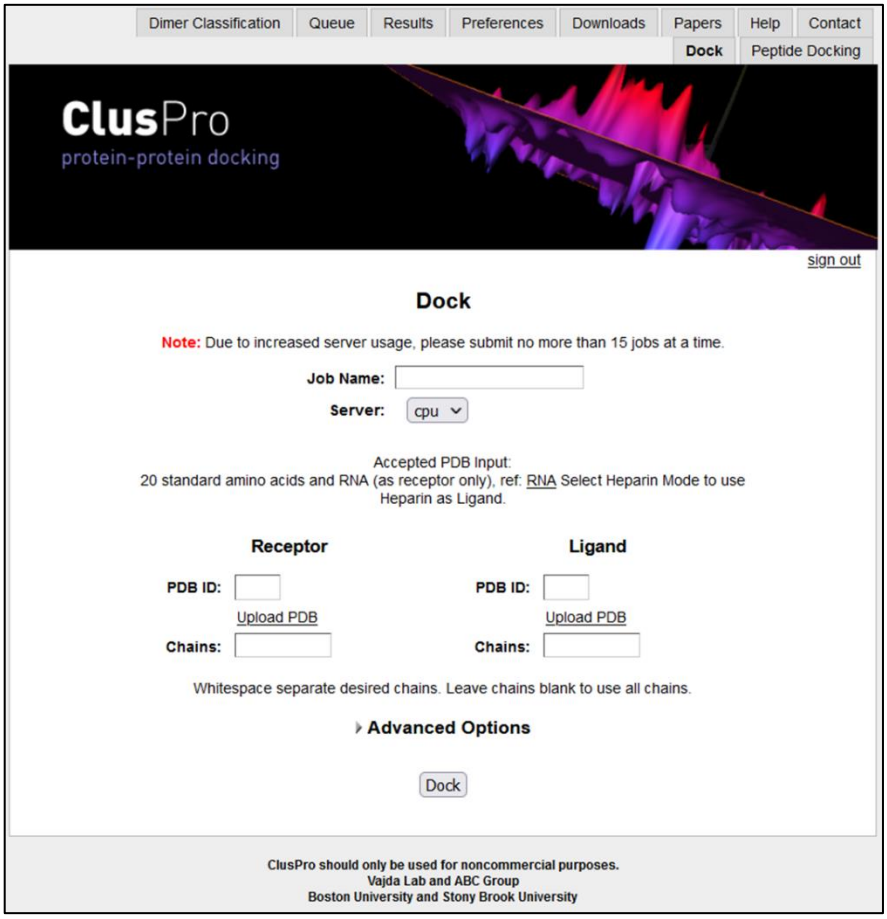
At this point, the RCSB-PDB codes selected by the highest similarity score have been employed to start the in silico 3D protein modelling. The RCSB-PDB codes of the proteins not mentioned in Supplementary Table I were directly obtained by searching them in this database.

In silico 3D protein modelling.

This method allows the generation of 3D structures of protein complexes by protein-protein docking modelling, based on the use of advanced ad hoc mathematical algorithms. To perform our in silico studies, we employed the protein-protein docking webtool ClusPro (https://cluspro.org/), hosted by the Boston University (Boston, MA, USA) and Stony Brooks University (New York, NY, USA) with the aid of the 3D protein viewer PyMol (downloadable at https://pymol.org/). ClusPro (Kozakov et al., 2017) is an automated and versatile protein docking server, based on a Fast Fourier Transform (FFT) software PIPER (Kozakov et al., 2006). By inputting the PDB file (user-generated or by codes from RCSB-PDB database) of two proteins, ClusPro easily allow to obtain a detailed 3D in silico model of the protein-protein interactions, fully viewable by a common 3D protein viewer (i.e., PyMol). Furthermore, ClusPro can elegantly be used in a sequential mode, where the first two proteins will be docked and the obtained in silico 3D model will be re-docked with another protein and so on. This procedure will allow to create advanced protein complex models, based on the interaction of multiple proteins (Figure 7).

In some specific cases, a web-based 3D protein viewer has been used to generate the 3D model of a particular protein. For example, to browse proteins in the UniProt and RCSB-PDB database web sites we exploited the default in-site 3D protein viewer.

To obtain the 3D in silico model showed in Figure 7B, a first protein-protein docking has been processed with the actin homotrimer (RCSB-PBD code: 7LRG) and kindlin-3 (Suppl. Figure 6).



Supplementary Figure 6. The protein-protein docking main page to insert data (RCSB-PDB codes or user-generated PDB files). Job name, receptor, ligand and chains to be used for each inputted protein must be added in this module. To start docking process the button "Dock" must be clicked.

Firstly, the codes were added in the appropriate Receptor/Ligand boxes. In this case, we used the actin homotrimer as receptor and the human kindlin-3 as ligand. No advanced option has been selected. At the end of the processing ClusPro provides alternative models (usually at least 10 models)

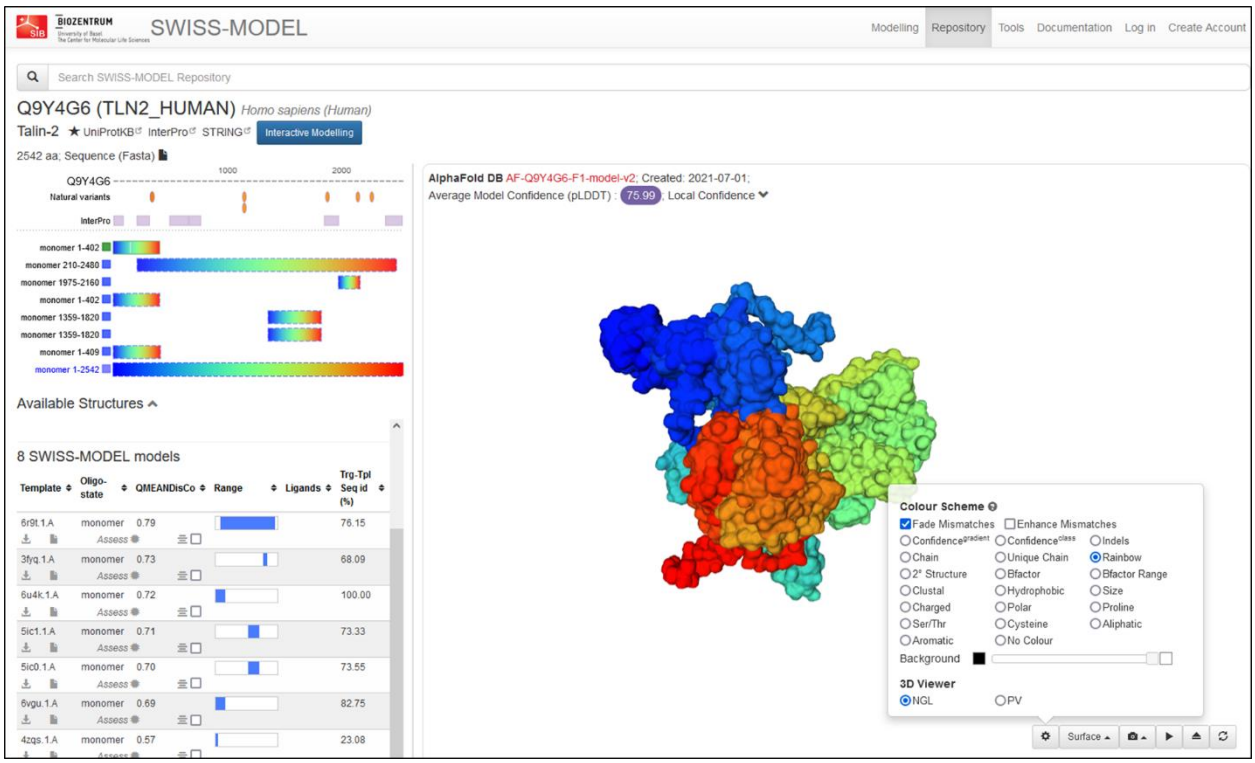
to be downloaded, based on type of interactions being interested to. We have chosen an *in silico* model where the interactions taken into consideration are mainly based on Van der Waals forces and electrostatic associations, recognised to be central for the formation of protein complexes. Our provisional model (Model-1) was then downloaded as user-generated PDB file. The Model-1 was used for a second docking analysis by ClusPro, where the input receptor (Suppl. Figure 6) was represented by this model and the Ligand was represented by human talin-2 as RCSB-PBD code (6R9T, Suppl. Table I). The resulting actin trimer/kindlin-3/talin-2 in silico 3D model (Model-2) has been employed for a third docking analysis, where the latter model was used as Ligand for the integrin alpha-M (7P2D, Suppl. Table I). For this docking analysis we excluded the chain B of the integrin. The resulting final *in silico* model (Model-3) is represented in Figure 7B, and strictly resembles to the hypothetic model drawn in Figure 3A. Then, thanks to the PIPER software of ClusPro, we were able to obtain a new *in silico* advanced protein complex by a sequential docking analysis starting by an actin homotrimer, which allowed to obtain the final Model-3 via the intermediates Model-1 and Model-2.

For the *in silico* 3D complex showing the association between fascin and actin homotrimer (Figure 8C), we employed the relative RCSB-PDB codes extrapolated from this database (1DFC and 7LRG, respectively). The in silico model defined by the association of BAIAP2 and the actin trimer (Figure 8D) has been yielded by using the same procedure with the relative RCSB-PDB codes (1Y2O and 7LRG, respectively).

The mean computing time to process models can vary depending on the traffic request inside the ClusPro website and on the complexity of inputted 3D protein models. This time could range from 5-10 hours to 1-5 days, with a mean of 2 days of computing time (Comeau et al., 2004).

Generation of the surface 3D model.

The model delineated in Figure 7A was obtained by using the Swiss-Model NGL protein viewer, available at the EXPASY website (https://swissmodel.expasy.org/). Swiss-Model is a protein viewer adapted for UniProt-derived proteins, and equipped with numerous and versatile view options.



Supplementary Figure 7. Swiss Model viewer interface. Example of use and browsing with the human talin-2 3D model layout by the in-site NGL viewer.

Each of the UniProt protein described in Figure 7A (integrin beta-2, integrin alpha-M, talin-2, kindlin-3 and actin monomers) was visualized via the NGL viewer by using their UniProt codes (P05107, P11215, Q9Y4G6, Q86UX7 and P68032, respectively). Each 3D model of these proteins was then visualized by selecting the "Surface" option of Swiss-Model (Suppl. Figure 7). At this point, the 3D models were properly rotated around their XYZ axes and the screenshots were then used to compose the proposed model in Figure 7A of the manuscript. Furthermore, the option "Rainbow" has been selected in order to map the C- and N-terminus of each 3D model (Suppl. Figure 7).