
*List of Equations



Mafalda Inês Apolinário Pereira

Bachelor of Sciences in Chemical and Biochemical Engineering

Design of Peptides to Interfere with the RANK-TRAF6 Pathway: An Integrated Approach

Dissertation to obtain the degree of Master in
Chemical and Biochemical Engineering

Adviser: Doctor João Domingos Galamba Correia, Instituto Superior Técnico, University of Lisbon

Co-adviser: Doctor Rita Lourenço Paiva de Melo, Instituto Superior Técnico, University of Lisbon

Examination Committee

Chairperson:

Raporteurs:



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

October, 2018

Mafalda Inês Apolinário Pereira

Bachelor of Sciences in Chemical and Biochemical Engineering



**FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA**

Design of Peptides to Interfere with the RANK-TRAF6 Pathway: an Integrated Approach

Dissertation to obtain the degree of Master in
Chemical and Biochemical Engineering

Adviser: Doctor João Domingos Galamba Correia

Co-adviser: Doctor Rita Lourenço Paiva de Melo

October, 2018

Design of Peptides to Interfere with the RANK-TRAF6 axis: an Integrated Approach

Copyright © Mafalda Inês Apolinário Pereira, Faculty of Sciences and Technology, NOVA University Lisbon.

The Faculty of Sciences and Technology and the NOVA University Lisbon have the right, perpetual and without geographical boundaries, to file and publish this dissertation through printed copies reproduced on paper or on digital form, or by any other means known or that may be invented, and to disseminate through scientific repositories and admit its copying and distribution for non-commercial, educational or research purposes, as long as credit is given to the author and editor.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Doctor João Domingos Galamba Correia and Doctor Rita Lourenço Paiva de Melo, for the efforts, support and knowledge devoted to my learning over the course of 6 months, and for letting me explore the investigation fields of Biochemistry and Computational Chemistry in their research home.

Thanks also to my laboratory colleague Ruben Silva for all the help provided, and also to Doctor Célia Fernandes and Doctor Francisco Silva, who took precious time to help me and teach me to work with HPLC and mass spectrometry.

A very special thank you to my boyfriend and best friend, whose support, help and love have been phenomenal and crucial for this thesis dissertation.

I thank my family and friends for their support and attention in this final and important stage of my course.

"When you are inspired by some great purpose, some extraordinary project, all your thoughts break their bonds; Your mind transcends limitations, your consciousness expands in every direction, and you find yourself in a new, great and wonderful world." —Patanjali

ABSTRACT

The RANK-TRAF6 pathway is involved in osteoporosis and in the development of breast cancer. One way to inhibit this pathway is to block TRAF6's binding to RANK using decoy peptides. It is known by experimental data that a cell-penetrating decoy peptide (cpp) based on TRAF6-binding site mTRANCE-R(1) effectively inhibits this pathway because of its high affinity to TRAF6, whereas the TRAF6-binding site mTRANCE-R(2) does not. One of the goals of this work is to validate by in silico studies the stability of TRAF6 with five peptides, containing TRAF6-binding sites in human (hPEP1, hPEP2, PepH2-hTRANCE1) and mice (mPEP1, mPEP2).

The stability of TRAF6 with the different peptides was accessed with two methods.

The first method uses the complexes buried surface area (BSA) values, as a higher BSA indicates a higher complex stability. The results obtained with this method are in agreement with the experimental data, suggesting that the peptide with the sequence mTRANCE-R(1) has a higher affinity to TRAF6 than the peptide with the sequence mTRANCE-R(2), indicating that it is more stable with TRAF6. $BSA(mTRAF6-mPEP1) > BSA(mTRAF6-mPEP2)$ and $BSA(hTRAF6-hPEP1) > BSA(hTRAF6-hPEP2)$. $BSA(hTRAF6-PepH2-hTRANCE1) > BSA(hTRAF6-hPEP1)$.

The other method uses an optimized HADDOCK score as a tool to predict binding affinity. These results are not in agreement with the literature and the ones determined by BSA studies, failing in validating the stability of TRAF6 with a peptide containing mTRANCE-R(1) and a peptide containing mTRANCE-R(2).

Peptides hTRANCE-R(1) and PepH2-hTRANCE1 (>99%) were prepared by manual Fmoc solid phase peptide synthesis followed by HPLC purification. In the PepH2-hTRANCE1 synthesis, besides the desired peptide, we also isolated a side product that was identified as the same peptide sequence with a Valine deletion. This peptide will also be used for future cancer cell studies. The secondary structure of the peptides was analysed by circular dichroism, showing the same fractions of β -sheet (0,4) and random coil (0,6) for all.

Keywords: BSA, Optimized HADDOCK score, Stability, Solid Phase Peptide Synthesis, Purification, Circular Dichorism

RESUMO

A via RANK-TRAF6 está envolvida na osteoporose e no desenvolvimento do cancro da mama. Uma maneira de inibir esta via é bloquear a ligação do TRAF6 ao RANK, usando pequenos péptidos. Dados experimentais indicam que um péptido penetrante (cpp) contendo o local de ligação do RANK ao TRAF6, mTRANCE-R(1), inibe eficazmente esta via devido à sua elevada afinidade para o TRAF6, enquanto que outro péptido contendo o local de ligação do RANK ao TRAF6 mTRANCE-R(2) não. Um dos objetivos deste trabalho é validar, através de estudos *in silico*, a estabilidade dos complexos formados entre o TRAF6 e cinco péptidos que contêm os locais de ligação do RANK ao TRAF6 em organismos de humano (hPEP1, hPEP2, PepH2-hTRANCE1) e de rato (mPEP1, mPEP2).

A estabilidade do TRAF6 com os diferentes péptidos foi estudada através de dois métodos.

O primeiro método utiliza os valores da área de superfície enterrada dos complexos (BSA), uma vez que uma BSA mais elevada indica uma maior estabilidade. Os resultados deste método estão de acordo com dados experimentais, sugerindo que o péptido com a sequência mTRANCE-R(1) tem uma maior afinidade para o TRAF6 do que o péptido com a sequência mTRANCE-R(2), indicando que forma um complexo mais estável com TRAF6. BSA (mTRAF6-mPEP1) > BSA (mTRAF6-mPEP2) e BSA (hTRAF6-hPEP1) > BSA (hTRAF6-hPEP2). BSA (hTRAF6-PepH2-hTRANCE1) > BSA (hTRAF6-hPEP1).

No segundo método é otimizada a equação utilizada pelo HADDOCK, servindo como uma ferramenta para prever a afinidade entre proteínas e péptidos. Estes resultados não estão de acordo com a literatura e com os resultados determinados através da BSA, não validando a estabilidade do TRAF6 com o péptido contendo mTRANCE-R(1) e o péptido contendo mTRANCE-R(2).

Os péptidos hTRANCE-R(1) e PepH2-hTRANCE1 (> 99% pureza) foram obtidos manualmente em fase sólida pela técnica Fmoc após purificação por HPLC. No caso do péptido PepH2-hTRANCE1, isolou-se ainda um produto que corresponde à mesma sequência sem uma Valina. Este péptido também será usado para futuros estudos com culturas celulares. A estrutura secundária dos péptidos foi analisada por dicorismo circular, mantendo-se as mesmas frações de folha- β (0,4) e *random coil* (0,6) para todos.

Palavras-chave: BSA, HADDOCK score otimizado, Estabilidade, Síntese de Peptídeos em Fase Sólida, Purificação, Dicroísmo Circular

CONTENTS

xvii

Abbreviations

xxv

1	Introduction	1
1.1	RANK-TRAF6 Pathway	1
1.1.1	TRAF6	3
1.1.2	Decoy Peptides	5
1.2	Protein Interactions	7
1.2.1	Interface and Hot Spots	7
1.2.2	Binding Affinity	9
1.3	Computational Chemistry	10
1.3.1	Tertiary Structure Prediction	10
1.3.2	Molecular Docking	13
1.4	Solid Phase Peptide Synthesis	18
1.4.1	Circular Dichroism	20
1.5	Objectives	21
2	Methods and Materials	23
2.1	TRAF6 Sequence Conservation	23
2.2	Interface & Hot Spots Prediction	24
2.3	Tertiary Structure Prediction	25
2.3.1	Structure Evaluation	26
2.4	Molecular Docking	27
2.4.1	HADDOCK Guru Protocol	27
2.4.2	Complex Structure Evaluation	28
2.5	Binding Affinity	29
2.6	Peptide Synthesis	30
2.6.1	Materials	30
2.6.2	Synthesis and Characterization of the Peptides	30
2.7	Peptide Purification	33
2.8	Peptide Concentration	33
2.9	Circular Dichroism	34

CONTENTS

3	Results and Discussion	35
3.1	TRAF6 Sequence Conservation	35
3.2	Active Residues	36
3.2.1	Computational Alanine Scanning	36
3.2.2	H-bonds	38
3.2.3	Interface & Active Residues	39
3.3	Tertiary Structure Prediction	40
3.4	Molecular Docking	41
3.4.1	hTRAF6-hPEP1 Cluster Analysis	41
3.4.2	hTRAF6-hPEP1 Model Selection	45
3.4.3	Complexes Model Selection	47
3.5	Binding Affinity	51
3.5.1	Buried Surface Area (BSA)	51
3.5.2	Optimized HADDOCK Score	53
3.6	Peptide Synthesis and Purification	54
3.6.1	Circular Dichroism	59
4	Conclusions and Perspectives	61
	Bibliography	65
	Annexes	77

1.1	Complexes buried surface area.	7
1.2	Overall change in the binding free energy.	8
1.3	K_d reaction equation and formula.	9
1.4	HADDOCK scoring function in the it0 stage.	15
1.5	HADDOCK scoring function in the it1 stage.	15
1.6	HADDOCK scoring function in the final stage.	15
1.7	Optimized HADDOCKscore.	16
1.8	Intermolecular energy equation for HADDOCK output models.	17
2.1	X^2 function to be minimized.	30
2.2	Peptide concentration.	33
3.1	Percentage of active residues conserved in the interface.	46
3.1	Optimized HADDOCK score for protein-peptide complexes.	53

LIST OF FIGURES

1.1	Schematic illustration of NF- κ B gene transcription.	1
1.2	Schematic overview of TRAF6 signalling pathway.	2
1.3	Non-scale illustration of human TRAF6 domain structure.	3
1.4	TRAF6-binding sites for TNFR family members and the TRAF6-binding motif.	4
1.5	Surface representation of TRAF6.	4
1.6	Crystal structure of TRAF-C domain of TRAF6 in complex with a human fragment of RANK.	5
1.7	Relationship between the $\log(K_d)$ and the BSA of protein-peptide and protein-protein complexes.	9
1.8	<i>Ab initio</i> and heuristic methods for protein structure prediction.	10
1.9	Important dihedral angles	11
1.10	Comparative Modelling Process	12
1.11	Example of an amino acid chain with blue N and C terminal amino acid residues.	18
1.12	SPPS	19
1.13	PS crosslinked with DVB	20
1.14	Nihidrine reaction with the peptide's free amine.	20
1.15	Poly-L-lysine circular dichroism spectra for the three main conformations.	21
2.1	SPPS assembly.	31
2.2	Resin in the filter reaction vessel.	31
2.3	Kaiser test trail in two resin samples.	32
3.1	BLAST result in pairwise sequence alignment between hTRAF6 and mTRAF6.	35
3.2	mTRAF6 and hTRAF6 sequence alignment, with the active residues in red.	37
3.3	H-bonds between hTRAF6 and RANK fragment.	39
3.4	Number of hydrogen bonds and participating residues in the 1LB5 complex.	39
3.5	hTRAF6 active and passive residues.	40
3.6	mTRAF6 tertiary structure.	41
3.7	HADDOCK score versus l-RMSDs for the hTRAF6-hPEP1 complex.	42
3.8	HADDOCK score versus i-RMSDs for the hTRAF6-hPEP1 complex.	43
3.9	Average intermolecular energy versus the 4 best clusters for the hTRAF6-hPEP1.	43

3.10	BSA ($\text{\AA}^2$) versus the HADDOCK score for the 16 study structures representing the hTRAF6-hPEP1 complex.	44
3.11	hTRAF6-hPEP1 fit of the first model from each cluster.	45
3.12	Model3_2 aligned to complex 1LB5.	46
3.13	hTRAF6-hPEP2 fitting of the best models from the two first clusters.	47
3.14	mTRAF6-mPEP1 fitting of the best models from the two first clusters.	48
3.15	mTRAF6-mPEP2 fitting of the best models from the two first clusters.	49
3.16	hTRAF6-PepH2-hPEP1 fitting of the best models from the two first clusters.	50
3.17	BSA ($\text{\AA}^2$) of the best chosen models.	52
3.18	Average BSA ($\text{\AA}^2$) of each complex.	52
3.19	Correlation between the HADDOCK scores and $\log(K_{\text{predicted}})$	54
3.20	TRANCE-R(1) ESI-MS before purification.	55
3.21	TRANCE-R(1) chromatogram after purification (analytical HPLC).	56
3.22	TRANCE-R(1) ESI-MS after purification.	56
3.23	PepH2-hPEP1 mass spectrum before purification.	57
3.24	PepH2-hPEP1 chromatogram after purification.	57
3.25	PepH2-hPEP1 mass spectrum after purification.	58
3.26	PepH2-hTRANCER1 (-V) mass spectrum after purification.	58
3.27	CD spectra for the three peptides synthesized.	60
A.1	Amino acids periodic chart.	77
A.2	Virtual alanine scanning from Robetta.	79
A.3	Virtual alanine scanning from KFC2.	81
A.4	Virtual alanine scanning from Hotpoint.	83
A.5	I-TASSER Model 1 quality score calculated by ProSA for peptide hPEP1.	89
A.6	I-TASSER Model 2 quality score calculated by ProSA for peptide hPEP1.	90
A.7	I-TASSER Model 3 quality score calculated by ProSA for peptide hPEP1.	90
A.8	I-TASSER Model 4 quality score calculated by ProSA for peptide hPEP1.	91
A.9	I-TASSER Model 5 quality score calculated by ProSA for peptide hPEP1.	91
A.10	PEP-FOLD3 Model 1 quality score calculated by ProSA for peptide hPEP1.	93
A.11	PEP-FOLD3 Model 2 quality score calculated by ProSA for peptide hPEP1.	94
A.12	PEP-FOLD3 Model 3 quality score calculated by ProSA for peptide hPEP1.	94
A.13	PEP-FOLD3 Model 4 quality score calculated by ProSA for peptide hPEP1.	95
A.14	PEP-FOLD3 Model 5 quality score calculated by ProSA for peptide hPEP1.	95
A.15	PEPstrMOD Model 1 quality score calculated by ProSA for peptide hPEP1.	97
A.16	PEPstrMOD Model 2 quality score calculated by ProSA for peptide hPEP1.	98
A.17	PEPstrMOD Model 3 quality score calculated by ProSA for peptide hPEP1.	98
A.18	PEPstrMOD Model 4 quality score calculated by ProSA for peptide hPEP1.	99
A.19	PEPstrMOD Model 5 quality score calculated by ProSA for peptide hPEP1.	99
A.20	HADDOCK results for the complex hTRAF6-hPEP1.	105
A.21	Intermolecular energy versus l-RMSD for the hTRAF6-hPEP1 complex.	109

A.22 Intermolecular energy versus i-RMSD for the hTRAF6-hPEP1 complex. . . .	109
A.23 I-TASSER Model 1 quality score calculated by ProSA for peptide hPEP2. . . .	113
A.24 I-TASSER Model 2 quality score calculated by ProSA for peptide hPEP2. . . .	113
A.25 I-TASSER Model 3 quality score calculated by ProSA for peptide hPEP2. . . .	114
A.26 I-TASSER Model 4 quality score calculated by ProSA for peptide hPEP2. . . .	114
A.27 I-TASSER Model 5 quality score calculated by ProSA for peptide hPEP2. . . .	115
A.28 I-TASSER Model 1 quality score calculated by ProSA for peptide mPEP1. . .	117
A.29 I-TASSER Model 2 quality score calculated by ProSA for peptide mPEP1. . .	117
A.30 I-TASSER Model 3 quality score calculated by ProSA for peptide mPEP1. . .	118
A.31 I-TASSER Model 4 quality score calculated by ProSA for peptide mPEP1. . .	118
A.32 I-TASSER Model 5 quality score calculated by ProSA for peptide mPEP1. . .	119
A.33 I-TASSER Model 1 quality score calculated by ProSA for peptide mPEP2. . .	121
A.34 I-TASSER Model 2 quality score calculated by ProSA for peptide mPEP2. . .	121
A.35 I-TASSER Model 3 quality score calculated by ProSA for peptide mPEP2. . .	122
A.36 I-TASSER Model 4 quality score calculated by ProSA for peptide mPEP2. . .	122
A.37 I-TASSER Model 5 quality score calculated by ProSA for peptide mPEP2. . .	123
A.38 I-TASSER Model 1 quality score calculated by ProSA for peptide PepH2-TRANCE1. 125	
A.39 I-TASSER Model 2 quality score calculated by ProSA for peptide PepH2-TRANCE1. 125	
A.40 I-TASSER Model 3 quality score calculated by ProSA for peptide PepH2-TRANCE1. 126	
A.41 I-TASSER Model 4 quality score calculated by ProSA for peptide PepH2-TRANCE1. 126	
A.42 I-TASSER Model 5 quality score calculated by ProSA for peptide PepH2-TRANCE1. 127	
A.43 HADDOCK results for the complex hTRAF6-hPEP2.	133
A.44 HADDOCK results for the complex mTRAF6-mPEP1.	141
A.45 HADDOCK results for the complex mTRAF6-mPEP2.	152
A.46 HADDOCK results for the complex hTRAF6-PepH2-hPEP1.	161
A.47 Peptide hTRANCE-R(1) chromatogram before purification (preparative HPLC) 167	
A.48 Peptide PepH2-hPEP1 chromatogram before purification (preparative HPLC). 169	

LIST OF TABLES

1.1	Potential therapeutic cell-penetrating peptides inhibit TRAF6 signaling . . .	6
1.2	CAPRI criteria for ranking models of protein-peptide complexes	17
2.1	Peptide sequences to be studied <i>in silico</i> in complex with TRAF6, for human and mouse complexes, and hTRANSE-R(1) (RANK's binding site I) to be synthesized	23
2.2	Peptides chemical formula and molecular weigh	32
3.1	hTRAF6 hot spots and binding residues in PDB:1LB5	36
3.2	hTRAF6 and mTRAF6 residues used as active in molecular docking	38
3.3	3 complexes considered "acceptable" from the first two clusters for hTRAF6-hPEP1	46
3.4	Models chosen for each complex, and the respective percentage of active residues conserved in the interface	47
3.5	BSA ($\text{\AA}^2$) of the best chosen models and average BSA ($\text{\AA}^2$)	51
3.6	K_d for each model selected in molecular docking.	54
3.7	Analytical data of the synthesized hTRANSE-R(1).	54
3.8	Analytical data of the synthesized PepH2-TRANCER1.	55
3.9	Peptide concentration and synthesis yield	59
3.10	Deconvolutions of the spectra obtained in CD	59
A.1	RANK hot spots and binding residues in complex with hTRAF6	85
A.2	hTRAF6 interface active residues for the PDB 1LB5	87
A.3	HADDOCK results of the first ranked cluster for complex hTRAF6-hPEP1 .	107
A.4	HADDOCK results of the second ranked cluster for complex hTRAF6-hPEP1	107
A.5	HADDOCK results of the third ranked cluster for complex hTRAF6-hPEP1 .	108
A.6	HADDOCK results of the fourth ranked cluster for complex hTRAF6-hPEP1	108
A.7	HADDOCK results of the fifth ranked cluster for complex hTRAF6-hPEP1 .	108
A.8	hTRAF6 interface residues divided by distances to the peptide hPEP1 for the best model	111
A.9	HADDOCK results of the first ranked cluster for complex hTRAF6-hPEP2 .	135
A.10	HADDOCK results of the second ranked cluster for complex hTRAF6-hPEP2	135
A.11	hTRAF6 interface active residues for the hTRAF6-hPEP2 best models	137

A.12 HADDOCK results of the first ranked cluster for complex mTRAF6-mPEP1 .	143
A.13 HADDOCK results of the second ranked cluster for complex mTRAF6-mPEP1	143
A.14 mTRAF6 interface active residues for the mTRAF6-mPEP1 best models . . .	145
A.15 HADDOCK results of the first ranked cluster for complex mTRAF6-mPEP2 .	153
A.16 HADDOCK results of the second ranked cluster for complex mTRAF6-mPEP2	153
A.17 mTRAF6 interface active residues for the mTRAF6-mPEP2 best models . . .	155
A.18 HADDOCK results of the first ranked cluster for the complex hTRAF6-PepH2- hPEP1	163
A.19 HADDOCK results of the second ranked cluster for the complex hTRAF6- PepH2-hPEP1	163
A.20 hTRAF6 interface active residues for the hTRAF6-PepH2-hPEP1 best models	165

ABBREVIATIONS

3D	Three dimensional.
ACN	Acetonitrile.
AIR	Ambiguous interaction restraints.
ASA	Accessible surface area.
BC	Breast cancer.
BLAST	Basic Local Alignment and Search Tool.
BSA	Buried surface area.
CAPRI	Critical Assessment of Predicted Interactions.
CD	Circular Dichroism.
Cpp	Cell-penetrating peptide.
DCM	Dichloromethane.
DIPEA	N,N-diisopropylethylamine.
DMF	N,N-dimethylformamide.
DVB	Divinylbenzene.
Eint	Intermolecular energie.
ESI-MS	Electrospray ionisation mass spectrometry.
FCC	Fraction of Common Contacts.
H-bonds	Hydrogen bonds.
HADDOCK	High Ambiguity Driven DOCKing.
HBTU	Tetramethyluronium hexafluorophosphate.
HF	Hydrofluoric acid.

ABBREVIATIONS

I κ B	kB inhibitor.
I-TASSER	Iterative Threading ASSEmbly Refinement.
i-RMSD	Interface root-mean-square-deviation.
iL-RMSD	Interface ligand RMSD.
ITC	Isothermal titration calorimetry.
K _d	Dissociation constant.
L-RMSD	Ligand root-mean-square-deviation.
NF- κ B	Transcription nuclear factor κ B.
OCs	Osteoclasts.
PA	Polyacrylamide.
PEG	Polyethylene glycol.
PPIs	Protein-protein interactions.
PR	Progesterone receptors.
ProSA-web	Protein Structure Analysis-web.
PS	Polystyrene.
RANK	Receptor activator of NF- κ B.
RANKL	RANK ligand.
RING	Really Interesting New Gene.
RP-HPLC	Reversed Phase High Performance Liquid Chromatography.
SA	Structural Alphabet.
SPPS	Solid phase peptide synthesis.
SVM	Support vector machine models.
TBM	Template based modeling.
TD	TRAF domain.
TFA	Trifluoroacetic acid.
TIS	Triisopropylsilane.
TNF	Tumour necrosis factor.
TNFR	TNF receptor.
TRAFs	TNF receptor-associated factors.

TRANCE-R TNF related activation-induced cytokine receptor.

INTRODUCTION

1.1 RANK-TRAF6 Pathway

Protein-protein interactions (PPIs) mediate most cellular processes in the cell, such as differentiation, proliferation, signal transduction and cell death [1].

The transcription nuclear factor kB (NF-kB) is a dimeric protein that, bond to its inhibitor IkB, is present inactively in the cell cytoplasm, being its pathway activated by an extracellular stimulus that allows it to enter the nucleus, bind to specific DNA genes and activate target genes expression that mediate cell differentiation, proliferation, survival, and the release of antimicrobial molecules and cytokines for immune response [2–4]. This differential gene expression control allows a variety of responses depending on the stimulus. The transcription factor NF-kB can be activated, among others, by many cytokines that signal through different cell surface receptors. Some cytokines are tumour necrosis factors (TNFs), multifunctional proinflammatory cytokines secreted predominantly by macrophages [5][6].

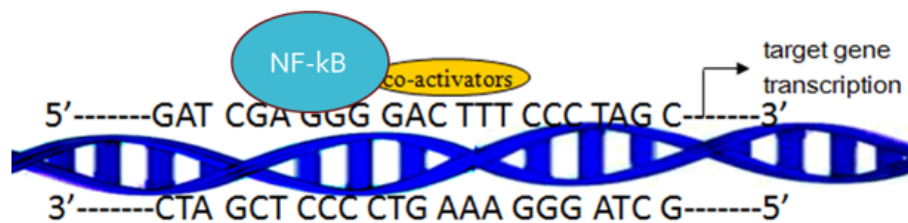


Figure 1.1: Schematic illustration of NF-kB gene transcription [7].

Among all TNF receptors (TNFRs), the transmembrane protein receptor activator of NF-kB (RANK), also known as TRANCE-R (TNF related activation-induced cytokine receptor), has an important role on the NF-kB activation [8]. The extracellular domain

RANK binds to its ligand RANKL, a member of TNF cytokine family, activating an intracellular pathway that culminates with NF- κ B activation [3][6][8].

RANK signalling can only activate NF- κ B when mediated by TNF receptor-associated factors (TRAFs), adaptor proteins that bind to the cytoplasmic tail of TNFRs and allow intracellular signalling pathways [9][10].

There are seven different TRAF members in human and mice (TRAF1-TRAF7) but only TRAF6 is necessary and sufficient for RANK-induced NF- κ B activation [9][11][12].

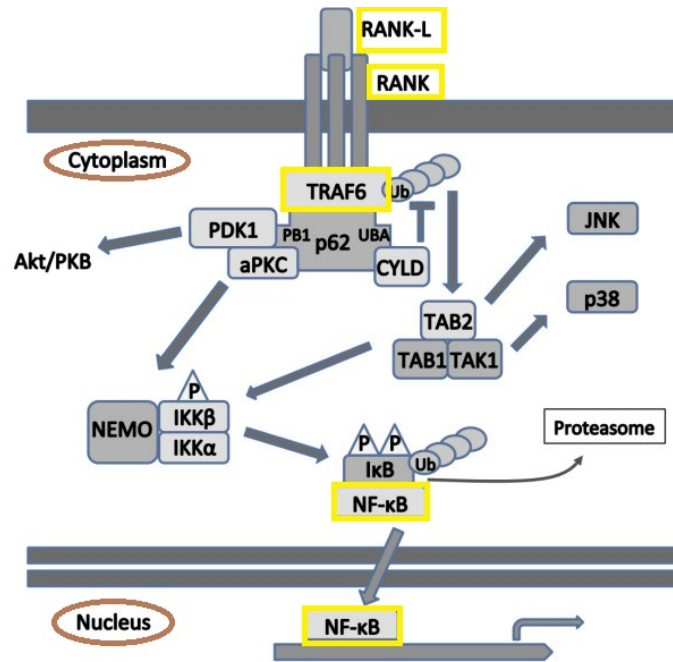


Figure 1.2: Schematic overview of TRAF6 signalling pathway [13].

In this way, RANKL signalling is followed by the RANK's interaction with TRAF6, which in turn recruits the NF- κ B inducing kinase IKK by intermediary factors [14].

This pathway controls osteoclastogenesis, inducing differentiation and activation of osteoclasts, responsible for bone destruction [15][16]. Thus, this pathway is required for the proper formation of cytoskeletal structures [9][17]. However, an increased osteoclast activation leads to osteoporosis, a decreased bone mass condition that causes skeletal abnormalities [18].

TRAF6 is also involved in the development of structures such as mammary glands [19–23]. Studies have shown that NF- κ B is involved in the development of carcinomas, as breast cancer (BC), by NF- κ B binding to DNA of the cancer cells [24–26]. Once NF- κ B has an anti-apoptotic action, its activity during inflammation or infection prevents the elimination of mutated cells and enhances its proliferation [10].

RANK-TRAF6 pathway is also involved in bone tropism of BC patients, as in many cases bone metastases represent a frequent complication, being BC one of the cancer types that preferentially metastasize to bone [15][24][25][27][28]. Cancer cells are chemo-attracted to cytokines (RANKL) and growth factors, which are in a high concentration in

the bone microenvironment. Since RANKL is highly expressed in the bone marrow, is thought that chemo-attracts BC cells [22][29][30].

According to American Cancer Society (2017), BC represents 25% of all new cancer diagnoses in woman globally, which translate to hundreds of millions of women. Although treatments have been improved, some types of BC are still very aggressive, and bone metastasis in cancer patients increase the morbidity and mortality rates [31][32].

1.1.1 TRAF6

Human TRAF6 is a 522-amino acid protein, illustrated in Figure 1.3. TRAF1 up to TRAF6 have in common a specific region near the C-terminal denominated TRAF domain (TD), containing 180 residues responsible for TRAFs interaction with TNFR members, other proteins and kinases. TRAF6 protein sequence and phylogenetic studies revealed that it is the less conserved TRAF family member, despite maintaining all the structural characteristics of the other members of the family, sharing the smallest amount of homology to the prototypical TD because of the TRAF-C, a domain located in the TD that determines protein function, TRAF6's TRAF-C shares only 30% sequence identity with the ones in TRAFs 1-5. Thus, TRAF6 does not share the same binding site as other TRAFs [10][27][33–36].

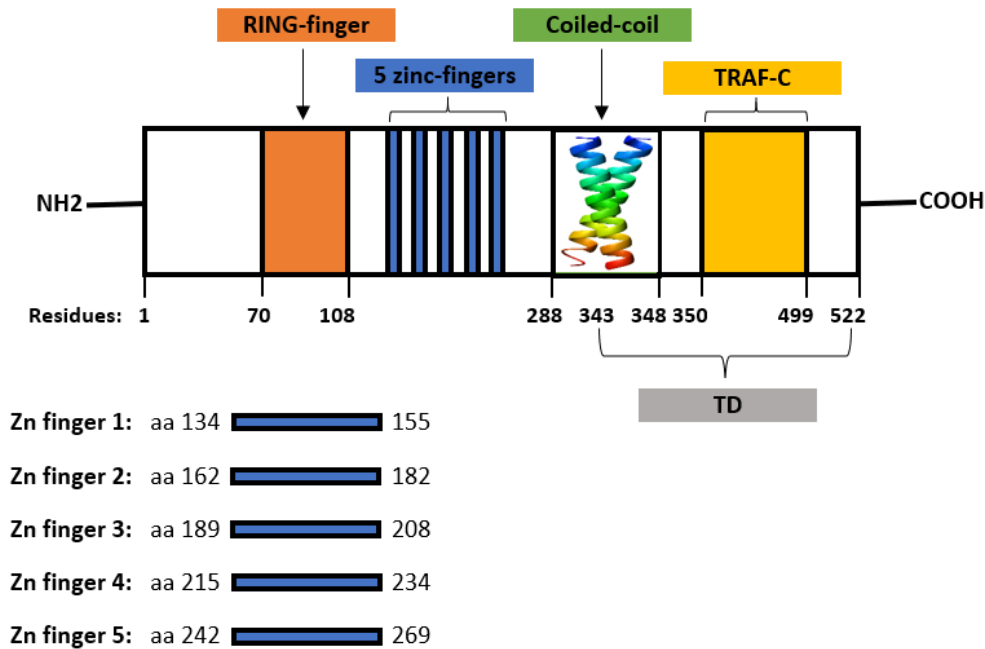


Figure 1.3: Non-scale illustration of human TRAF6 domain structure [19][37].

There are three TRAF6-binding sites in RANK, being all involved in TRAF6 interaction and signal transduction [19]. The presence of three binding sites increases the TRAF6 recruitment and explains the dominant role of TRAF6 in RANK signalling [18][19].

Structure sequence alignment of human and mouse TRAF6-binding sites presented in Figure 1.4, allowed to conclude that in RANK (TRANCE-R), and other members of the TNFR family, the TRAF6-binding motif, an amino acid sequence pattern that is widespread, is Pro-X-Glu-X-X-(aromatic/acid residue) [19][38–40]. Analysing RANK’s residues that contact TRAF6 when in a complex, the buried surface area (BSA, Å) is greater in the motif residues mentioned above. These residues are considered “hot spots”, which are essential for maintaining the integrity of the interface. The residue Glu is determined to be in the position P_0 , being the remaining hot spots in positions P_{-2} (Pro) and P_3 (Ar/Ac) [19]. Isothermal titration calorimetry (ITC) measurements showed that peptides with acidic residues at P_1 and P_2 positions possess higher affinity to TRAF6 owing to their complementarity to the basic TRAF6 surface in that region (Figure 1.5) [19].

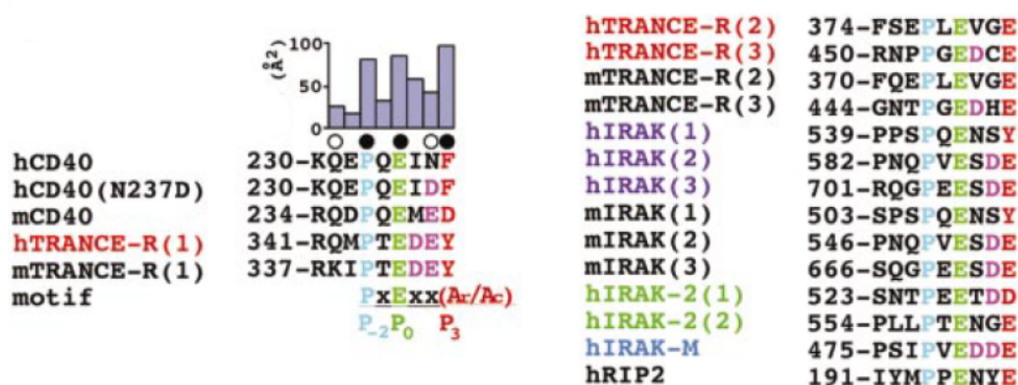


Figure 1.4: TRAF6-binding sites for TNFR family members in human (h) and mouse (m) and the TRAF6-binding motif Pro-X-Glu-X-X-(aromatic/acid residue) [19].

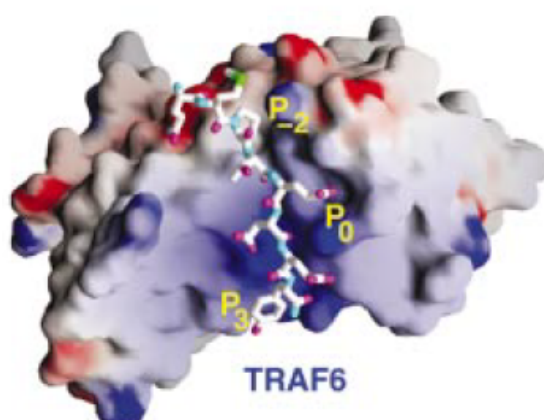


Figure 1.5: Surface representation of TRAF6, coloured based on electrostatic potential, bounded with a TRANCE-R peptide [19].

The TRAF6-RANK complex represented in the Figure 1.6 is related to the TRAF-C domain of TRAF6 binding with hTRANCE-R(1).

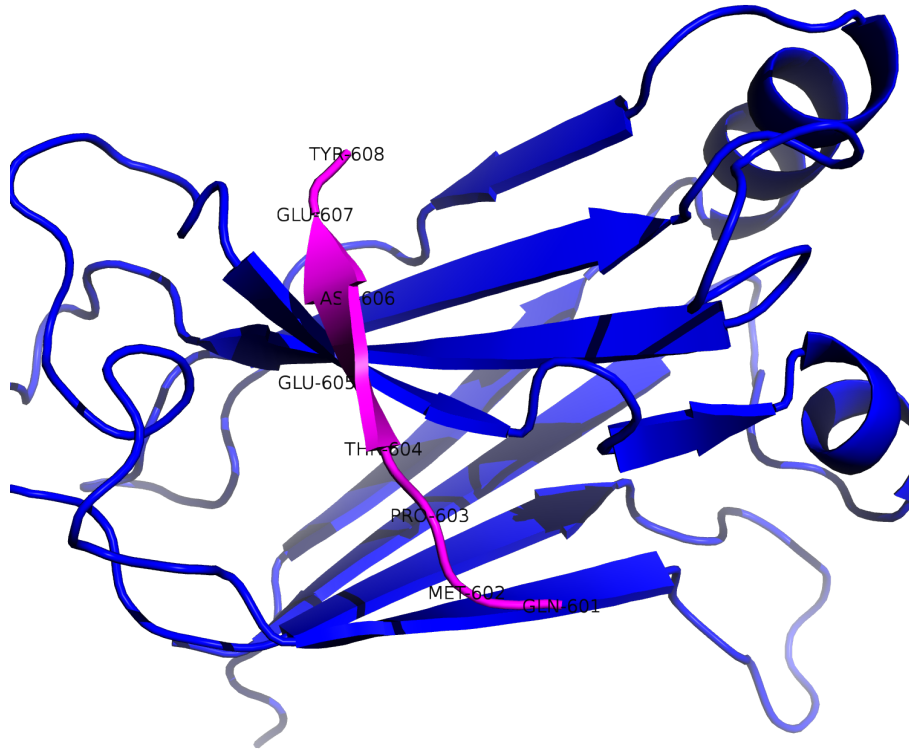


Figure 1.6: Crystal structure of TRAF-C domain (blue) of TRAF6 in complex with a human fragment of RANK (PDB:1LB5) [41].

In the human TRAF6-RANK complex, Hong Ye, et al. described residues Arg392 and Pro468-Tyr473 of TRAF6 as important residues involved in the interaction. Moreover, residues Arg392, Phe471 and Tyr473 are considered hot spots.

1.1.2 Decoy Peptides

Inhibition of TRAF6 or a cell-penetrating TRAF6 decoy peptide decreases NF- κ B signalling pathway activity, reduces breast cancer proliferation, enhances apoptosis of cancer cells, and reduces RANKL-mediated osteoclastogenesis and bone resorption [42][43]. Given the importance of TRAF6, it is rather important to design TRAF6 inhibitors in order to develop therapeutics for controlling inflammation and the discussed diseases.

The use of decoy peptides as PPI inhibitors is supposed to catalyze the development of novel therapeutics for diseases, such as cancer [44]. Based on this, protein-peptide interactions have been exploited in various biotechnological and pharmaceutical applications, by using peptides inhibitors as therapeutic agents [45].

Different cell-penetrating decoy peptides based on RANK-TRAF6 binding sites (Figure 1.4) were tested *in vitro* to potentially inhibit TRAF6 signalling (Table 1.1), and some of them effectively inhibit NF- κ B activation by blocking TRAF6 binding to RANK [19][46][47][48].

Table 1.1: Potential therapeutic cell-penetrating peptides inhibit TRAF6 signaling

TRAF6-binding site	Peptide sequence	Effectiveness	References
mTRANCE-R(1)	AAVALLPAVLLALLAP-RKIPTEDEYTD RPSQPST	High	[19]
mTRANCE-R(2)	AAVALLPAVLLALLAP-IPPFQE PLEV GEND	Low	[19]
mTRANCE-R(1)	AAVALLPAVLLALLAP-RKIPTEDEYTD RPSQPST	High	[46]
mTRANCE-R(1)	J-RKIPTEDEYTD RPSQ	Medium	[47]
mTRANCE-R(1)	J-RKIPTEDEYTD R	Medium	[47]
mTRANCE-R(1)	J-RKIPTEDEY	High	[47]
mTRANCE-R(1)	J-PTEDEY	Null	[47]
mTRANCE-R(1)	J-PTEDEYTD R	Null	[47]
hTRANCE-R(1)	AVALLPAVLLALLAP-RQM PTEDEY	High	[48]

The peptide sequences comprise two parts: the first residues belong to the cell-penetrating sequence and the second part contains the residues of the TRAF6-binding site. Two different cell-penetrating peptides (cpp) were tested: AAVALLPAVLLALLAP, a hydrophobic sequence of the Kaposi fibroblast growth factor signal peptide; and J which represents palmitic acid, linked to the N-terminus as a way to internalize the peptide by making a smaller scaffold.

Ye et al. showed that only mTRANCE-R(1) is an effective inhibitor, not registering NF- κ B activation and decreasing osteoclasts formation [19]. As mentioned in the subsection 1.1.1, this is explained by the degree of complementarity between TRAF6's basic surface with the peptides P_1 and P_2 positions, whose binding to TRAF6 is stronger if occupied by acidic residues. In fact, the residues at the positions P_1 (Valine) and P_2 (Glycine) in mTRANCE-R(2) are not acid, different from the residues at the positions P_1 (Aspartic Acid) and P_2 (Glutamic Acid) of mTRANCE-R(1) [19].

Studies with the peptides containing palmitic acid (Table 1.1) tested RANKL-mediated osteoclast differentiation and allowed to identify the minimal region needed to effectively prevent TRAF6 signaling, corresponding to the sequence RKIPTEDEY. Accordingly, the first three amino acids prior to the residue P_2 are indispensable for a significant osteoclast differentiation inhibition, and longer peptides were able to inhibit osteoclastogenesis but

with less extent [47].

1.2 Protein Interactions

Most cellular proteins form complexes, requiring specific interactions that became the basis for almost all biological process as signal transduction [49]. PPIs can be homomeric or heteromeric (constituted by identical or different molecules, respectively), and can be classified as either obligate - strong and long lived - or non-obligate - weaker and transient [50]. Various factors influence proteins interactions, like shape complementarity, electrostatic complementarity, hydrophobicity and flexibility [50–54].

1.2.1 Interface and Hot Spots

PPI studies commonly requires a deeper understanding of their structure by regions, such as the surface, the interior and the interface, the latter defined as a set of amino acids which represents a region that links two chains by non-covalent interactions [55].

During the protein folding process, the physical process by which a polypeptide folds into its characteristic and functional tertiary structure, hydrophobic residues shield themselves from the solvent, resulting in a hydrophobic protein interior and a hydrophilic surface [56]. The basic description of the protein surface is the atomic layer of exposed residues [57]. This distinction between interior and surface residues has profound implications in stability and evolution, as residues at the surface of proteins are more likely to mutate than those that are buried, once mutations in the protein interior are more destabilizing to its structure and function [53][55][58–63].

This surface-interior contradistinction coming from protein folding can be applied to protein-protein or protein-peptide interfaces, where the BSA (equation 1.1) is defined as the surface buried away from the solvent when a complex is formed [64–67].

$$BSA = \sum_{n=1}^{N_{comp}} ASA_{free}^n - ASA_{complex} \quad (1.1)$$

Equation 1.1: Complexes buried surface area ($\text{\AA}^2$).

Where:

ASA_{free}^n indicates the accessible surface area of the unbound molecules ($\text{\AA}^2$);

$ASA_{complex}$ indicates the accessible surface area of the complex ($\text{\AA}^2$);

N_{comp} indicates the total number of free components in the complex;

n is relative to the free reactant.

The interface residues are the intermolecular residues present in the binding pocket of a complex, the region encompassed by important residues, where the interface surface is called the rim and its interior is called the core [65] [66]. At the interface core, hot spots

are a small subset of the buried residues that mostly contribute to the binding affinity, defined as the strength of these interactions [68]. Not only are hot spots structurally conserved, mutating at a slower rate when compared to other surface residues, core residues in general are more evolutionary conserved than those at the rim [69–76]. To know which residues are structurally and evolutionary conserved, sequence alignment is fundamental to infer homology (common ancestry) between two proteins. If two sequences are in alignment and part or all the polypeptides pattern match, then they are similar and may be homologous. Moreover, if the sequence of a protein significantly matches another sequence with a known structure and function, then the molecules may share the same structure and function as well [77].

PPI interfaces differ from their type of obligation, as non-obligate PPI interfaces are typically smaller than obligate PPI interfaces [78]. An area between 1150–1200Å² is considered small, standard-size interfaces are roughly 1600Å² (+/- 400Å²) and a large interface bury 2000 to 4660Å² [66][79].

A part of structural analysis of a 3D macromolecular complex is the annotation of the intermolecular interactions present in the interface [80]. The cut-offs for inter-residue interactions range from 5 to 14Å, depending on the literature, so non-covalent interactions do not have a consensus on geometrical definition [81–84]. The only consensus is in hydrophobic contacts, defined between atoms within 5Å from each other [85]. Besides hydrophobic contacts, hydrogen bonds (H-bonds) are the most important in defining a stable contact between a protein and its binding partner [86]. Because interfaces consist in a central desolvated region surrounded by energetically less important residues in contact with water, an appropriate approach to obtain information on which residues belong to the interface is mutagenesis, the mutation of a residue followed by the observation of its consequence in protein stability [66]. One example of this approach is Alanine scanning, which is based on the principle that when a residue is mutated by this amino acid, getting rid of side chain atoms beyond C β , a loss of binding affinity indicates that the mutated residue mediates the interaction between molecules, and therefore is present in the interface [87]. Alanine is used because is viewed as neutral, once its relatively inert methyl functional group does not contribute additional flexibility [88–90].

For all proteins, the energy distribution is not uniform across the interface, as a small subset of residues will have a more significant contribution to the binding [68][73][74][91–95]. Hot spots have been defined as sites where Alanine mutations cause a significant decrease in the binding energy ($\Delta\Delta G$). In equation 1.2, the ΔG^{mut} is the binding energy with the alanine substituted and ΔG^{wt} is referent to the wild type residues in the protein complex [91][96].

$$\Delta\Delta G = \Delta G^{mut} - \Delta G^{wt} \quad (1.2)$$

Equation 1.2: Overall change in the binding free energy [91][96].

1.2.2 Binding Affinity

Protein interactions play an important role in living cells, and their stability and affinity affects the cell's function [58][97].

The binding affinity is associated with the dissociation constant (K_d), an experimental measure that determines whether an interaction occurs in solution (low K_d) or not (high K_d). K_d equation (Equation 1.3) consists on the dissociation of a complex (PL) into its unbound molecules, the ligand (L) and the protein (P) [98].

$$\begin{aligned}
 & PL \xrightarrow{K_d} P + L \\
 & K_d = \frac{[P][L]}{[PL]}
 \end{aligned}
 \tag{1.3}$$

Equation 1.3: K_d reaction equation and formula.

Where:

K_d is the dissociation constant (M);

$[P]$ is the concentration of the protein (M);

$[L]$ is the concentration of the ligand (M);

$[PL]$ is the concentration of the complex (M).

Jieming Chen et al. demonstrated that for protein-protein and protein-peptide complexes there is a correlation between binding affinity and the amount of buried surface area at the interface (Figure 1.7), indicating that as the BSA increases, the K_d decreases, which means that the affinity increases [97]. In conclusion, as BSA increases the binding affinity also increases. This correlation will be used as a tool to molecular docking analysis.

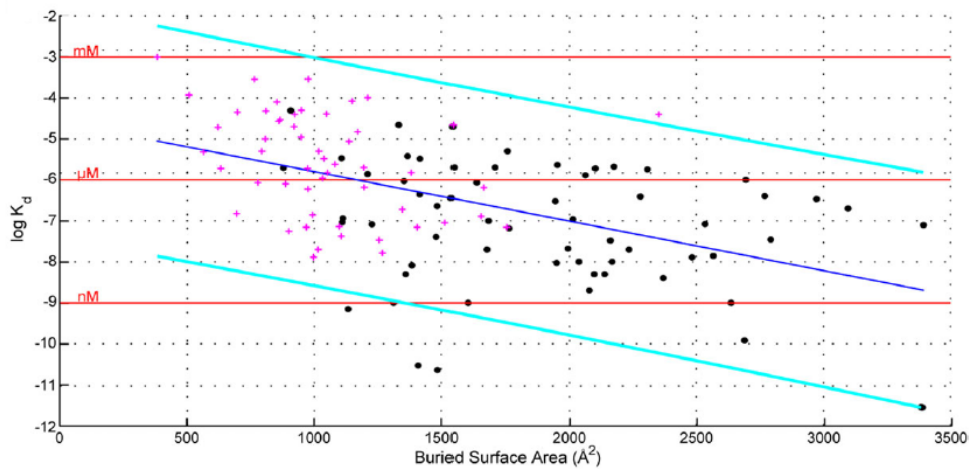


Figure 1.7: Relationship between the $\log(K_d)$ and the BSA of protein-peptide (magenta crosses) and protein-protein (black circles) complexes of 113 heterodimers complexes. The horizontal red lines indicate K_d values of 1 mM, 1 μ M, and 1 nM. [97].

1.3 Computational Chemistry

Computational chemistry combines chemical theory and model experimental data, using mathematical algorithms, statistics and databases. It is an important field to validate results and for the design of new experiments and materials. For example, computational chemistry is used to simulate and identify protein sites that most probably bind to a new molecule as drug, to determinate properties that are not accessible experimentally and to interpret experimental data [99].

The analysis and the prediction of the protein-peptide mechanism, and the consequent design of novel protein-peptide complexes, have been applied in the fields of biology, medicine and pharmaceutical sciences [100].

The modelling of protein-peptide complexes is most often approached in two steps, firstly by identifying the peptide binding site on a protein and lastly determining the native pose of the peptide [100].

1.3.1 Tertiary Structure Prediction

Mapping the molecular details of individual proteins and their interactions with each other and with small molecules is crucial to understand life at the molecular level. Drug design and protein engineering are highly related to protein interactions and structure–function relationships, which can only be understood if proteins 3D structure is determined [101].

There are two main computational alternatives to predict tertiary structures of proteins and peptides from sequence data (Figure 1.8) [77].

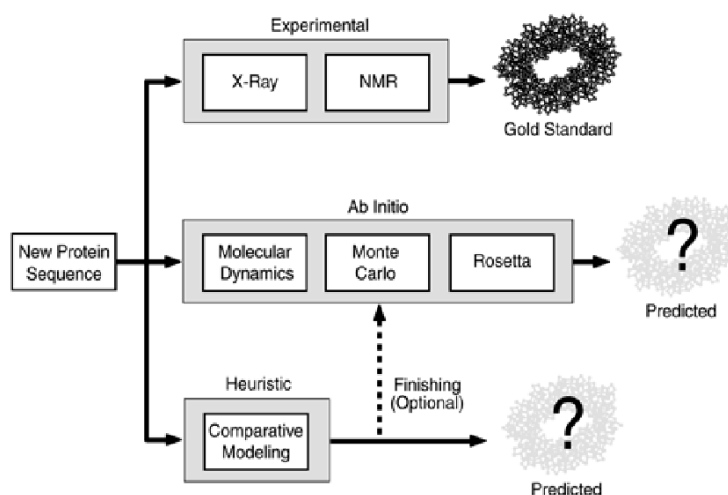


Figure 1.8: *Ab initio* and heuristic methods for protein structure prediction versus experimental protein structure determination methods [77].

Ab Initio

This approach involves methods that rely on molecular physics without knowledge of any biologic relationship between the target structure and other proteins/peptides [97].

The shape of a protein/peptide can be defined as a function of the amino acid sequence (primary structure), the temperature, the pressure, the pH, among others. Having only the primary structure, the secondary structure is designed having into account bond lengths, bond angles and dihedral angles [77].

Dihedral angles, refer to the torsion of two planes defined by the two atoms that form the bond, being in this way four atoms involved in total (2 atoms in each plane). There are three dihedral angles for each residue that define the backbone of a polypeptide (Figure 1.9) [77].

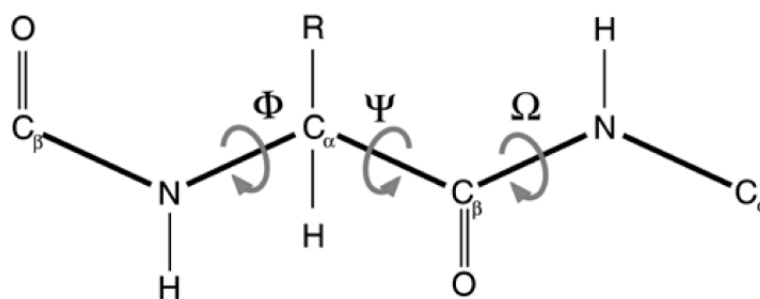


Figure 1.9: Important dihedral angles [77].

The omega (Ω) $C\beta$ and N angle assumed constant at 180° ; the psi (ψ) $C\alpha$ and $C\beta$ angle, restricted due to the interaction of side chain residues; and the phi (ϕ) N and $C\alpha$ angle, restricted for the same reason as the psi angle.

Furthermore, to generate the tertiary structure techniques such as molecular dynamics, Monte Carlo and Rosetta are needed. The first uses Newtonian physics to calculate atom force and moves that atom in a short fraction of time, so the final structure has the most stable molecular conformation. The Monte Carlo method identifies structural combinations with the lowest free energy, by repeating structure construction simulations, clustering the resulting structures and choosing the one that is in the centre of the largest clusters. Finally, Rosetta is a hybrid method, where the target is broken into chains of three and nine residues, and the conformations of these segments are limited and created with fragment libraries extracted from protein structure databases. The combination of local conformations with the lowest overall energy is taken as a candidate structure [77].

Ab initio modelling will be used to predict the tertiary structure of the five inhibitor peptides.

Heuristic methods

Heuristic methods, which are more accurate, are based on pattern matching, using

the knowledge of existing protein structures combined in databases as templates, and assuming that changes in amino acid sequences may not alter protein structure as similarity in sequences implies similarity in structure [77].

Heuristic methods are based on comparative modelling, which finds similarities in amino acid sequence not having into account the protein's lineage. Homology modelling is another heuristic method similar to comparative modelling but takes into account the protein's lineage and assumes that proteins from the same families share folding motifs, even if they have different sequences [77]. Comparative modelling will be the one considered for this project, once it is the main method used.

Figure 1.10 shows the main phases of comparative modelling, given protein sequence data.

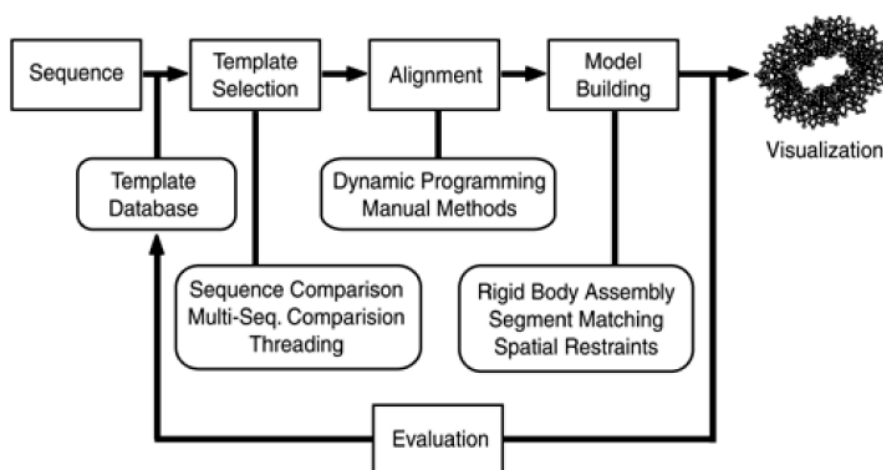


Figure 1.10: Comparative Modelling Process [77].

The search for a template in a database that could be a match to the target structure is based on the amino acid sequence. The template could have similar amino acid sections with the target, and so the structure of those sections is considered the same for the target [77].

The comparison threading involves the sequence alignment of the target with the 3D structure of the template and compare the spatial and chemical similarity. Only the templates that fit these parameters are potential templates. One way to do threading is based on contact potential, where the position of amino acids and the resulting association energy are evaluated, being only considered energetically stable conformations for target structure representation [77].

Models are built having into the account all the chosen templates above. Probably the template and the target have different backbone sequences, thus, a minor change in a bond angle could result in a major conformational change. Because of this, it is better to brake the template in several parts and then fit those parts individually in the target. This technique is the rigid body assembly, which is improved if the intended assembly is

formed by structural areas coming from different templates, once in this way each area of the target assembly has the most similar conformation possible to the sequence [77].

Finally, the models created are evaluated given their final accuracy. The greater the similarity between the model and the template, the more likely the model is an accurate structure prediction.

Comparative modelling will be used to predict the tertiary structure of the mice TRAF6's TRAF-C. The ModBase¹ database of comparative protein structure models will be used to this end, where comparative modeling is only applicable when the target sequence is detectably related to a known protein structure, and it is carried out in four sequential steps [102]:

- Identifying known structures (templates) related to the sequence to be modelled (target);
- Aligning the target sequence with the templates;
- Building model;
- Assessing the models.

1.3.2 Molecular Docking

Molecular docking is a computational method that aims to study complexes formed between two or more molecules, starting from several possible conformations, which can be either experimentally determined or predicted [100]. The docking method is based on the complementarity between interacting molecules, which could be geometric, electrostatic or hydrophobic, or all three [57]. It can be divided in two main steps [1][103][104]:

1. Sampling: Exploration of binding poses;
2. Scoring: Identification of the best biological models.

This methodology allows a pre-selection of the most potential compounds to be synthesized, which involve techniques known to be time-consuming and resource-intensive [105].

The experimental determination of high-resolution biomolecular complexes atomic structures is usually time-consuming and expensive. Moreover, flexibility of the peptides and their transient interactions, along with the fact that they do not adopt a well-defined conformation when unbound, make the recognition of these complex mechanisms difficult to study experimentally [106]. Therefore, integrated with x-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, molecular docking has become an

¹<http://salilab.org/modbase>

important tool in designing novel potential drugs, allowing the modulation of three dimensional (3D) complexes constituted by a target (protein or nucleic acids) and a ligand (peptide, protein or small organic molecule), and study their interactions [100][103][107].

In molecular docking, there are two main strategies that can be followed: *ab initio* and template-based modelling (TBM) [108]. Although most docking methods are *ab initio*, which only considers the coordinates of the starting structures disregarding any experimental knowledge about the system, TBM, whenever 3D structures of complexes with the homologs of the structures to be docked are available, is the best alternative [109][110]. Moreover, the conservation of interfaces across homologous complexes can improve docking [111].

Furthermore, it is also possible to do molecular docking with residue restraints, evidencing their presence in the interface or the distance between them [112].

1.3.2.1 High Ambiguity Driven DOCKing

HADDOCK (High Ambiguity Driven DOCKing) is a web server that incorporates experimental data as restraints and uses them to guide the docking process alongside traditional energetics and shape complementarity [113].

One of the most important parts in HADDOCK is the ambiguous interaction restraints (AIR), which are defined by active and passive residues [1]:

1. Active residues: are those of main importance for the interaction, such as residues whose absence enables the interaction between molecules or those where the chemical shift perturbation is higher. Throughout the molecular docking process, active residues are restrained to be part of the interface;
2. Passive residues: are those that contribute for the interaction but are less important and only might be at the interface.

AIRs are created between each active residue of a partner and the combination of active and passive residues of the other partner. The scoring penalty is a critical tool to define conveniently active and passive residues, directly influencing the molecular docking success if active residues are defined and that do not belong to the interface. Definition of a very strict set of AIRs leads to a very narrow sampling of the conformational space, meaning that the generated poses will be very similar. Conversely, very sparse restraints (e.g. the entire surface of a partner) will result in very different solutions, displaying greater variability in the region of binding [1]. So, if the restraints are wrong, the true interface might never be sampled.

Besides AIRs-guided docking, the interfacial distance restraint alpha carbon–alpha carbon guided docking have been used for the prediction of some targets, imposing stricter restraints on the orientation (i.e docking poses) of interaction [104][114][115]. For the purpose of this work only AIR restraints will be applied, once the available information is the interacting active residues.

To allow the molecules to experience varying degrees of flexibility and chemical environments, HADDOCK protocol consists of three stages [1][106]:

1. "it0": Rigid-body docking by energy minimization from random orientations of the starting conformations;

$$HADDOCKscore - it0 = 0.01Evdw + 1.0Eelec + 1.0Edesol + 0.01Eair - 0.01BSA \quad (1.4)$$

Equation 1.4: HADDOCK scoring function in the it0 stage.

2. "it1": Semi-flexible refinement of the interface, being the protein interface and the ligand treated as flexible;

$$HADDOCKscore - it1 = 1.0Evdw + 1.0Eelec + 1.0Edesol + 0.1Eair - 0.01BSA \quad (1.5)$$

Equation 1.5: HADDOCK scoring function in the it1 stage.

3. "water": Molecular dynamics refinement in solvent to further improve scoring, solvating the complex in a shell of water molecules.

$$HADDOCKscore - water = 1.0Evdw + 0.2Eelec + 1.0Edesol + 0.1Eair \quad (1.6)$$

Equation 1.6: HADDOCK scoring function in the water (final) stage.

Where:

Evdw is the non-bonded intermolecular van der Waals energy (adimensional);

Eelec is the non-bonded intermolecular electrostatic energy (adimensional);

Edesol is the empirical desolvation energy term (adimensional);

Eair is the AIR energy (adimensional);

BSA is the buried surface area ($\text{\AA}^2$).

HADDOCK ranks the output clusters with the scoring function 1.6, based on the average score of the top 4 members of each cluster [116]. The cluster numbering reflects the size of the cluster, with cluster 1 being the most populated cluster [1][113].

In protein-peptide interfaces, the recognition of long-range electrostatic interactions is very important [117]. Therefore, it should be specified charged N and C terminals when working with naturally occurring peptides, or uncharged terminal(s) when the peptide is a protein fragment, avoiding undesired interaction with the terminal(s) [106].

The final models are automatically clustered based on two different clustering algorithms. One is the positional interface ligand RMSD (iL-RMSD), a specific similarity

measure that captures conformational changes about the interface by fitting on the interface of the receptor and calculating the RMSDs on the interface of the smaller partner. The other is the Fraction of Common Contacts (FCC), grouping models based on the similarity of their contact networks at the interface [1]. By default, a cluster is defined if a minimum of 4 models fall into the same group.

There is also statistics analysis of energetic terms and other structural measures for each cluster as the z-score, a quality measure parameter given by HADDOCK that indicates how many standard deviations from the average HADDOCK score a cluster is located, so the more negative the z-score, the better it is [106][113].

HADDOCK has also been used as a tool to predict binding affinity. This approach consists on a multiple linear regression analysis used to create an optimized HADDOCK score (1.7) [118].

$$\begin{aligned} \log(K_{pred}) &= \text{HADDOCKscore} \\ &= \beta_1 E_{vdw} + \beta_2 E_{elec} + \beta_3 E_{desol} + \beta_4 E_{air} + \beta_5 BSA + f \end{aligned} \quad (1.7)$$

Equation 1.7: Optimized HADDOCKscore.

Where:

$\log(K_{pred})$ is the logarithm base 10 of the predicted dissociation constant of the peptides.

E_{vdw} is the non-bonded intermolecular van der Waals energy (adimensional);

E_{elec} is the non-bonded intermolecular electrostatic energy (adimensional);

E_{desol} is the empirical desolvation energy term (adimensional);

E_{air} is the AIR energy (adimensional);

BSA is the buried surface area ($\text{\AA}^2$);

β_i and f are constants.

1.3.2.2 Docking assessment

The difference among clusters considering the HADDOCK score and z-score allows a quick assessment of the quality of the generated models and the overall molecular docking simulation [1][113].

The docked models that have a reference structure are qualitative evaluated according to the CAPRI (Critical Assessment of Predicted Interactions) parameters. CAPRI is a communitywide experiment on the comparative evaluation of protein-ligand docking for structure prediction, where typically given the structure or the sequence of the unbound partners the aim is to predict the correct 3D structure of an unknown biomolecular protein-ligand complex in a limited time [57][116][119]. For peptide-protein complexes the parameters are [120]:

1. Ligand root-mean-square-deviation (L-RMSD): RMSD of the ligand backbone atoms after fitting on receptor backbone;

2. Interface root-mean-square-deviation (i-RMSD): calculated on backbone atoms of both protein and peptide residues which are within 8Å from each other in the reference crystal structure of the complex;

In bioinformatics, the root-mean-square deviation (RMSD) of atomic positions is the measure of the average distance between the atoms (usually the backbone atoms) of superimposed proteins. CAPRI criteria for ranking models of protein-peptide complexes is present in the Table 1.2 [120]. This previous analysis is very important since it is possible to produce false positives, i.e., structures with good scores but high RMSDs [121].

Table 1.2: CAPRI criteria for ranking models of protein-peptide complexes

	Incorrect	Acceptable	Medium	High
L-RMSD (Å)	>5.0	5.0 - 2.0	2.0 - 1.0	1.0 ≥
i-RMSD (Å)	>2.0	2.0 - 1.0	1.0 - 0.5	0.5 ≥

Although the RMSD value provides a quantitative measure of fit, visualizing the overlap of structure pairs is more intuitive and provides information on the difference in what the RMSD value represents in each case [77].

It is also possible to evaluate and rank clusters considering its intermolecular energies, considering equation 1.8 [115].

$$E_{int} = E_{elec} + E_{vdw} + E_{air} \quad (1.8)$$

Equation 1.8: Intermolecular energy equation for HADDOCK output models.

Where:

E_{int} is the intermolecular energie (adimensional);

E_{elec} is the non-bonded intermolecular electrostatic energy (adimensional);

E_{vdw} is the non-bonded intermolecular van der Waals energy (adimensional);

E_{air} is the AIR energy (adimensional).

1.4 Solid Phase Peptide Synthesis

Peptides are short amino acid sequences up to 30 or 50 amino acids, depending on the literature source. They result from sequential reactions of the α -carboxyl group of one amino acid with the α -amino group of another amino acid in a successive manner, forming peptidic bonds (Figure 1.11). What distinguishes a peptide from a protein is not only the number of amino acids but also its physicochemical properties [122][123].

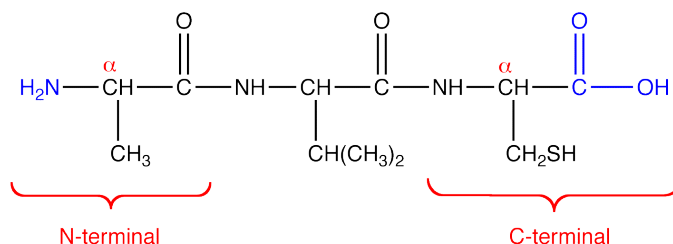


Figure 1.11: Example of an amino acid chain with blue N and C terminal amino acid residues, both on α carbon [124].

Solid phase peptide synthesis (SPPS) is a method based on the sequential addition of protected amino acid residues to an insoluble polymeric matrix (resin), through the carboxyl group of the first amino acid anchored, via a *linker*, to the resin in a covalent manner. Hence, the peptide grows in the C to N direction [125]. Two different strategies for SPPS can be followed: the Boc-SPPS which resorts to the use of a t-butyloxycarbonyl (Boc) to protect the α -amino group of the amino acids, and the Fmoc-SPPS in which the α -amino is protected with a fluorenylmethoxycarbonyl group (Fmoc). The Boc group is removed by an acid, normally hydrofluoric acid (HF), whereas the Fmoc group is removed with a base [126]. The Boc-SPPS methodology was the first to be used, but the Fmoc-SPPS is the most used nowadays, as it does not require the use of HF, which makes it safer and easier to automatize [125]. In this thesis we have used Fmoc-SPPS and piperidine for Fmoc deprotection [127]. The overall process is shown in Figure 1.12.

The side chains of some amino acids are protected as well but in a different manner from the α -amino groups, while the protection of the α -amino group is temporary, so the next amino acid could be linked to one already in the resin, the protection of the side chains is "permanent", being only removed when the peptide is cleaved from the solid support (Figure 1.12).

The synthesis consists in a set of cyclic steps [126]:

- Cleavage of the α -amino protecting groups;
- Washing to remove the cleavage reagent;
- Coupling of the protected amino acid;
- Washing to remove non-reacting material.

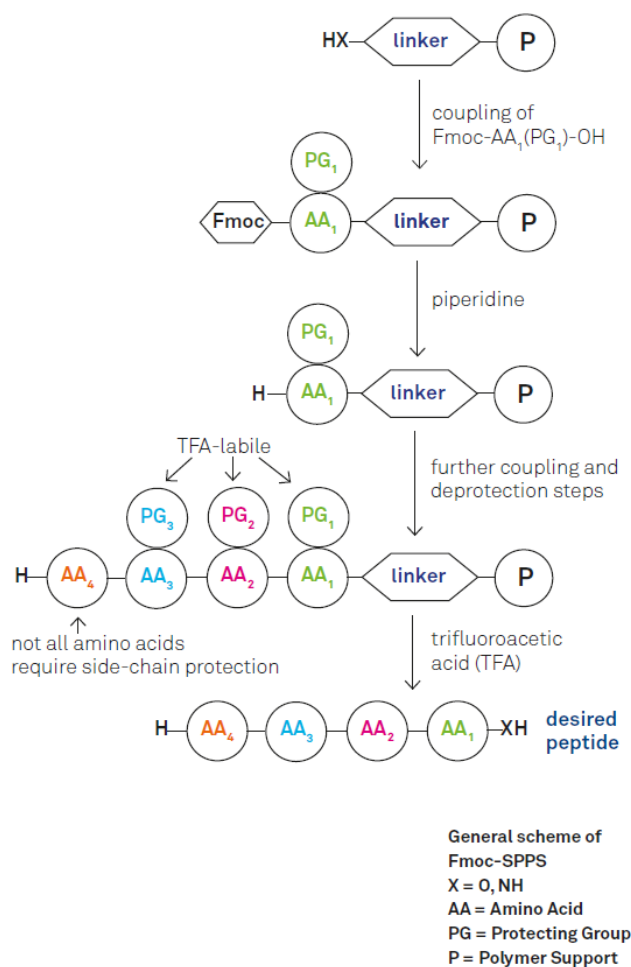


Figure 1.12: SPPS [127].

Dichloromethane (DCM) and N,N-dimethylformamide (DMF) are the primary solvents used for resin deprotection, coupling and washing [128].

The final step is the cleavage of the peptide from the resin, usually with trifluoroacetic acid (TFA). The resulting peptide could have a carboxamide group or carboxylic acid at its C-terminus, or other functional group like a alcohol group, depending on the *linker* or the type of the resin used. The final product contains the desired peptide and impurities that must be removed, such as deletion peptides, incompletely deprotected peptides, modified peptides, scavengers and by-products [128].

Normally, the resin and the side chains protecting groups are chosen as a combo so they could be removed simultaneously in the peptide cleavage [125]. For peptide synthesis are recommended resins with small particles and low *cross-linking*, so the reagents diffuse rapidly between the polymeric resin, which swelling enables the accommodation of the growing peptide chain. Most resins are composed by polystyrene (PS), polyacrylamide (PA), polyethylene glycol (PEG), or a combination between these [129].

PS crosslinked with 1% divinylbenzene (DVB) (Figure 1.13) is the most used resin in SPPS, because it is chemical inert, allows the introduction of a large variety of anchoring groups and swells easily with the right solvents [127]. Having this into account, and because polystyrene resins are normally employed for batch-wise synthesis (the one performed in this project), this type of resin was used for SPPS [128].

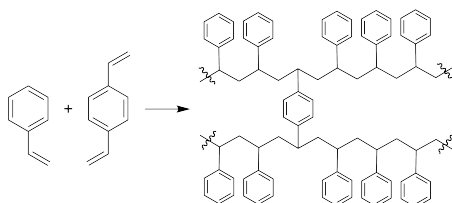


Figure 1.13: [130].

To control the manual SPPS, we have performed a qualitative color test to monitor the completeness of amino acid coupling and the presence of free amine after deprotection [131][132]. This test (Kaiser test) consists in the reaction of ninhydrin with free primary amines, generating an intensive blue color, which indicates that the amine group in free and another amino acid can be coupled. By the contrary, when an amino acid is coupled and the amine group is not free, the color of the solution stays yellow, indicating that the amino acid coupling is complete [133]. The test does not yield the typical dark blue color with serine, asparagine, aspartic acid, and proline (that as a secondary amine) [132]. Only the latter two exist in the peptides that are going to be synthesized.

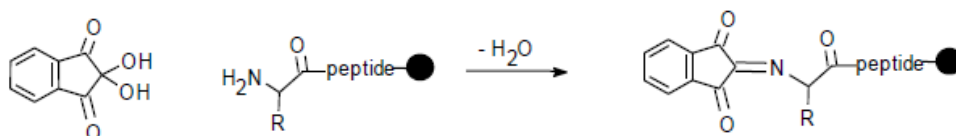


Figure 1.14: Nihidrine (first molecule) reaction with the peptide's free amine [132].

1.4.1 Circular Dichroism

Circular Dichroism (CD) is a spectrophotometric method used to analyze the secondary structure of peptides and proteins. The main advantage is the fact that chiral molecules, that occur in two nonsuperimposable mirror images forms, or molecules in chiral media interact differently with polarized light circularly to the left and right. The CD spectrum is sensible to the chiral molecules 3D characteristics, and so, it is possible to analyze the peptide structures [134]. The poli-L-Lysine CD spectra for the three principal secondary structures (Figure 1.15) have been used as a reference [135]. These spectra were used in this work for the approximate calculation of the fraction of each secondary structure in each peptide solution.

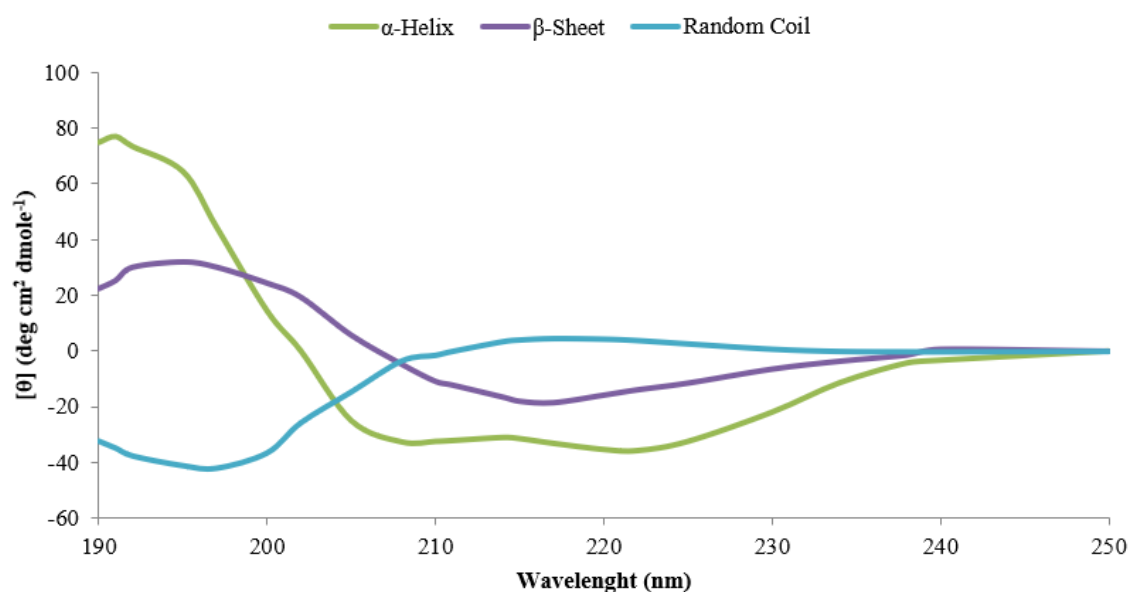


Figure 1.15: Poly-L-lysine circular dichroism spectra for the three main conformations [135].

1.5 Objectives

The work presented herein aims at the design of innovative peptides able to interfere with the RANK-TRAF6 interaction, which is a relevant PPI, with a key role in the NF- κ B regulation.

As already mentioned in the subsection 1.1.2, only the cell-penetrating peptide with mTRANCE-R(1) inhibits effectively the TRAF6-RANK pathway. A computational framework that combines *ab initio* modelling, comparative modelling and molecular docking will be performed to clarify peptide-TRAF6 interactions by structural and energetic analysis. The goal is to validate by *in silico* studies the stability of TRAF6 with a peptide containing mTRANCE-R(1) and a peptide containing mTRANCE-R(2). Throughout this document, this peptide will be referred as mPEP1 and mPEP2, respectively, and the complexes as mTRAF6-mPEP1 and mTRAF6-mPEP2, (m) representing "mouse". The stability of TRAF6 with a peptide containing hTRANCE-R(1) and a peptide containing hTRANCE-R(2) will also be evaluated. Throughout this document, these peptides will be referred as hPEP1 and hPEP2, respectively, and the complexes as hTRAF6-hPEP1 and hTRAF6-hPEP2, (h) representing "human".

Firstly we will model the peptide sequences and determine the active residues present in the TRAF6-RANK complex, necessary for molecular docking, which will be made using HADDOCK. The stability of TRAF6 with the different peptides will be accessed with two methods, one using the BSA values of the complexes and the other using the HADDOCK optimized score as a tool to predict binding affinity.

A cell-penetrating peptide derived from dengue virus capsid protein that was reported to have high cellular internalization, heavily associating to cellular membranes

and accumulating inside cells. This peptide, designated PepH2, will also be linked to hTRANCE-R(1), giving the fusion peptide PepH2-hPEP1. A comparative study will be performed between this peptide and hPEP1, as it would be interesting to compare by *in silico* studies the stability of TRAF6 with a peptide containing PepH2 towards hTRAF6 with the peptide containing the Kaposi fibroblast growth factor signal peptide towards hTRAF6.

Among the human peptide sequences proposed (hPEP1, hPEP2 and PepH2-hPEP1), the one that gives the most stable complex with TRAF6 will be synthesized and evaluated biologically.

METHODS AND MATERIALS

Among the human peptide sequences proposed (hPEP1, hPEP2 and PepH2-hPEP1), the one that gives the most stable complex with TRAF6 *in silico* will be synthesized, along with hTRANSE-R(1) (RANK's binding site I). The Table 2.1 lists the five peptide sequences that are going to be studied *in silico* in complex with TRAF6, and also hTRANSE-R(1).

Table 2.1: Peptide sequences to be studied *in silico* in complex with TRAF6, for human and mouse complexes, and hTRANSE-R(1) (RANK's binding site I) to be synthesized

Peptide nomenclature	Peptide sequence	Complex
hPEP1	AAVALLPAVLLALLAP-RQMPTED EY	hTRAF6-hPEP1
hPEP2	AAVALLPAVLLALLAP-FSEPLEVGE	hTRAF6-hPEP2
mPEP1	AAVALLPAVLLALLAP-RKIPTED EY	mTRAF6-mPEP1
mPEP2	AAVALLPAVLLALLAP-FQEPLEVGE	mTRAF6-mPEP2
PepH2-hPEP1	KLFMALVAFLRFLT-RQMPTED EY	hTRAF6-PepH2-hPEP1
hTRANSE-R(1)	RQMPTED EY	-

2.1 TRAF6 Sequence Conservation

The human TRAF-C fragment structure of TRAF6 (hTRAF6) is available at PDBID 1LB5 [41]. As experimental studies were developed using mice cells, and with the scope of an *in silico* validation, it is necessary to predict mice TRAF-C's TRAF6 (mTRAF6) structure.

Aiming to evaluate sequence conservation (homology) between both sequences (mTRAF6 and hTRAF6), an alignment of both sequences was performed. This method involves the matching of two sequences and finding the best alignment of two sequences with some degree of similarity. The chosen alignment is typically based on a score that reflects the number of paired characters in the two sequences and the number and length of gaps required to adjust the sequences [77]. For this purpose, the web-based heuristic tool BLAST¹ (Basic Local Alignment and Search Tool) was used.

2.2 Interface & Hot Spots Prediction

In order to identify the hot spots and important residues, besides the ones already mentioned in the subsection 1.1.1, and to validate that information as well, computational alanine scanning was performed with the RANK-TRAF6 3D crystal structure retrieved from the PDB (PDB Code: 1LB5) [19]. Several web servers were used to get a consensus result in hot spots prediction and important interfacial residues identification:

- **Robetta Beta**²

Robetta is an energy-based computational method that considers Van der Waal potential, Coloumb electrostatics, solvation energy and hydrogen bonds [136].

As an output, positive values for the $\Delta\Delta G_{complex}$ indicate that alanine mutation is predicted to destabilize the complex, which means that the residues are important for the binding, and $\Delta\Delta G_{complex}$ negative values predict a stabilizing effect. Hot spot residues have a predicted $\Delta\Delta G_{complex} > 1$ kcal/mol [96].

This analysis allows knowing besides which residues are hot spots, but also other residues that may be involved in the binding site. Robetta cannot identify hot spots that are involved in water-mediated hydrogen bonds, and when waters can compensate for a mutation, it will incorrectly predict mutations as hot spots, so the input PDB cannot have water molecules [96].

- **KFC2**³

KFC2 is a machine learning-based method that incorporates two support vector machine models (SVM): KFC2a and KFC2b. KFC2a is composed by eight features primarily related to solvent accessibility, interface position, packing density, and local plasticity, and the specific features of KFC2b are residue size and hydrophobicity. KFC2b has a higher specificity as it predicts less number of hot spot residues but with higher accuracy, and KFC2a predicts a larger number of hot spot residues but with lower accuracy [137].

Thus, KFC2a has higher reported sensitivity and accuracy when compared to KFC2b, but KFC2b has higher specificity [96].

¹<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

²<http://robetta.bakerlab.org/>

³https://mitchell-lab.biochem.wisc.edu/KFC_Server/index.php

These structural features return a binary answer as to whether a residue is a hot spot or an interface residue, and a confidence score with each prediction is given. KFC2 labels a residue a hot spot if the prediction confidence is greater than zero. When KFC is combined with Robetta's computational alanine scanning technique described previously, the predictive accuracy increases significantly [96].

- **Hotpoint**⁴

Hotpoint is an empirical model that characterizes the interface and hot spot residues on the basis of accessible surface area (ASA) and knowledge-based pairwise residue contact potential of the interface, labelling a hot spot if the relative accessible surface area is up to 20% and the total pair potential is greater or equal to 18.0 [96][138].

2.3 Tertiary Structure Prediction

The peptides structures used in this study are unknown. Therefore their tertiary structures were predicted by *ab initio* methods, relying on three web-servers picked based on their accuracy, diversity and capacity to accept small amino acid sequences (23 and 25 residues) as an input.

The selected web-servers are briefly explained below:

- **I-TASSER**⁵

Iterative Threading ASSEmbly Refinement (I-TASSER) is a hybrid approach for protein structure prediction based on the comparison modelling and *ab initio* Monte Carlo folding [139][140]. Firstly, I-TASSER identifies structural templates from the PDB based on sequence alignments. The threading-aligned regions are built by template fragment assembly simulations, and the threading-unaligned regions are constructed by an optimized knowledge-based force field. Finally, the structures are clustered and the lowest five free-energy conformations are identified. These models are refined to remove steric clashes [139].

These five models were used to make the docking ensemble.

- **PEP-FOLD3**⁶

PEP-FOLD3 is also a hybrid approach, with a rigid body assembly component and an *ab initio* Monte Carlo final step. This 3D structure peptide prediction web-server divides the structure into a series of fragments of four amino acids in aqueous solution, and describes them as a Structural Alphabet (SA) composed by patterns of a number of

⁴<http://prism.cccb.ku.edu.tr/hotpoint>

⁵<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>

⁶<http://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::PEP-FOLD3>

consecutive residues, where the 3D protein structures can be related to a library of 3D fragments [141][142].

This method is divided in [143]:

1. Prediction of SA letters from the amino acid sequence;
2. Alignment and a probability profile is returned for each SA letter.
3. 3D assembly of the prototype fragments associated with the letters selected. This step is repeated 100 times starting from various positions in the sequence;
4. Monte-Carlo method for the final refinement of the 100 simulations and for clustering, using the RMSD.

The obtained five models were used to make the docking ensemble.

- **PEPstrMOD**⁷

PEPstrMOD is an *ab initio* method developed specifically for predicting the tertiary structure of peptides containing natural, non-natural and modified residues, given the sequence as an input. The software integrates structure prediction with force field libraries and simulations [144].

For structure prediction PEPstrMOD uses ideal dihedral angles as restraints for secondary structure prediction [144][145]:

- Helix: $\phi=-60^\circ$ and $\psi=-40^\circ$
- Strand: $\phi=-60^\circ$ and $\psi=-40^\circ$

The initial structure is energetically minimized followed by molecular dynamics simulation to generate a final tertiary peptide structure [144]. Since molecular dynamics simulation was used, the software also gives a five cluster output with different peptides conformations, each with five models. Without any criteria, cluster 1 was used to make the docking ensemble, so five models were used by each web-server.

The peptide environment was set as hydrophilic.

2.3.1 Structure Evaluation

The assessment of the accuracy and reliability of the predicted peptide structures was made by ProSA-web⁸ (Protein Structure Analysis-web), a validation tool used to check errors in 3D protein structures [146].

This platform gives a final z-score score to an input structure and displays a plot with the scores of all experimentally determined structures that are in the database PDB, relating the two. Z-scores outside a characteristic range for native proteins of similar size indicate incorrect structures [147].

⁷http://osddlinux.osdd.net/raghava/pepstrmod/comb_ds.php

⁸<https://prosa.services.came.sbg.ac.at/prosa.php>

2.4 Molecular Docking

A web browser, an internet connection and registration to the HADDOCK web server are the only pre-requisites to access HADDOCK. The HADDOCK Guru interface along with a docking protocol were used [1].

It was made an ensemble docking, that refers to the use of multiple starting conformations for one or more of the binding partners within the same docking run. In this case the first complex to be docked was hTRAF6-hPEP1. For that, the ensemble contains 15 starting conformations of the peptide hPEP1, 5 from each structure predictor web server (I-TASSER, PEP-FOLD and PEPstrMOD), and one conformation of the protein (hTRAF6), which was selected from the Protein Data Bank (PDB:1LB5) [148]. Five different models from each web server were evaluated in complex with hTRAF6, allowing HADDOCK to adopt the most stable and probable bound peptide conformation. The web-server that predicted the peptide tertiary structure present in the chosen hTRAF6-hPEP1 model was the web-server used to predicted also the tertiary structures of the peptides hPEP2, mPEP1, mPEP2 and PepH2-hTRANCE1. For the molecular docking of the complexes with these peptides it was performed an ensemble docking with 5 conformations of each peptide, given by the selected web-server.

In order to avoid HADDOCK docking errors, the input PDBs need to comply with three requirements [1][113]:

1. Remove any irrelevant water and other small molecules from the PDB files;
2. The PDB files must have an END statement as a last line;
3. When submitting an ensemble, all members must contain exactly the same atoms.

2.4.1 HADDOCK Guru Protocol

In the Guru interface, the following was taken into account [1][106]:

1. Is important for the final clustering that the PDB file of the largest molecule is entered first, because in HADDOCK the structures are first fitted on the interface residues of the first molecule and then the RMSD is calculated on the interface residues of the second molecule;
2. Passive residues were selected automatically by HADDOCK;
3. Peptide residues were considered fully-flexible, contrarily to protein residues;
4. The C-terminal of the peptides was considered as charged. The N-terminal was not since the first part of the peptides contains the cell-penetrating sequence that is not expected to bind to the protein;

5. For hTRAF6 and mTRAF6, both N and C terminals were not considered as charged because the PDB used for docking only contains only the TRAF-C fragment of TRAF6;
6. Once the goal for the complex hTRAF6-hPEP1 is to be similar to the complex in 1LB5 but with a cell penetrating sequence linked to the peptide, and the list of active residues is the only information available, for input data the lists of protein and peptide active residues were used (AIR-guided docking). Once, as mentioned in section 3.4, conservation of interfaces across homologous complexes can improve docking, the same approach will be used for mice complexes;
7. The *Number of structures for rigid body docking* (it0) was increased from 1000 to 10000 (maximum). In this way, each conformation is sampled at least 666 times in the rigid body stage;
8. The *Number of structures for semi-flexible refinement* was increased to 400;
9. The *Number of structures for the explicit solvent refinement* was increased to 400;
10. In order to ensure conformational homogeneity, for peptide and small molecules iL-RMSD clustering is recommend with a cut-off of 5Å instead of the usual 7.5Å applied for proteins;
11. It was used the option to automatically define dihedral angle restraints based on the structure;
12. The number of steps for all four stages of the semi-flexible refinement were doubled;
13. The number of models to analyse were even to the number of models generated at the water refinement stage.

2.4.2 Complex Structure Evaluation

The first screening of results was performed according to HADDOCK score and z-score.

The complexes resulting from hTRAF6-hPEP1 were compared to the native structure 1LB5 by l-RMSD and i-RMSD. Since there is no other complex structure for comparison, the other models hTRAF6-hPEP2, mTRAF6-mhPEP1, mTRAF6-mPEP2 and hTRAF6-PepH2-hTRANCE1 were evaluated having into account the results from the hTRAF6-hPEP1 models. The goal was to find a correlation between the HADDOCK score and the RMSDs, and between the intermolecular energy and the RMSDs, for the complex hTRAF6-hPEP1, to strict the choice of the best models within the best clusters for the other complexes that do not have a reference structure.

The calculation of l-RMSD between a model and a reference was performed using *Pymol* by fitting the protein of the model onto the protein of the reference and calculate

the peptide RMSD of the backbone atoms between the docked models and the reference structure [113]. The calculation of i-RMSD between a model and a reference was also done in *Pymol*, by fitting the protein of the model onto the protein of the reference and calculating the positional RMSD between the docked models and the reference structure for the backbone atoms of the interface residues, aligning the interface residues [106].

The overlapping of structure pairs was also taken into account, providing information about peptide orientation. Active residue conservation was also accessed.

2.5 Binding Affinity

One of the methods to evaluate the stability of TRAF6 with the different peptides uses the complexes BSA values, predicting the binding affinity using the method described in subsection 1.2.2. The buried surface area (BSA) values of each complex were in the HADDOCK output files.

The other method used for the same purpose consists in the HADDOCK optimized score as a tool to predict binding affinity. Because the method is related to K_d , to obtain structural and thermodynamic protein-peptide data, a series of filters were applied in PDB advanced search, starting from the dataset of structures and based on Jieming Chen et al. (2013). It was considered:

1. X-ray crystal structures, since the crystal structure from PDB:1LB5 was determined by X-ray crystallography;
2. Structure with a resolution better than 3Å;
3. Structures annotated with the keywords "complex" and "peptide";
4. Peptides that have a length up to 30 amino acids;
5. Structures that contained two different protein entities;
6. The removal of complexes that contain DNA or RNA in detriment of protein complexes;
7. Inclusion of only protein complexes in which each protein has less than 100% sequence identity to another protein.

After these filters, 65 complex structures remained in the dataset. Then, a search in the PDBbind⁹ database for thermodynamic data on these complexes [149] is made. This database contains dissociation constants for 21 of the complex structures.

⁹<http://www.pdbbind-cn.org/>

Finally, it will be made a list of structures and dissociation constants by scanning the literature to ensure that:

1. The K_d listed in PDBbind matches the PPI in the PDB structure;
2. Small molecules or ions are not involved in the interaction;
3. The biological assembly of the complex has not been demonstrated to be an oligomer with order >2 .

The β_i and c coefficients were optimized using the tool *Solver*, by minimizing the X^2 function 2.1, where K_{exp} corresponds to the experimentally measured K_d and K_{pred} corresponds to the predicted K_d .

$$\chi^2 = \sum_{i=1}^{N_{complex}} [\log K_{pred}(i) - \log K_{exp}(i)]^2 \quad (2.1)$$

Equation 2.1: X^2 function to be minimized.

2.6 Peptide Synthesis

2.6.1 Materials

Rink Amide MBHA resin LL (302,3 mg) from *Novabiochem*®, with a low substitution (molar substitution=0,38 mmol/g). Peptide cleavage was performed with 95% TFA, following a standard literature procedures [128].

All the amino acids and the activator 2-(1H-Benzotriazole-1-yl)-1,3,3-tetramethyluronium hexafluorophosphate (HBTU) ($\geq 99\%$) were acquired from *Novabiochem*®.

The activating base *N,N*-diisopropylethylamine (DIPEA) (99,5%), piperidine ($\geq 99,5\%$), triisopropylsilane (TIS) (98%), DCM (99,8%), and TFA were acquired from SIGMA-ALDRICH®. The solvent DMF is from CARLO ERBA REAGENTS®.

For synthesis control, it was used a Kaiser teste kit from SIGMA-ALDRICH®, containing phenol, 80% in ethanol; KCN in H₂O/pyridine; and Ninhydrin 6% in ethanol.

2.6.2 Synthesis and Characterization of the Peptides

The peptides were synthesized manually in a batch-wise manner following the methodology Fmoc-SPPS. The resin added to a reaction vessel with a filter and reagents are added and removed under manual control. The following peptides were synthesized: the 9 residue peptide hTRANSE-R(1) (RQMPTED EY) and the 23 residue peptide PepH2-hTRANSE1 (KLFMALVAFLRFLTRQMPTED EY).

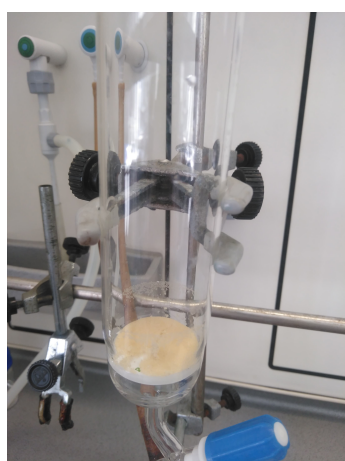
Briefly, the first step was the pre-swelling of the polystyrene-based resin with DMF for 30 min, followed by the resin washing thoroughly with DCM. Next, for Fmoc deprotection a 10 mL solution of 20% piperidine was used.



Figure 2.1: SPPS assembly, consisting on a filter reaction vessel, round bottom flask, neck glass adapter and a nitrogen stream.

HBTU (3 equiv) and the first amino acid (3 equiv) were dissolved up to 10 mL DMF, and before adding this solution to the resin, DIPEA (10 equiv) was also mixed into it. This procedure was done for the other amino acid solutions. For Fmoc deprotection, a 10 mL solution of 20% piperidine in DMF was used between the amino acids additions.

For the peptide PepH2-hTRANCE1 tenth amino acid the coupling times doubled. In the Figure 2.2 it is represented the used resin.



a Resin swelled (first step)

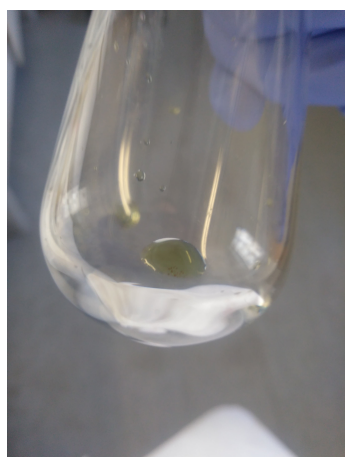


b Resin with the peptide attached

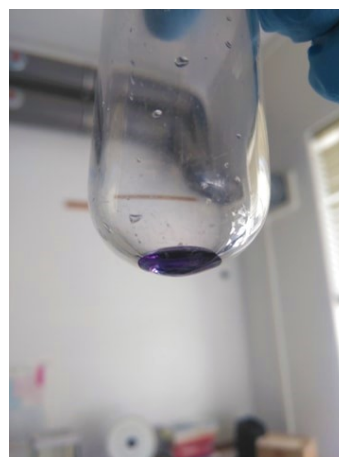
Figure 2.2: Resin in the filter reaction vessel.

The synthesis was controlled with the Kaiser test kit, following the method provided by SIGMA-ALDRICH® [132].

The resin beads and the solution turn dark blue/purple when the primary amine is present. In the opposite, in an expected result after successful coupling, the resin beads remain their colour and the solution stays yellow when no free primary amine is present. A re-coupling or deprotection step is necessary when the result is not the expected. The results observed in Figure 2.3 were present for all amino acids but for aspartic acid and proline, as expected.



a Yellow solution - amine group protected



b Dark blue/purple solution - amine group deprotected

Figure 2.3: Kaiser test trail in two resin samples, one with the amine group protected (a) and other with the amine group deprotected (b).

Prior to cleavage, the resin containing the peptide was washed repeatedly (4 times) with DMF, and finally with DCM. The cleavage was performed with 10 mL of 95% TFA / 2.5% TIS / 2.5% H₂O solution [150]. The mixture with resin and cleavage solution was gently shaken with rising stream of nitrogen for 4h. After cleavage, the resulting solvent, which has the peptide, was recovered and the resin was washed with DCM. The solvent resulting from washing was added to the solvent initially recovered, and the solution was concentrated with a stream of nitrogen to facilitate peptide precipitation. The peptides were precipitated with the addition of cold ethyl ether and centrifugation at 6000 rpm for 10 min. The obtained solid was washed twice with ethyl ether, and collected by centrifugation.

Table 2.2: Peptides chemical formula and molecular weigh (Mw) without and with (final) the carboxamide group in the C-terminus

	hTRANCE-R(1)	PepH2-hTRANCE1
Chemical Formula	$C_{48}H_{73}N_{13}O_{19}S$	$C_{131}H_{203}N_{31}O_{34}S_2$
Peptide Mw (g/mol)	1168,2	2820,4
Chemical Formula w/ CONH2	$C_{48}H_{74}N_{14}O_{18}S$	$C_{131}H_{204}N_{32}O_{33}S_2$
Peptide Mw w/ CONH2 (g/mol)	1167,2	2819,4

2.7 Peptide Purification

Peptide purification was performed by Reversed Phase High Performance Liquid Chromatography (RP-HPLC). The solvents acetonitrile (ACN), from CARLO ERBA REAGENTS®, and filtered water were used in the purification.

• Preparative Peptide Purification

The RP-HPLC system was composed by the *Waters 2535 Quaternary Gradient Module* and *Waters 2998 Photodiode Array Detector* system.

Column: Semi-preparative, *Supelco Analytical 567209-U* (C_{18} , 25 cm x 10 mm, 10 μ m, SIGMA-ALDRICH®); **UV detection:** $\lambda=210-220$ nm; **Eluents:** A- TFA 0,1% in H_2O ; TFA 0,1% in ACN.

HPLC Method for hPEP1: 0-5 min, 10%B; 5-25 min, 55%B; 25-30 min, 100%B; 30-32 min, 10%B.

HPLC Method for PepH2-TRANCE1: 0-5 min, 5%B; 5-10 min, 50%B; 10-32 min, 100%B; 32-34 min, 5%B.

• Analytical Peptide Control

The analytic RP-HPLC was a *PerkinElmer® Series 200 pump* with a *PerkinElmer® Series 200 UV/Vis Detector*.

Column: Analytical, *Supelco Analytical 568223-U* (C_{18} , 25 cm x 4,6 mm, 5 μ m, SIGMA-ALDRICH®); **UV detection:** $\lambda=210-220$ nm; **Eluents:** A- TFA 0,1% in H_2O ; TFA 0,1% in ACN.

HPLC Method for hPEP1: 0-5 min, 10%B; 5-25 min, 55%B; 25-30 min, 100%B; 30-32 min, 10%B.

HPLC Method for PepH2-TRANCE1: 0-5 min, 5%B; 5-10 min, 50%B; 10-32 min, 100%B; 32-34 min, 5%B.

Fractions containing pure target peptide, as determined by analytical RP-HPLC, were lyophilized.

2.8 Peptide Concentration

Peptide concentration can be estimated based on whether there are tryptophan or tyrosine residues present in the sequence, using UV spectrophotometry and applying the equation 2.2. The synthesized peptides in this project have one tyrosine.

$$[Peptide] = \frac{A_{280} \cdot DF \cdot M_w}{\epsilon} \quad (2.2)$$

Equation 2.2: Peptide concentration (mg ml^{-1}) [151].

Where:

A_{280} is the absorbance of the peptide solution at 280 nm in a 1 cm cell

DF is the dilution factor

M_w (g mol^{-1}) is the molecular weight of the peptide

ϵ ($\text{M}^{-1} \text{cm}^{-1}$) is the molar extinction coefficient of tyrosine at 280 nm

The purified peptides were dissolved in MiliQ water and its concentration was determined using UV spectrophotometry. The concentration of the samples was calculated using the extinction coefficient of tyrosine ($\epsilon = 1280 \text{M}^{-1} \text{cm}^{-1}$) and $\lambda = 280 \text{ nm}$ [151]. The peptide samples were frozen in MiliQ water and stored at -80°C [152].

2.9 Circular Dichroism

The circular dichroism study was performed on the Chirascan qCD spectrophotometer, Paralab, coupled with temperature controller TC125, QUANTUM NORTHWEST. These devices were available at the BioLab group, FCT/UNL. The 1 mm quartz cells were used and the CD spectra were collected between 190 nm and 260 nm.

RESULTS AND DISCUSSION

3.1 TRAF6 Sequence Conservation

The TRAF-C domains hTRAF6 and mTRAF6 have **95% sequence identity** (figure 3.1). Thus, these two proteins are homologous.

It is not available the crystal structure of the mTRAF6-mRANK complex, and so are not known the interfacial residues related to this interaction.

Laying on the fact that important binding residues are conserved throughout specimens, a putative interface can be inferred for the mouse complexes if hTRAF6 hot spots and other binding residues present in the hTRAF6-hPEP1 interaction complex are conserved in mTRAF6.

NW Score	Identities	Positives	Gaps
812	148/155(95%)	152/155(98%)	0/155(0%)
Query 1	QQCNGIYIWKIGNFGMHLKSQEEERPVIHSPGFYTGKPGYKLCMRLHLQLPTAQRRCANY		60
Sbjct 1	QQCNGIYIWKIGNFGMHLK QEEE+PVIHSPGFYTG+PGYKLCMRLHLQLPTAQRRCANY		60
Query 61	ISLFVHTMQGEYDShLPWPFQGTIRLTILDQSEALIRQNHEEVMDAKPELLAFQRPTIPR		120
Sbjct 61	ISLFVHTMQGEYDShLPWPFQGTIRLTILDQSEA +RQNHEE+MDAKPELLAFQRPTIPR		120
Query 121	NPKGFGYVTFMHLEALRQGTFIKDDTLVRCVST		155
Sbjct 121	NPKGFGYVTFMHLEALRQ TFIKDDTLVRCVST		155

Figure 3.1: BLAST result in pairwise sequence alignment between hTRAF6 (Sbjct) and mTRAF6 (Query).

Because is known that the two proteins share the same function, and the protein's sequence significantly match, then they probably share a very similar structure.

3.2 Active Residues

3.2.1 Computational Alanine Scanning

The table 3.1 summarizes the results by web-server, having into account that hot spots are labelled in Robetta when the $\Delta\Delta G > 1$ kcal/mol, in KFC2 when the confidence score is > 0 , and in Hotpoint when the pair potential is > 18 .

Table 3.1: hTRAF6 literature hot spots (blue) and binding residues in complex with RANK's fragment (PDB:1LB5) from Rosetta (hot spots $\Delta\Delta G > 1$ kcal/mol), KFC2 (hot spots confidence score > 0) and Hotpoint (hot spots pair potential > 18)

Residues	Robetta $\Delta\Delta G$ (kcal/mol)	KFC2a Score	KFC2b Score	Hotpoint Potencial
Val374	0,4	-1,9	-1,0	12,4
His376	0,4	-2,4	-1,0	18,1
Arg392	3,8	0,3	0,2	25,6
His394	-	-2,0	-1,0	-
Phe410	0,8	0,2	0,3	28,6
His412	0,4	0,0	0,1	-
Glu448	-0,0	-2,8	-1,0	-
Ile449	-	-2,4	-1,0	-
Met450	0,1	-1,4	-0,3	33,5
Asp451	-	-2,9	-1,0	-
Lys453	-0,1	-1,9	-1,0	-
Leu456	0,5	-1,0	-0,8	11,0
Leu457	-	-2,8	-1,0	7,6
Ala458	-	-0,3	-0,8	13,8
Arg466	0,1	-1,9	-0,8	-
Pro468	-	-0,5	-0,8	4,8
Lys469	0,6	-0,4	-0,6	8,9
Gly470	-	-0,2	-0,5	12,8
Phe471	2,1	0,6	0,3	37,6
Gly472	-	-0,4	-0,6	16,2
Tyr473	1,7	-0,5	-0,0	17,2
Val474	0,1	-1,0	-1,0	-
Thr475	-	-2,0	-1,0	-

All three web-servers predicted the hTRAF6 residues Arg392, Phe471 and Tyr473 as very important to the binding with RANK, as described in literature.

Robetta provided a very accurate prediction of the TRAF6 hot spot residues Arg392, Phe471 and Tyr473 (Annex 2). The other residues with a positive $\Delta\Delta G$ are also important since they have a destabilizing effect when mutated, so they must belong to the complex interface and are considered active residues. The residue Phe410 has a $\Delta\Delta G$ of 0,76 kcal/mol, so if not essential to the binding, it is also very important.

KFC2a and KFC2b (Annex 3) predicted both the same hot spots (Arg392, Phe471, Phe410 and His412), incorrectly labelling the residue Phe410 and His412. As seen in

Robetta, although the residue Phe410 is not a hot spot, it is also very important. Concerning His412, the prediction confidence is near zero (0,02 for KFC2a and 0,07 for KFC2b) so this residue is also very important and any error influences the result as "hot spot". Besides this, neither KFC2a or KFC2b predicted Tyr473 as a hot spot, but its prediction confidence is the closest to zero of all the other residues in KFC2b (-0,03).

Hotpoint (Annex 4) predicted Arg392 and Phe471 as hot spots. Tyr473 was not considered as a hot spot as the potential was not greater than 18.0, although it was close (17.22). Hotpoint's false positives (Phe410 and Met450) exist only from the high potential of those residues, not from the ASA factors. Both these residues were experimentally shown not to be hot spots [19].

Interestingly, Phe410 was highly scored in each of the used webserver, which could indicate that its mutation to alanine in TRAF6 heavily destabilizes the complex, but not enough for it to affect the complex binding affinity.

For molecular docking, these residues will be considered as "active residues" for complexes hTRAF6-hPEP1 and hTRAF6-hPEP2, apart from residue 449, since only appears in the KFC2 analysis, has one of the lowest confidence levels both on KFC2a and KFC2b and it is not a conserved residue in mTRAF6 (see Figure 3.1), which is meaningful because important binding residues are conserved throughout specimens. Because all the other residues are conserved in mTRAF6, they most probably are active residues in mTRAF6-mPEP1 and mTRAF6-mPEP2. In Figure 3.2 is the alignment of mTRAF6 and hTRAF6 sequences, where can be seen that the hTRAF6 active residues are conserved in mTRAF6.

	360	370	380	390	400	410	
mTRAF6	EMEAQQCNGI	YIWKIGNFGM	HLKSQEEERP	VVIHSPGFYT	GRPGYKLCMR	LHLQLPTAQR	
	350	360	370	380	390	400	
hTRAF6	VAEIEAQQCN	GIYIWKIGNF	GMHLKCQEEE	KPVVIHSPGF	YTGKPGYKLC	MRLHLQLPTA	
	420	430	440	450	460	470	
mTRAF6	CANYISL	FVH	TMQGEYDHL	PWPFQGTIRL	TILDQSEALI	RQNHEEVM	DAKPELLAFQRP
	410	420	430	440	450	460	
hTRAF6	QRCANYISL	F	VHTMQGEYDS	HLPWPFQGTI	RLTILDQSEA	PVRQNHEEIM	DAKPELLAFQ
	480	490	500	510			
mTRAF6	TIP	RNP	KGFG	YVT	FMHLEAL	RQGTFIKDDT	LLVRCEVSTR
	470	480	490	500			
hTRAF6	RPTIP	RNP	KG	FGYVT	FMHLE	ALRQRTFIKD	DTLLVRCEVS

Figure 3.2: mTRAF6 and hTRAF6 sequence alignment, with the active residues in red.

The TRAF6 residues used as active in molecular docking for human and mouse complexes are present in Table 3.2.

Table 3.2: hTRAF6 and mTRAF6 residues used as active in molecular docking for human and mouse complexes, respectively. Hot spots are in blue

Residue	hTRAF6	mTRAF6
Val	374	382
His	376	384
Arg	392	400
His	394	402
Phe	410	418
His	412	420
Glu	448	456
Met	450	457
Asp	451	458
Lys	453	459
Leu	456	461
Leu	457	464
Ala	458	465
Arg	466	474
Pro	468	476
Lys	469	477
Gly	470	478
Phe	471	479
Gly	472	480
Tyr	473	481
Val	474	482
Trt	475	483

Robetta and KFC2a also provided information about RANK's fragment hot spots and other binding residues (Annex 5). As already mentioned in the subsection 1.1.2, all the residues are needed for RANK's binding to TRAF6, and therefore all residues will be considered as active for molecular docking.

3.2.2 H-bonds

H-Bonds analysis identified 10 H-bonds, in Figure 3.3, between TRAF6 and RANK fragment peptide, all within 3Å, between the residues presented in Figure 3.4.

Intermolecular H-bond analysis reveals that the TRAF6 hot spot Arg392 contributed to three H-bond, and the other seven H-bond formations are due to residues marked as interface residues by the KFC2 and Hotpoint web-servers. Thus, Arg392 could be crucial not only for complex stability but also specificity. Surprisingly only one hot spot is involved in H-bond formation, corroborating that not only these residues are necessary for a successful complex interaction and all the residues listed in table 3.1 should be considered as active in molecular docking.

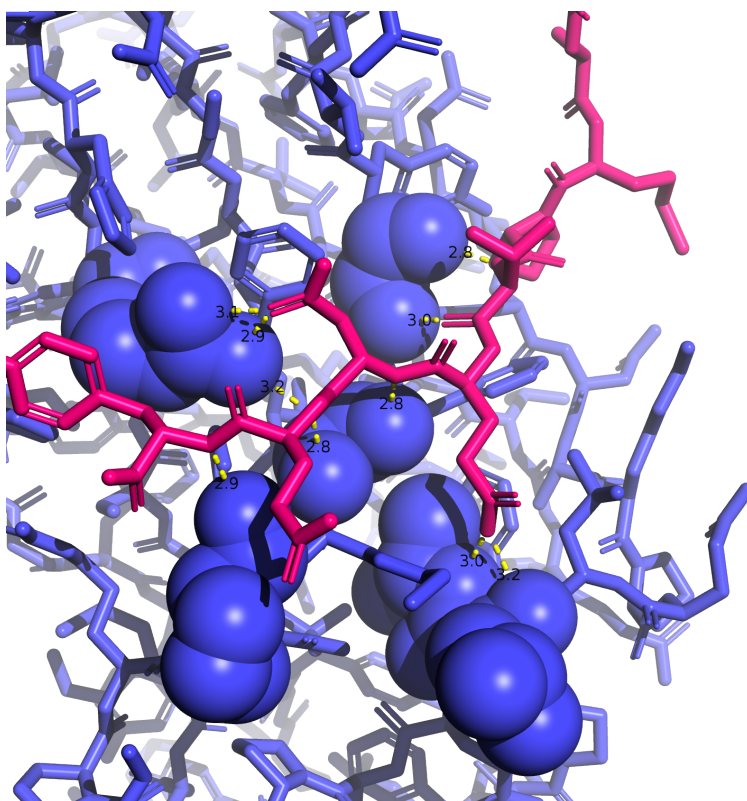


Figure 3.3: H-bonds (yellow dotted lines) visualization between hTRAF6 (blue) and RANK fragment peptide residues (pink) (PDB:1LB5).

1.	721	NH2	ARG	392	A	<-->	2594	O	ASP	606	B
2.	721	NH2	ARG	392	A	<-->	2597	OD1	ASP	606	B
3.	720	NH1	ARG	392	A	<-->	2597	OD1	ASP	606	B
4.	1762	N	LEU	457	A	<-->	2584	OE2	GLU	605	B
5.	1781	N	ALA	458	A	<-->	2584	OE2	GLU	605	B
6.	1954	O	PRO	468	A	<-->	2618	N	TYR	608	B
7.	1990	O	GLY	470	A	<-->	2591	N	ASP	606	B
8.	1987	N	GLY	470	A	<-->	2594	O	ASP	606	B
9.	2017	O	GLY	472	A	<-->	2562	N	THR	604	B
10.	2014	N	GLY	472	A	<-->	2565	O	THR	604	B

Figure 3.4: Number of hydrogen bonds and participating residues in the 1LB5 complex.

3.2.3 Interface & Active Residues

As mentioned in subsection 1.3.2.2, CAPRI defines a protein-peptide complex interface as the residues within a 8Å distance from each other, counting for active and passive residues.

The cut-offs for inter-residue interactions range from 5Å to 14Å, thus it is interesting to know the distance between the residues considered active listed in table 3.2. Once these residues necessarily belong to the interface, for the other complexes it is possible to know if this “active interface” is maintained as a comparison method.

As already mentioned, almost all H-bonds are within 3Å, and the active residues are within a maximum of 6Å (Annex 6), most being in the 5Å range, meaning that most active

residues are hydrophobic, as expected.

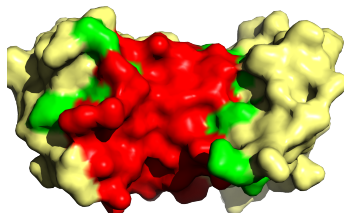


Figure 3.5: hTRAF6 active residues (within 6Å), in red, and passive residues (within 8Å), in green.

3.3 Tertiary Structure Prediction

For the tertiary structure prediction of the peptides, I-TASSER and PEP-FOLD3 gave a output of five models, and PEPstrMOD generate five clusters, each with five models. As mentioned in the subsection 2.3, in the ensemble docking the five models from I-TASSER, PEP-FOLD3, and PEPstrMOD were used.

ProSA was used for hPEP1 structure evaluation. The overall quality z-scores calculated by ProSA for all input structures are divided by web-servers in Annex 7, Annex 8 and Annex 9, along with the plot with the z-scores of all experimentally determined structures that are in the PDB database. The z-scores of all predicted peptide structures are within the range of scores typically found for native proteins with sequence length of 25 amino acids, suggesting that the structures predicted by I-TASSER, PEP-FOLD3 and PEPstrMOD probably do not contains errors and are going to be used in the ensemble molecular docking.

The structure presented in Figure 3.6 was given by ModBase as an output. The template chosen by ModBase corresponds to the the human TRAF6's TRAF-C, which has 95% sequence identity with mice TRAF6's TRAF-C as already known.

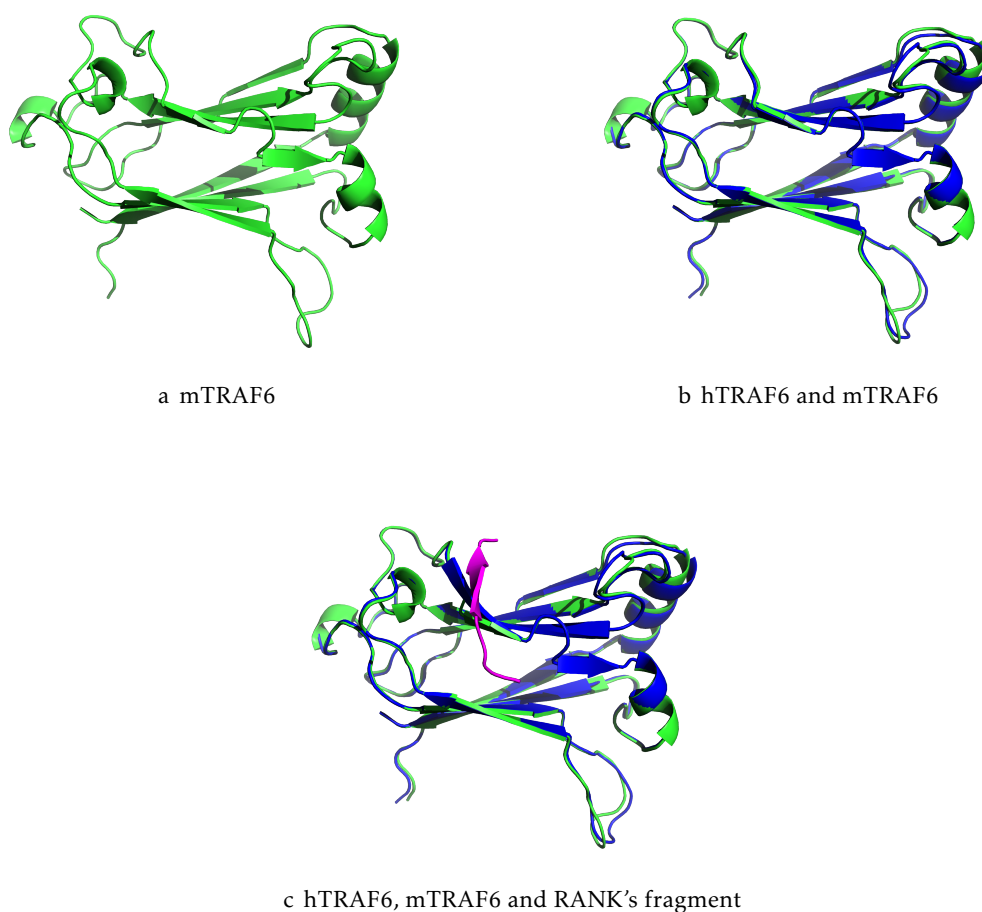


Figure 3.6: mTRAF6 tertiary structure.

As expected, the mTRAF6 tertiary structure seems to be very similar to hTRAF6, being the RMSD between them 1,2Å.

3.4 Molecular Docking

3.4.1 hTRAF6-hPEP1 Cluster Analysis

The HADDOCK results for the complex hTRAF6-hPEP1 are presented in Annex 10, where the best clusters are sorted in ascending order, along with the best 4 representatives complexes for each cluster. The HADDOCK scores are all very close, with the scores of the clusters overlapping within standard deviation from each other. Thus, the selection of the best cluster is harder, it is impossible to choose the best cluster based on the HADDOCK score. However, in terms of z-score, which the more negative the better, it is only negative until cluster 5, so clusters beyond that will not be considered. This first evaluation leaves 4 clusters to analyze and the best 4 quaternary structures from each one, so 16 structures in total.

A correlation between the l-RMSD and i-RMSD with the other parameters was performed, in order to help choosing the best models for the other four docked complexes that do not have any reference model to be compared. Annex 11 presents, for each docked model, the calculated values for l-RMSD, i-RMSD, HADDOCK score and intermolecular energy. The l-RMSD and i-RMSD were calculate for each complex as mentioned in subsection 2.4.2, the HADDOCK score with the equation 1.4 and the intermolecular energy with the equation 1.8. The BSA for each structure is specified as well, taken from the PDB files.

Figure 3.7 represents the HADDOCK score as a function of the l-RMSD and Figure 3.8 represents the HADDOCK score as a function of the i-RMSD taking the PDB 1LB5 as template for the 16 calculated structures. Both Figures show that for similar HADDOCK scores, the two first clusters ranked by HADDOCK have significantly lower l-RMSD and i-RMSD values, indicating that the 8 models that represent them are the closest to the PDB 1LB5. Furthermore, as the HADDOCK score increases, the l-RMSD and i-RMSD values increase as well. The main conclusion is that the first two clusters best represent the reference structure.

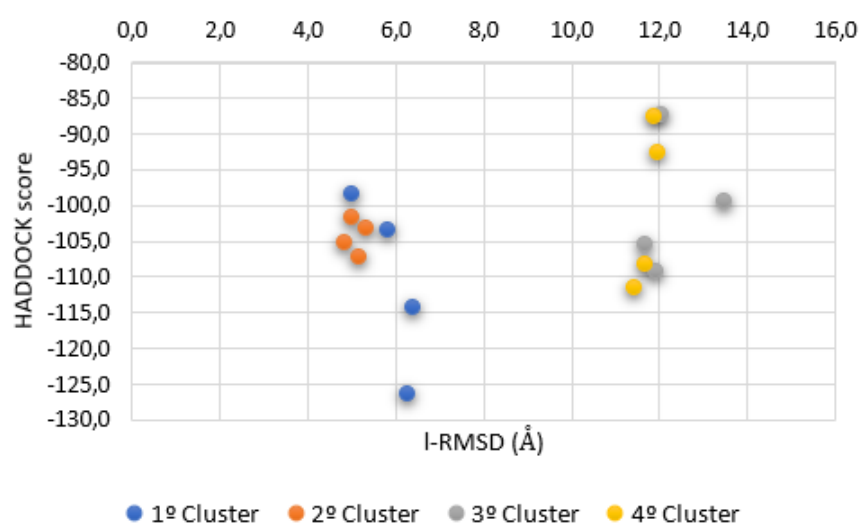


Figure 3.7: HADDOCK score versus l-RMSDs for the hTRAF6-hPEP1 complex.

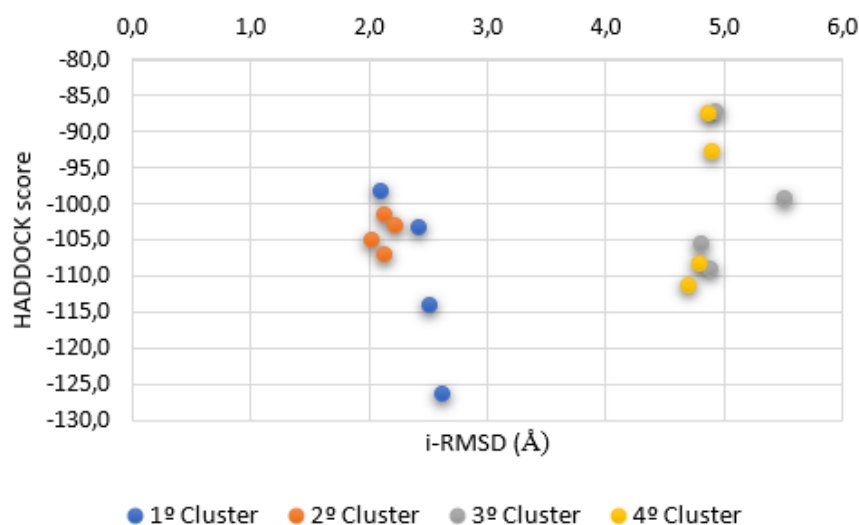


Figure 3.8: HADDOCK score versus i-RMSDs for the hTRAF6-hPEP1 complex.

Figure 3.9 represents the ranking of the clusters based on their intermolecular energies. The results are in agreement with what was mentioned above as cluster 2 has the lower average intermolecular energy, increasing up to cluster 5. Considering the intermolecular energy, the first two clusters best represent the reference structure (Annex 12). Being the conclusion taken the same as the one taken with the HADDOCK score, it also means that the dissociation energy does not have influence in the HADDOCK score, since the intermolecular energy equation does not have the dissociation energy parcel and the results are the same as the HADDOCK score results. Again, the main conclusion is that the first two clusters best represent the reference structure, and that the models with the highest intermolecular energy correspond to the highest l-RMSD and i-RMSD values.



Figure 3.9: Average intermolecular energy versus the 4 best clusters for the hTRAF6-hPEP1.

The hypothesis of a correlation between the BSA and the HADDOCK score would be

also valid, since the affinity increases with the BSA. However, there is not any correlation, since the values of BSA are all very close for the 4 clusters (figure 3.10).

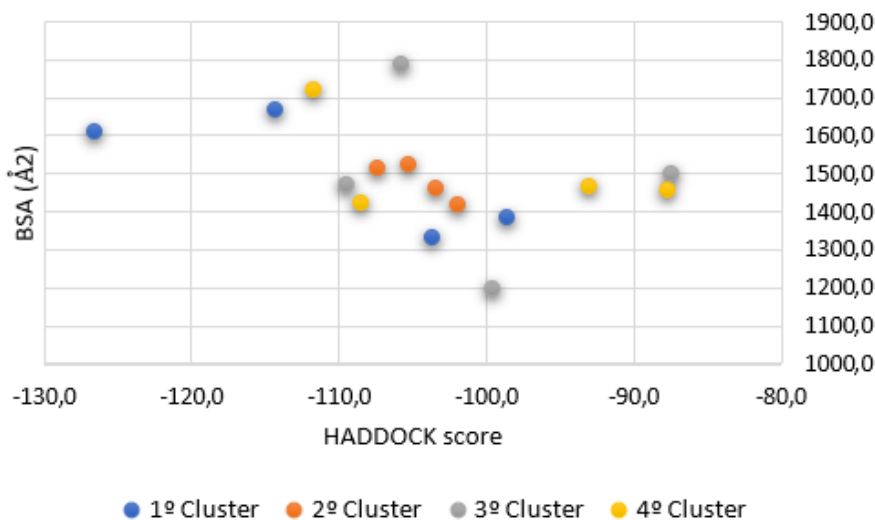


Figure 3.10: BSA ($\text{\AA}^2$) versus the HADDOCK score for the 16 study structures representing the hTRAF6-hPEP1 complex.

Visualizing the overlap of structure pairs allowed to conclude that only the models from clusters 2 and 3, which are the best clusters according to HADDOCK ranking and the RMSD analysis, are the only ones with the peptides in the same orientation as the RANK's fragment in PDB 1LB5 (figure 3.11), and so, as mentioned in section 3.4, structures ranked at the top are in the correct orientation.

Since the best docked models are concentrated in the best two clusters, the analysis of the other four docked complexes will be made between them two.

All the models from the best two and first clusters have the peptide hPEP1 predicted by I-TASSER, so the structures of the peptides from the other four complexes are going to be predicted by I-TASSER as well.

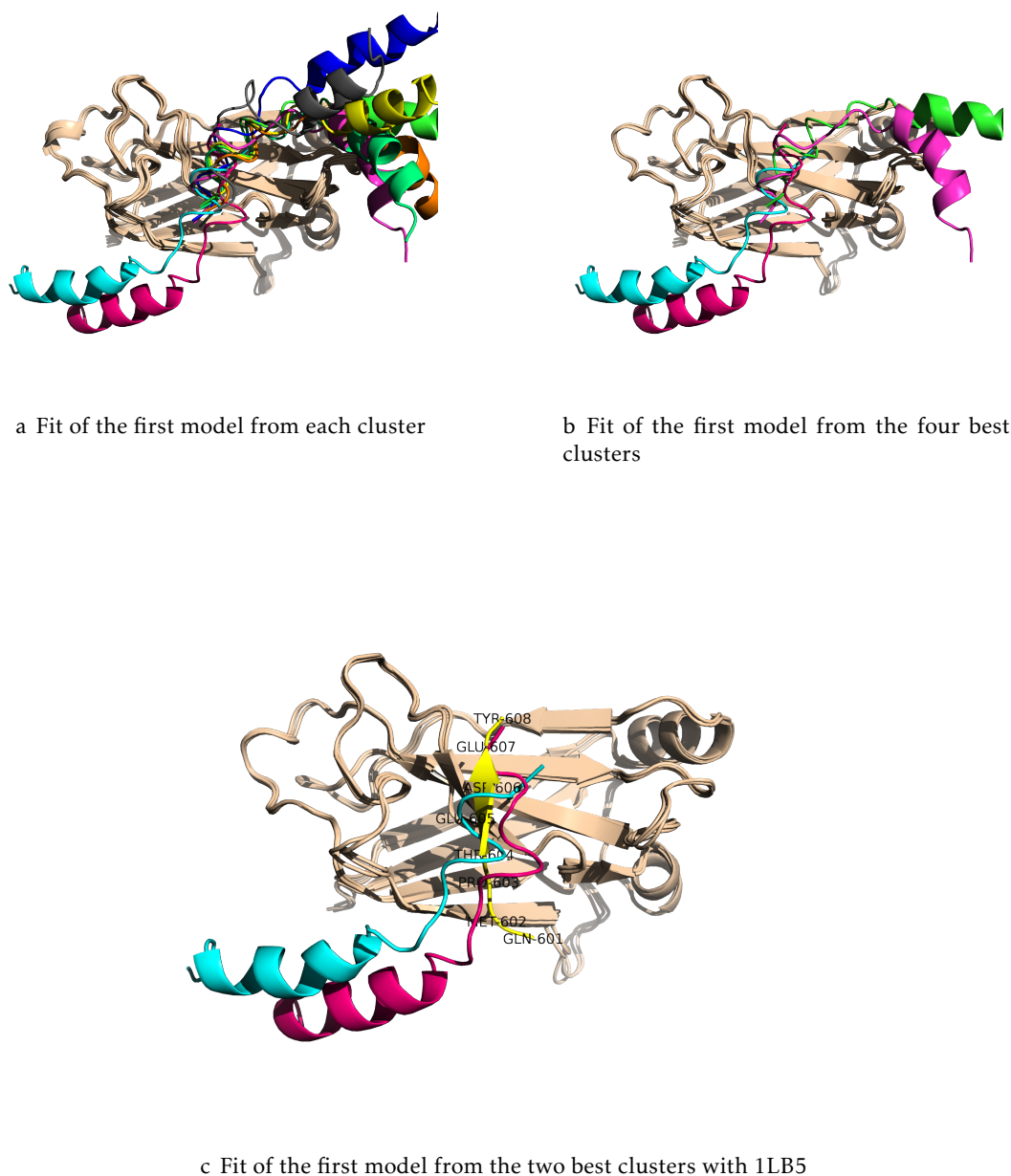


Figure 3.11: hTRAF6-hPEP1 fit of the first model from each cluster. Cluster 1 - light green; Cluster 2 - light blue; Cluster 3 - dark pink; Cluster 4 - yellow; Cluster 5 - light pink; Cluster 6 - grey; Cluster 7 - dark blue; Cluster 8 - orange; Cluster 9 - dark green.

3.4.2 hTRAF6-hPEP1 Model Selection

Considering the CAPRI evaluation data in Table 1.2, the l-RMSD and i-RMSD are very high, with only 3 complexes considered “acceptable” from the first two clusters, present in Table 3.3.

Table 3.3: 3 complexes considered “acceptable” from the first two clusters for hTRAF6-hPEP1

Cluster	Model	Acceptable RMSD (Å)
2	model2_4	l-RMSD=5,0 (5,0-2,0)
3	model3_2	l-RMSD=4,8 (5,0-2,0) & i-RMSD=2,0 (2,0-1,0)
3	model3_4	l-RMSD=5,0 (5,0-2,0)

This demonstrates that any change in one of the unbound structures, even with the docking restrictions, makes a difference in the final complex. In this case the change made was the cell-penetrating peptide to TRAF6-binding site TRANCE-R(1).

The chosen model to represent the complex hTRAF6-hPEP1 is the **model3_2**, since it complies with both criteria for l-RMSD and i-RMSD. In Figure 3.12, it is possible to check that the part of the peptide that is meant to interact with TRAF6, is fitted with the residues in RANK’s fragment at the PDB 1LB5.

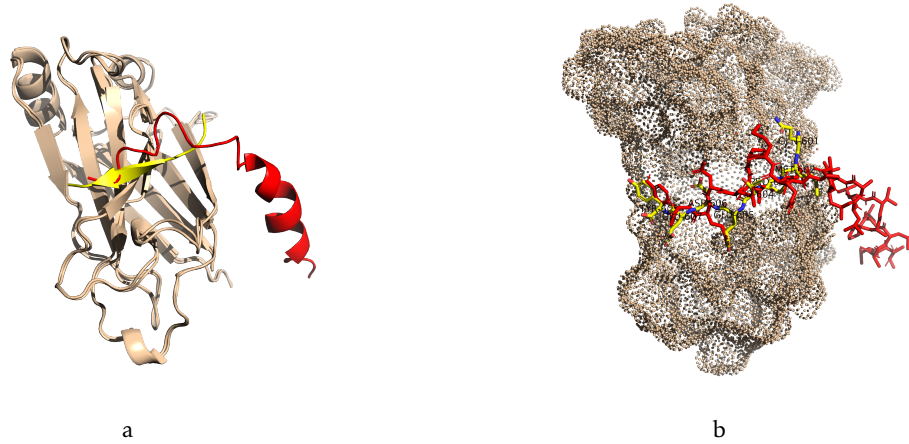


Figure 3.12: Model3_2 aligned to complex 1LB5. RANK’s fragment in yellow and the peptide predicted by I-TASSER in red.

It is important to know if the residues determined in HADDOCK as active are conserved in the docked complexes. Calculated by equation 3.1 and for a cut-off of 6Å, the percentage of active residues conserved in the interface ($\%AR_{conserved}$) is 100%. Thus, the 22 residues determined as active in hTRAF6 for the complex 1LB5 are present in this interface (Annex 13). There are 7 extra residues that are natural to appear since the docked peptide is not equal to the peptide in the complex 1LB5.

$$\%AR_{conserved} = \frac{N_{ActiveResidues}}{N_{TotalResidues}} \times 100 = \frac{N_{ActiveResidues}}{22} \times 100 \quad (3.1)$$

Equation 3.1: Percentage of active residues conserved in the interface.

Thus, the model 3_2, from the second best cluster ranked by HADDOCK (cluster 3),

could be considered a good model to represent hTRAF6-hPEP1 complex.

3.4.3 Complexes Model Selection

To evaluate the tertiary structure of the other peptides the ProSA web-server was also used. The results are in Annex 14 for hPEP2, Annex 15 for mPEP1, Annex 16 for mPEP2 and Annex 17 for PepH2-hPEP1. The structures predicted by I-TASSER will be used in the ensemble molecular docking, in which five peptide models and one protein are docked.

In Table 3.4 are the models chosen for each complex, and the respective percentage of residues conserved in the interface that were determined as active in HADDOCK ($\%AR_{conserved}$).

Table 3.4: Models chosen for each complex, and the respective percentage of active residues conserved in the interface

Complex	Model selected	%ARconserved
hTRAF6-hPEP2	model2_1	83,3%
mTRAF6-mPEP1	model1_1	95,5%
mTRAF6-mPEP2	model2_1	90,9%
hTRAF6-PepH2-hPEP1	model8_1	100%

3.4.3.1 hTRAF6-hPEP2

The HADDOCK results for the complex hTRAF6-hPEP2 are shown in Annex 18, ordered by the HADDOCK ranking. The first two clusters have the HADDOCK scores values far enough apart, and the z-score negative. The models with the best HADDOCKs score and intermolecular energy from each cluster (Annex 19) were overlapped (Figure 3.13) to analyze the peptides orientation, which is unknown since the binding peptide is different from the one studied above.

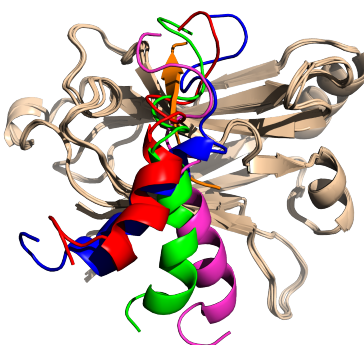


Figure 3.13: hTRAF6-hPEP2 fitting of the best two models from each cluster (model2_1, green; model2_3, pink - from cluster 2 - and model3_1, blue; model3_2, red - from cluster 3). RANK's fragment, orange.

To choose among these four models, the conservation of the hTRAF6 active residues will be analyzed, comparing with table 3.2, being the results in Annex 20.

In the first cluster (cluster 2), for the peptide ranked in first by the HADDOCK score (model2_1), that all the active residues expected are present in the interface. For the peptide ranked in first having into account the intermolecular energy (model2_3), there are missing 4 residues that should be maintained as active, and so the $\%AR_{conserved}$ is 83,3%.

In the second cluster (cluster 3), for the peptide ranked in first by the HADDOCK score (model3_1), it is verified that once again there are missing 4 residues that should be maintained as active, and so the $\%AR_{conserved}$ is 83,3%. For the peptide ranked in first having into account the intermolecular energy (model3_2), it is verified the same thing.

Thus, the model2_1 seems to be the best choice to represent the complex hTRAF6-hPEP2, which is, by HADDOCK score, the best ranked model in the best ranked cluster.

3.4.3.2 mTRAF6-mPEP1

The HADDOCK results for the complex mTRAF6-mPEP1 are in Annex 21, ordered by the HADDOCK ranking. The first two clusters, both with the same value of negative z-score, have the HADDOCK scores values within standard deviation from each other, so it is expected that all the models are very similar.

The models with the best HADDOCK scores and intermolecular energy from each cluster, in Annex 22, were overlapped in figure 3.14, to analyze the peptides orientation which is expected to be the same as the one presented in the complex hTRAF6-hPEP1 since the binding peptide hTRANSE-R(1) is very similar to the peptide mTRANSE-R(1), only differing in 2 residues.

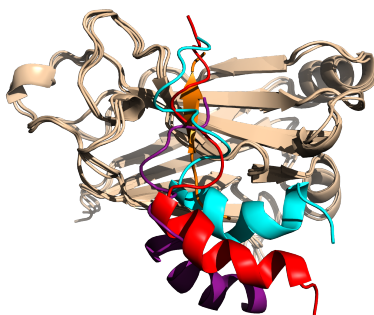


Figure 3.14: mTRAF6-mPEP1 fitting of the best two models from each cluster (model4_1, purple - from cluster 4 - and model1_1, red; model1_2, blue - from cluster 1). RANK's fragment, orange.

For cluster 4 the model4_1 has both the best HADDOCK score and intermolecular energy. All the three peptides are in the same orientation, which is the same as in the

1LB5 PDB, as expected. However, the peptide from cluster 4 seems dislocated comparing to the other peptides.

To choose between these three models, the conservation of the mTRAF6 active residues, in Table 3.2, will be analyzed, being the results in Annex 23.

In the first cluster (cluster 4), for the peptide ranked in first by the HADDOCK score and intermolecular energy (model4_1), it is verified that all the active residues expected are present in the interface except for 1, and so the $\%AR_{conserved}$ is 95,5%.

In the second cluster (cluster 1), for the peptide ranked in first by the HADDOCK score (model1_1), it is verified that once again it is missing 1 residue, and so the $\%AR_{conserved}$ is the same as for the model4_1. For the peptide ranked in first having into account the intermolecular energy (model1_2), it is missing 4 residues, which translate to 81,8% of conservation of the residues.

As expected, the results are very close. Observing the results in Annex 21, the RMSD from the overall lowest-energy structure is better for cluster 1 than cluster 4. Because of the better RMSD result the model1_1 is the best one to represent the complex mTRAF6-mPEP1, having the best HADDOCK score from the second best ranked cluster.

3.4.3.3 mTRAF6-mPEP2

The HADDOCK results for the complex mTRAF6-mPEP2 are shown in Annex 24, ordered by the HADDOCK ranking. The first two clusters, both with the close negative z-score values, have the HADDOCK scores values within standard deviation from each other, so it is expected that all the models are very similar. The models with the best HADDOCK scores and intermolecular energy from each cluster, in Annex 25, were overlapped in figure 3.15, to analyze the peptides orientation which is expected to be the same as the one presented in the complex hTRAF6-hPEP2, and the others since the peptide orientation is always the same, since the binding peptide hTRANCE-R(2) is very similar to the peptide mTRANCE-R(2), only differing in 2 residues.

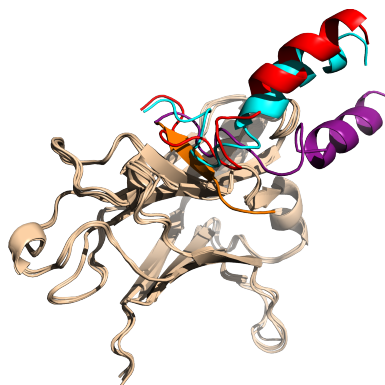


Figure 3.15: mTRAF6-mPEP2 fitting of the best models from the two first clusters (model2_1, purple - from cluster 2 - and model1_1, red; model1_2, blue - from cluster 1). RANK's fragment, orange.

For cluster 2 the model2_1 has both the best HADDOCK score and intermolecular energy. All the three peptides are in the same orientation, which is the same as the expected.

To choose between these three models, the conservation of the mTRAF6 active residues, in table 3.2, will be analyzed, being the results in Annex 26.

In the first cluster (cluster 2), for the peptide ranked in first by the HADDOCK score and intermolecular energy (model2_1), it is that are missing 2 active residues, and so the $\%AR_{conserved}$ is 90,9%.

In the second cluster (cluster 1), for the peptide ranked in first by the HADDOCK score (model1_1), it is verified that are missing 3 active residues, and so the $\%AR_{conserved}$ is 86,4%. For the peptide ranked in first having into account the intermolecular energy (model1_2), it is missing 4 residues, which translate to 81,8% conservation of the active residues.

Thus, the model2_1 is the best choice to represent de complex mTRAF6-mPEP2, which is, by HADDOCK score, the best ranked model in the best ranked cluster.

3.4.3.4 hTRAF6-PepH2-hPEP1

The HADDOCK results for the complex mTRAF6-mPEP2 are presented in Annex 27, ordered by the HADDOCK ranking. The first two clusters, both with the close negative z-score values, have the HADDOCK scores within standard deviation from each other, so it is expected that all the models are very similar.

The models with the best HADDOCK scores and intermolecular energy from each cluster, in Annex 28, were overlapped in figure 3.16, to analyze the peptides orientation which is expected to be the same as the one presented in the complex hTRAF6-hPEP1, and the others since the peptide orientation is always the same, since the binding peptide hTRANCE-R(1) is shared by these two complexes.



Figure 3.16: hTRAF6-PepH2-hPEP1 fitting of the best models from the two first clusters (model8_1, blue; model8_2, red - from cluster 8 - and model5_1, purple - from cluster 5. RANK's fragment, orange).

For cluster 5 the model5_1 has both the best HADDOCK score and intermolecular energy. All the three peptides are in the same orientation, which is the same as the expected.

To choose between these three models, the conservation of the mTRAF6 active residues, in table 3.2, will be analyzed, being the results in Annex 29.

In the first cluster (cluster 8), the models 8_1 and 8_2 have both the same highest HADDOCK score, and the latter has the highest intermolecular energy. The model8_1 has all the supposed active residues, and so has a $\%AR_{conserved}$ of 100%. For the model8_2, it is missing 2 residues, which translate to 90,9% conservation of the residues.

In the second cluster (cluster 5), for the peptide ranked in first by the HADDOCK score and intermolecular energy (model5_1), are missing 2 active residues as well, with a $\%AR_{conserved}$ of 90,9%.

Thus, the model8_1 is the best choice to represent hTRAF6-PepH2-hPEP1, which is, by HADDOCK score, the best ranked model in the best ranked cluster.

3.5 Binding Affinity

3.5.1 Buried Surface Area (BSA)

The binding affinity will be evaluated having into account the BSA results, given by HADDOCK, of the chosen 8 models, for each complex, and the best model chosen to represent each one (table 3.5).

Taking into consideration the data presented in Table 3.5 and Figure 3.17, when comparing the BSA of the chosen models, the complex mTRAF6-mPEP1 has a bigger BSA than the complex mTRAF6-mPEP2, indicating that mPEP1 has a higher affinity (lower K_d) for mTRAF6 than mPEP2 (subsection 1.1.2). The same behavior is observed between the complexes hTRAF6-hPEP1 and hTRAF6-hPEP2.

Table 3.5: BSA ($\text{\AA}^2$) of the best chosen models, for each complex, and average BSA ($\text{\AA}^2$) of each complex, having into account the first two clusters ranked by HADDOCK (8 models). The difference between BSA is also displayed

Complexes	BSA ($\text{\AA}^2$)	
	Models Average	Chosen Structure
1) mTRAF6-mPEP1	1402,2	1645,1
2) mTRAF6-mPEP2	1415,3	1543,0
$\Delta 1 - 2 $	13,2	102,2
3) hTRAF6-hPEP1	1485,8	1520,5
4) hTRAF6-hPEP2	1499,4	1439,9
$\Delta 3 - 4 $	13,6	80,5
5) hTRAF6-PepH2-hPEP1	1582,8	1662,3
$\Delta 5 - 3 $	97,0	141,8

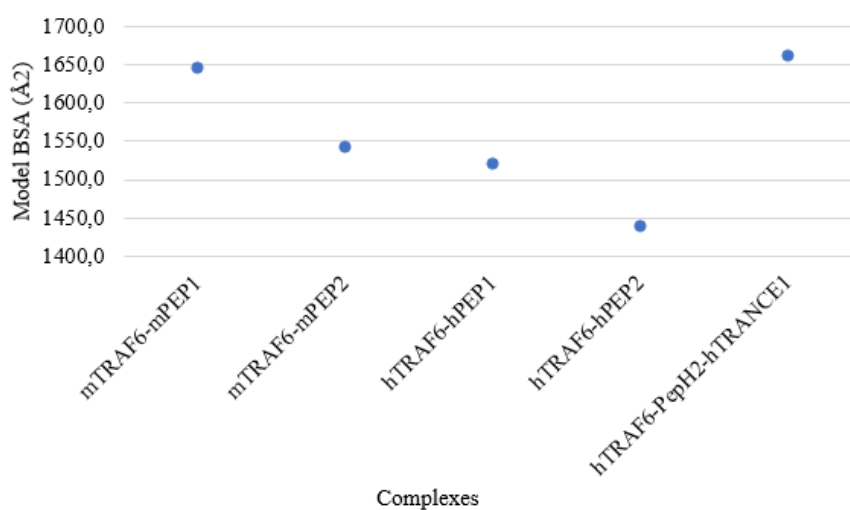


Figure 3.17: BSA ($\text{\AA}^2$) of the best chosen model, for each complex.

However, there is not the same BSA behavior when it is considered an average between all the BSAs of the 8 complexes coming from the best two HADDOCK clusters, as presented in Figure 3.18 and Table 3.5. The complexes mTRAF6-mPEP1 and mTRAF6-mPEP2 have only a difference of about $13\text{\AA}^2$, having the latter complex a bigger BSA, and that the same occurs for human complexes. As the difference value is not noteworthy, in this case the peptides seem to have the same affinity for TRAF6.

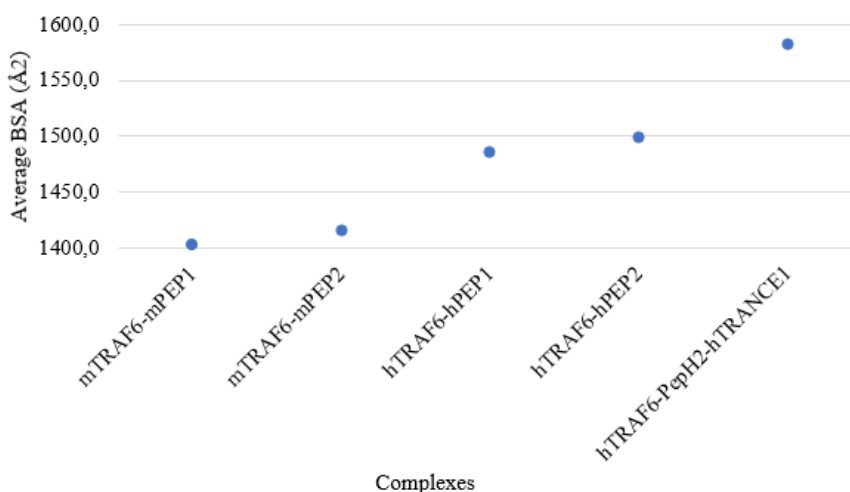


Figure 3.18: Average BSA ($\text{\AA}^2$) of each complex, having into account the first two clusters ranked by HADDOCK (8 models).

Although the models for each cluster were chosen without having into account the BSA, the BSA value of the chosen structure is higher than the average BSA. Therefore, a bigger BSA is associated with an increased complex stability, reinforcing that the proposed models are stable and good representations of the complexes, which suggests that the BSA data for the chosen models is reliable and can be used to study the complexes binding

affinity.

The complex hTRAF6-PepH2-hPEP1 is the one with the highest BSA when dealing with the chosen structure or when doing the average BSA. These results suggest that when the peptide hTRANCE-R(1) is linked to the cpp PepH2 instead of the cpp from the Kaposi fibroblast growth factor signal, its affinity towards hTRAF6 would increase and the complex would be much more stable. It is expected that the K_d of the peptide PepH2-hPEP1 with hTRAF6 is lower than the K_d of hPEP1, indicating that as far as peptide affinity, it is better to use the cpp PepH2 when the goal is to dock the peptide with hTRAF6. PepH2-hPEP1 will be synthesized, along with hTRANCE-R(1) (RANK's binding site without any cpp).

As far as the interface area, all the BSA values are within the standard-size range ($1600\text{\AA}^2 \pm 400\text{\AA}^2$).

This computational approach seems a good methodology to predict the stability of protein-peptide complexes. With this methodology is possible to do a pre-screening of the most potential compounds to be synthesized, avoiding time- and resource-consuming methodologies.

3.5.2 Optimized HADDOCK Score

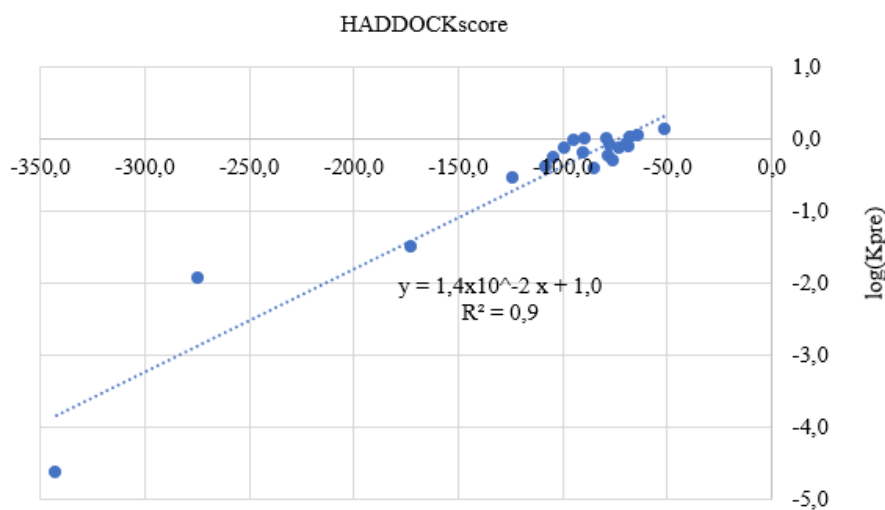
The optimized HADDOCK score for protein-peptide complexes is presented in Equation 3.1.

$$\begin{aligned} \log(K_{pred}) &= \text{HADDOCKscore} \\ &= 1,8 \times 10^{-3} E_{vdw} + 3,9 \times 10^{-3} E_{elec} + 7,2 \times 10^{-3} E_{desol} + 2,4 \times 10^{-6} BSA + 4,1 \times 10^{-1} \end{aligned}$$

Equation 3.1: Optimized HADDOCK score for protein-peptide complexes.

In fact, as seen in Figure 3.19, the HADDOCK score could be correlated with $\log(K_{pred})$, indicating that the lower the HADDOCK score, the higher the affinity of the peptide to the protein is.

This optimized equation was tested for the chosen models in molecular docking, with the results in Table 3.6. However, the results are not the expected, as the complex hTRAF6-hPEP1 has a K_d higher than the complex hTRAF6-hPEP2, and the complex mTRAF6-mPEP1 has also a higher K_d than the complex mTRAF6-mPEP2, indicating that peptides hPEP2 and mPEP2 have more affinity to hTRAF6 and mTRAF6, respectively. These results are not in agreement with the literature and the ones determined by BSA studies. This unsucces might have been a consequence of an error made in the method used to its development, such as the lack of filters applied in PDB advanced search that allowed different complexes to be included in the dataset.

Figure 3.19: Correlation between the HADDOCK scores and $\log(K_{\text{predicted}})$.Table 3.6: K_d for each model selected in molecular docking.

Peptide	Model	$\log(K_{\text{pred}})$	K_d
hTRAF6-hPEP1	Cluster3_2	-0,8	0,2
hTRAF6-hPEP2	Cluster2_1	-0,9	0,1
mTRAF6-mPEP1	Cluster1_1	-0,5	0,3
mTRAF6-mPEP2	Cluster2_1	-0,8	0,2
hTRAF6-PepH2-hTRANCER1	Cluster8_1	-1,0	0,1

3.6 Peptide Synthesis and Purification

The peptides were assembled to a MBHA Rink Amide resin by Fmoc-based manual SPPS. The peptides were purified by semi-preparative RP-HPLC and characterized by ESI-MS (Tables 3.7 and 3.8).

Table 3.7: Analytical data of the synthesized hTRANCE-R(1).

Peptide	hTRANCE-R(1)
Sequence	RQMPTED EY
Formula	$C_{48}H_{74}N_{14}O_{18}S$
Calculated exact mass (Da)	1167,2
Found [ion]	584,8 $[M + 2H]^{2+}$ & 1167,8 $[M + H]^+$
* t_R (min) (Purity (%))	12,2 (>99%)

* t_R : Retention time for the analytical HPLC system.

Table 3.8: Analytical data of the synthesized PepH2-TRANCER1.

Peptide	PepH2-hPEP1	PepH2-hPEP1(-V)
Sequence	KLFMALVAFLRFLT RQMPTDEY	KLFMALAFLRFLT RQMPTDEY
Formula	$C_{131}H_{204}N_{32}O_{33}S_2$	-
Calculated exact mass (Da)	2819,4	-
Found [ion]	994,8 $[M + 3H + ACN]^{3+}$	928,2 $[M + 2H + K + ACN]^{3+}$
* t_R (min) (Purity (%))	14,6 (>99%)	14,9 (>99%)

* t_R : Retention time for the analytical HPLC system.

The HPLC chromatogram of hTRANCE-R(1) before purification is presented in Annex 30, and the corresponding ESI-MS spectrum is shown in Figure 3.20. After purification, (> 99%) the peptide was controlled by analytical HPLC (Figure 3.21) and characterized by ESI-MS (Figure 3.22).

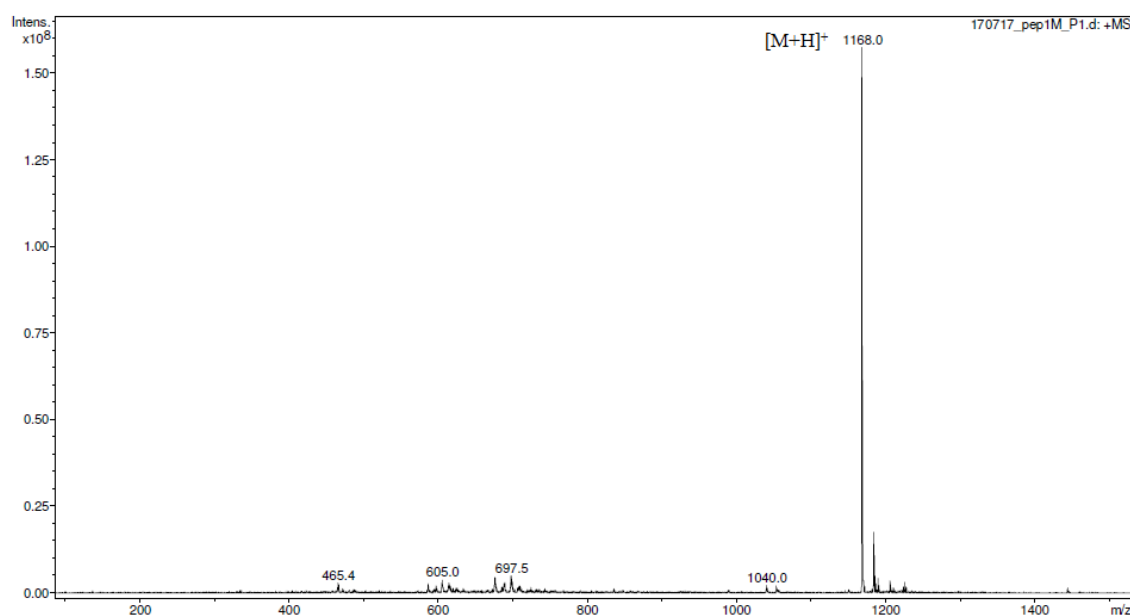


Figure 3.20: TRANCE-R(1) ESI-MS before purification.

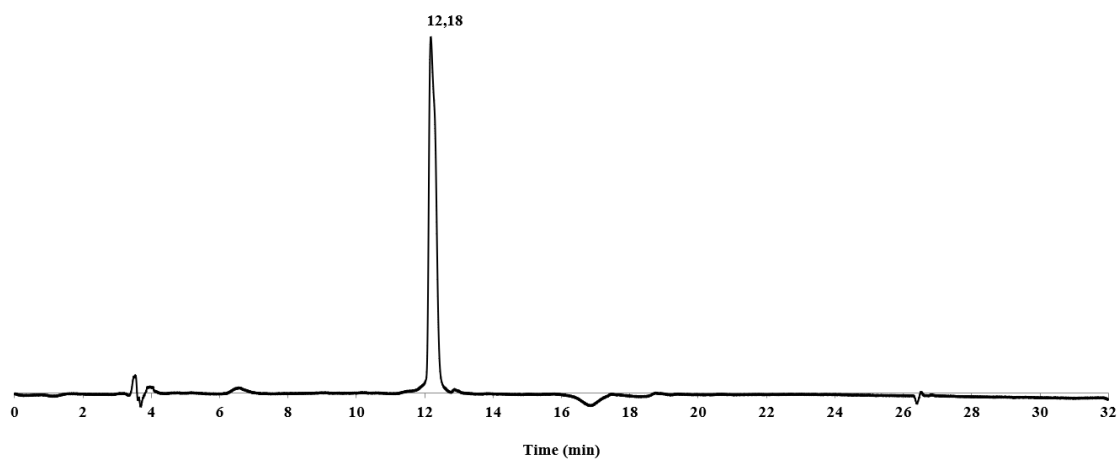


Figure 3.21: TRANCE-R(1) chromatogram after purification (analytical HPLC).

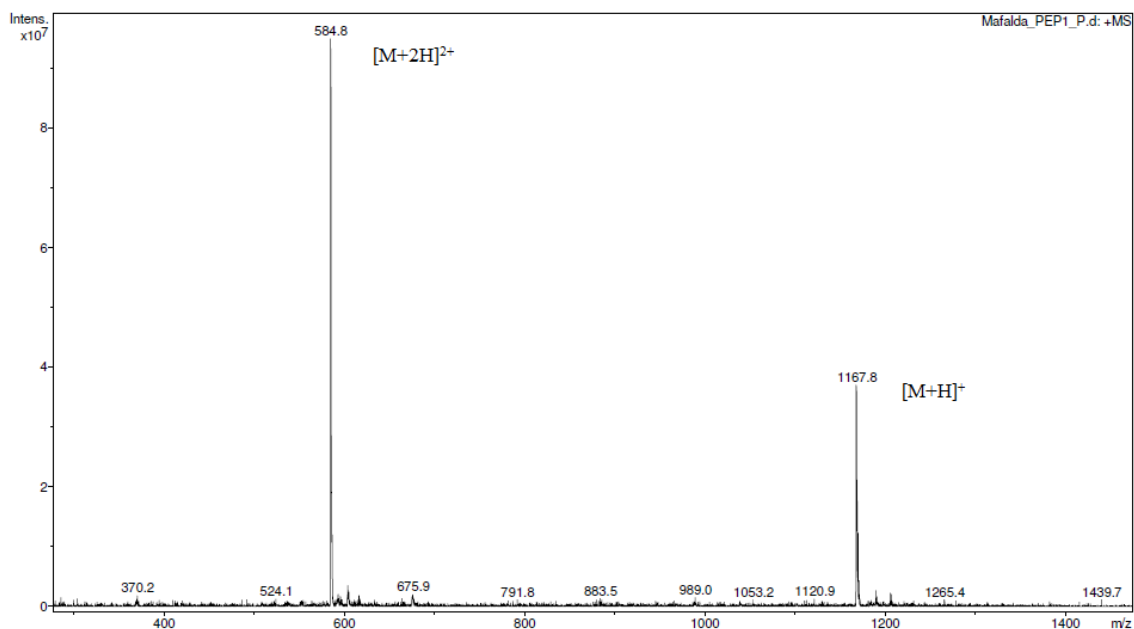


Figure 3.22: TRANCE-R(1) ESI-MS after purification.

The HPLC chromatogram of PepH2-hPEP1 before purification is presented in Annex 31, and the corresponding ESI-MS spectrum is shown in Figure 3.23, which is not conclusive.

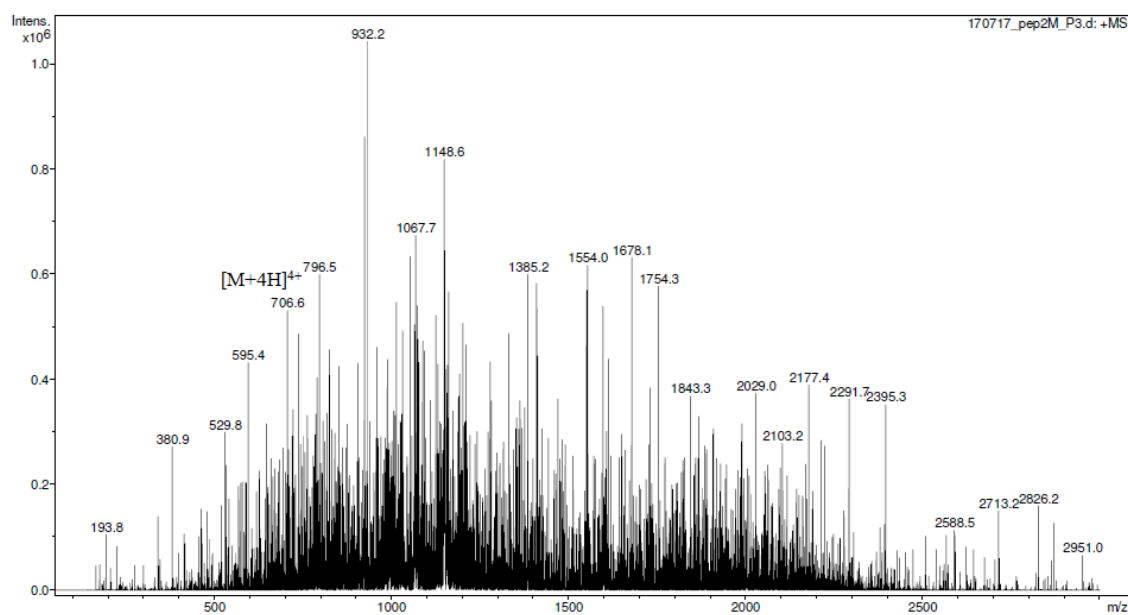


Figure 3.23: PepH2-hPEP1 mass spectrum before purification.

The PepH2-hPEP1 chromatogram before purification shows two main peaks (retention time of 20,63 min and 20,91 min). The peaks were collected by HPLC purification and each fraction was analysed by analytical HPLC (Figure 3.24) (purity > 99%) and characterized by ESI-MS (Figures 3.25 and 3.26).

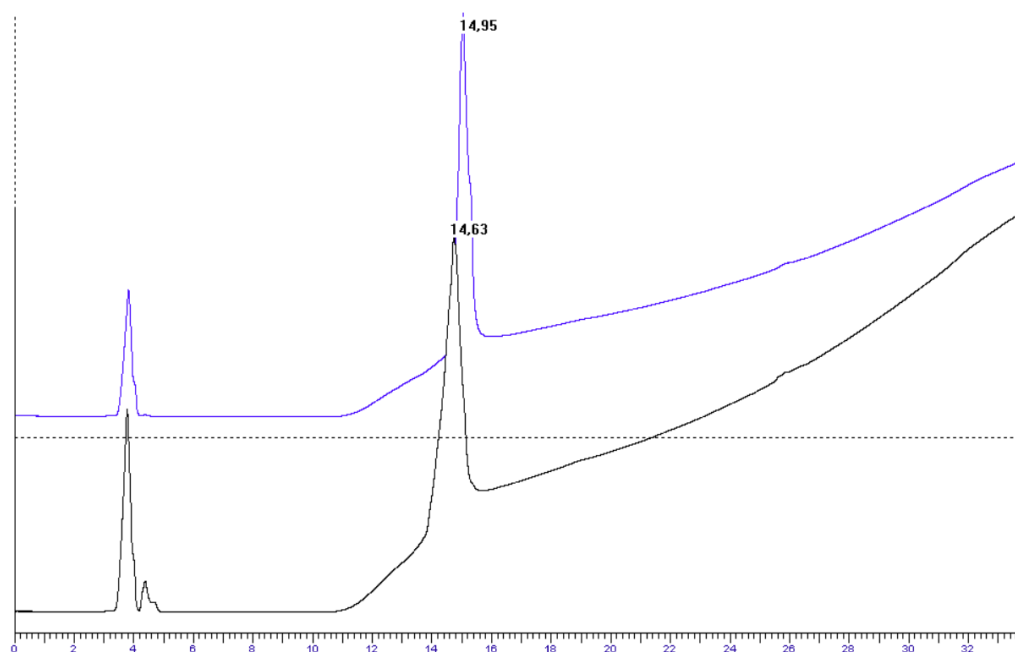


Figure 3.24: PepH2-hPEP1 chromatogram after purification of the first fraction with the entire peptide (black line) and the fraction with the peptide without the residue Valine (blue line) (analytical HPLC).

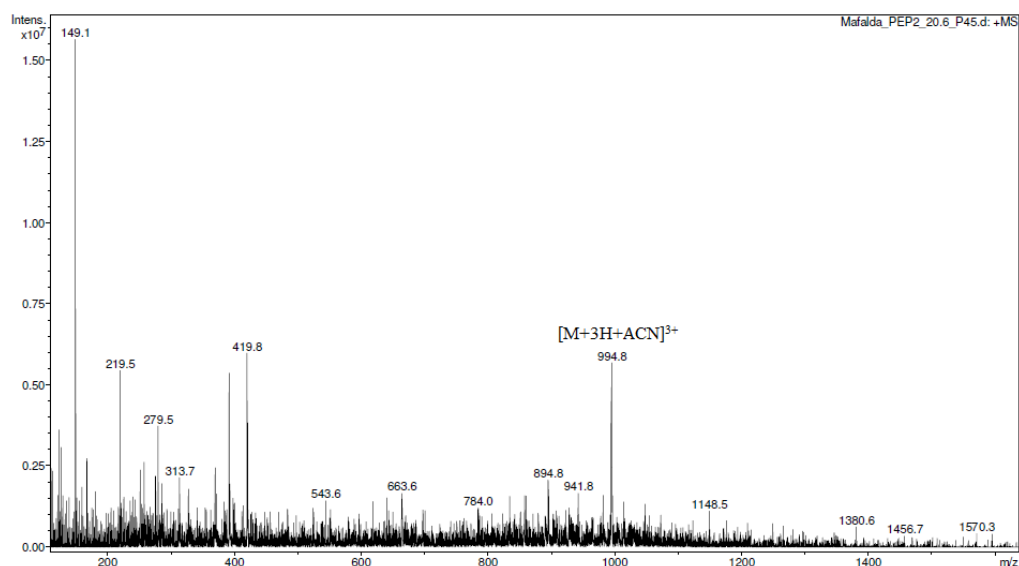


Figure 3.25: PepH2-hPEP1 mass spectrum after purification.

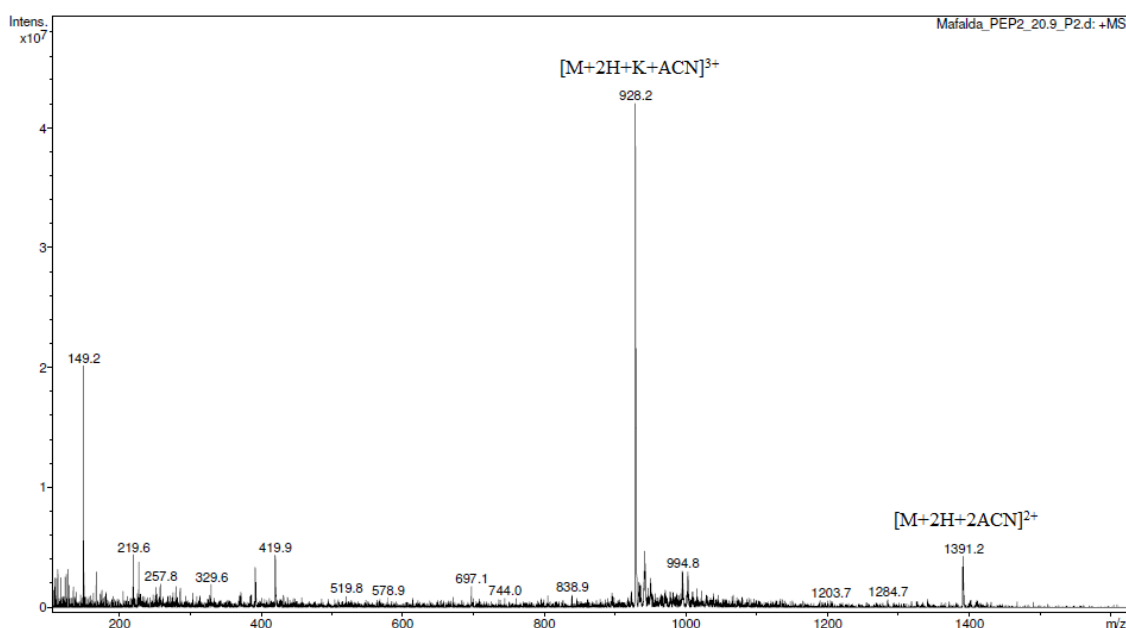


Figure 3.26: PepH2-hTRANCER1 (-V) mass spectrum after purification.

The mass spectrum of the first fraction reveals a molecular ion M/z of 994,8, which seems to be the peptide. The mass spectrum of the second fraction reveals a molecular ion M/z of 928,2 which is compatible with the mass of the peptide with less a Valine. The molecular ion M/z of 1391,2 is also compatible with the mass of the peptide with less a Valine residue. This result can be explained because during the synthesis of the peptide PepH2-hTRANCER1 by the time Valine was added to the growing peptide on the resin, we realized that piperidine was contaminated. The piperidine was purified by distillation and the synthesis continued without further problems.

Table 3.9: Peptide concentration and synthesis yield

Peptide	hTRANCE-R(1)	PepH2-hPEP1	PepH2-hPEP1 (-V)
Absorbance	0,5	0,3	0,4
Concentration (mg/ml)	0,5	0,6	0,9
Mass (mg)	0,1	0,2	0,3
Synthesis yield (%)	3,1	1,6	2,6

3.6.1 Circular Dichroism

Solutions of the three peptides hTRANCE-R(1), PepH2-hPEP1 and PepH2-hPEP1(-V), with a dilution factor of 2 were analyzed by CD. The results are presented in Table 3.10 and Figure 3.27.

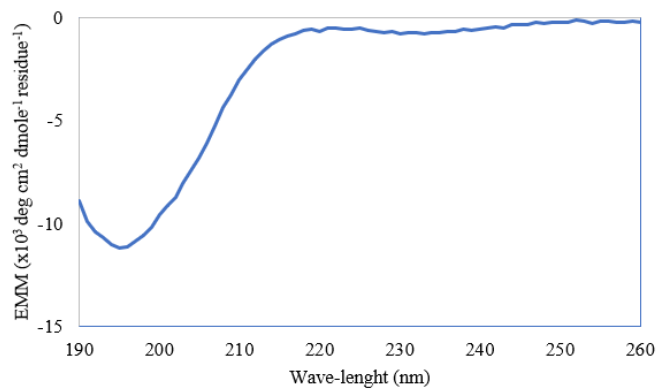
Table 3.10: Deconvolutions of the spectra obtained in CD. Fractions of the three major secondary structures

	hTRANCE-R(1)	PepH2-hPEP1	PepH2-hPEP1(-V)
α -Helix	-	-	-
β -Sheet	0,4	0,4	0,4
R.coil	0,6	0,6	0,6

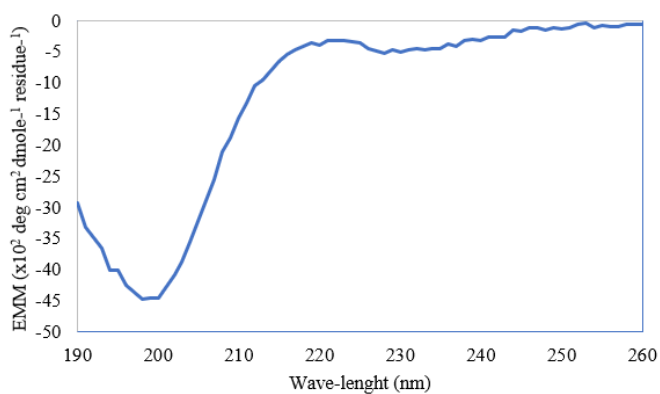
Comparing the hTRANCE-R(1) secondary structure represented in Figure 1.6, composed by a fraction of β -sheet and random coil, the CD results for this peptide were the expected.

The cell-penetrating peptide PepH2 has a helical structure [153]. However, although it was expected that the cpp PepH2 would introduce a α -helix fraction to the peptide, this fact did not occur and the peptide fractions composed by β -sheet and random coil stayed the same for PepH2-hPEP1 and PepH2-hPEP1(-V). One explanation for this unexpected conformation could be the fact that the cpp PepH2 linked to the peptide hTRANCE-R(1) changes conformation because of the peptide size. In this sense, there are factors that influence the conformational equilibria of peptide chains, such as:

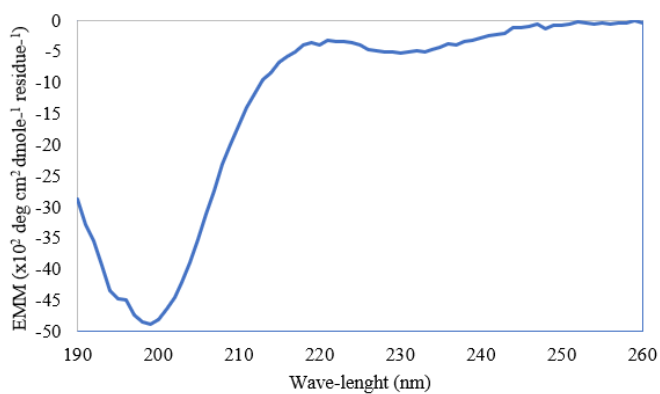
- The planarity of peptide bonds, being the conformations defined by dihedral angles;
- Steric crowding of neighboring groups;
- Repulsion and attraction of charged groups.



a hTRANSE-R(1)



b PepH2-hPEP1



c PepH2-hPEP1 (-V)

Figure 3.27: CD spectra for the three peptides synthesized.

CONCLUSIONS AND PERSPECTIVES

Sequence alignment allowed inferring homology (common ancestry) between mTRAF6 and hTRAF6. The predicted mTRAF6 tertiary structure was successfully predicted, once its structure fits hTRAF6.

The computational Alanine scanning has proved to be very useful and reliable when using different web-servers with different characteristics to predict hot spots and binding residues, as all the hot spots and important residues involved in the TRAF6-RANK (PDB:1LB5) interaction are in agreement with the literature. Moreover, the residues predicted were conserved both in hTRAF6 and in mTRAF6. These residues are active, which means that must be present in the interface and mostly contribute to the binding affinity. Molecular docking simulations between protein-peptide complexes were fulfilled, by firstly identifying the peptide binding site on a protein and further determining the native pose of the peptide. Thus, it is safe to use the docked complexes for the study of the interaction of TRAF6 with the peptides of interest, in order to validate by *in silico* studies the stability of the complexes formed.

The buried surface area (BSA) of the complexes was used to assess the stability of TRAF6 with the different peptides studied. The results obtained suggest that the peptide with the sequence mTRANCE-R(1) has a higher affinity to mTRAF6 than the peptide with the sequence mTRANCE-R(2), once mTRANCE-R(1) has a higher BSA value than mTRANCE-R(2), suggesting that it is more stable in complex with mTRAF6. This result is in agreement with experimental data.

The stability of TRAF6 with the peptide containing the sequence hTRANCE-R(1) and the peptide containing the sequence hTRANCE-R(2) was also studied, being the results the same as above, in which hTRANCE-R(1) has a higher BSA value than hTRANCE-R(2). Although there is no experimental data regarding TRAF6 in complex with these peptides, ITC measurements of mTRAF6-mTRANCE-R(1) and mTRAF6-mTRANCE-R(2)

showed that peptides with acidic residues at P_1 and P_2 positions display higher affinity to TRAF6 owing to their complementarity to the basic TRAF6 surface in that region, which is the case of mTRANCE-R(1) and hTRANCE-R(1) that present a acidic residues in those positions [19].

The complex hTRAF6-PepH2-hPEP1 is the one with the highest BSA value, suggesting that when hTRANCE-R(1) is conjugated to the cpp PepH2 instead of the cpp from the Kaposi fibroblast growth factor signal, its affinity towards hTRAF6 increases and the complex is much more stable. It is expected that the K_d of the peptide PepH2-hPEP1 with hTRAF6 is lower than the K_d of hPEP1, indicating that as far as peptide affinity, it is better to use the cpp PepH3 when the goal is to dock the peptide with hTRAF6.

The fact that the BSA values are all in the same range - standard-size interfaces with $1600\text{\AA}^2 \pm 400\text{\AA}^2$ - not existing a large disparity in interface areas, it is a factor that helps validate the quality of the molecular docking and the confidence in these results.

Using this method, the goal of validating by in silico studies the stability of TRAF6 with a peptide containing mTRANCE-R(1) and a peptide containing mTRANCE-R(2) was completed and successful.

This computational approach seems a good methodology to predict the stability of protein-peptide complexes. With this methodology is possible to do a pre-screening of the most potential compounds to be synthesized, avoiding time- and resource-consuming methodologies.

The evaluation of the complexes obtained from molecular docking allowed to conclude that structures ranked at the top are generally good fitted, thus being the HADDOCK score a good parameter to study TRAF6 interactions with the peptides of interest. In this sense, the other method used to access the stability of TRAF6 with the different peptides was the optimized HADDOCK score as a tool to predict binding affinity. However, although it was possible to make a correlation between the HADDOCK score and the $\log(K_{pred})$, indicating that the lower the HADDOCK score the more affinity the peptide has to the protein, the results are not in agreement with the literature and the ones determined by BSA studies. The complex hTRAF6-hPEP1 has a K_d higher than the complex hTRAF6-hPEP2, and the complex mTRAF6-mPEP1 has also a higher K_d than the complex mTRAF6-mPEP2, indicating that peptides hPEP2 and mPEP2 have more affinity to hTRAF6 and mTRAF6, respectively. For that reason, this method seems to not validate the stability of TRAF6 with a peptide containing mTRANCE-R(1) and a peptide containing mTRANCE-R(2).

The other goal of this work was to synthesized the peptides hTRANCE-R(1) and PepH2-hPEP1. The peptide synthesis was successful once it was verified by ESI-MS that the peptides were in fact synthesized, meaning that the synthesis steps were followed and well done, and its control was also strict. However, the synthesis produced a great amount of impurities along with the peptides, that resulted from deletion peptides (PepH2-hPEP1(-V)), incompletely deprotected peptides, modified peptides, scavengers and by-products. This resulted in low peptide concentrations and synthesis yields. As

expected, it is more difficult to perform manual synthesis of longer peptides.

The peptides were obtained with a purity > 95% which is adequate for further in vitro and in vivo studies.

The CD hTRANCE-R(1) results showed the expected secondary structure.

Supposedly, the cpp PepH2 would introduce a helix- α fraction to the peptides, but this did not occur and the peptide fractions composed by sheet- β and random coil stayed the same for hTRANCE-R(1), PepH2-hPEP1 and PepH2-hPEP1(-V).

In the near future we intend to evaluate the ability of the synthesized peptides to interact with TRAF6 and, consequently, inhibit NF- κ B activation in various cancer cell lines. In this way, it would be possible to propose the PPI RANK-TRAF6 as a druggable target to be exploited in the future for cancer therapy.

BIBLIOGRAPHY

- [1] A. M.J. R. Bonvin. *HADDOCKing of the p53 N-terminal peptide to MDM2*. 2018. URL: <http://www.bonvinlab.org/education/molmod/docking/> (visited on 06/07/2018).
- [2] R. G. Baker, M. S. Hayden, and S. Ghosh. "NF- κ B, inflammation, and metabolic disease." In: *Cell Metabolism* 13.1 (2011), pp. 11–22.
- [3] M. S. Hayden and S. Ghosh. "Shared Principles in NF- κ B Signaling." In: *Cell* 132.3 (2008), pp. 344–362.
- [4] Boston University. *NF- κ B Transcription Factors*. URL: <http://www.bu.edu/nf-kb/physiological-mediators/diseases/> (visited on 06/04/2018).
- [5] P. J.V.A. M. Converse. *TUMOR NECROSIS FACTOR; TNF*. 2016. URL: <https://www.omim.org/entry/191160> (visited on 06/06/2018).
- [6] G. E.V.A. M. Tiller. *TNF RECEPTOR-ASSOCIATED FACTOR 6; TRAF6*. 2017. URL: <https://www.omim.org/entry/602355> (visited on 06/06/2018).
- [7] D. De Stefano. "Oligonucleotides decoy to NF-kappaB: becoming a reality?" In: *Discovery medicine* 12.63 (2011), pp. 97–105.
- [8] D. M. Anderson, E. Maraskovsky, W. L. Billingsley, W. C. Dougall, M. E. Tometsko, E. R. Roux, M. C. Teepe, R. F. DuBose, D. Cosman, and L. Galibert. "A homologue of the TNF receptor and its ligand enhance T-cell growth and dendritic-cell function." In: *Nature* 390.6656 (1997), pp. 175–179.
- [9] B. G. Darnay, V Haridas, J Ni, P. A. Moore, and B. B. Aggarwal. "Characterization of the intracellular domain of receptor activator of NF- κ B (RANK)." In: *Journal of Biological Chemistry* 273.32 (1998), pp. 20551–20555.
- [10] Z. Cao, J. Xiong, M. Takeuchi, T. Kurama, and D. V. Goeddel. "TRAF6 is a signal transducer for interleukin-1." In: *Nature* 383.6599 (1996), pp. 443–446.
- [11] L. V. Pham, H.-j. Zhou, A. T. Tamayo, L. C. Yoshimura, L. Fu, B. G. Darnay, and R. J. Ford. "Nuclear Tumor Necrosis Factor Receptor-associated Factor 6 in Lymphoid Cells Negatively Regulates c-Myb-mediated Transactivation through Small Ubiquitin-related Modifier-1 Modification." In: *Journal of Biological Chemistry* 283.8 (2008), pp. 5081–5089.

- [12] T. Wada, T. Nakashima, N. Hiroshi, and J. M. Penninger. "RANKL-RANK signaling in osteoclastogenesis and bone disease." In: *Trends in Molecular Medicine* 12.1 (2006), pp. 17–25.
- [13] S. McManus and S. Roux. "The adaptor protein p62/SQSTM1 in osteoclast signaling pathways." In: *Journal of Molecular Signaling* 7.1 (2012), p. 1.
- [14] L. Deng, C. Wang, E. Spencer, L. Yang, A. Braun, J. You, C. Slaughter, C. Pickart, and Z. J. Chen. "Activation of the I κ B Kinase Complex by TRAF6 Requires a Dimeric Ubiquitin-Conjugating Enzyme Complex and a Unique Polyubiquitin Chain." In: *Cell* 103.2 (2000), pp. 351–361.
- [15] A. Leibbrandt and J. M. Penninger. "RANK/RANKL: Regulators of immune responses and bone physiology." In: *Annals of the New York Academy of Sciences* 1143.1 (2008), pp. 123–150.
- [16] S. Kohli and V. Kohli. "Role of RANKL-RANK/osteoprotegerin molecular complex in bone remodeling and its immunopathologic implications." In: *Indian Journal of Endocrinology and Metabolism* 15.3 (2011), p. 175.
- [17] A. P. Armstrong, M. E. Tometsko, M. Glaccum, C. L. Sutherland, D. Cosman, and W. C. Dougall. "A RANK/TRAF6-dependent signal transduction pathway is essential for osteoclast cytoskeletal organization and resorptive function." In: *Journal of Biological Chemistry* 277.46 (2002), pp. 44347–44356.
- [18] M. A. Lomaga, W. C. Yeh, I. Sarosi, G. S. Duncan, C. Furlonger, A. Ho, S. Morony, C. Capparelli, G. Van, S. Kaufman, A. Van Der Heiden, A. Itie, A. Wakeham, W. Khoo, T. Sasaki, Z. Cao, J. M. Penninger, C. J. Paige, D. L. Lacey, C. R. Dunstan, W. J. Boyle, D. V. Goeddel, and T. W. Mak. "TRAF6 deficiency results in osteopetrosis and defective interleukin-1, CD40, and LPS signaling." In: *Genes and Development* 13.8 (1999), pp. 1015–1024.
- [19] H. Ye, J. R. Arron, B. Lamothe, M. Cirilli, T. Kobayashi, N. K. Shevde, D. Segal, O. K. Dzivenu, M. Vologodskaya, M. Yim, K. Du, S. Singh, J. W. Pike, B. G. Darnay, Y. Choi, and H. Wu. "Distinct molecular mechanism for initiating TRAF6 signalling." In: *Nature* 418.6896 (2002), pp. 443–447.
- [20] H. Wu and J. R. Arron. "TRAF6, a molecular bridge spanning adaptive immunity, innate immunity and osteoimmunology." In: *BioEssays* 25.11 (2003), pp. 1096–1105.
- [21] E. Tsuda, M. Goto, S. I. Mochizuki, K. Yano, F. Kobayashi, T. Morinaga, and K. Higashio. "Isolation of a novel cytokine from human fibroblasts that specifically inhibits osteoclastogenesis." In: *Biochemical and Biophysical Research Communications* 234.1 (1997), pp. 137–142.
- [22] V. Sigl and J. M. Penninger. "RANKL/RANK - From bone physiology to breast cancer." In: *Cytokine and Growth Factor Reviews* 25.2 (2014), pp. 205–214.

-
- [23] S. Rao, S. J. Cronin, V. Sigl, and J. M. Penninger. "RANKL and RANK: From Mammalian Physiology to Cancer Treatment." In: *Trends in Cell Biology* 28.3 (2017), pp. 213–223.
- [24] P. C. Cogswell, D. C. Guttridge, W. K. Funkhouser, and A. S. Baldwin Jr. "Selective activation of NF-kappa B subunits in human breast cancer: potential roles for NF-kappa B2/p52 and for Bcl-3." In: *Oncogene* 19.9 (2000), pp. 1123–1131.
- [25] M. A. Sovak, R. E. Bellas, D. W. Kim, G. J. Zanieski, A. E. Rogers, and A. M. Traish. "Aberrant Nuclear Factor-kB / Rel Expression and the Pathogenesis of Breast Cancer." In: *Journal of Clinical Investigation* 100.12 (1997), pp. 2952–2960.
- [26] H Nakshatri, P Bhat-Nakshatri, D. a. Martin, R. J. Goulet, and G. W. Sledge. "Constitutive activation of NF-kappa B during progression of breast cancer to hormone-independent growth." In: *Mol Cell Biol* 17.7 (1998), pp. 3629–3639.
- [27] N.-K. Lee and S.-Y. Lee. "Modulation of Life and Death by the Tumor Necrosis Factor Receptor-Associated Factors (TRAFs)." In: *BMB Reports* 35.1 (2002), pp. 61–66.
- [28] M. Shackleton, F. Vaillant, K. J. Simpson, J. Stingl, G. K. Smyth, M. L. Asselin-Labat, L. Wu, G. J. Lindeman, and J. E. Visvader. "Generation of a functional mammary gland from a single stem cell." In: *Nature* 439.7072 (2006), pp. 84–88.
- [29] P. S. "Distribution of secondary growths in cancer of the breast." In: *Cancer Microenvironment* 7.3 (2014), pp. 91–93.
- [30] D. H. Jones, T. Nakashima, O. H. Sanchez, I. Kozieradzki, S. V. Komarova, I. Sarosi, S. Morony, E. Rubin, R. Sarao, C. V. Hojilla, V. Komnenovic, Y. Y. Kong, M. Schreiber, S. J. Dixon, S. M. Sims, R. Khokha, T. Wada, and J. M. Penninger. "Regulation of cancer cell migration and bone metastasis by RANKL." In: *Nature* 440.7084 (2006), pp. 692–696.
- [31] J. Berry. *Medical News Today- Worldwide statistics on breast cancer*. 2017. URL: <https://www.medicalnewstoday.com/articles/317135.php> (visited on 05/07/2018).
- [32] G. Christofori. "New signals from the invasive front." In: *Nature* 441.7092 (2006), pp. 444–450.
- [33] J. R. Bradley and J. S. Pober. "Tumor necrosis factor receptor-associated factors (TRAFs)." In: *Oncogene* 20.44 (2001), pp. 6482–91.
- [34] L. G. Xu, L. Y. Li, and H. B. Shu. "TRAF7 Potentiates MEKK3-induced AP1 and CHOP Activation and Induces Apoptosis." In: *Journal of Biological Chemistry* 279.17 (2004), pp. 17278–17282.
- [35] H. Ye, Y. C. Park, M. Kreishman, E. Kieff, and H. Wu. "The Structural Basis for the Recognition of Diverse Receptor Sequences by TRAF2." In: 4 (1999), pp. 321–330.

- [36] C. Biology and B. I. Pharmaceuticals. “Crystallographic analysis of CD40 recognition and signaling by human TRAF2.” In: *Proceedings of the National Academy of Sciences* 96.July (1999), pp. 8408–8413.
- [37] Uniprot. *UniProtKB - Q9Y4K3 (TRAF6_HUMAN)*. 2018. URL: <http://www.uniprot.org/uniprot/Q9Y4K3> (visited on 08/09/2018).
- [38] T Ishida, S Mizushima, S Azuma, N Kobayashi, T Tojo, K Suzuki, S Aizawa, T Watanabe, G Mosialos, E Kieff, T Yamamoto, and J Inoue. “Identification of TRAF6, a novel tumor necrosis factor receptor- associated factor protein that mediates signaling from an amino- terminal domain of the CD40 cytoplasmic region.” In: *The Journal of Biological Chemistry* 271.15 (1996), pp. 28745–28748.
- [39] W. E. Miller, J. L. Cheshire, and N Raab-Traub. “Interaction of tumor necrosis factor receptor-associated factor signaling proteins with the latent membrane protein 1 PXQXT motif is essential for induction of epidermal growth factor receptor expression.” In: *Molecular and cellular biology* 18.5 (1998), pp. 2835–2844.
- [40] T. Ishida, T. Tojo, T. Aoki, N. Kobayashi, T. Ohishi, T. Watanabe, T. Yamamoto, and J.-I. Inoue. “TRAF5, a novel tumor necrosis factor receptor-associated factor family protein, mediates CD40 signaling.” In: *Proceedings of the National Academy of Sciences of the United States of America* 93.18 (1996), pp. 9437–9442.
- [41] H. Ye, H., Arron, J.R., Lamothe, B., Cirilli, M., Kobayashi, T., Shevde, N.K., Segal, D., Dzivenu, O., Vologodskaya, M., Yim, M., Du, K., Singh, S., Pike, J.W., Darnay, B.G., Choi, Y., Wu. *TRAF6-RANK Complex: PDB 1LB5*. 2002. URL: <http://www.rcsb.org/structure/1LB5> (visited on 09/04/2018).
- [42] A. Liu, P. Gong, S. W. Hyun, K. Z. Wang, E. A. Cates, D. Perkins, D. D. Bannerman, A. C. Puché, V. Y. Toshchakov, S. Fang, P. E. Auron, S. N. Vogel, and S. E. Goldblum. “TRAF6 protein couples toll-like receptor 4 signaling to Src family kinase activation and opening of paracellular pathway in human lung microvascular endothelia.” In: *Journal of Biological Chemistry* 287.20 (2012), pp. 16132–16145.
- [43] H. Zhang, C. Wu, L. E. Matesic, X. Li, Z. Wang, B. F. Boyce, and L. Xing. “Ubiquitin E3 ligase itch negatively regulates osteoclast formation by promoting deubiquitination of tumor necrosis factor (TNF) receptor-associated factor 6.” In: *Journal of Biological Chemistry* 288.31 (2013), pp. 22359–22368.
- [44] M. Vidal, M. E. Cusick, A.-L. Barabási, I. Networks, and H. Disease. “Interactome networks and human disease.” In: *Cell* 144.6 (2011), pp. 986–998.
- [45] P. Vlieghe, V. Lisowski, J. Martinez, and M. Khrestchatisky. *Synthetic therapeutic peptides: science and market*. 2010.

-
- [46] A. C. Bharti, Y. Takada, and B. B. Shishodia Shishir Aggarwal. "Evidence That Receptor Activator of Nuclear Factor (NF) - kB Ligand Can Suppress Cell Proliferation and Induce Apoptosis through Activation of a NF-kB-independent and TRAF6-dependent Mechanism." In: *Journal of Biological Chemistry* 279.7 (2004), pp. 6065–6076.
- [47] A. T. Poblenz, J. J. Jacoby, S. Singh, and B. G. Darnay. "Inhibition of RANKL-mediated osteoclast differentiation by selective TRAF6 decoy peptides." In: *Biochemical and Biophysical Research Communications* 359.3 (2007), pp. 510–515.
- [48] H. Chen, M. Li, E. Sanchez, C. S. Wang, T. Lee, C. M. Soof, C. E. Casas, J. Cao, C. Xie, K. A. Udd, K. Decorso, G. Y. Tang, T. M. Spektor, and J. R. Berenson. "Combined TRAF6 Targeting and Proteasome Blockade Has Anti-myeloma and Anti-Bone Resorptive Effects." In: *Mol Cancer Res.* 15.5 (2017), pp. 598–610.
- [49] Gavin AC1, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M. "Functional organization of the yeast proteome by systematic analysis of protein complexes." In: *Nature* 415.6868 (2002), pp. 141–147.
- [50] S. Jones and J. M. Thornton. "Principles of protein-protein interactions." In: *PNAS* 93.1 (1996), pp. 13–20.
- [51] J. Janin. "Elusive affinities." In: *Proteins: Structure, Function, and Bioinformatics* 21.1 (1995), pp. 30–39.
- [52] J. Janin. "Protein-protein recognition." In: *Progress in Biophysics and Molecular Biology* 64.2-3 (1995), pp. 145–166.
- [53] C. Chothia and A. M. Lesk. "The evolution of protein structures." In: *Cold Spring Harbor Symposia on Quantitative Biology* 52 (1987), pp. 399–405.
- [54] C. J. Tsai and R. Nussinov. "Hydrophobic folding units at protein-protein interfaces: Implications to protein folding and to protein-protein association." In: *Protein Science* 6.7 (1997), pp. 1426–1437.
- [55] E. D. Levy. "A Simple Definition of Structural Regions in Proteins and Its Use in Analyzing Interface Evolution." In: *Journal of Molecular Biology* 403.4 (2010), pp. 660–670.
- [56] M. J. Gething and J. Sambrook. "Protein folding in the cell." In: *Nature* 355.6355 (1992), pp. 33–45.
- [57] I. S. Moreira, P. A. Fernandes, and M. J. Ramos. "Protein-protein docking dealing with the unknown." In: *Journal of Computational Chemistry* 31.2 (2010), pp. 317–342.
- [58] D. A. Drummond and C. O. Wilke. "Mistranslation-Induced Protein Misfolding as a Dominant Constraint on Coding-Sequence Evolution." In: *Cell* 134.2 (2008), pp. 341–352.

- [59] M. F. Perutz, J. C. Kendrew, and H. C. Watson. "Structure and function of haemoglobin: II. Some relations between polypeptide chain configuration and amino acid sequence." In: *Journal of Molecular Biology* 13.3 (1965), pp. 669–678.
- [60] E. A. Franzosa and Y. Xia. "Structural determinants of protein evolution are context-sensitive at the residue level." In: *Molecular Biology and Evolution* 26.10 (2009), pp. 2387–2395.
- [61] G. C. Conant and P. F. Stadler. "Solvent exposure imparts similar selective pressures across a range of yeast proteins." In: *Molecular Biology and Evolution* 26.5 (2009), pp. 1155–1161.
- [62] T. Zhou, D. A. Drummond, and C. O. Wilke. "Contact density affects protein evolutionary rate from bacteria to animals." In: *Journal of Molecular Evolution* 66.4 (2008), pp. 395–404.
- [63] R. Sasidharan and C. Chothia. "The selection of acceptable protein mutations." In: *Proceedings of the National Academy of Sciences* 104.24 (2007), pp. 10080–10085.
- [64] B. Lee and F. M. Richards. "The interpretation of protein structures: Estimation of static accessibility." In: *Journal of Molecular Biology* 55.3 (1971), pp. 379–400.
- [65] R. P. Bahadur, P. Chakrabarti, F. Rodier, and J. Janin. "Dissecting Subunit Interfaces in Homodimeric Proteins." In: *Proteins: Structure, Function and Genetics* 53.3 (2003), pp. 708–719.
- [66] L. L. Conte, C. Chothia, and J. Janin. "The atomic structure of protein-protein recognition sites." In: *Journal of Molecular Biology* 285.5 (1999), pp. 2177–2198.
- [67] P. L. Kastitis and A. M.J. J. Bonvin. "On the binding affinity of macromolecular interactions : daring to ask why proteins interact on the binding affinity of macromolecular interactions : daring to ask why proteins interact." In: 10.79 (2013), pp. 1–27.
- [68] A. A. Bogan and K. S. Thorn. "Anatomy of hot spots in protein interfaces." In: *Journal of Molecular Biology* 280.1 (1998), pp. 1–9.
- [69] M. Eames and T. Kortemme. "Structural Mapping of Protein Interactions Reveals Differences in Evolutionary Pressures Correlated to mRNA Level and Protein Abundance." In: *Structure* 15.11 (2007), pp. 1442–1451.
- [70] J. Mintseris and Z. Weng. "Structure, function, and evolution of transient and obligate protein-protein interactions." In: *Proceedings of the National Academy of Sciences* 102.31 (2005), pp. 10930–10935.
- [71] M. Guharoy and P. Chakrabarti. "Conservation and relative importance of residues across protein-protein interfaces." In: *Proceedings of the National Academy of Sciences* 102.43 (2005), pp. 15447–15452.
- [72] D. R. Caffrey. "Are protein-protein interfaces more conserved in sequence than the rest of the protein surface?" In: *Protein Science* 13.1 (2004), pp. 190–202.

- [73] O. Keskin, B. Ma, and R. Nussinov. "Hot regions in protein-protein interactions: The organization and contribution of structurally conserved hot spot residues." In: *Journal of Molecular Biology* 345.5 (2005), pp. 1281–1294.
- [74] G. Schreiber and A. R. Fersht. "Energetics of protein-protein interactions: Analysis of the Barnase-Barstar interface by single mutations and double mutant cycles." In: *Journal of Molecular Biology* 248.2 (1995), pp. 478–486.
- [75] S. W. Lockless and R. Ranganathan. "Evolutionarily conserved pathways of energetic connectivity in protein families." In: *Science* 286.5438 (1999), pp. 295–299.
- [76] H. B. Fraser, A. E. Hirsh, L. M. Steinmetz, C. Scharfe, and M. W. Feldman. "Evolutionary rate in the protein interaction network." In: *Science (New York, N.Y.)* 296.5568 (2002), pp. 750–820.
- [77] B. Bergeron. *On the Horizon: Bioinformatics Computing*. Vol. 36. 5. 2002, pp. 5–14.
- [78] O. Keskin, A. Gursoy, B. Ma, and R. Nussinov. *Principles of protein-protein interactions: What are the preferred ways for proteins to interact?* 2008.
- [79] N. Horton and M. Lewis. "Calculation of the free energy of association for protein complexes." In: *Protein Science* 1.1 (1992), pp. 169–181.
- [80] P. Chakrabarti and R. Bhattacharyya. "Geometry of nonbonded interactions involving planar groups in proteins." In: *Progress in Biophysics and Molecular Biology* 95.1-3 (2007), pp. 83–137.
- [81] S. V. NARAYANA and P. ARGOS. "Residue contacts in protein structures and implications for protein folding." In: *International Journal of Peptide and Protein Research* 24.1 (1984), pp. 25–39.
- [82] K. Nishikawa and T. Ooi. "Radial locations of amino acid residues in a globular protein: Correlation with the sequence." In: *Journal of Biochemistry* 100.4 (1986), pp. 1043–1047.
- [83] I. Bahar and R. L. Jernigan. "Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation." In: *Journal of Molecular Biology* 266.1 (1997), pp. 195–214.
- [84] C. Zhang and S.-H. Kim. "Environment-dependent residue contact energies for proteins." In: *Proceedings of the National Academy of Sciences* 97.6 (2000), pp. 2550–2555.
- [85] G. R. Bickerton, A. P. Higuieruelo, and T. L. Blundell. "Comprehensive, atomic-level characterization of structurally characterized protein-protein interactions: The PICCOLO database." In: *BMC Bioinformatics* 12.303 (2011), pp. 1–15.
- [86] R. E. Hubbard and M. Kamran Haider. *Encyclopedia of Life Sciences- Hydrogen Bonds in Proteins: Role and Strength*. 2010.

- [87] “Chemical Cross-Linking: Reagents and Problems in Studies of Membrane Structure.” In: *Annual Review of Biochemistry* 46.1 (1977), pp. 523–551.
- [88] J Skolnick, J. S. Fetrow, and a Kolinski. “Structural genomics and its importance for gene function analysis.” In: *Nature biotechnology* 18.3 (2000), pp. 283–287.
- [89] J. A. Wells. “Additivity of Mutational Effects in Proteins.” In: *Biochemistry* 29.37 (1990), pp. 8509–8517.
- [90] B. C. Cunningham and J. A. Wells. “High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis.” In: *Science* 244.4908 (1989), pp. 1081–1085.
- [91] T Clackson and J. A. Wells. “A Hot-Spot of Binding-Energy in a Hormone-Receptor Interface.” In: *Science* 267.5196 (1995), pp. 383–386.
- [92] J. A. Wells. “Systematic Mutational Analyses of Protein-Protein Interfaces.” In: *Methods in Enzymology* 202.C (1991), pp. 390–411.
- [93] T Clackson, M. H. Ultsch, J. a. Wells, and a. M. de Vos. “Structural and functional analysis of the 1:1 growth hormone:receptor complex reveals the molecular basis for receptor affinity.” In: *Journal of molecular biology* 277.5 (1998), pp. 1111–1128.
- [94] Z. Hu, B. Ma, H. Wolfson, and R. Nussinov. “Conservation of polar residues as hot spots at protein interfaces.” In: *Proteins: Structure, Function and Genetics* 39.4 (2000), pp. 331–342.
- [95] J. L. K. Kouadio, J. R. Horn, G. Pal, and A. A. Kossiakoff. “Shotgun alanine scanning shows that growth hormone can bind productively to its receptor through a drastically minimized interface.” In: *Journal of Biological Chemistry* 280.27 (2005), pp. 25524–25532.
- [96] J. Kenneth Morrow and S. Zhang. “Computational Prediction of Protein Hot Spot Residues.” In: *Current Pharmaceutical Design* 18.9 (2012), pp. 1255–1265.
- [97] J. Chen, N. Sawyer, and L. Regan. “Protein-protein interactions: General trends in the relationship between binding affinity and interfacial buried surface area.” In: *Protein Science* 22.4 (2013), pp. 510–515.
- [98] S. Lee. “Characterization of Protein Interactions by ITC , SPR and BLI.” Suwon, 2016.
- [99] *Computational Chemistry*. URL: <https://www.acs.org/content/acs/en/careers/college-to-career/chemistry-careers/computational-chemistry.html> (visited on 07/08/2018).
- [100] D. Rognan. “Proteome-scale docking: Myth and reality.” In: *Drug Discovery Today: Technologies* 10.3 (2013), pp. 403–409.
- [101] *Why Structure Prediction Matters*. 2015. URL: <https://www.dnastar.com/blog/structural-biology/why-structure-prediction-matters/> (visited on 09/08/2018).

-
- [102] U. Pieper, B. M. Webb, G. Q. Dong, D. Schneidman-Duhovny, H. Fan, S. J. Kim, N. Khuri, Y. G. Spill, P. Weinkam, M. Hammel, J. A. Tainer, M. Nilges, and A. Sali. "ModBase, a database of annotated comparative protein structure models and associated resources." In: *Nucleic Acids Research* 42.1 (2014), pp. 1–6.
- [103] L. Ferreira, R. dos Santos, G. Oliva, and A. Andricopulo. *Molecular Docking and Structure-Based Drug Design Strategies*. Vol. 20. 7. 2015, pp. 13384–13421.
- [104] E. Karaca and A. M. Bonvin. "Advances in integrative modeling of biomolecular complexes." In: *Methods* 59.3 (2013), pp. 372–381.
- [105] R. Buyya, K. Branson, and J. Giddy. "The Virtual Laboratory : a toolset to enable distributed molecular modelling for drug design on the World-Wide Grid." In: *CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE* 25.March 2002 (2003), pp. 1–25.
- [106] M. Trellet, A. S. J. Melquiond, and A. M.J. J. Bonvin. "Information-driven modeling of protein-peptide complexes." In: *Computational Peptidology*. 2015, pp. 221–239.
- [107] N. Andrusier, E. Mashiach, R. Nussinov, and H. J. Wolfson. "Principles of flexible protein-protein docking." In: *Proteins: Structure, Function and Genetics* 73.2 (2008), pp. 271–289.
- [108] S. J. De Vries, M. Van Dijk, and A. M. Bonvin. "The HADDOCK web server for data-driven biomolecular docking." In: *Nature Protocols* 5.5 (2010), pp. 883–897.
- [109] E. Karaca, A. S. J. Melquiond, S. J. de Vries, P. L. Kastitis, and A. M.J. J. Bonvin. "Building Macromolecular Assemblies by Information-driven Docking." In: *Molecular & Cellular Proteomics* 9.8 (2010), pp. 1784–1794.
- [110] L. C. Xue, R. A. Jordan, E. M. Yasser, D. Dobbs, and V. Honavar. "DockRank: Ranking docked conformations using partner-specific sequence homology-based protein interface prediction." In: *Proteins: Structure, Function and Bioinformatics* 82.2 (2014), pp. 250–267.
- [111] R. Sinha, P. J. Kundrotas, and I. A. Vakser. "Docking by structural similarity at protein-protein interfaces." In: *Proteins: Structure, Function and Bioinformatics* 78.15 (2010), pp. 3235–3241.
- [112] L. C. Xue, J. P. Rodrigues, D. Dobbs, V. Honavar, and A. M. Bonvin. "Template-based protein-protein docking exploiting pairwise interfacial residue restraints." In: *Briefings in Bioinformatics* 18.3 (2017), pp. 458–466.
- [113] Alexandre M.J.J. João Rodrigues. *HADDOCK basic protein-protein docking tutorial*. 2018. URL: <http://www.bonvinlab.org/education/HADDOCK-protein-protein-basic/> (visited on 05/06/2018).

- [114] M. F. Lensink and S. J. Wodak. "Docking, scoring, and affinity prediction in CAPRI." In: *Proteins: Structure, Function and Bioinformatics* 81.12 (2013), pp. 2082–2095.
- [115] C. Dominguez, R. Boelens, and A. M. Bonvin. "HADDOCK: A protein-protein docking approach based on biochemical or biophysical information." In: *Journal of the American Chemical Society* 125.7 (2003), pp. 1731–1737.
- [116] A. Vangone, J. P. Rodrigues, L. C. Xue, G. C. van Zundert, C. Geng, Z. Kurkcuoglu, M. Nellen, S. Narasimhan, E. Karaca, M. van Dijk, A. S. Melquiond, K. M. Visscher, M. Trellet, P. L. Kastritis, and A. M. Bonvin. "Sense and simplicity in HADDOCK scoring: Lessons from CASP-CAPRI round 1." In: *Proteins: Structure, Function and Bioinformatics* 85.3 (2017), pp. 417–423.
- [117] O. Dagliyan, E. A. Proctor, K. M. D'Auria, F. Ding, and N. V. Dokholyan. "Structural and dynamic determinants of protein-peptide recognition." In: *Structure* 19.12 (2011), pp. 1837–1845.
- [118] P. L. Kastritis, J. P.G.L. M. Rodrigues, and A. M.J. J. Bonvin. "HADDOCK2P2I: A biophysical model for predicting the binding affinity of protein-protein interaction inhibitors." In: *Journal of Chemical Information and Modeling* 54.3 (2014), pp. 826–836.
- [119] J. Janin, K. Henrick, J. Moult, L. T. Eyck, M. J. Sternberg, S. Vajda, I. Vakser, and S. J. Wodak. "CAPRI: A critical assessment of PRedicted interactions." In: *Proteins: Structure, Function and Genetics* 52.1 (2003), pp. 2–9.
- [120] M. F. Lensink, S. Velankar, and S. J. Wodak. "Modeling protein–protein and protein–peptide complexes: CAPRI 6th edition." In: *Proteins: Structure, Function and Bioinformatics* 85.3 (2017), pp. 359–377.
- [121] C. J. Camacho and D. W. Gatchell. "Successful discrimination of protein interactions." In: *Proteins: Structure, Function and Genetics* 52.1 (2003), pp. 92–97.
- [122] A Quintas, A. P. Freire, and M. J. Halpern. *Bioquímica: organiza molecular da vida*. Vol. Edi Lidel, Lisboa. 7. 2008, pp. 141–159.
- [123] D. L. Nelson and M. M. Cox. *Lehninger principios de bioquímica*. Vol. 3. 7. 2002, pp. 89–122.
- [124] G. Gunawardena. *N-Terminal*. 2015. URL: <https://chem.libretexts.org/Reference/Organic\Chemistry\Glossary/N-Terminal> (visited on 05/06/2018).
- [125] V. A. Zhemkov, A. A. Kulinskaya, I. B. Bezprozvanny, and M. Kim. "The 2.2-Angstrom resolution crystal structure of the carboxy-terminal region of ataxin-3." In: *FEBS Open Bio* 6.3 (2016), pp. 168–178.
- [126] R. Behrendt, P. White, and J. Offer. "Advances in Fmoc solid-phase peptide synthesis." In: *Journal of Peptide Science* 22.1 (2016), pp. 4–27.

- [127] Bachem. "Bachem Peptide Guide." In: *BioConnect* (2016), pp. 1–60.
- [128] *Peptide Synthesis*. 1st ed. Vol. 1. MERCK, 2000. Chap. Solid Phase Peptide Synthesis, pp. 1.1–3.27.
- [129] W. C. Chan and P. D. White. *Fmoc Solid Phase Peptide Synthesis: A practical Approach*. Ed. by PAS. 1st ed. Oxford: Oxford University Press, 2000, pp. 1–288.
- [130] DMacks. *Poly(styrene-co-divinylbenzene)*. 2011. URL: [https://en.wikipedia.org/wiki/File:Poly\(styrene-co-divinylbenzene\).png](https://en.wikipedia.org/wiki/File:Poly(styrene-co-divinylbenzene).png) (visited on 09/06/2018).
- [131] "Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides." In: *Analytical Biochemistry* 34.2 (1970), pp. 595–598.
- [132] S.-A. Co LLC. "60017 Kaiser test kit." In: *Analytical Biochemistry Analytical Biochemistry Peptide Research* 34.4 (1970), pp. 595–147.
- [133] G. T. G. Reis. *Trabalho Ninhidrina*. URL: <http://www.ebah.pt/content/ABAAABWUwAD/trabalho-ninhidrina> (visited on 07/05/2018).
- [134] R. W. Woody. "Circular dichroism." In: *Methods of Enzymology*. 1995, pp. 34–71.
- [135] N. Greenfield and G. D. Fasman. "Computed Circular Dichroism Spectra for the Evaluation of Protein Conformation." In: *Biochemistry* 8.10 (1969), pp. 4108–4116.
- [136] T. Kortemme and D. Baker. "A simple physical model for binding energy hot spots in protein-protein complexes." In: *Proceedings of the National Academy of Sciences* 99.22 (2002), pp. 14116–14121.
- [137] X. Zhu and J. C. Mitchell. "KFC2: A knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features." In: *Proteins: Structure, Function and Bioinformatics* 79.9 (2011), pp. 2671–2683.
- [138] N. Tuncbag, O. Keskin, and A. Gursoy. "HotPoint: Hot spot prediction server for protein interfaces." In: *Nucleic Acids Research* 38.SUPPL. 2 (2010), pp. 1–5.
- [139] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang. "The I-TASSER suite: Protein structure and function prediction." In: *Nature Methods* 12.1 (2014), pp. 7–8.
- [140] A. Roy, A. Kucukural, and Y. Zhang. "I-TASSER: A unified platform for automated protein structure and function prediction." In: *Nature Protocols* 5.4 (2010), pp. 725–738.
- [141] A. Lamiable, P. Thévenet, J. Rey, M. Vavrusa, P. Derreumaux, and P. Tufféry. "PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex." In: *Nucleic Acids Research* 44.1 (2016), pp. 449–454.
- [142] A. C. Camproux, R. Gautier, and P. Tufféry. "A hidden Markov model derived structural alphabet for proteins." In: *Journal of Molecular Biology* 339.3 (2004), pp. 591–605.

- [143] P. Thevenet, Y. Shen, J. Maupetit, F. Guyon, P. Derreumaux, and P. Tuffery. "PEP-FOLD: An updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides." In: *Nucleic Acids Research* 40.1 (2012), pp. 288–293.
- [144] S. Singh, H. Singh, A. Tuknait, K. Chaudhary, B. Singh, S. Kumaran, and G. P. Raghava. "PEPstrMOD: Structure prediction of peptides containing natural, non-natural and modified residues." In: *Biology Direct* 10.1 (2015), pp. 1–19.
- [145] Harpreet Kaur, Aarti Garg, and G.P.S. Raghava. "PEPstr: A de novo Method for Tertiary Structure Prediction of Small Bioactive Peptides." In: *Protein & Peptide Letters* 14.7 (2007), pp. 626–631.
- [146] M. J. Sippl. "Recognition of errors in three-dimensional structures of proteins." In: *Proteins: Structure, Function, and Bioinformatics* 17.4 (1993), pp. 355–362.
- [147] M. Wiederstein and M. J. Sippl. "ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins." In: *Nucleic Acids Research* 35.SUPPL.2 (2007), pp. 407–410.
- [148] L. C. Xue and A. Vangone. "Performance of HADDOCK and a simple contact-based protein – ligand binding affinity predictor in the D3R Grand Challenge 2." In: *Journal of Computer-Aided Molecular Design* 32.1 (2018), pp. 175–185.
- [149] R. Wang, X. Fang, Y. Lu, and S. Wang. "The PDBbind database: Collection of binding affinities for protein-ligand complexes with known three-dimensional structures." In: *Journal of Medicinal Chemistry* 47.12 (2004), pp. 2977–2980.
- [150] F. Albericio and S. A. Kates. *Solid-Phase Synthesis*. Ed. by C. Press. 1st ed. New York, 2000.
- [151] TkinkPeptides. *Measuring Peptide Concentration*. URL: <http://www.thinkpeptides.com/peptidemeasuring.html> (visited on 08/06/2018).
- [152] TkinkPeptides. *Measuring Peptide Concentration*. URL: <http://www.thinkpeptides.com/peptidemeasuring.html> (visited on 07/08/2018).
- [153] V. Neves, M. M. Frederico Aires-da-Silva, L. Gano, E. Ribeiro, A. Pinto, S. Aguiar, D. Gaspar, C. Fernandes, J. D. G. Correia, and M. A.R. B. Castanho. "Novel Peptides Derived from Dengue Virus Capsid Protein Translocate Reversibly the Blood Brain Barrier through a Receptor- Free Mechanism." In: *ACS Chemical Biology* 12 (2017), pp. 1257–1268.



Annex 1

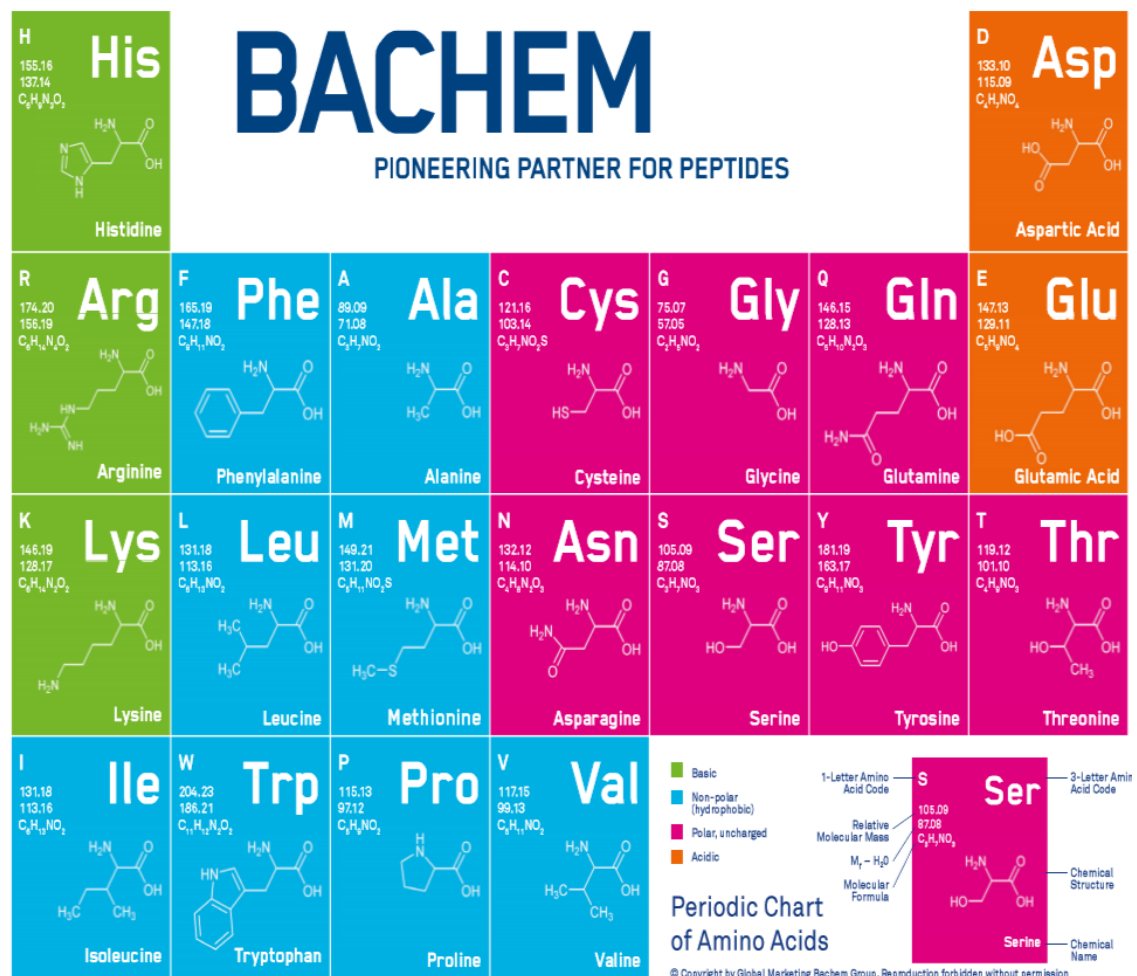


Figure A.1: Amino acids periodic chart [127].

Annex 2

Virtual alanine scanning with the web-server Robetta for hot-spot and binding residue prediction for hTRAF6 (374-474) and RANK (602-608).

```
Mon May 28 08:56:07 PDT 2018
virtual alanine scanning, TRAF6-RANK.alascan
```

pdb#	chain	int_id	res#	aa	DDG (complex)	DDG (complex, obs)	DG (partner)
374	A	1	28	18	0.36	0.00	1.04
376	A	1	30	7	0.44	0.00	0.32
392	A	1	46	15	3.76	0.00	1.51
410	A	1	64	5	0.76	0.00	3.21
412	A	1	66	7	0.40	0.00	2.61
448	A	0	102	4	-0.03	0.00	-0.12
450	A	1	104	11	0.13	0.00	2.38
453	A	0	107	9	-0.12	0.00	0.08
456	A	1	110	10	0.48	0.00	1.34
466	A	1	120	15	0.07	0.00	0.94
469	A	1	123	9	0.62	0.00	-0.14
471	A	1	125	5	2.05	0.00	1.65
473	A	1	127	20	1.73	0.00	1.80
474	A	0	128	18	0.07	0.00	0.43
602	B	0	157	11	0.04	0.00	0.12
604	B	1	159	17	0.01	0.00	0.67
605	B	1	160	4	1.63	0.00	-0.34
606	B	1	161	3	2.85	0.00	-0.88
607	B	1	162	4	0.61	0.00	-0.35
608	B	1	163	20	1.78	0.00	-0.69

Figure A.2: Virtual alanine scanning from Robetta with PDB 1LB5 for hTRAF6 (374-474) and RANK (602-608).

Annex 3

Virtual alanine scanning with the web-server KFC2 for hot-spot and binding residue prediction for hTRAF6 (374-475) and RANK (601-608).

KFC2 Hot Spot Prediction Server @mitchell-lab.org from Mon, 23 Apr 2018 08:00:40 CDT
JobId: 40f43ca47187e70b3f8e45c43d186732 JobName: TRAF6-RANK ChainSet1: A ChainSet2: B

Chain	Res	Num	KFC2-A Class	KFC2-A Conf	KFC2-B Class	KFC2-B Conf	ConSurf Class	ConSu Value	Rosetta Class	Roset DDG	Exper Class	Exper Value
A	VAL	374	-----	-1.88	-----	-0.99	-----	---	-----	---	-----	---
A	HIS	376	-----	-2.43	-----	-0.98	-----	---	-----	---	-----	---
A	ARG	392	Hotspot	0.30	Hotspot	0.16	-----	---	-----	---	-----	---
A	HIS	394	-----	-2.00	-----	-0.95	-----	---	-----	---	-----	---
A	PHE	410	Hotspot	0.24	Hotspot	0.25	-----	---	-----	---	-----	---
A	HIS	412	Hotspot	0.02	Hotspot	0.07	-----	---	-----	---	-----	---
A	GLU	448	-----	-2.78	-----	-0.99	-----	---	-----	---	-----	---
A	ILE	449	-----	-2.39	-----	-0.92	-----	---	-----	---	-----	---
A	MET	450	-----	-1.37	-----	-0.25	-----	---	-----	---	-----	---
A	ASP	451	-----	-2.86	-----	-0.96	-----	---	-----	---	-----	---
A	LYS	453	-----	-1.87	-----	-0.96	-----	---	-----	---	-----	---
A	LEU	456	-----	-0.99	-----	-0.82	-----	---	-----	---	-----	---
A	LEU	457	-----	-2.84	-----	-0.94	-----	---	-----	---	-----	---
A	ALA	458	-----	-0.33	-----	-0.76	-----	---	-----	---	-----	---
A	ARG	466	-----	-1.90	-----	-0.79	-----	---	-----	---	-----	---
A	PRO	468	-----	-0.50	-----	-0.80	-----	---	-----	---	-----	---
A	LYS	469	-----	-0.40	-----	-0.55	-----	---	-----	---	-----	---
A	GLY	470	-----	-0.22	-----	-0.46	-----	---	-----	---	-----	---
A	PHE	471	Hotspot	0.63	Hotspot	0.30	-----	---	-----	---	-----	---
A	GLY	472	-----	-0.38	-----	-0.59	-----	---	-----	---	-----	---
A	TYR	473	-----	-0.52	-----	-0.03	-----	---	-----	---	-----	---
A	VAL	474	-----	-1.01	-----	-0.96	-----	---	-----	---	-----	---
A	THR	475	-----	-2.00	-----	-0.99	-----	---	-----	---	-----	---
B	GLN	601	-----	-2.02	-----	-1.00	-----	---	-----	---	-----	---
B	MET	602	-----	-2.98	-----	-0.93	-----	---	-----	---	-----	---
B	PRO	603	Hotspot	0.84	-----	-0.47	-----	---	-----	---	-----	---
B	THR	604	-----	-1.04	-----	-0.83	-----	---	-----	---	-----	---
B	GLU	605	Hotspot	0.69	-----	-0.35	-----	---	-----	---	-----	---
B	ASP	606	Hotspot	0.88	-----	-0.42	-----	---	-----	---	-----	---
B	GLU	607	-----	-1.21	-----	-0.80	-----	---	-----	---	-----	---
B	TYR	608	Hotspot	0.23	-----	-0.43	-----	---	-----	---	-----	---

Figure A.3: Virtual alanine scanning from KFC2 with PDB 1LB5 for hTRAF6 (374-475) and RANK (601-608).

Annex 4

Virtual alanine scanning with the web-server Hotpoit for hot-spot and binding residue prediction for hTRAF6 (374-473) and RANK (603-608).

Residue Number	Residue Name	Chain	RelCompASA	RelMonomerASA	Potential	Prediction
374	V	A	36.17	50.36	12.38	NH
376	H	A	29.78	35.89	18.13	NH
392	R	A	1.34	22.39	25.64	H
410	F	A	1.94	11.86	28.55	H
450	M	A	1.76	6.09	33.49	H
456	L	A	21.77	35.29	11.00	NH
457	L	A	61.09	62.37	7.59	NH
458	A	A	2.41	5.39	13.80	NH
468	P	A	37.32	74.16	4.79	NH
469	K	A	25.35	53.96	8.90	NH
470	G	A	0.19	29.14	12.76	NH
471	F	A	1.22	22.90	37.55	H
472	G	A	0.45	35.78	16.16	NH
473	Y	A	7.52	27.91	17.22	NH
603	P	B	0.00	86.45	11.00	NH
604	T	B	52.75	87.17	12.38	NH
605	E	B	33.29	102.80	15.80	NH
606	D	B	27.00	88.60	13.65	NH
607	E	B	65.22	98.29	4.53	NH
608	Y	B	59.49	126.14	20.74	NH

Figure A.4: Virtual alanine scanning from KFC2 with PDB 1LB5 for hTRAF6 (374-473) and RANK (603-608).

Annex 5

Virtual alanine scanning with the web-servers Robetta, KFC2 and Hotpoint for hot-spot and binding residue prediction for RANK, in complex with hTRAF6. It was used the PDB 1LB5.

Is mentioned in literature that the residues 603, 605 and 608 are hot spots.

Table A.1: RANK hot spots (grey) and binding residues in complex with hTRAF6 (PDB:1LB5) from Rosetta, KFC2 and Hotpoint

RANK's fragment	Robetta	KFC2a	KFC2b	Hotpoint
Residues	-	601	-	-
	602	602	-	-
	-	603	-	-
	604	604	-	-
	605	605	-	-
	606	606	-	-
	607	607	-	-
	608	608	-	-

Annex 6

Distance of hTRAF6 residues to RANK's fragment, in the PDB 1LB5. The active residues have a 6Å cut-off.

Table A.2: hTRAF6 interface active residues for the PDB 1LB5. Visualization with *Pymol*

hTRAF6				
3A	4A	5A	6A	8A
392	410	374	449	373
458	412	376	451	375
468	450	394	453	393
470	456	448	459	396
472	457	468		402
	466	474		406
	469	475		408
	471			409
	468			411
	469			430
	473			432
				460
				464
				467

Annex 7

ProSA was used for hPEP1 structure evaluation. It is represented the overall quality z-scores calculated by ProSA for each one of the 5 structures predicted by I-TASSER, along with the plot with the z-scores of all experimentally determined structures that are in the PDB database. The z-scores of all predicted peptide structures are within the range of scores typically found for native proteins with sequence length of 25 amino acids, suggesting that the structures predicted by I-TASSER probably do not contains errors and are going to be used in the ensemble molecular docking.

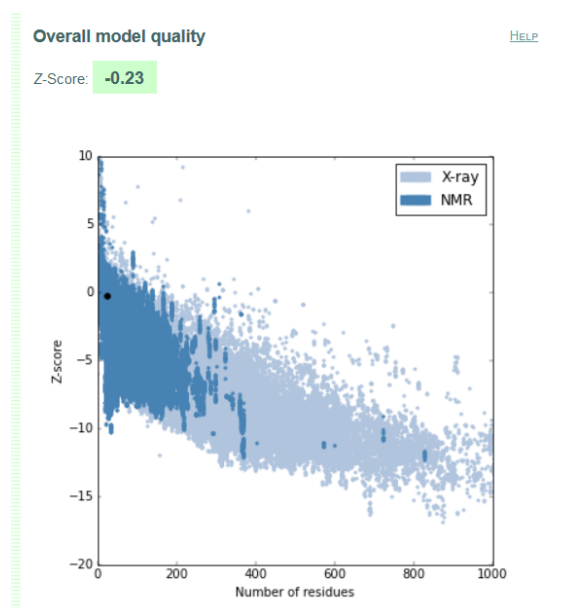


Figure A.5: I-TASSER Model 1 quality score calculated by ProSA for peptide hPEP1.

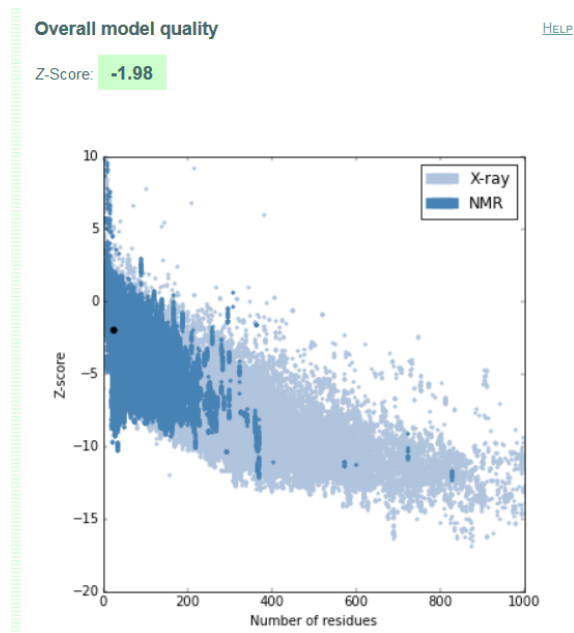


Figure A.6: I-TASSER Model 2 quality score calculated by ProSA for peptide hPEP1.

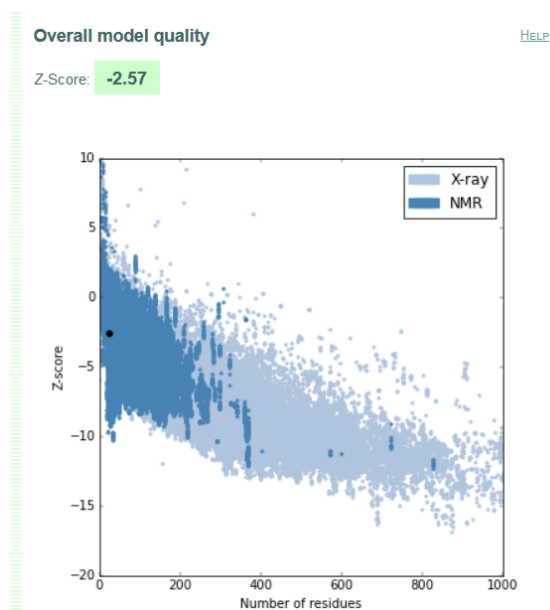


Figure A.7: I-TASSER Model 3 quality score calculated by ProSA for peptide hPEP1.

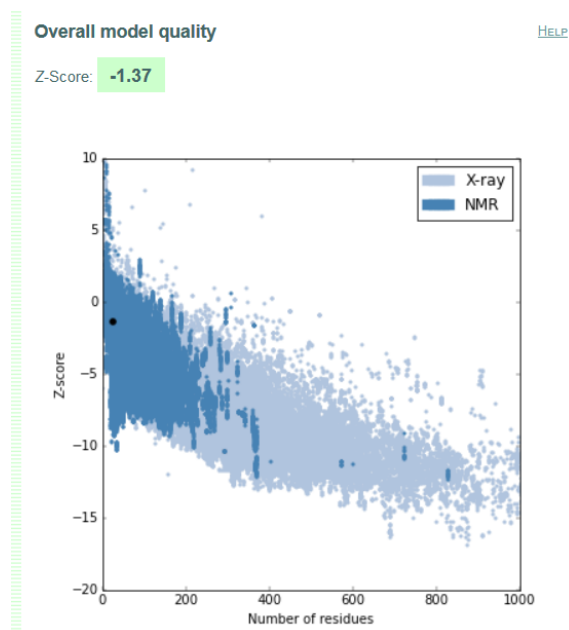


Figure A.8: I-TASSER Model 4 quality score calculated by ProSA for peptide hPEP1.

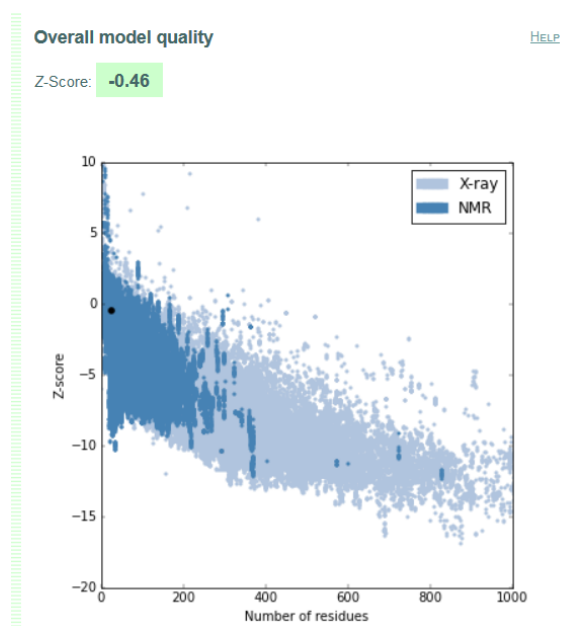


Figure A.9: I-TASSER Model 5 quality score calculated by ProSA for peptide hPEP1.

Annex 8

ProSA was used for hPEP1 structure evaluation. It is represented the overall quality z-scores calculated by ProSA for each one of the 5 structures predicted by PEP-FOLD3, along with the plot with the z-scores of all experimentally determined structures that are in the PDB database. The z-scores of all predicted peptide structures are within the range of scores typically found for native proteins with sequence length of 25 amino acids, suggesting that the structures predicted by PEP-FOLD3 probably do not contains errors and are going to be used in the ensemble molecular docking.

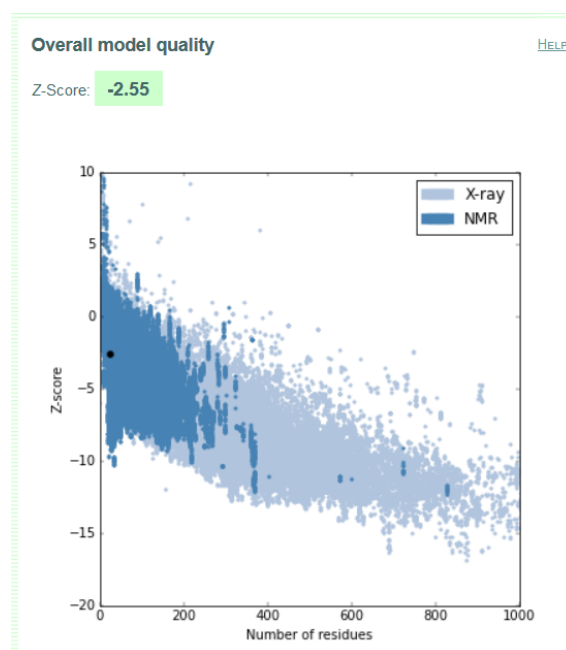


Figure A.10: PEP-FOLD3 Model 1 quality score calculated by ProSA for peptide hPEP1.

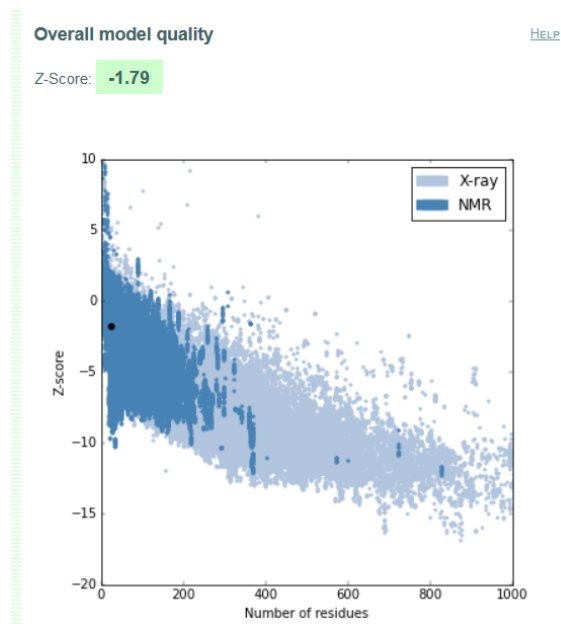


Figure A.11: PEP-FOLD3 Model 2 quality score calculated by ProSA for peptide hPEP1.

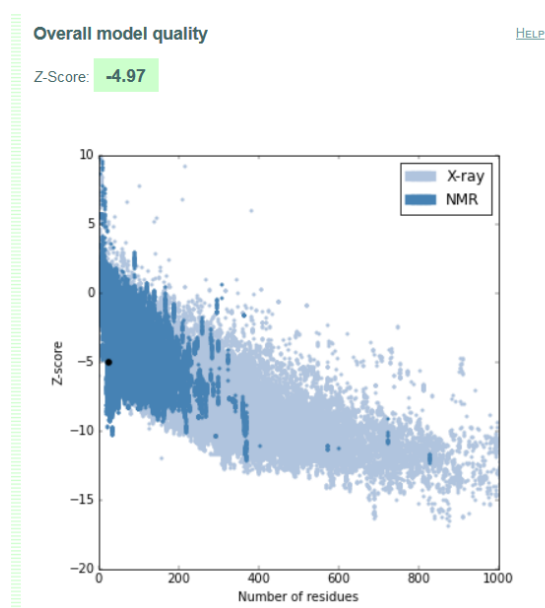


Figure A.12: PEP-FOLD3 Model 3 quality score calculated by ProSA for peptide hPEP1.

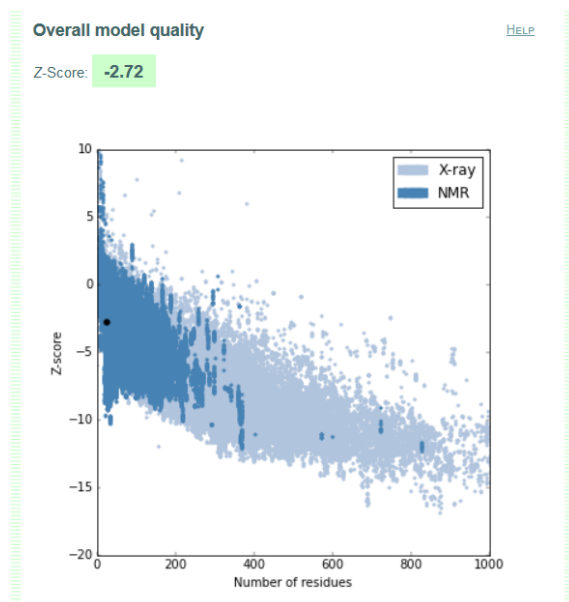


Figure A.13: PEP-FOLD3 Model 4 quality score calculated by ProSA for peptide hPEP1.

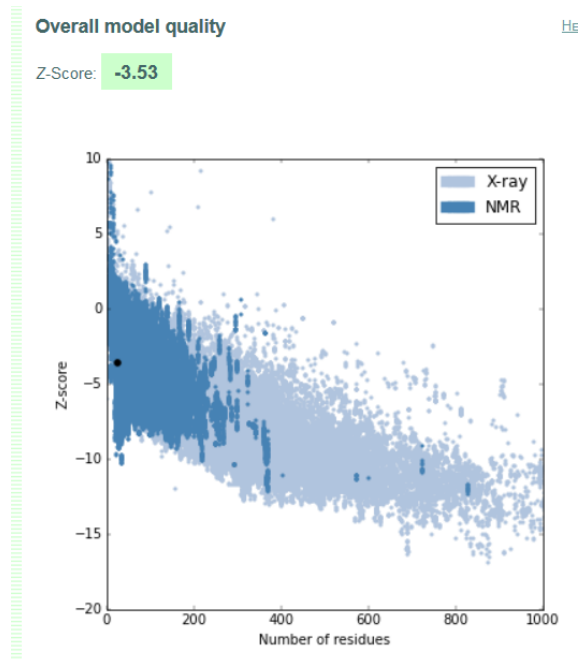


Figure A.14: PEP-FOLD3 Model 5 quality score calculated by ProSA for peptide hPEP1.

Annex 9

ProSA was used for hPEP1 structure evaluation. It is represented the overall quality z-scores calculated by ProSA for each one of the 5 structures predicted by PEPstrMOD, along with the plot with the z-scores of all experimentally determined structures that are in the PDB database. The z-scores of all predicted peptide structures are within the range of scores typically found for native proteins with sequence length of 25 amino acids, suggesting that the structures predicted by PEPstrMOD probably do not contains errors and are going to be used in the ensemble molecular docking.

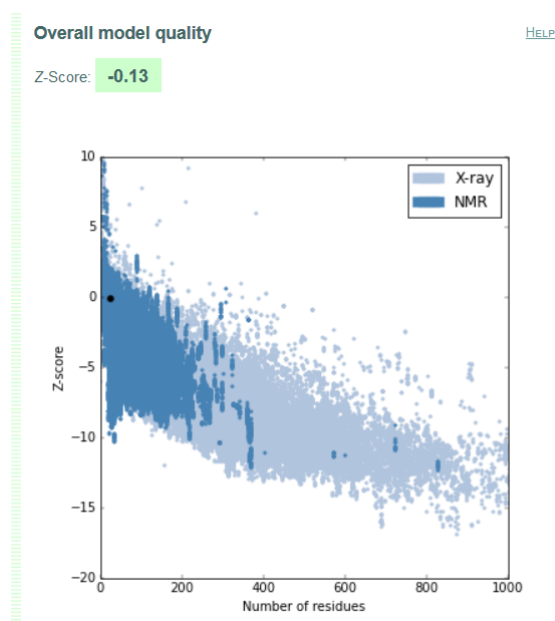


Figure A.15: PEPstrMOD Model 1 quality score calculated by ProSA for peptide hPEP1.

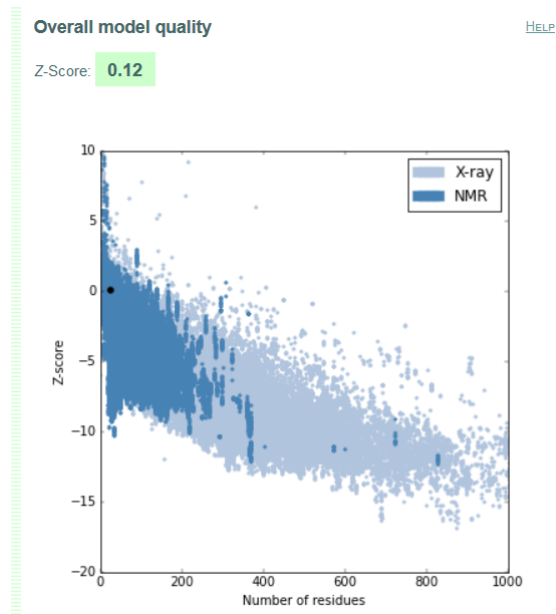


Figure A.16: PEPstrMOD Model 2 quality score calculated by ProSA for peptide hPEP1.

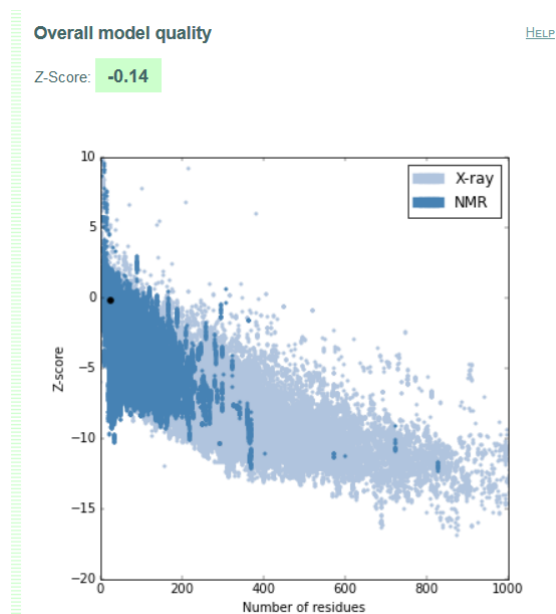


Figure A.17: PEPstrMOD Model 3 quality score calculated by ProSA for peptide hPEP1.

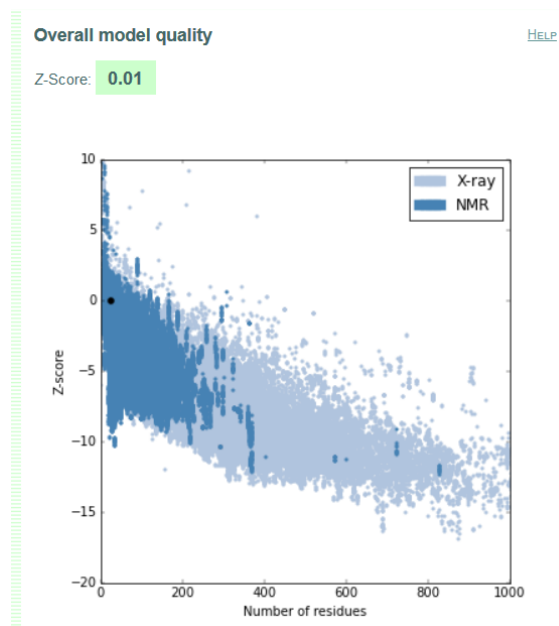


Figure A.18: PEPstrMOD Model 4 quality score calculated by ProSA for peptide hPEP1.

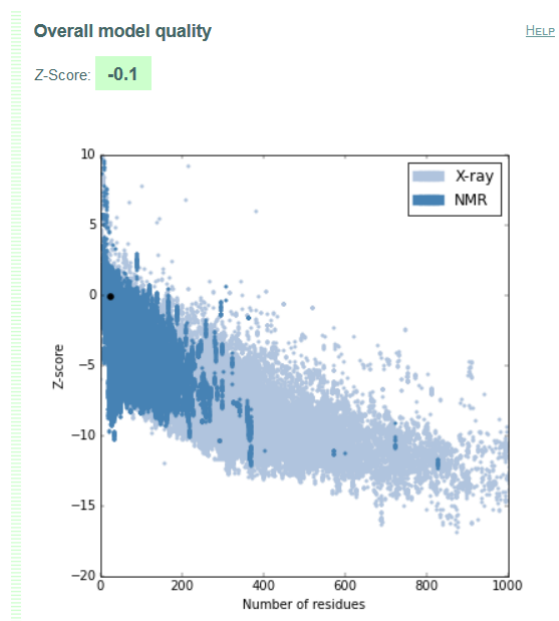


Figure A.19: PEPstrMOD Model 5 quality score calculated by ProSA for peptide hPEP1.

Annex 10

HADDOCK results for the docking of the complex hTRAF6-hPEP1, with the clusters ordered in descending order (best to worst having into account the HADDOCK score). The cluster ranking is made having into account the best 4 models from each one.

CLUSTER 2

HADDOCK score	-110.7 +/- 10.7
Cluster size	6
RMSD from the overall lowest-energy structure	10.6 +/- 0.4
Van der Waals energy	-48.8 +/- 3.5
Electrostatic energy	-327.6 +/- 47.9
Desolvation energy	-2.0 +/- 2.8
Restraints violation energy	55.6 +/- 21.39
Buried Surface Area	1494.8 +/- 141.5
Z-Score	-1.5

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 3

HADDOCK score	-104.4 +/- 2.0
Cluster size	6
RMSD from the overall lowest-energy structure	9.1 +/- 0.5
Van der Waals energy	-48.0 +/- 2.3
Electrostatic energy	-281.0 +/- 17.2
Desolvation energy	-2.3 +/- 3.7
Restraints violation energy	20.6 +/- 12.42
Buried Surface Area	1476.7 +/- 42.5
Z-Score	-1.1

CLUSTER 1

HADDOCK score	-100.5 +/- 8.3
Cluster size	8
RMSD from the overall lowest-energy structure	6.5 +/- 0.4
Van der Waals energy	-43.8 +/- 8.2
Electrostatic energy	-281.7 +/- 67.9
Desolvation energy	-3.9 +/- 4.2
Restraints violation energy	36.0 +/- 27.14
Buried Surface Area	1485.0 +/- 209.1
Z-Score	-0.7

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 5

HADDOCK score	-100.1 +/- 10.1
Cluster size	5
RMSD from the overall lowest-energy structure	5.2 +/- 0.1
Van der Waals energy	-46.9 +/- 6.3
Electrostatic energy	-302.6 +/- 49.6
Desolvation energy	-1.8 +/- 6.5
Restraints violation energy	90.2 +/- 30.41
Buried Surface Area	1513.0 +/- 118.9
Z-Score	-0.7

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 4

HADDOCK score	-88.3 +/- 13.7
Cluster size	5
RMSD from the overall lowest-energy structure	6.5 +/- 0.2
Van der Waals energy	-35.0 +/- 6.3
Electrostatic energy	-292.2 +/- 36.4
Desolvation energy	-4.5 +/- 0.6
Restraints violation energy	97.1 +/- 22.48
Buried Surface Area	1168.6 +/- 27.7
Z-Score	0.2

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 9

HADDOCK score	-86.4 +/- 13.2
Cluster size	4
RMSD from the overall lowest-energy structure	5.7 +/- 0.2
Van der Waals energy	-40.0 +/- 3.2
Electrostatic energy	-275.1 +/- 58.0
Desolvation energy	-0.7 +/- 3.5
Restraints violation energy	93.1 +/- 19.25
Buried Surface Area	1251.5 +/- 124.4
Z-Score	0.4

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 7

HADDOCK score	-82.3 +/- 15.0
Cluster size	5
RMSD from the overall lowest-energy structure	7.3 +/- 0.3
Van der Waals energy	-37.3 +/- 4.9
Electrostatic energy	-252.9 +/- 31.2
Desolvation energy	-3.2 +/- 8.4
Restraints violation energy	87.3 +/- 38.88
Buried Surface Area	1281.6 +/- 138.6
Z-Score	0.7

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 6

HADDOCK score	-75.2 +/- 10.3
Cluster size	5
RMSD from the overall lowest-energy structure	7.2 +/- 0.4
Van der Waals energy	-36.4 +/- 5.7
Electrostatic energy	-215.9 +/- 15.1
Desolvation energy	-0.6 +/- 6.9
Restraints violation energy	49.7 +/- 27.10
Buried Surface Area	1152.3 +/- 83.1
Z-Score	1.3

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 8

HADDOCK score	-72.6 +/- 19.0
Cluster size	4
RMSD from the overall lowest-energy structure	5.9 +/- 0.2
Van der Waals energy	-38.7 +/- 8.0
Electrostatic energy	-241.6 +/- 52.5
Desolvation energy	2.5 +/- 6.8
Restraints violation energy	119.2 +/- 22.66
Buried Surface Area	1463.3 +/- 164.8
Z-Score	1.5

Nr 1 best structure	Download structure	
Nr 2 best structure	Download structure	
Nr 3 best structure	Download structure	
Nr 4 best structure	Download structure	

Figure A.20: HADDOCK results for the complex hTRAF6-hPEP1. The clusters are sorted by descending order, so the best is in top along with the best 4 representative docked models.

Annex 11

The parameters of the 4 best models from each cluster, coming from the hTRAF6-hPEP1 docking with HADDOCK.

Table A.3: HADDOCK results of the first ranked cluster for complex hTRAF6-hPEP1

	1° Cluster - 2			
	Cluster2_1	Cluster2_2	Cluster2_3	Cluster2_4
Webserver	I-TASSER	I-TASSER	I-TASSER	I-TASSER
l-RMSD (Å)	6,3	6,4	5,8	5,0
iRMSD (Å)	2,6	2,5	2,4	2,1
Capri assessment	Incorrect	Incorrect	Incorrect	Acceptable
HADDOCKscore	-126,5	-114,3	-103,6	-98,7
Eair	44,4	53,9	33,6	90,5
Evdw	-53,1	-50,5	-43,6	-48,1
Eelec	-396,0	-347,8	-292,7	-274,1
Edesol	1,3	0,4	-4,8	-4,7
Eint	-404,7	-344,5	-302,7	-231,6
BSA (Å²)	1604,8	1662,4	1329,5	1382,5

Table A.4: HADDOCK results of the second ranked cluster for complex hTRAF6-hPEP1

	2° Cluster - 3			
	Cluster3_1	Cluster3_2	Cluster3_3	Cluster3_4
Webserver	I-TASSER	I-TASSER	I-TASSER	I-TASSER
l-RMSD (Å)	5,2	4,8	5,4	5,0
iRMSD (Å)	2,2	2,0	2,4	2,2
Capri assessment	Incorrect	Acceptable	Incorrect	Acceptable
HADDOCKscore	-107,2	-105,3	-103,4	-101,8
Eair	10,4	16,0	41,9	14,2
Evdw	-46,5	-51,4	-48,9	-45,3
Eelec	-301,2	-294,9	-265,9	-262,2
Edesol	-1,5	3,5	-5,4	-5,5
Eint	-337,3	-330,2	-273,0	-293,2
BSA (Å²)	1512,3	1520,5	1458,3	1415,8

Table A.5: HADDOCK results of the third ranked cluster for complex hTRAF6-hPEP1

	3° Cluster - 1			
	Cluster1_1	Cluster1_2	Cluster1_3	Cluster1_4
Webserver	I-TASSER	I-TASSER	I-TASSER	I-TASSER
l-RMSD (Å)	12,0	11,7	13,5	12,0
iRMSD (Å)	5,0	4,8	5,4	4,9
Capri assessment	Incorrect	Incorrect	Incorrect	Incorrect
HADDOCKscore	-109,4	-105,7	-99,5	-87,4
Eair	64,2	10,1	7,7	62,1
Evdw	-38,6	-51,7	-33,2	-52,0
Eelec	-357,3	-291,5	-306,1	-171,9
Edesol	-5,7	3,3	-5,8	-7,3
Eint	-331,7	-333,1	-331,6	-161,8
BSA (Å²)	1470,3	1782,9	1192,0	1494,9

Table A.6: HADDOCK results of the fourth ranked cluster for complex hTRAF6-hPEP1

	4° Cluster - 5			
	Cluster5_1	Cluster5_2	Cluster5_3	Cluster5_4
Webserver	I-TASSER	I-TASSER	I-TASSER	I-TASSER
l-RMSD (Å)	11,4	11,7	12,0	11,9
iRMSD (Å)	4,8	4,9	5,0	5,0
Capri assessment	Incorrect	Incorrect	Incorrect	Incorrect
HADDOCKscore	-111,6	-108,5	-92,9	-87,6
Eair	74,5	50,0	129,8	106,4
Evdw	-53,1	-38,4	-52,8	-43,1
Eelec	-298,3	-372,9	-306,5	-232,7
Edesol	-6,3	-0,5	8,2	-8,6
Eint	-277,0	-361,3	-229,4	-169,4
BSA (Å²)	1717,0	1418,6	1462,5	1454,0

Table A.7: HADDOCK results of the fifth ranked cluster for complex hTRAF6-hPEP1

	5° Cluster - 4			
	Cluster4_1	Cluster4_2	Cluster4_3	Cluster4_4
Webserver	I-TASSER	I-TASSER	I-TASSER	I-TASSER
l-RMSD (Å)	11,9	11,8	11,9	12,0
iRMSD (Å)	4,9	4,9	4,8	4,8
Capri assessment	Incorrect	Incorrect	Incorrect	Incorrect
HADDOCKscore	-104,1	-99,2	-77,9	-71,7
Eair	78,0	71,8	115,1	123,4
Evdw	-45,2	-33,3	-33,8	-27,8
Eelec	-306,4	-345,3	-259,0	-258,0
Edesol	-5,4	-4,1	-3,8	-4,7
Eint	-273,6	-306,7	-177,7	-162,5
BSA (Å²)	1171,5	1204,0	1126,4	1172,5

Annex 12

Visualization of cluster behaviour having into account two RMSD parameters and the intermolecular energy of the best 4 clusters that resulted from the hTRAF6-hPEP1 docking.

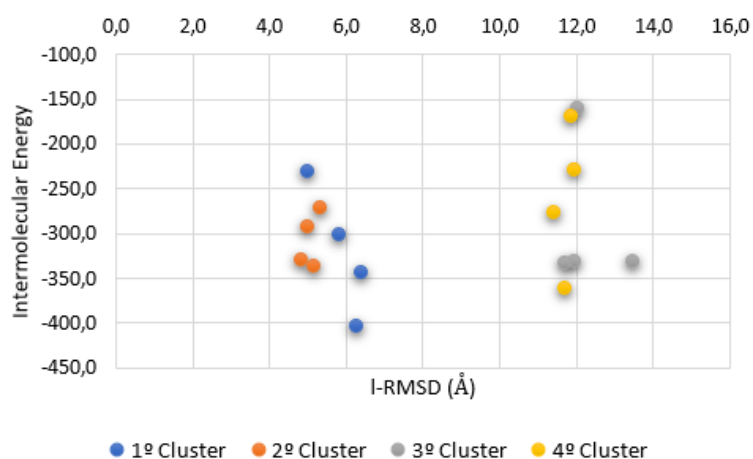


Figure A.21: Intermolecular energy versus l-RMSD for the hTRAF6-hPEP1 complex.

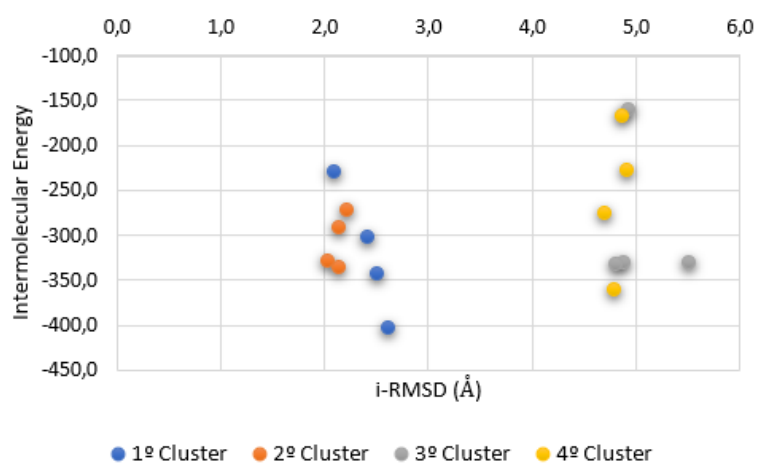


Figure A.22: Intermolecular energy versus i-RMSD for the hTRAF6-hPEP1 complex.

Annex 13

Distance of hTRAF6 residues to the predicted peptide hPEP1 for the best model that represents the complex hTRAF6-hPEP1.

Table A.8: hTRAF6 interface residues divided by distances to the peptide hPEP1 for the best model. In gray are the residues maintained from the reference structure 1LB5. Visualization with *Pymol*

Model 3_2						
3Å	4Å	5Å	6Å			
394	392	470	374	457	374	456
448	394	471	376	458	376	457
449	410	472	392	459	392	458
450	412	473	394	466	394	459
451	431	474	402	468	402	466
453	448	475	410	469	410	467
457	449		412	470	412	468
458	450		429	471	429	469
466	451		431	472	431	470
468	453		448	473	447	471
469	456		449	474	448	472
470	457		450	475	449	473
472	458		451	500	450	474
474	466		453		451	475
	468		455		453	500
	469		456		455	

Annex 14

The z-scores of all predicted peptide structures are within the range of scores typically found for native proteins with sequences lengths of 25 amino acids.

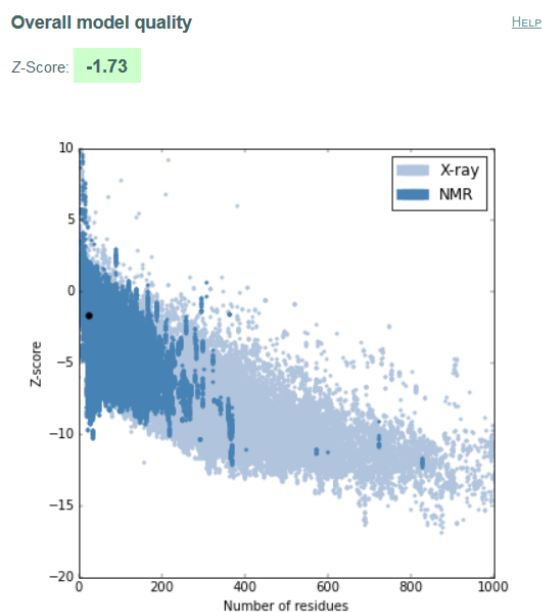


Figure A.23: I-TASSER Model 1 quality score calculated by ProSA for peptide hPEP2.

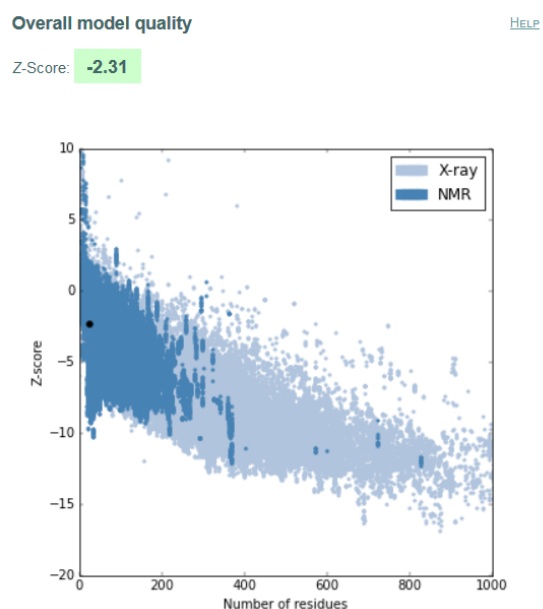


Figure A.24: I-TASSER Model 2 quality score calculated by ProSA for peptide hPEP2.

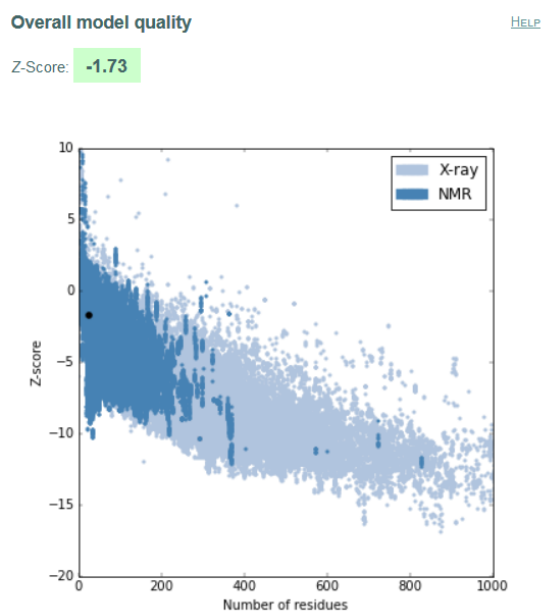


Figure A.25: I-TASSER Model 3 quality score calculated by ProSA for peptide hPEP2.

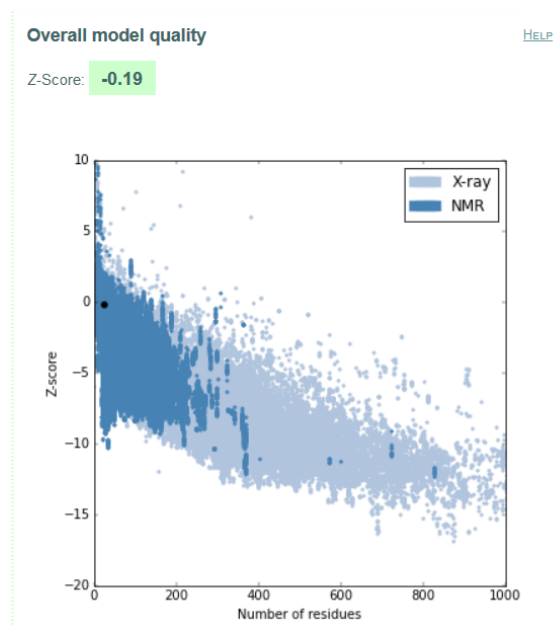


Figure A.26: I-TASSER Model 4 quality score calculated by ProSA for peptide hPEP2.

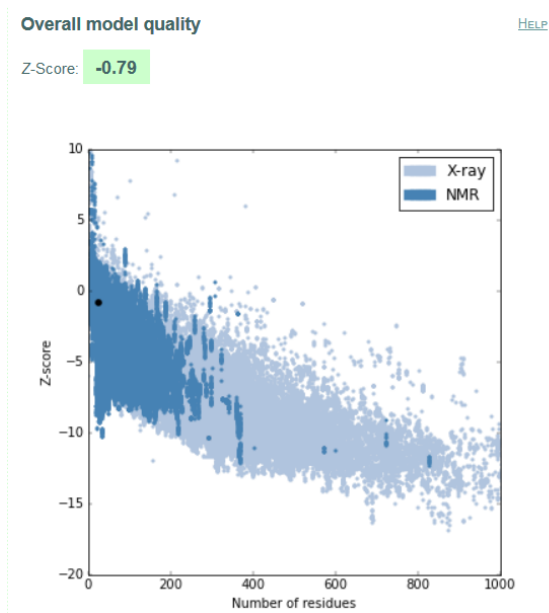


Figure A.27: I-TASSER Model 5 quality score calculated by ProSA for peptide hPEP2.

Annex 15

The z-scores of all predicted peptide structures are within the range of scores typically found for native proteins with sequences lengths of 25 amino acids.

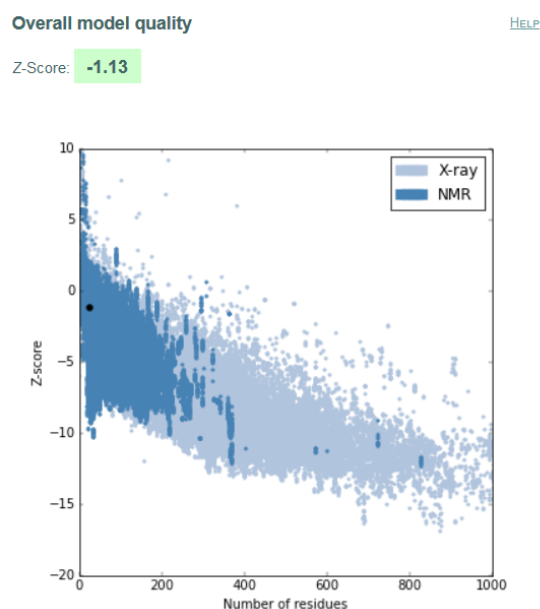


Figure A.28: I-TASSER Model 1 quality score calculated by ProSA for peptide mPEP1.

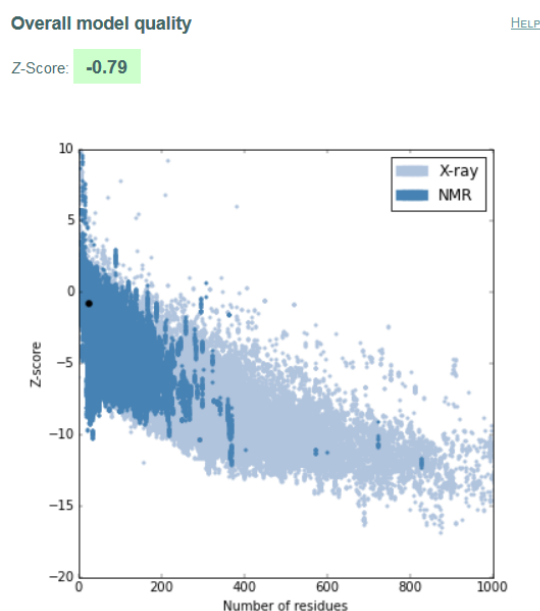


Figure A.29: I-TASSER Model 2 quality score calculated by ProSA for peptide mPEP1.

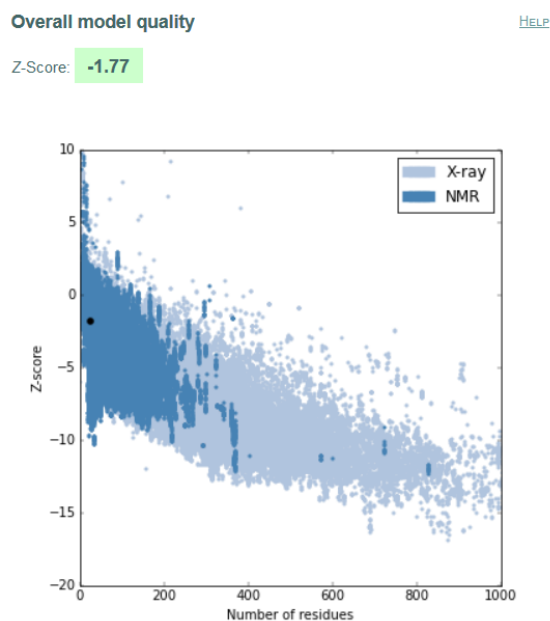


Figure A.30: I-TASSER Model 3 quality score calculated by ProSA for peptide mPEP1.

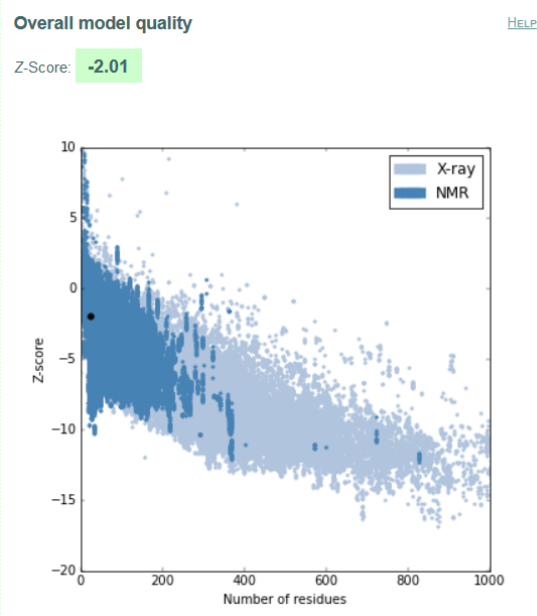


Figure A.31: I-TASSER Model 4 quality score calculated by ProSA for peptide mPEP1.

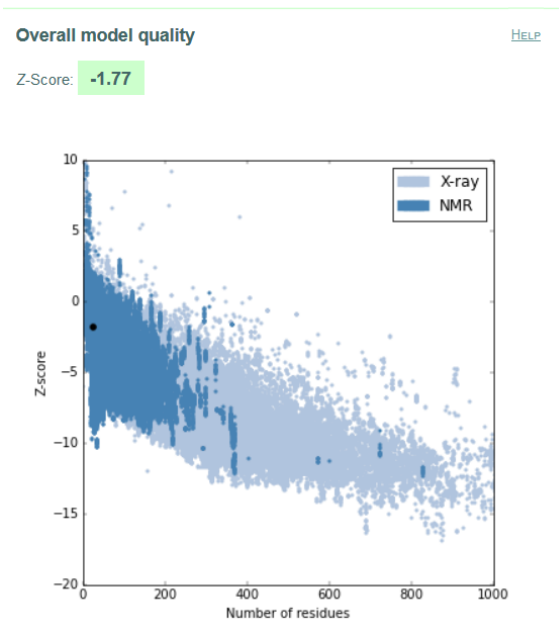


Figure A.32: I-TASSER Model 5 quality score calculated by ProSA for peptide mPEP1.

Annex 16

The z-scores of all predicted peptide structures are within the range of scores typically found for native proteins with sequences lengths of 25 amino acids.

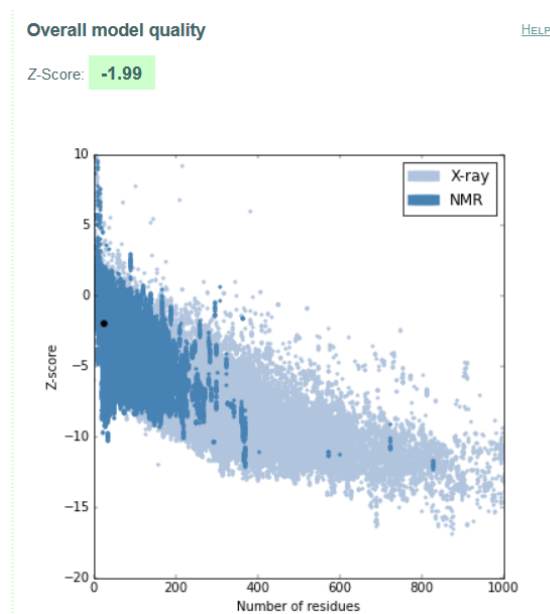


Figure A.33: I-TASSER Model 1 quality score calculated by ProSA for peptide mPEP2.

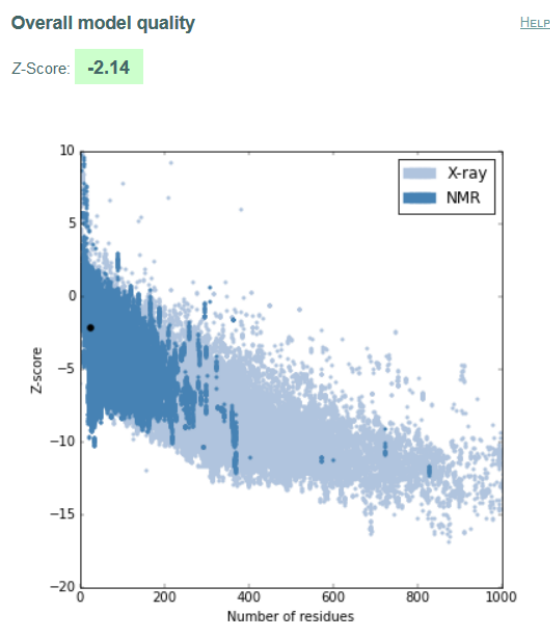


Figure A.34: I-TASSER Model 2 quality score calculated by ProSA for peptide mPEP2.

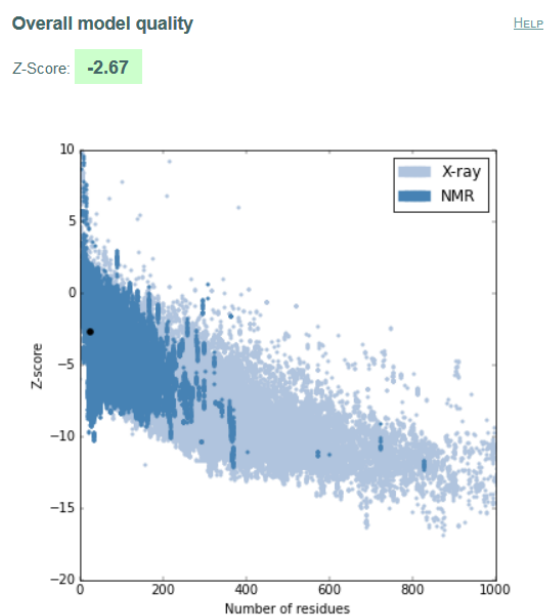


Figure A.35: I-TASSER Model 3 quality score calculated by ProSA for peptide mPEP2.

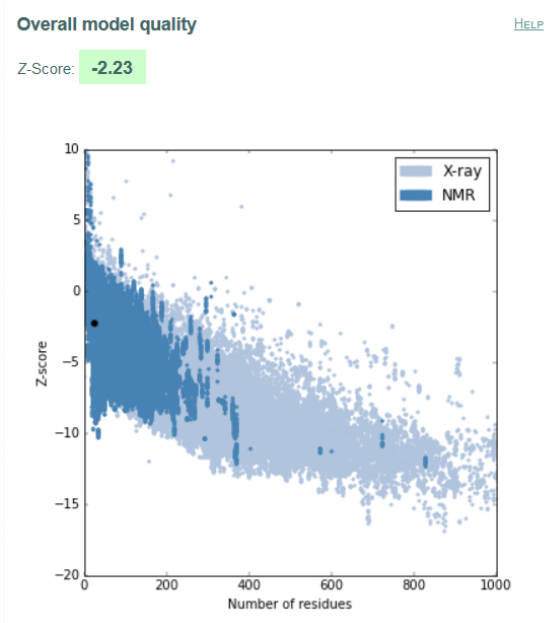


Figure A.36: I-TASSER Model 4 quality score calculated by ProSA for peptide mPEP2.

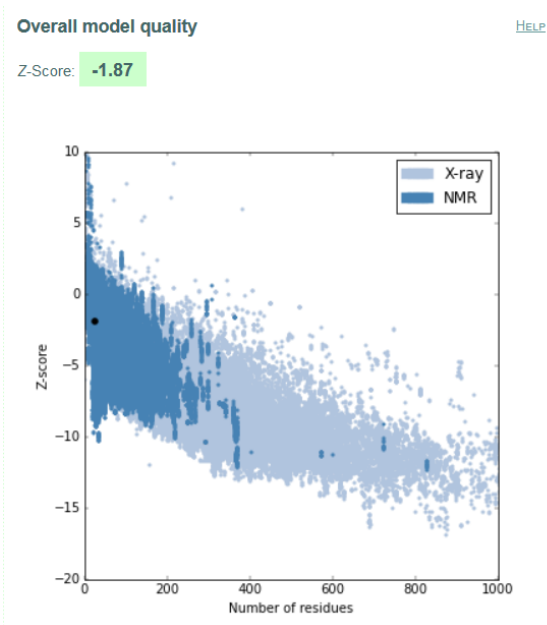


Figure A.37: I-TASSER Model 5 quality score calculated by ProSA for peptide mPEP2.

Annex 17

The z-scores of all predicted peptide structures are within the range of scores typically found for native proteins with sequences lengths of 23 amino acids.

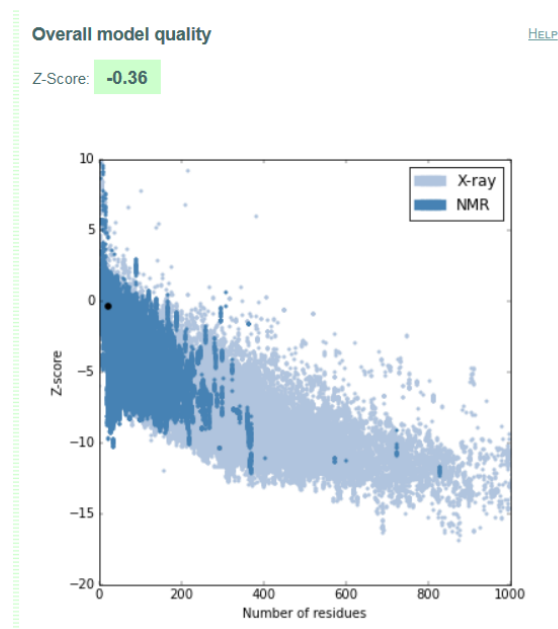


Figure A.38: I-TASSER Model 1 quality score calculated by ProSA for peptide PepH2-TRANCE1.

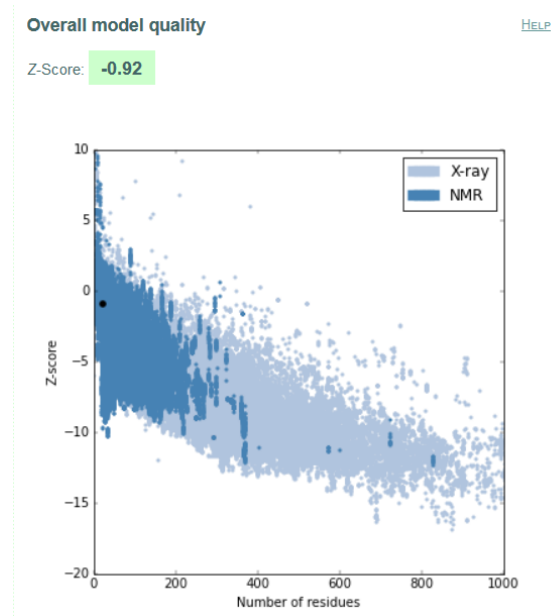


Figure A.39: I-TASSER Model 2 quality score calculated by ProSA for peptide PepH2-TRANCE1.

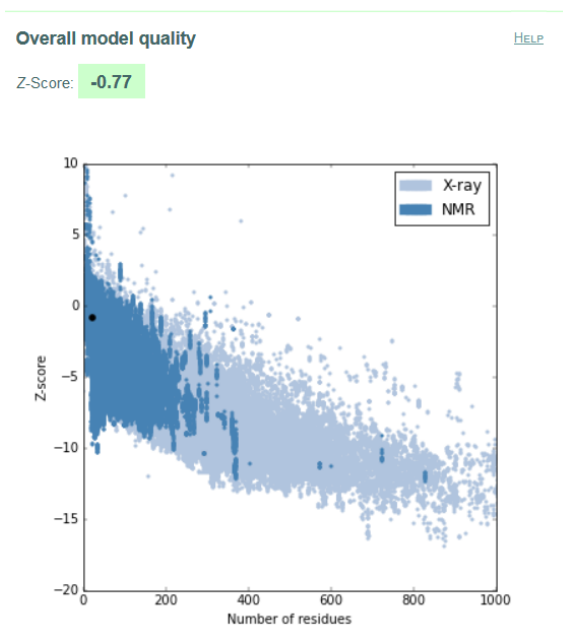


Figure A.40: I-TASSER Model 3 quality score calculated by ProSA for peptide PepH2-TRANCE1.

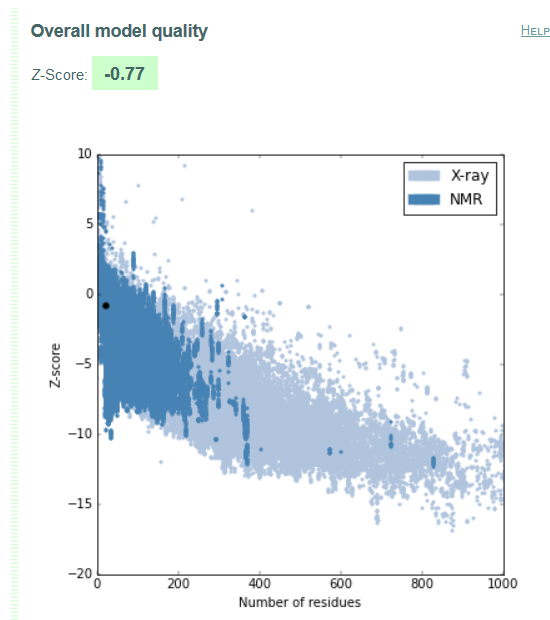


Figure A.41: I-TASSER Model 4 quality score calculated by ProSA for peptide PepH2-TRANCE1.

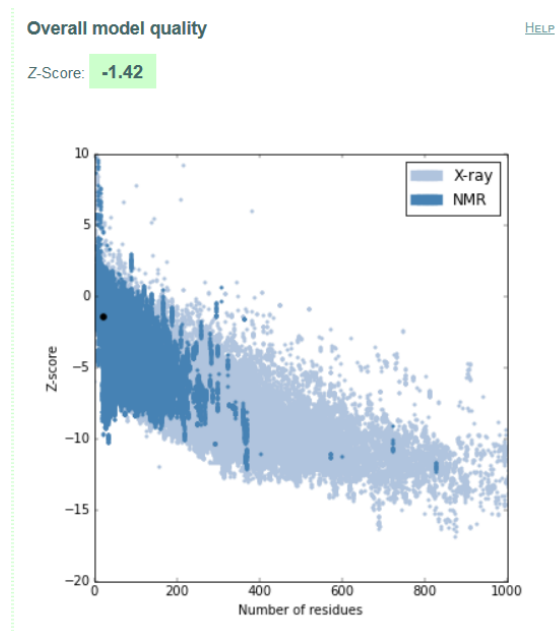


Figure A.42: I-TASSER Model 5 quality score calculated by ProSA for peptide PepH2-TRANCE1.

Annex 18

HADDOCK results for the docking of the complex hTRAF6-hPEP2, with the clusters ordered in descending order (best to worst having into account the HADDOCK score). The cluster ranking is made having into account the best 4 models from each one.

CLUSTER 2

HADDOCK score	-109.5 +/- 3.0
Cluster size	11
RMSD from the overall lowest-energy structure	6.3 +/- 0.4
Van der Waals energy	-46.4 +/- 2.8
Electrostatic energy	-305.1 +/- 41.6
Desolvation energy	-7.9 +/- 7.3
Restraints violation energy	59.1 +/- 27.80
Buried Surface Area	1565.3 +/- 79.9
Z-Score	-1.5

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 3

HADDOCK score	-104.9 +/- 0.4
Cluster size	10
RMSD from the overall lowest-energy structure	3.4 +/- 0.3
Van der Waals energy	-46.1 +/- 3.6
Electrostatic energy	-262.5 +/- 35.6
Desolvation energy	-11.4 +/- 5.6
Restraints violation energy	51.4 +/- 11.05
Buried Surface Area	1433.5 +/- 139.3
Z-Score	-0.9

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 5

HADDOCK score	-104.7 +/- 6.8
Cluster size	7
RMSD from the overall lowest-energy structure	2.4 +/- 0.4
Van der Waals energy	-45.0 +/- 2.4
Electrostatic energy	-254.4 +/- 33.0
Desolvation energy	-16.5 +/- 5.3
Restraints violation energy	76.1 +/- 10.15
Buried Surface Area	1558.7 +/- 46.3
Z-Score	-0.9

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 1

HADDOCK score	-102.9 +/- 5.6
Cluster size	12
RMSD from the overall lowest-energy structure	4.2 +/- 0.2
Van der Waals energy	-37.4 +/- 2.8
Electrostatic energy	-319.6 +/- 32.9
Desolvation energy	-10.2 +/- 2.6
Restraints violation energy	86.1 +/- 21.76
Buried Surface Area	1392.9 +/- 50.6
Z-Score	-0.7

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 12

HADDOCK score	-99.3 +/- 17.7
Cluster size	4
RMSD from the overall lowest-energy structure	1.1 +/- 0.7
Van der Waals energy	-46.1 +/- 8.5
Electrostatic energy	-246.3 +/- 62.5
Desolvation energy	-14.5 +/- 3.4
Restraints violation energy	105.9 +/- 21.42
Buried Surface Area	1448.0 +/- 81.2
Z-Score	-0.3

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 7

HADDOCK score	-97.9 +/- 7.1
Cluster size	6
RMSD from the overall lowest-energy structure	4.8 +/- 0.2
Van der Waals energy	-43.7 +/- 3.0
Electrostatic energy	-246.3 +/- 37.2
Desolvation energy	-11.1 +/- 4.0
Restraints violation energy	61.5 +/- 32.49
Buried Surface Area	1487.6 +/- 41.2
Z-Score	-0.1

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 9

HADDOCK score	-92.5 +/- 12.5
Cluster size	5
RMSD from the overall lowest-energy structure	6.4 +/- 0.5
Van der Waals energy	-37.2 +/- 5.7
Electrostatic energy	-243.2 +/- 25.6
Desolvation energy	-12.9 +/- 6.0
Restraints violation energy	63.3 +/- 38.46
Buried Surface Area	1305.6 +/- 191.8
Z-Score	0.6

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 13

HADDOCK score	-90.9 +/- 20.7
Cluster size	4
RMSD from the overall lowest-energy structure	5.2 +/- 0.6
Van der Waals energy	-35.6 +/- 4.7
Electrostatic energy	-245.1 +/- 73.7
Desolvation energy	-15.5 +/- 5.4
Restraints violation energy	92.0 +/- 45.06
Buried Surface Area	1310.9 +/- 117.0
Z-Score	0.8

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 4

HADDOCK score	-85.7 +/- 6.8
Cluster size	8
RMSD from the overall lowest-energy structure	3.2 +/- 0.2
Van der Waals energy	-30.5 +/- 4.2
Electrostatic energy	-280.1 +/- 49.7
Desolvation energy	-9.7 +/- 5.8
Restraints violation energy	105.2 +/- 42.29
Buried Surface Area	1137.7 +/- 60.9
Z-Score	1.4

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 14

HADDOCK score	-83.3 +/- 9.9
Cluster size	4
RMSD from the overall lowest-energy structure	5.9 +/- 0.2
Van der Waals energy	-44.4 +/- 3.6
Electrostatic energy	-152.2 +/- 26.8
Desolvation energy	-21.9 +/- 3.4
Restraints violation energy	133.9 +/- 19.39
Buried Surface Area	1405.0 +/- 133.3
Z-Score	1.7

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

Figure A.43: HADDOCK results for the complex hTRAF6-hPEP2. The clusters are sorted by descending order, so the best is in top along with the best 4 representative docked models.

Annex 19

The parameters of the 2 best models from each cluster, coming from the hTRAF6-hPEP2 docking with HADDOCK.

Table A.9: HADDOCK results of the first ranked cluster for complex hTRAF6-hPEP2

	1° - Cluster 2			
	Cluster2_1	Cluster2_2	Cluster2_3	Cluster2_4
Webserver	I-TASSER			
HADDOCKscore	-113,4	-111,3	-107,1	-106,0
Eair	53,5	102,5	25,1	55,1
Evdw	-47,9	-49,8	-42,5	-45,4
Eelec	-301,9	-340,8	-339,6	-238,3
Edesol	-10,4	-3,6	0,8	-18,4
Eint	-296,3	-288,1	-356,9	-228,6
BSA ($\text{\AA}^2$)	1439,9	1576,2	1582,7	1662,3

Table A.10: HADDOCK results of the second ranked cluster for complex hTRAF6-hPEP2

	2° - Cluster 3			
	Cluster3_1	Cluster3_2	Cluster3_3	Cluster3_4
Webserver	I-TASSER			
HADDOCKscore	-105,4	-105,1	-104,7	-104,3
Eair	42,0	46,2	70,2	47,3
Evdw	-52,2	-43,7	-43,4	-45,2
Eelec	-231,6	-317,0	-271,4	-230,0
Edesol	-11,2	-2,6	-14,1	-17,8
Eint	-241,8	-314,5	-244,5	-227,9
BSA ($\text{\AA}^2$)	1484,1	1294,7	1638,6	1316,5

Annex 20

Percentage of conservation of the residues determined as active in HADDOCK to perform molecular docking of the complex hTRAF6-hPEP2. This percentage was used as a comparative parameter between the best models previous from the best 2 first clusters.

Table A.11: hTRAF6 interface active residues for the hTRAF6-hPEP2 best models. In gray are the residues not present in 1LB5 PDB. Visualization with *Pymol* with a 6Å cut-off

% Active residues conservation			
86,4%	100,0%	86,4%	86,4%
2_3	2_1	3_2	3_1
6			
	374		374
	376		376
392	392	392	392
	394	394	394
396		396	396
			401
402	402	402	402
		403	403
		408	408
410	410	410	410
		411	
412	412	412	412
	431		
444			
445			
446			
447			
448	448		448

% Active residues conservation			
86,4%	100,0%	86,4%	86,4%
2_3	2_1	3_2	3_1
6			
449	449		449
450	450	450	450
451	451	451	451
		452	
453	453	453	453
455	455	455	455
456	456	456	
457	457	457	
458	458	458	
	459	459	
466	466	466	466
467	467	467	
468	468	468	468
469	469	469	469
470	470	470	470
471	471	471	471
472	472	472	472
473	473	473	473
474	474	474	474
475	475	475	475
476			476

Annex 21

HADDOCK results for the docking of the complex mTRAF6-mPEP1, with the clusters ordered in descending order (best to worst having into account the HADDOCK score). The cluster ranking is made having into account the best 4 models from each one.

CLUSTER 4

HADDOCK score	-73.7 +/- 15.3
Cluster size	4
RMSD from the overall lowest-energy structure	5.7 +/- 0.1
Van der Waals energy	-42.8 +/- 11.9
Electrostatic energy	-195.6 +/- 42.8
Desolvation energy	-1.6 +/- 4.4
Restraints violation energy	98.3 +/- 21.24
Buried Surface Area	1415.7 +/- 151.2
Z-Score	-0.7

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 1

HADDOCK score	-73.6 +/- 8.1
Cluster size	4
RMSD from the overall lowest-energy structure	4.3 +/- 0.1
Van der Waals energy	-41.8 +/- 7.1
Electrostatic energy	-180.9 +/- 49.7
Desolvation energy	-9.1 +/- 4.4
Restraints violation energy	134.2 +/- 8.00
Buried Surface Area	1388.6 +/- 164.2
Z-Score	-0.7

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 3

HADDOCK score	-68.5 +/- 14.7
Cluster size	4
RMSD from the overall lowest-energy structure	6.6 +/- 0.2
Van der Waals energy	-28.5 +/- 6.6
Electrostatic energy	-253.2 +/- 38.1
Desolvation energy	-0.2 +/- 3.8
Restraints violation energy	108.3 +/- 25.85
Buried Surface Area	1128.5 +/- 76.6
Z-Score	-0.4

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 2

HADDOCK score	-61.3 +/- 0.6
Cluster size	4
RMSD from the overall lowest-energy structure	4.5 +/- 0.5
Van der Waals energy	-32.9 +/- 5.0
Electrostatic energy	-204.2 +/- 45.8
Desolvation energy	4.9 +/- 2.9
Restraints violation energy	74.0 +/- 44.87
Buried Surface Area	1132.1 +/- 72.9
Z-Score	-0.0

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 5

HADDOCK score	-28.0 +/- 14.6
Cluster size	4
RMSD from the overall lowest-energy structure	4.8 +/- 0.2
Van der Waals energy	-29.8 +/- 8.7
Electrostatic energy	-65.4 +/- 49.1
Desolvation energy	-1.7 +/- 5.1
Restraints violation energy	165.2 +/- 53.54
Buried Surface Area	1209.4 +/- 101.6
Z-Score	1.9

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

Figure A.44: HADDOCK results for the complex mTRAF6-mPEP1. The clusters are sorted by descending order, so the best is in top along with the best 4 representative docked models.

Annex 22

The parameters of the 2 best models from each cluster, coming from the mTRAF6-mPEP1 docking with HADDOCK.

Table A.12: HADDOCK results of the first ranked cluster for complex mTRAF6-mPEP1

	1° - Cluster 4			
	Cluster4_1	Cluster4_2	Cluster4_3	Cluster4_4
Webserver	I-TASSER			
HADDOCKscore	-93,5	-81,2	-67,8	-52,3
Eair	94,5	82,9	134,0	81,7
Evdw	-53,0	-48,2	-47,6	-22,5
Eelec	-242,4	-233,8	-147,3	-158,8
Edesol	-1,4	5,4	-4,2	-6,2
Eint	-200,9	-199,1	-60,9	-99,6
BSA ($\text{\AA}^2$)	1460,5	1535,7	1508,7	1158,0

Table A.13: HADDOCK results of the second ranked cluster for complex mTRAF6-mPEP1

	2° - Cluster 1			
	Cluster1_1	Cluster1_2	Cluster1_3	Cluster1_4
Webserver	I-TASSER			
HADDOCKscore	-85,4	-74,7	-71,7	-62,6
Eair	147,8	127,3	129,8	132,1
Evdw	-52,7	-32,9	-41,7	-39,8
Eelec	-189,5	-258,8	-138,2	-137,1
Edesol	-9,6	-2,8	-15,3	-8,6
Eint	-94,4	-164,4	-50,1	-44,8
BSA ($\text{\AA}^2$)	1645,1	1225,7	1416,5	1267,1

Annex 23

Percentage of conservation of the residues determined as active in HADDOCK to perform molecular docking of the complex mTRAF6-mPEP1. This percentage was used as a comparative parameter between the best models previous from the best 2 first clusters.

Table A.14: mTRAF6 interface active residues for the mTRAF6-mPEP1 best models. In gray are the residues not considered active. Visualization with *Pymol* with a 6Å cut-off

% Active residues conservation		
95,5%	95,5%	81,8%
4_1	1_1	1_2
382	382	
384	384	384
400	400	400
402	402	402
404		404
	410	410
	411	411
	414	
418	418	418
		419
420	420	420
437		
452		
453		
454		
455		
456	456	

% Active residues conservation		
95,5%	95,5%	81,8%
4_1	1_1	1_2
457	457	
458	458	
459	459	
460	460	
461	461	461
462		
463		
464	464	464
465	465	465
466	466	466
467	467	467
		472
	474	474
	475	475
476	476	476
477	477	477
478	478	478
479	479	479
480	480	480
481	481	481
482	482	482
483	483	483
		484

Annex 24

HADDOCK results for the docking of the complex mTRAF6-mPEP2, with the clusters ordered in descending order (best to worst having into account the HADDOCK score). The cluster ranking is made having into account the best 4 models from each one.

CLUSTER 2

HADDOCK score	-92.7 +/- 4.5
Cluster size	5
RMSD from the overall lowest-energy structure	3.3 +/- 0.3
Van der Waals energy	-47.8 +/- 6.5
Electrostatic energy	-246.1 +/- 64.8
Desolvation energy	-5.5 +/- 8.3
Restraints violation energy	98.0 +/- 33.42
Buried Surface Area	1519.3 +/- 103.5
Z-Score	-1.4

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 1

HADDOCK score	-91.3 +/- 7.6
Cluster size	7
RMSD from the overall lowest-energy structure	6.1 +/- 0.6
Van der Waals energy	-39.4 +/- 6.2
Electrostatic energy	-304.9 +/- 44.5
Desolvation energy	1.5 +/- 7.4
Restraints violation energy	76.3 +/- 9.47
Buried Surface Area	1311.3 +/- 86.4
Z-Score	-1.3

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 3

HADDOCK score	-87.0 +/- 14.4
Cluster size	5
RMSD from the overall lowest-energy structure	5.8 +/- 0.5
Van der Waals energy	-35.3 +/- 5.0
Electrostatic energy	-278.8 +/- 45.7
Desolvation energy	-5.3 +/- 5.0
Restraints violation energy	93.4 +/- 28.16
Buried Surface Area	1250.5 +/- 126.6
Z-Score	-1.0

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 4

HADDOCK score	-79.8 +/- 5.3
Cluster size	4
RMSD from the overall lowest-energy structure	9.4 +/- 0.5
Van der Waals energy	-33.6 +/- 4.7
Electrostatic energy	-281.7 +/- 41.7
Desolvation energy	5.5 +/- 9.8
Restraints violation energy	45.5 +/- 27.53
Buried Surface Area	1159.2 +/- 43.5
Z-Score	-0.5

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 6

HADDOCK score	-74.7 +/- 5.4
Cluster size	4
RMSD from the overall lowest-energy structure	7.5 +/- 0.3
Van der Waals energy	-38.1 +/- 4.4
Electrostatic energy	-246.0 +/- 21.4
Desolvation energy	6.0 +/- 4.3
Restraints violation energy	66.7 +/- 15.79
Buried Surface Area	1169.0 +/- 140.3
Z-Score	-0.2

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 5

HADDOCK score	-66.6 +/- 11.9
Cluster size	4
RMSD from the overall lowest-energy structure	11.1 +/- 0.3
Van der Waals energy	-34.9 +/- 4.7
Electrostatic energy	-221.9 +/- 39.3
Desolvation energy	3.7 +/- 5.8
Restraints violation energy	89.1 +/- 32.78
Buried Surface Area	1100.7 +/- 116.1
Z-Score	0.3

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 7

HADDOCK score	-66.4 +/- 8.1
Cluster size	4
RMSD from the overall lowest-energy structure	7.8 +/- 0.6
Van der Waals energy	-38.0 +/- 3.9
Electrostatic energy	-218.1 +/- 23.6
Desolvation energy	8.0 +/- 11.5
Restraints violation energy	72.2 +/- 23.58
Buried Surface Area	1139.1 +/- 51.3
Z-Score	0.4

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 8

HADDOCK score	-62.0 +/- 6.4
Cluster size	4
RMSD from the overall lowest-energy structure	8.1 +/- 0.2
Van der Waals energy	-37.0 +/- 1.4
Electrostatic energy	-104.7 +/- 25.4
Desolvation energy	-18.0 +/- 6.7
Restraints violation energy	139.6 +/- 18.60
Buried Surface Area	1218.1 +/- 35.3
Z-Score	0.6

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 10

HADDOCK score	-53.5 +/- 22.3
Cluster size	4
RMSD from the overall lowest-energy structure	12.2 +/- 0.2
Van der Waals energy	-31.1 +/- 4.7
Electrostatic energy	-192.0 +/- 32.0
Desolvation energy	4.9 +/- 9.7
Restraints violation energy	109.8 +/- 19.20
Buried Surface Area	1127.0 +/- 70.9
Z-Score	1.2

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 9

HADDOCK score	-44.8 +/- 19.3
Cluster size	4
RMSD from the overall lowest-energy structure	8.9 +/- 0.4
Van der Waals energy	-27.4 +/- 11.6
Electrostatic energy	-195.2 +/- 35.9
Desolvation energy	6.7 +/- 3.2
Restraints violation energy	148.9 +/- 20.00
Buried Surface Area	879.4 +/- 125.1
Z-Score	1.8

Nr 1 best structure	Download structure	
Nr 2 best structure	Download structure	
Nr 3 best structure	Download structure	
Nr 4 best structure	Download structure	

Figure A.45: HADDOCK results for the complex mTRAF6-mPEP2. The clusters are sorted by descending order, so the best is in top along with the best 4 representative docked models.

Annex 25

The parameters of the 2 best models from each cluster, coming from the mTRAF6-mPEP2 docking with HADDOCK.

Table A.15: HADDOCK results of the first ranked cluster for complex mTRAF6-mPEP2

	1° - Cluster 2			
	Cluster2_1	Cluster2_2	Cluster2_3	Cluster2_4
Webserver	I-TASSER			
HADDOCKscore	-98,5	-95,6	-89,4	-87,5
Eair	41,2	108,8	114,5	127,2
Evdw	-46,9	-55,3	-51,3	-37,7
Eelec	-291,5	-189,9	-175,6	-327,3
Edesol	2,6	-13,1	-14,4	2,9
Eint	-297,2	-136,5	-112,4	-237,8
BSA ($\text{\AA}^2$)	1543,0	1652,9	1518,5	1363,0

Table A.16: HADDOCK results of the second ranked cluster for complex mTRAF6-mPEP2

	2° - Cluster 1			
	Cluster1_1	Cluster1_2	Cluster1_3	Cluster1_4
Webserver	I-TASSER			
HADDOCKscore	-102,1	-91,9	-90,4	-80,6
Eair	85,3	60,5	77,9	81,3
Evdw	-30,0	-37,7	-45,8	-44,1
Eelec	-358,8	-337,6	-251,7	-271,4
Edesol	-8,9	7,2	-2,1	9,6
Eint	-303,5	-314,8	-219,6	-234,1
BSA ($\text{\AA}^2$)	1197,7	1260,2	1413,7	1373,6

Annex 26

Percentage of conservation of the residues determined as active in HADDOCK to perform molecular docking of the complex mTRAF6-mPEP2. This percentage was used as a comparative parameter between the best models previous from the best 2 first clusters.

Table A.17: mTRAF6 interface active residues for the mTRAF6-mPEP2 best models. In gray are the residues not considered active. Visualization with *Pymol* with a 6Å cut-off

% Active residues conservation		
90,9%	86,4%	81,8%
2_1	1_1	1_2
	382	382
	383	
384	384	384
400	400	400
402	402	402
	404	404
		409
410	410	410
		411
414		
		416
418	418	418
420	420	420
456		

% Active residues conservation		
90,9%	86,4%	81,8%
2_1	1_1	1_2
457		
458		
459		
460		
461	461	461
	463	463
464	464	464
465	465	465
466	466	466
467	467	467
	468	
474	474	474
475	475	
476	476	476
477	477	477
478	478	478
479	479	479
480	480	480
481	481	481
482	482	482
483	483	483

Annex 27

HADDOCK results for the docking of the complex hTRAF6-PepH2-hPEP1, with the clusters ordered in descending order (best to worst having into account the HADDOCK score). The cluster ranking is made having into account the best 4 models from each one.

CLUSTER 8

HADDOCK score	-121.2 +/- 2.6
Cluster size	7
RMSD from the overall lowest-energy structure	3.9 +/- 0.6
Van der Waals energy	-49.0 +/- 5.1
Electrostatic energy	-337.8 +/- 27.7
Desolvation energy	-8.5 +/- 8.2
Restraints violation energy	38.1 +/- 14.88
Buried Surface Area	1578.9 +/- 156.9
Z-Score	-1.4

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 5

HADDOCK score	-120.7 +/- 8.9
Cluster size	8
RMSD from the overall lowest-energy structure	1.1 +/- 0.7
Van der Waals energy	-45.9 +/- 4.5
Electrostatic energy	-365.1 +/- 39.2
Desolvation energy	-6.6 +/- 5.2
Restraints violation energy	48.1 +/- 20.87
Buried Surface Area	1586.6 +/- 83.5
Z-Score	-1.3

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 10

HADDOCK score	-118.0 +/- 6.5
Cluster size	6
RMSD from the overall lowest-energy structure	5.7 +/- 0.6
Van der Waals energy	-42.1 +/- 4.8
Electrostatic energy	-391.5 +/- 50.9
Desolvation energy	0.4 +/- 6.0
Restraints violation energy	21.1 +/- 14.13
Buried Surface Area	1465.3 +/- 138.4
Z-Score	-1.1

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 1

HADDOCK score	-116.1 +/- 5.4
Cluster size	17
RMSD from the overall lowest-energy structure	10.4 +/- 0.1
Van der Waals energy	-48.0 +/- 9.5
Electrostatic energy	-276.6 +/- 69.0
Desolvation energy	-16.4 +/- 1.8
Restraints violation energy	36.8 +/- 17.86
Buried Surface Area	1579.3 +/- 106.8
Z-Score	-0.9

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 7

HADDOCK score	-103.7 +/- 11.0
Cluster size	7
RMSD from the overall lowest-energy structure	9.2 +/- 0.4
Van der Waals energy	-47.0 +/- 7.5
Electrostatic energy	-224.3 +/- 31.8
Desolvation energy	-15.8 +/- 7.8
Restraints violation energy	39.9 +/- 20.62
Buried Surface Area	1557.6 +/- 117.1
Z-Score	0.3

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 3

HADDOCK score	-103.1 +/- 5.1
Cluster size	11
RMSD from the overall lowest-energy structure	9.8 +/- 0.3
Van der Waals energy	-44.8 +/- 5.5
Electrostatic energy	-260.4 +/- 38.2
Desolvation energy	-10.4 +/- 5.7
Restraints violation energy	41.6 +/- 35.52
Buried Surface Area	1473.9 +/- 38.8
Z-Score	0.3

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 9

HADDOCK score	-101.4 +/- 11.8
Cluster size	6
RMSD from the overall lowest-energy structure	9.3 +/- 0.2
Van der Waals energy	-51.1 +/- 1.1
Electrostatic energy	-267.4 +/- 43.0
Desolvation energy	-5.4 +/- 6.2
Restraints violation energy	85.9 +/- 29.37
Buried Surface Area	1631.3 +/- 114.8
Z-Score	0.5

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 6

HADDOCK score	-94.6 +/- 10.1
Cluster size	8
RMSD from the overall lowest-energy structure	8.6 +/- 0.3
Van der Waals energy	-41.8 +/- 4.5
Electrostatic energy	-211.5 +/- 37.2
Desolvation energy	-13.5 +/- 12.9
Restraints violation energy	30.3 +/- 18.23
Buried Surface Area	1265.2 +/- 135.8
Z-Score	1.1

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 16

HADDOCK score	-94.3 +/- 9.6
Cluster size	4
RMSD from the overall lowest-energy structure	8.9 +/- 0.6
Van der Waals energy	-53.1 +/- 6.3
Electrostatic energy	-197.1 +/- 21.0
Desolvation energy	-11.1 +/- 5.8
Restraints violation energy	92.9 +/- 62.18
Buried Surface Area	1580.3 +/- 137.5
Z-Score	1.1

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

CLUSTER 14

HADDOCK score	-92.5 +/- 14.2
Cluster size	5
RMSD from the overall lowest-energy structure	8.1 +/- 0.2
Van der Waals energy	-42.2 +/- 5.4
Electrostatic energy	-218.4 +/- 31.8
Desolvation energy	-10.5 +/- 5.6
Restraints violation energy	37.7 +/- 22.56
Buried Surface Area	1339.6 +/- 112.7
Z-Score	1.3

Nr 1 best structure [Download structure](#)



Nr 2 best structure [Download structure](#)

Nr 3 best structure [Download structure](#)

Nr 4 best structure [Download structure](#)

Figure A.46: HADDOCK results for the complex hTRAF6-PepH2-hPEP1. The clusters are sorted by descending order, so the best is in top along with the best 4 representative docked models.

Annex 28

The parameters of the 2 best models from each cluster, coming from the hTRAF6-PepH2-hPEP1 docking with HADDOCK.

Table A.18: HADDOCK results of the first ranked cluster for the complex hTRAF6-PepH2-hPEP1

	1° - Cluster 8			
	Cluster8_1	Cluster8_2	Cluster8_3	Cluster8_4
Webserver	I-TASSER			
HADDOCKscore	-123,8	-123,8	-119,2	-118,1
Eair	40,1	15,0	40,8	56,5
Evdw	-47,2	-56,3	-50,2	-42,3
Eelec	-293,1	-337,8	-354,0	-366,4
Edesol	-22,0	-1,4	-2,2	-8,2
Eint	-300,3	-379,1	-363,4	-352,2
BSA ($\text{\AA}^2$)	1662,3	1790,7	1392,4	1470,2

Table A.19: HADDOCK results of the second ranked cluster for the complex hTRAF6-PepH2-hPEP1

	2° - Cluster 5			
	Cluster8_1	Cluster8_2	Cluster8_3	Cluster8_4
Webserver	I-TASSER			
HADDOCKscore	-136,1	-116,6	-115,0	-115,0
Eair	17,4	75,7	45,5	54,0
Evdw	-46,3	-44,6	-52,7	-40,1
Eelec	-384,9	-395,2	-297,6	-382,7
Edesol	-14,6	-0,5	-7,3	-3,8
Eint	-413,7	-364,1	-304,9	-368,7
BSA ($\text{\AA}^2$)	1662,5	1623,1	1615,5	1445,4

Annex 29

Percentage of conservation of the residues determined as active in HADDOCK to perform molecular docking of the complex hTRAF6-PepH2-hPEP1. This percentage was used as a comparative parameter between the best models previous from the best 2 first clusters.

Table A.20: hTRAF6 interface active residues for the hTRAF6-PepH2-hPEP1 best models. In gray are the residues not considered active. Visualization with *Pymol* with a 6Å cut-off

% Active residues conservation		
100,0%	90,9%	90,9%
8_1	8_2	5_1
372		
373		
374		
376		
392	392	392
394	394	394
396	396	396
	397	
	399	
	400	
401	401	
402	402	402
403	403	403
406	406	
	408	408
410	410	410
412	412	412
		429
		430
431		
	444	

% Active residues conservation		
100,0%	90,9%	90,9%
8_1	8_2	5_1
	445	
446	446	446
447	447	
448	448	448
449	449	449
450	450	450
451	451	451
		452
453	453	453
455	455	455
456	456	456
457	457	457
458	458	458
	459	459
466	466	466
467	467	467
468	468	468
469	469	469
470	470	470
471	471	471
472	472	472
473	473	473
474	474	474
475	475	475
476	476	476
477		
478	478	
	481	
	484	
485		
	487	

Annex 30

The RP-HPLC chromatogram of hTRANCE-R(1) before purification.

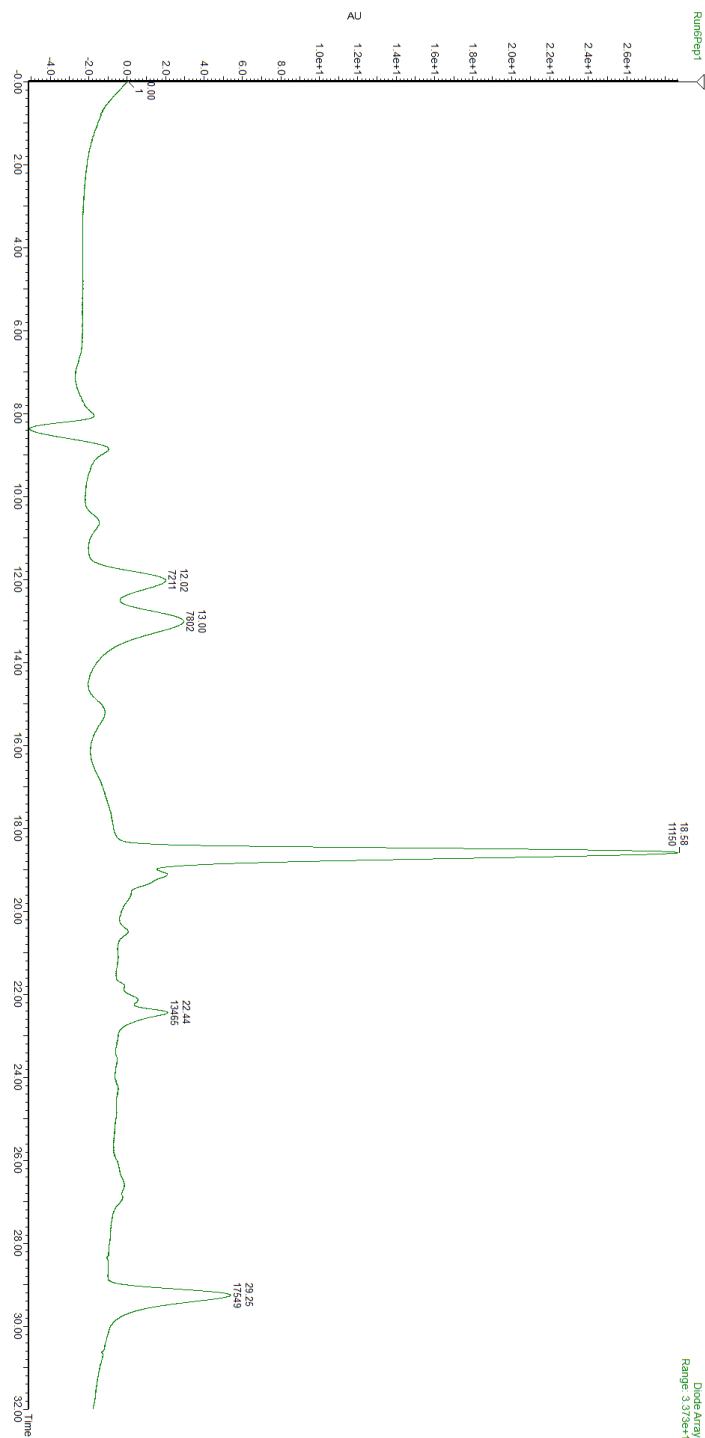


Figure A.47: Peptide hTRANCE-R(1) chromatogram before purification (preparative HPLC)

Annex 31

The RP-HPLC chromatogram of PepH2-hPEP1 before purification.

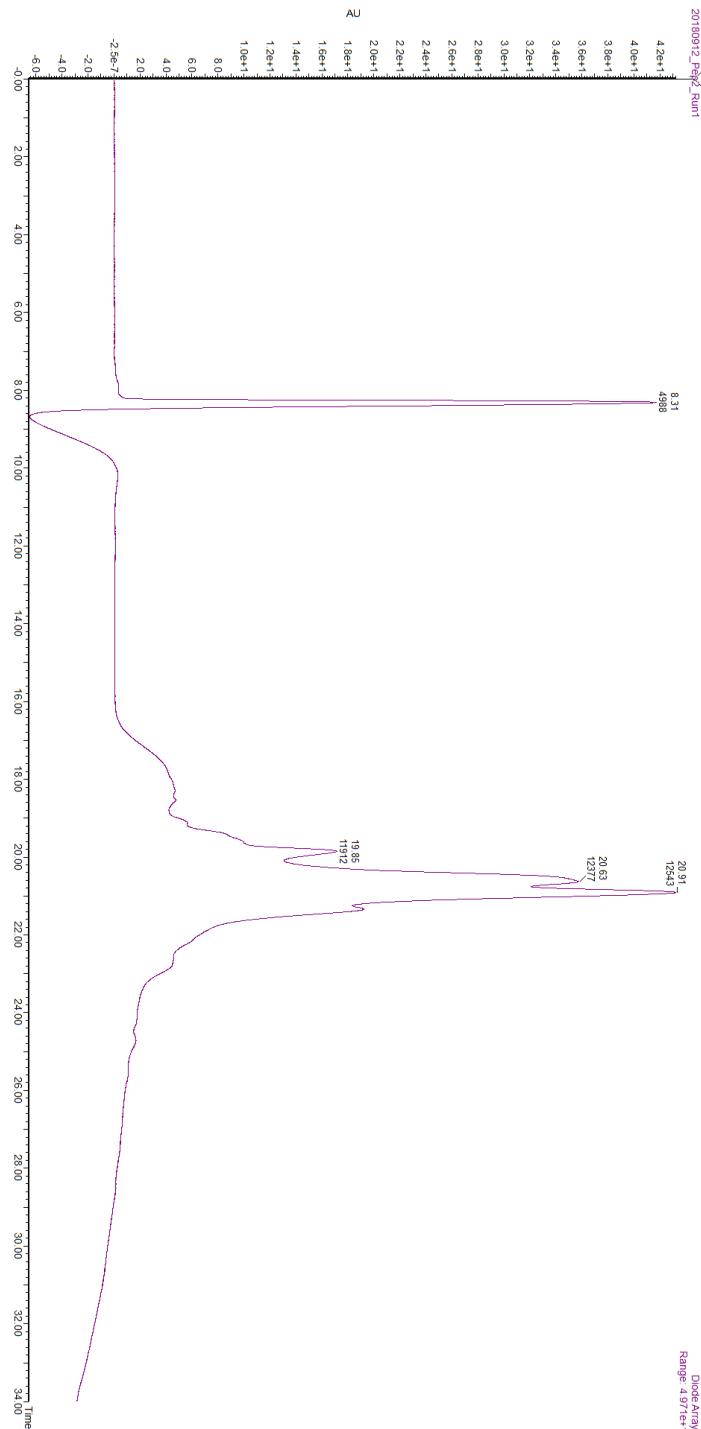


Figure A.48: Peptide PepH2-hPEP1 chromatogram before purification (preparative HPLC).