UNIVERSITY OF CATANIA

DEPARTMENT OF CHEMICAL SCIENCES

INTERNATIONAL PhD IN CHEMICAL SCIENCES – XXXI CYCLE

Giuseppe Floresta

Design and synthesis of novel compounds as fatty acids binding protein inhibitors and as gallium-68 chelators for positron emission tomography

PhD Thesis

Tutor:
Prof. Antonio Rescifina

PhD Coordinator:
Prof. Salvatore Sortino

Table of content

Abstract

Finding new small molecules targets as well as improving the diagnosis methodologies are two of the most important areas in which the researchers are spending efforts to improve our arsenal to fight cancer and other diseases. In this thesis, two different chapters are discussed. In the first one, the design of new Fatty acid binding protein 4 inhibitors is discussed. In the second one, the design of two targeting peptide bioconjugates for the detection of cancers is reported.

Fatty acid binding proteins are a class of proteins involved particularly in the transport of fatty acids in human. Recently it comes out that the Fatty acid binding proteins are an interesting molecular target for the treatment of type 2 diabetes, other metabolic diseases and some type of cancers. In this chapter (first), three new molecules inhibitors of the Fatty acid binding protein 4 are designed, using computer-aided drug design methodologies, and synthesized. The three molecules, AST_1–3, were synthesized and tested against the target protein and showed an IC_{50} between 3.70 and 5.59 μM. Moreover, a huge number of different other molecules were theorized to be as effective as the three synthesized. Among them, some are derived from a virtual screening of an FDA approved drugs database and some from the bioisosteric scaffold-hopping analysis of a note inhibitor of the fatty acid binding protein 4 (BMS309403).

In the second chapter, two different targeting peptides, against two proteins involved in particular types of cancer (c-Met and GLP-1), were synthesized and then conjugated to a small molecule able to chelate gallium, for their potential applicability as PET tracers. The two compounds were then evaluated as effective 68-gallium chelating compounds and the result showed their capability in the binding of the ^{68}Ga.

Chapter 1. Design, synthesis, and biological evaluation of small molecule inhibitors of fatty acid binding protein 4

1.1. Introduction

Fatty acids (FAs) are a class of carboxylic acids with a long aliphatic tail. The different FAs carry out many different vital functions in the organism.[1] They are one important source of energy, which is stored in triacylglycerol and produced in muscles and liver. They are used for the biosynthesis of complex lipids, such as phospholipids and cholesterol, and they are also hormones and signaling compounds. These essential nutrients are normally obtained from the diet, released from the storage in adipocytes, or synthesized from glucose in the liver. There are different pieces of evidence showing that chronically elevated plasma fatty acid leads to some physiological disorders.[2] The elevated fatty acid levels in circulation are associated with the pathogenesis of type 2 diabetes,[3] obesity,[4] and atherosclerosis.[5] Obviously, because the fatty acids are slightly soluble in water their trafficking requires a cluster of specific carrier proteins; normally, the different FAs are associated with albumin, lipocalins and fatty acid-binding proteins (FABPs), which highly increases their water solubility and thus facilitates their transport.[6] FABPs are members of the superfamily of lipid-binding proteins (LBP). So far, nine different FABPs have been identified. The different FABPs have tissue-specific distribution, among them: L (liver), I (intestinal), H (muscle and heart), A (adipocyte), E (epidermal), Il (ileal), B (brain), M (myelin), and T (testis). Furthermore, the nine different types of FABPs can be divided into two groups: those associated with the plasma membrane ($FABP_{PM}$) and the ones localized in the cytoplasm ($FABP_C$).[7] The adipocyte FABP (A-FABP), also called aP2 or FABP4, is highly expressed in adipocytes and it is regulated by peroxisome-proliferator-activated receptor-c (PPARc) agonists, by the level of insulin and by fatty acids.[8] It was reported that animal models with a deficiency of FABP4 are protected against the development of insulin resistance.[9] These mice also show better performances in both insulin and glucose

tolerance tests. Studies in FABP4-deficient mice[10, 11] have shown that FABP4 has a significant role in many aspects of the metabolic syndrome. The lack of FABP4 partially protects mice against the progress of insulin resistance associated with genetic or with obesity. The adipocytes of these knockout animal models have a reduced efficiency of lipid transport both *in vitro* and *in vivo*. Interestingly, recent studies demonstrated that FABP4 is highly expressed in macrophages and regulated by the phorbol 12-myristate 13-acetate, by lipopolysaccharides, by oxidized low-density lipoproteins, and by PPARc ligands.[12] The macrophage is a fundamental site of FABP action, and total or macrophage-specific FABP4-deficiency leads to a marked protection against early and advanced atherosclerosis.[13] These findings show a significant part for FABP4 in the progress of different components of the metabolic syndrome through its different actions in adipocytes and macrophages. Aside from genetic approaches, pharmacological agents that modify FABP function can potentially mimic the phenotype of FABP4-deficient mice. Therefore, small molecules that inhibit FABP4-mediated responses might serve as potential candidates for many components of metabolic syndrome, such as insulin resistance, type 2 diabetes, and atherosclerosis. FABPs play also an important role in carcinogenesis.[14] Modified FABPs expression patterns were described for prostate, bladder, renal cell carcinoma and other types of cancer cells,[15-17] but the biological function of FABPs in cancer remains unclear.[18] The aim of this project is to identify novel potent inhibitors of FABP4. Knowing the fundamental interactions of recognized inhibitors with the protein, new compounds were designed, with the help of computer-aided drug design, synthesized and biologically evaluated.

1.1.1. Physiological properties and functions of FABPs

Transport of FAs into cells is divided into three steps: (1) adsorption-binding to the outer leaflet of the plasma membrane; (2) crossing the membrane; (3) desorption-leaving the cytosolic leaflet of the plasma membrane. To allow an effective FAs flow proteins must catalyze each

step. There are two different types of FAs translocation across plasma membranes: (1) simple diffusion and (2) protein-mediated translocation. One of the classes of the proteins involved in the FAs translocation is a 43-kDa $FABP_{PM}$, which was purified from hepatocytes, adipocytes, jejunal enterocytes, and cardiac myocytes. Clarke et al. provided the first evidence that overexpression of $FABP_{PM}$ in a mammalian tissue increase the FA transport. On the other hand, the $FABP_c$ can accelerate FA uptake in different ways: increasing the rate of dissociation from membranes by raising the solubility of FA,[19] and/or improving FA transfer to acceptor membranes by direct interaction with the phospholipid bilayer or by an aqueous diffusion-mediated process. $FABP_c$ proteins facilitate not only FA desorption, but also the cytoplasmic diffusion. A role of FABPs in intracellular FA movement was supported by a series of experiments; thus, FABPs can be defined as transport proteins.[6]

Regulation of the lipid metabolism, due to the importance of energy homeostasis, involves many control systems, which need a precise and correct coordination.[20] Response to different signals triggers the activation of specific transcription factors and the FAs can act such as signaling molecules. Several mechanisms of gene transcription mediated by FAs have been described. They include a series of events that lead to the modification of a transcription factor, for example: by the binding to a particular transcription factor and the subsequent activation of the transcription factor, by modifying the mRNA stability, and also by influencing the transcription factor expression.[21] It has been revealed that increased FAs concentration start the process of preadipocyte differentiation and the expression of terminal differentiation-related genes, among them FABP4.[22, 23] One important group of the transcription factors cooperating with FABPs is the peroxisome proliferator-activated receptor (PPAR) family.[24] PPAR family members are nuclear receptors proteins well-knows for the regulation of the transcription of many genes involved in lipid metabolism.[25] FABPs may also have an effect on ligand-dependent transactivation of PPARs. Members of the FABP family are involved in

modulation of different cells growth and proliferation. L-FABP induced by two classes of peroxisome proliferators promotes mitogenesis of hepatocytes. H-FABP regulates cardiomyocyte growth and differentiation in neonatal mouse hearts and stimulates an increase in cell surface area, leading to cardiac myocyte hypertrophy. Furthermore, it specifically inhibits the growth of normal mammary epithelial mouse cells and is proposed as the breast tumor suppressor gene.[26]

1.1.2. The gene and its expression and regulation

Despite their large number, wide tissue distribution, and sequence variations, all the fatty acid binding proteins presumably derive from a single ancestral gene. The FABPs gene is found as a single copy and is similar to that of other different intracellular lipid binding proteins (iLBPs). The gene is composed of four exons separated by three introns.[27] The gene regulatory elements upstream of the gene differ greatly within the family. An enhancer element is thought to be necessary and sufficient for FABP expression in differentiated adipocytes. This region may bind heterodimers of peroxisome proliferator-activated receptors and retinoid X receptors, which in turn bind a range of metabolites and drugs.[28] Other FABP gene regulatory elements include a glucocorticoid response element for positive regulation, a CCAAT/enhancer binding protein site that responds to a decrease in insulin, and an activator protein 1 site that binds c-fos/c-jun heterodimers. A negative regulatory element also overlays part of this site.[28, 29]

1.1.3. Protein structure

Although the amino acid sequence identities of the iLBPs range from 20% to 70%, all known proteins structures, including FABPs, have a common tertiary structure as revealed on the basis of X-ray crystallography and nuclear magnetic resonance.[27] The tertiary structure is common to all the family members and is characterized by a twisted barrel surrounding a hydrophobic core. The barrel is composed of ten antiparallel strands, organized into two sheets

oriented almost orthogonally, and on one end the barrel is capped by a little helix-turn-helix pattern.[30] Inside the barrel, there is a large water-filled cavity lined with polar and hydrophobic amino acids. Interestingly, only a small portion of the cavity volume (1/3 to 1/2) is occupied by the single bound ligand. The carboxyl group of the FAs is oriented inwards and coordinated through electrostatic interactions with a tyrosine and two arginine residues. Helical *N*-terminus seems to be significant for the effective interplay between FABPs and membranes and dictates the overall rate of FA transfer. A comparison between the amino acids sequences of different FABPs revealed a wide variance in the primary structure. Based on amino acids sequence, FABPs can be divided into 3 groups: (1) L-FABP and Il-FABP; (2) H-FABP, B-FABP, E-FABP, M-FABP, A-FABP and T-FABP; and (3) I-FABP. Obviously, this division is reflected in the type of natural ligands of different classes of proteins. All proteins of group 1 (L-FABP and Il-FABP) are capable of binding FAs and bulky ligands, such as bile salts, cholesterol, and heme. Members of group 2 (H-FABP, B-FABP, E-FABP, M-FABP, A-FABP, and T-FABP) bind FAs and additionally retinoids and eicosanoids. I-FABP, the only member of the group 3, binds solely FAs, but in a different conformation than other FABPs.

1.1.4. The binding cavity

Since the first apo-FABP crystal structure appeared in 1992,[31] many holo-FABP structures with a range of ligands and several mutant structures have been solved and studied. The side chain variations generally conserve their hydrophobic character and therefore may only affect the inner cavity shape. The side chains engage a hydrogen bond to the carboxylate end of FAs, R126, and Y128, as well as R106, through a conserved water molecule, and have been shown to play a critical role in FABP ligand binding specificity and affinity.[32] Figure 1.1 shows the internal FABP4 binding site along with an internal water network that has been found in many of the FABPs crystal structures. Nevertheless, R106 is not strictly conserved; variation may occur with another basic residue, lysine. Since R126 and Y128 are found in many

iLBPs that bind different ligands, these residues are involved with binding affinity but do not exclusively determine the specificity of ligand binding.

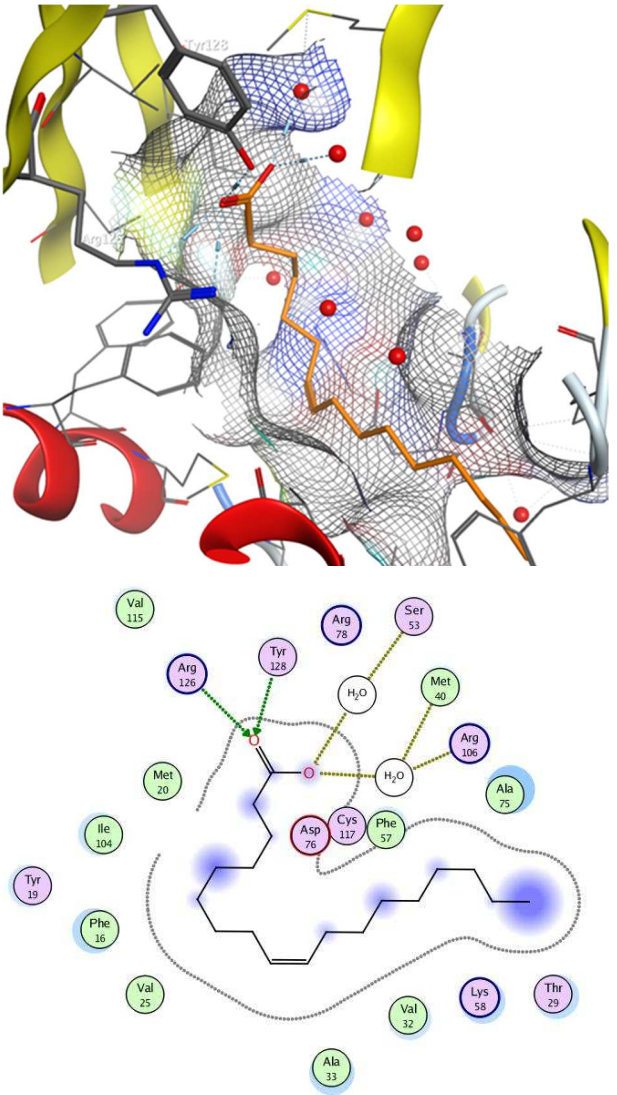


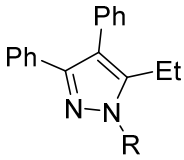
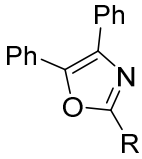
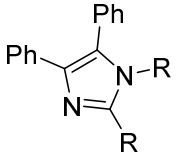
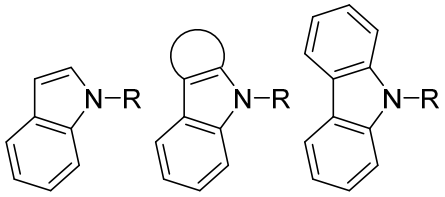
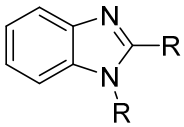
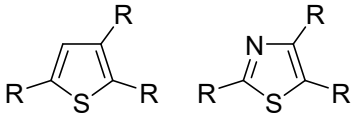
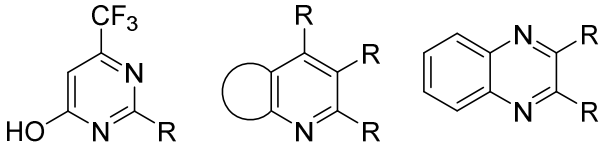
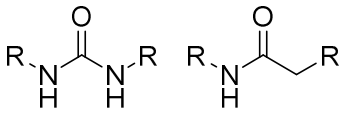
Figure 1.1. Crystal structures and interaction of the apoprotein with a natural ligand, oleic acid. PDB structure: 1LID.

1.1.5. State of art

In the past years, a variety of effective FABP inhibitors have been synthesized for antiatherosclerosis and anti-diabetes treatments, including derivatives of niacin, quinoxaline, aryl-quinoline, bicyclic

pyridine, urea, aromatic compounds, and other heterocyclic compounds.[33] Because FABP3 (H-FABP) and FABP4 are highly similar in their primary amino acid sequence and also in their three-dimensional structures, the issue of selectivity towards FABP3 and FABP4 must be considered during the research on FABP inhibitors. At present, no FABP4 inhibitors are in the clinical research phase. Several potent small molecules have been identified as FABP4 inhibitors, such as BMS309403 and HTS01037, but only BMS309403 has been systematically studied in both *in vitro* and *in vivo* animal diabetic models. Although the antidiabetic effects of BMS309403 have been well validated in mouse models, it has yet to be approved for clinical use. In Table 1.1 are shown the various classes of FABP4 inhibitors, among them: pyrazole derivatives, oxazole derivatives, imidazole derivatives, indole derivatives, benzimidazole derivatives, thiophene and thiazoles derivatives, pyrimidine, bicyclic pyridine and quinoxaline derivatives, and urea and carbamoyl derivatives. In Figure 1.2 are reported some representative example of the different classes of inhibitors designed and synthesized until now whereas in Table 1.2 there are reported their biological evaluations. Inspection of their structural features highlighted that the five-membered heterocyclic ring and the carboxylic group seem to be crucial for anti FABP4 potency.

Table 1.1. Various chemical classes of the FABP inhibitor.

Chemical Classification	**Chemical Structure**
Pyrazole derivatives	
Oxazole derivatives	
Imidazole derivatives	
Indole derivatives	
Benzimidazole derivatives	
Thiophene and Thiazoles derivatives	
Pyrimidine, Bicyclic Pyridine and Quinoxaline derivatives	
Urea and carbamoyl derivatives	

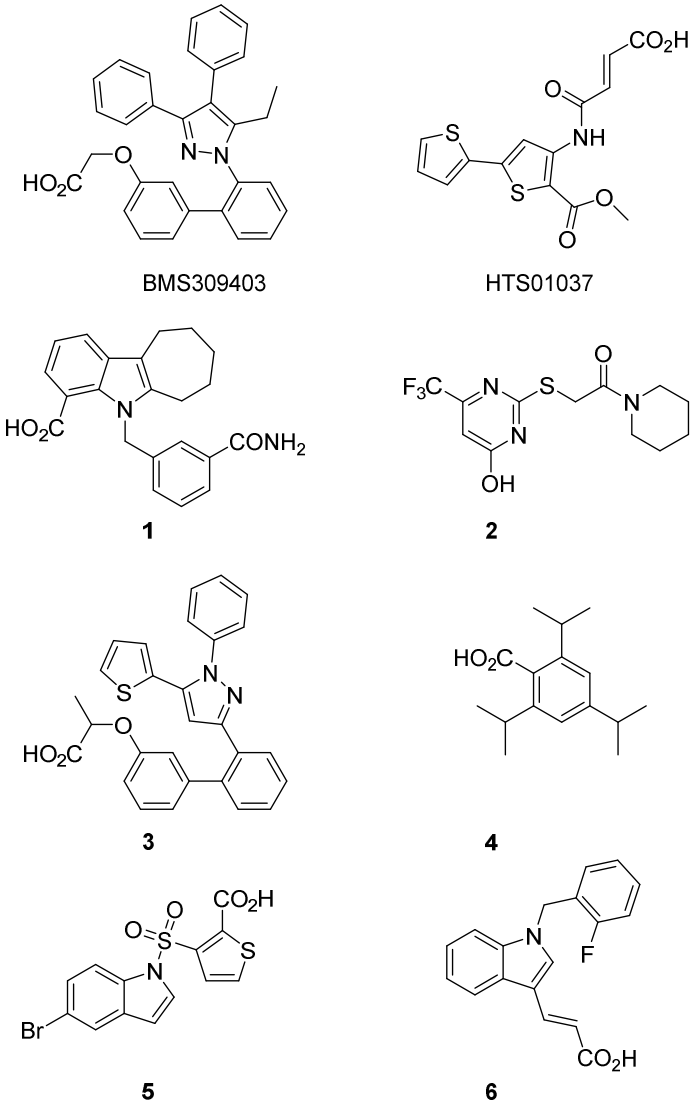


Figure 1.2. Structures of different FABP4 selective inhibitors.

Table 1.2. Measured biological activity of the FABP4 inhibitors.

Compound	IC_{50} or K_i μM
BMS309403[34]	0.71*
HTS01037[35]	0.67**
1[36]	0.45*
2[37]	1.00*
3[38]	0.001**
4[39]	4.00*
5[40]	1.30*
6[41]	0.033**

*Measured IC_{50}; **Measured K_i

Three different virtual screening studies with libraries of different compounds have been already published and these have led to interesting results.[42, 43] Wang et al. have screened two different libraries, one of natural compounds and one of FDA approved drugs, against FABP4. After the virtual screening and the biological test of the best-scored compounds, they have found that the quercetin, other flavonoids, and levofloxacin - an FDA already approved drug for the treatment of different bacterial infections - shoved a good percentage of inhibition of FABP4. Moreover, during the second year of Ph.D., I have collected and reviewed all of the compounds acting as FABP4 inhibitors.[33] The database of the collected compounds was then used as a dataset for the building of a 3D-QSAR model, Section 1.4.

1.2. Computational approach for FABP4 inhibitor design

Computer-aided molecular design is an important aspect of drug design and discovery. In general, these methods for drug discovery and design can be divided into two categories: structure-based and ligand-based drug design.[44] In the first category, using available 3D structural and other important biological information concerning the target protein, the binding energy of small molecule interacting with proteins targets are calculated. The 3D structural information is obtained either from experiment (X-ray or NMR determination of the protein

structure) or from reliable computational methods (i.e. homology modeling). Molecular mechanical (MM) as well as quantum mechanical (QM) approaches may be used to identify small molecules with binding affinity, and then to refine the structure of these small molecules toward higher binding affinity ones.[45-48] Often, the 3D structural information of the target protein is not available and there is not a good template for homology modeling. In this case, the ligand-based methods such as QSAR are preferentially used. From the dataset obtained from a series of lead compounds, 2D or 3D descriptors are calculated. QSAR equations are derived based on the 2D descriptors and, usually, a pharmacophore model is created from the 3D descriptors. The QSAR equation and the pharmacophore model are used to suggest new compounds with improved activity.[49]

For hit identification and optimization the docking of small molecule and the scoring of their potential complementarity to the binding site of FABP4 were used as a preferential tool in the first part of the project. Later, a 3D-QSAR equation was developed to describe the pharmacophore of the FAPB4 and to virtually score different compounds. Among them: the best-scored compounds derived from the structure-based approach, but also a library of compounds derived from a scaffold hopping study of BMS309403 and a library of FDA approved drugs.

1.3. Structure-based approach for FABP4 inhibitor design

In the first part of the design of novel FABP4 inhibitors, the crystal structure of the protein complexed with BMS309403 (PDB code: 2NNQ) was used, by means of docking calculations.[50] After an initial study of the binding site, the ligand structure was modified to better fit the target.

As shown in Figure 1.3, some free areas can be localized close to the two phenyl groups, in position 3 and 4 of BMS309403. Therefore, it was decided to design the first set of molecules with different substituents in these positions. The BMS309403 structure was used as a starting point for a systematic study of chemical modifications, to

improve the potency, and to explore the chemical space inside the binding pocket. The different poses were generated with AutoDock Vina and scored with the scoring function present in the same software. The results and the relative energy scored values are listed in Tables 1.3 and 1.4.

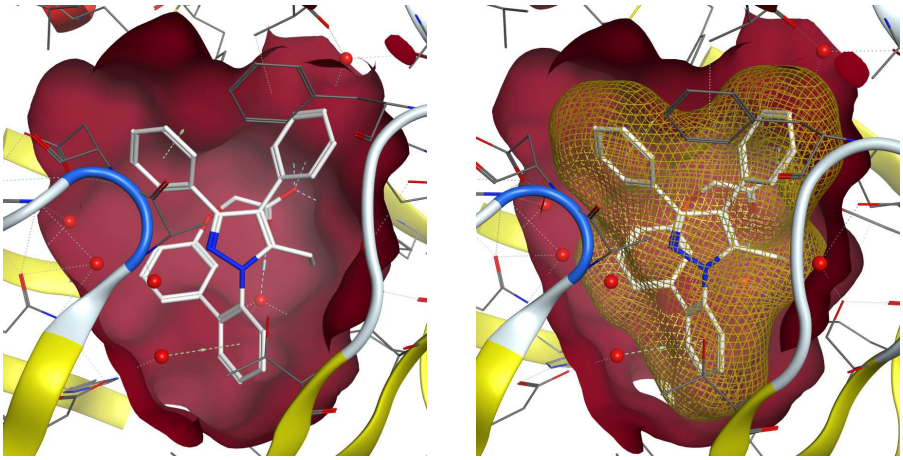
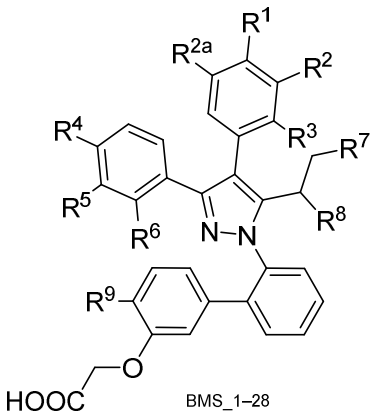


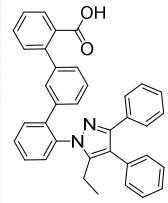
Figure 1.3. Binding site of FABP4 in complex with BMS309403 (PDB code: 2NNQ).

Table 1.3. Score values of BMS_1–28, modified structures of BMS309403.

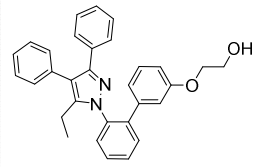


BMS_1–28

		kcal/mol
BMS309403		–10.5
BMS_1	R^1 = F	–10.7
BMS_2	R^1 = Cl	–10.0
BMS_3	R^2 = COOH	–10.4
BMS_4	R^2 = $CONH_2$	–10.2
BMS_5	R^3 = COOH	–9.8
BMS_6	R^3 = $CONH_2$	–11.4
BMS_7	R^4 = F	–10.9
BMS_8	R^4 = Cl	–9.9
BMS_9	R^5 = COOH	–11.0
BMS_10	R^5 = $CONH_2$	–10.8
BMS_11	R^6 = COOH	–10.2
BMS_12	R^6 = $CONH_2$	–10.6
BMS_13	R^2 = F	–11.0
BMS_14	R^2 = Cl	–10.9
BMS_15	R^3 = CH_2OH	–10.4
BMS_16	R^7 = SO_3H	–10.7
BMS_17	R^7 = COOH	–10.3
BMS_18	R^8 = COOH	–11.0
BMS_19	R^8 = SO_3H	–11.0
BMS_20	R^7 = Ph	–10.6
BMS_21		–10.1
BMS_22		–9.9
BMS_23	R^9 = F	–10.8
BMS_24	R^9 = Cl	–10.2



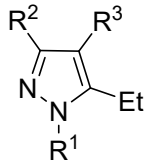
BMS_21



BMS_22

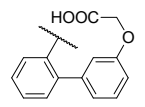
		kcal/mol
BMS_25	R^{2a} = F R^3 = $CONH_2$	–11.3
BMS_26	R^{2a} = F R^5 = COOH	–11.5
BMS_27	R^3 = $CONH_2$ R^5 = COOH	–11.4
BMS_28	R^{2a} = F R^3 = $CONH_2$ R^5 = COOH	–10.8

Table 1.4. Score values of BMS_29–36, modified structures of BMS309403.

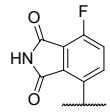


BMS_29–36

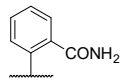
R =



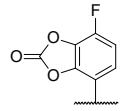
a



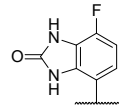
d



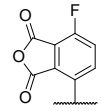
b



e



c



f

	R^1	R^2	R^3	kcal/mol
BMS_29	a	b	c	–11.8
BMS_30	a	c	b	–12.1
BMS_31	a	b	d	–10.9
BMS_32	a	d	b	–11.8
BMS_33	a	b	e	–11.3
BMS_34	a	e	b	–12.7
BMS_35	a	b	f	–10.9
BMS_36	a	f	b	–11.9

As shown by the docking results (Tables 1.3 and 1.4), in quite a number of cases different molecules have a better score than the starting compound (BMS309403), so, presumably, they should possess improved biological activity.

Moreover, the diversity of the amino acids residues present in the active site of the FABP4, compared to the ones of FABP3,[51] and involved in the interaction with the ligand, can be exploited to increase the selectivity of the novel compounds against the FABP4 (Figures 1.4 and 1.5).

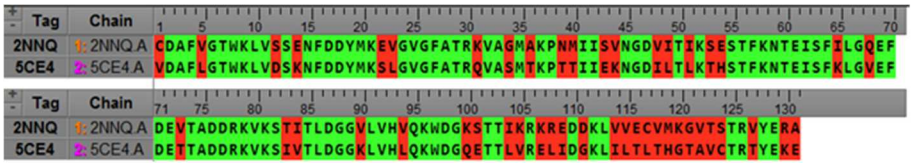


Figure 1.4. Sequence alignment of human FABP4 (2NNQ) and FABP3 (5CE4). The sequence identity is 64% (green-highlighted).

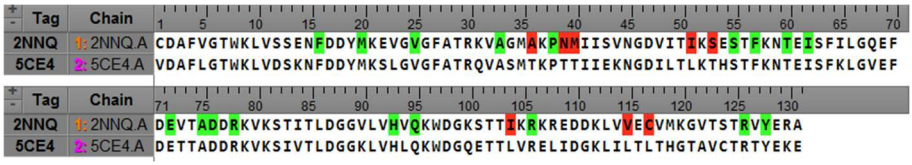


Figure 1.5. Sequence alignment of human FABP4 (2NNQ) and FABP3 (5CE4). The amino acids of the binding site are marked; those in red are different between the two proteins.

As shown in Figure 1.5 the active sites of the two different proteins are formed by 26 amino acid residues, but only eight are different between the two different forms. These differences of the amino acids expression in the active sites could be used for enhancing the activity against FABP4 compared with FABP3.

Analyzing the first set of designed molecules, BMS_1–24, we can see that they are placed in a similar manner to the original ligand (Figure 1.6). A halogen in the *para*-position of the two aromatic rings is tolerated, but only if it is fluorine; in fact, the score is worse with atoms with a larger radius. Positive results are obtained by the substitution of different position with a carboxyl and/or amide group. Both the carboxyl and the amide group are well tolerated when

located in the R^2 position. The substitution of the R^3 position with an amide group leads to BMS_6 that is the molecule of the series with the best score. The two groups can still be inserted in the position R^5 and R^6 and result again in a score better than the starting molecule. The substitution at other positions of the molecule, R^7–R^9, has not given promising results. In fact, in these positions, the resulting molecules are hardly able to create new interactions with the protein.

Concluding, the most interesting result of the series of molecules BMS_1–24 is the one of BMS_6; in fact, the molecule presents an excellent score compared to the starting compound (–11.4 *vs*. –10.5 kcal/mol). The better-calculated binding energy is due to a hydrogen bond with SER53 of the active site; importantly, this serine is replaced by THR53 in FABP3. Therefore, presumably, this interaction could enhance both the power and the selectivity of the molecule (Figure 1.7).

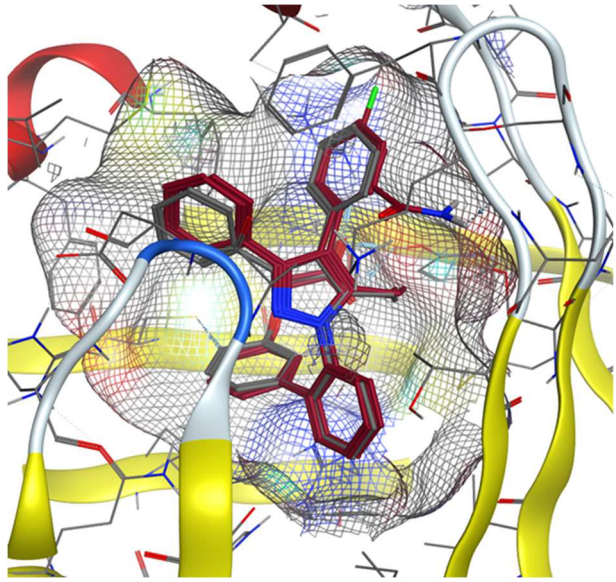


Figure 1.6. Poses of BMS309403 (grey), BMS_1 and BMS_6 (both in burgundy).

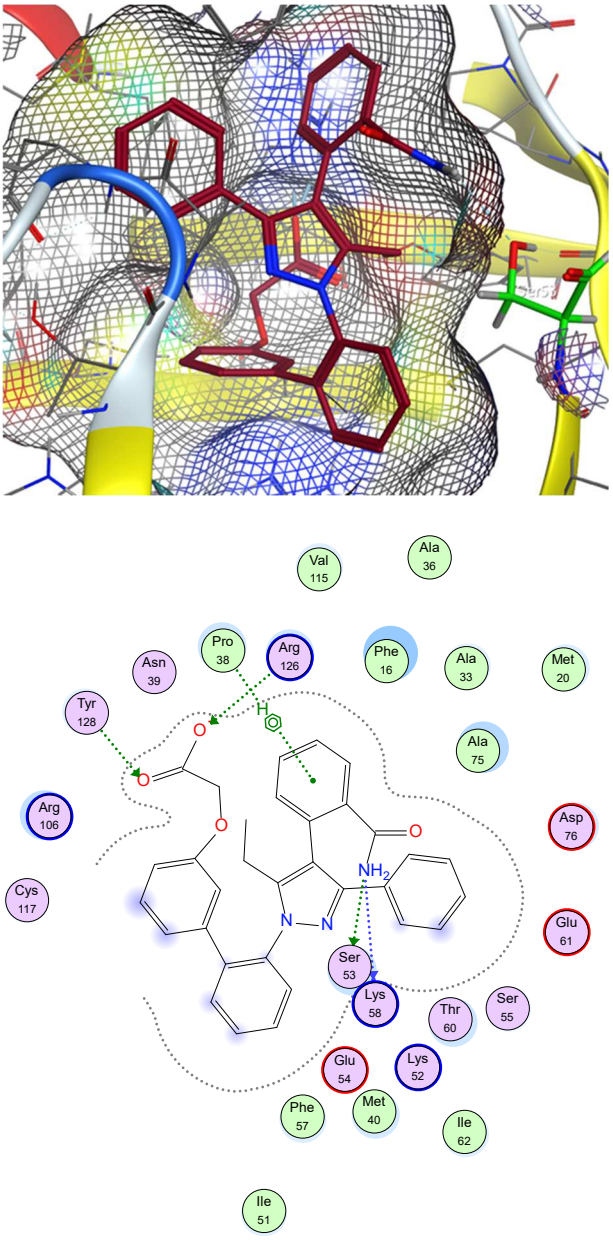


Figure 1.7. Interaction of BMS_6.

Better results were achieved by the second set of molecules, BMS_25–28, where two or three positions were substituted. With the third set of compounds, BMS_29–36, the steric hindrance of the molecules in one of the aromatic rings was increased, with interesting results. Apparently, the binding cavity tolerates a newly condensed

heterocycle in one of the two phenyl groups in position 3 and 4. The new bulky group can be easily accommodated in the binding pocket; this results in an improvement of the binding energy (Figure 1.8). Several molecules have been evaluated and in all of them, the amide group is always able to interact with the SER53. The best molecule of this series, BMS_34, has a score of –12.4 kcal/mol due to the synergism between the hydrogen bond with SER53 and the new interactions promoted by the heterocycle condensed to phenyl at the position 3.

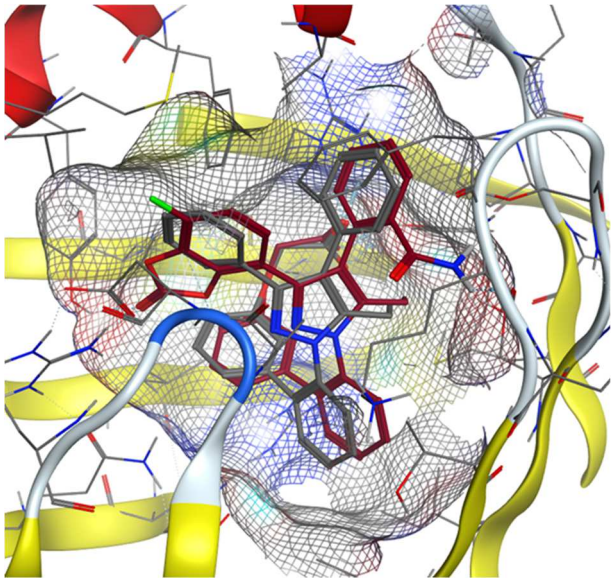
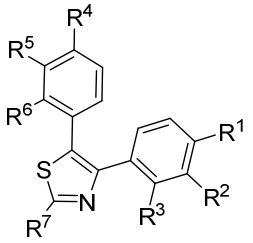


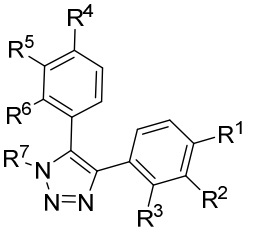
Figure 1.8. Poses of BMS309403 (grey), and BMS_34 (burgundy).

1.3.1. Design of new inhibitors with thiazole and triazole core

The new information acquired during the first part of the project, dedicated to the study of the binding cavity and where new possibilities for selective interaction were assessed, have been used in the design of new compounds with thiazole and triazole core (Figure 1.9). The central core of the BMS309403 has not been already reported to be substituted with the selected thiazole and triazole scaffolds.



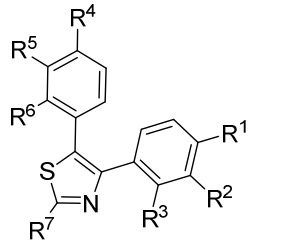
SMS_0–11



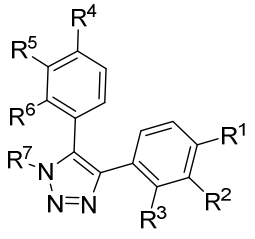
TAR_0–11

Figure 1.9. Newly designed compounds with thiazole and triazole core.

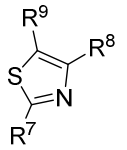
Table 1.5. Structures and scores of SMS and TAR series.



		kcal/mol
SMS_0	R^{1-6}= H	–10.9
SMS_1	R^3 = $CONH_2$	–11.4
SMS_2	R^6 = $CONH_2$	–11.4
SMS_3	R^5 = COOH	–10.7
SMS_4	R^2 = COOH	–11.5
SMS_5	R^1 = F	–11.2
SMS_6	R^4 = F	–10.5
SMS_7		–11.8
SMS_8		–12.7
SMS_9		–11.4
SMS_10		–12.7
SMS_11		–11.7

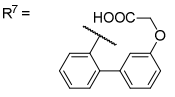


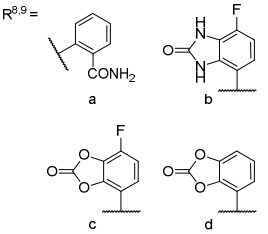
		kcal/mol
TAR_0	R^{1-6}= H	–10.2
TAR_1	R^3 = $CONH_2$	–11.1
TAR_2	R^6 = $CONH_2$	–10.6
TAR_3	R^5 = COOH	–10.9
TAR_4	R^2 = COOH	–10.7
TAR_5	R^1 = F	–10.0
TAR_6	R^4 = F	–10.5
TAR_7		–10.9
TAR_8		–10.9
TAR_9		–12.2
TAR_10		–10.9
TAR_11		–10.9



SMS_7–11

TAR_7–11





	R^8	R^9
SMS_7, TAR_7	a	b
SMS_8, TAR_8	b	a
SMS_9, TAR_9	a	c
SMS_10, TAR_10	c	a
SMS_11, TAR_11	d	Ph

The design started by drawing and scoring new compounds with thiazole (SMS_0–11) and triazole central core (TAR_0–11). The designed molecules are shown in Table 1.5. Even in this case, the crystal structure of the FABP4 in complex with BMS309403 (2NNQ) was used as a starting point for docking studies of this set of virtual compounds.

As we can see from the docking results, different series of molecules show better score compared to the starting ligand BMS309403. The analysis of the generated poses shows that when the aromatic substituents to the central core are positioned in 3- or 4-position (as in BMS309403) the best-scored pose is very similar to the inhibitor in the crystal form (Figure 1.10).

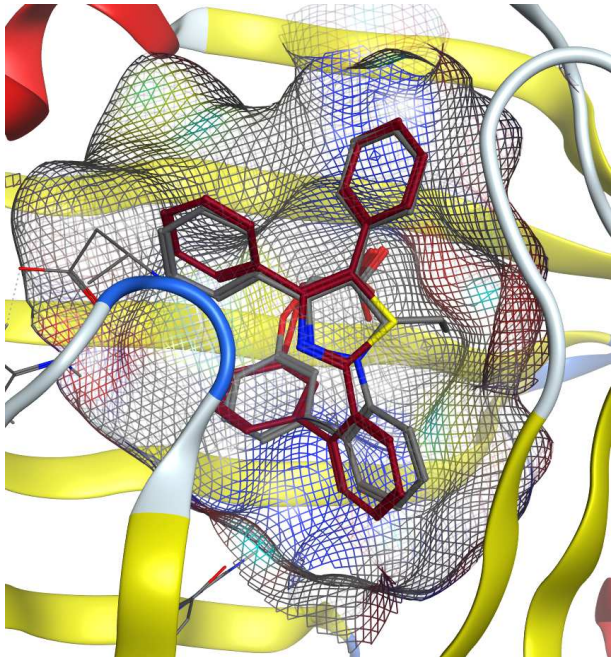


Figure 1.10. Poses of BMS309403 (grey), and SMS_0 (burgundy).

However, when the substituents are in different positions the poses and the generated interactions are completely different (Figure 1.11).

As shown in Figure 1.10, the compounds of the SMS series are able to get similar interaction to that of BMS309403. Otherwise, the presence of the two phenyl groups in a different position, TAR series, leads to a completely different pose of the molecule and to the loss of

fundamental interactions necessary to the inhibitory activity. As a matter of fact, some triazole derivatives, like TAR_1 and TAR_9, have shown interesting results, but the interactions with the binding site are completely different to that of the other series of designed compounds and so as the interaction of BMS309403.

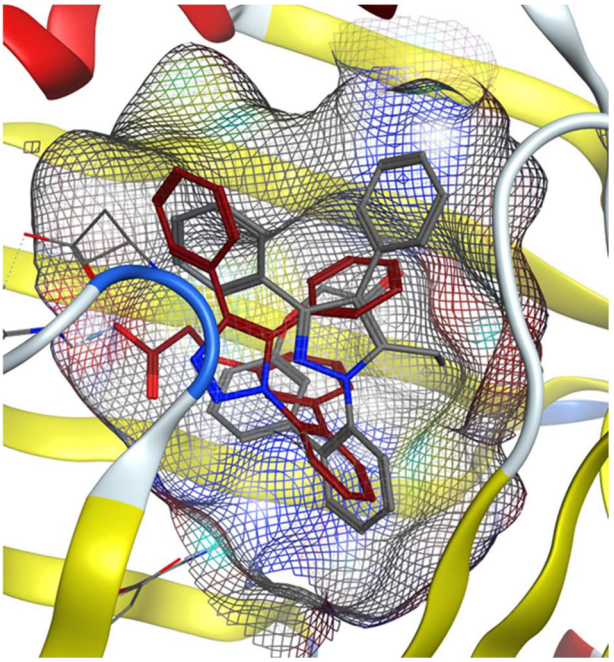


Figure 1.11. Poses of BMS309403 (grey), and TAR_0 (burgundy).

Given the excellent results of the SMS series, it was decided to proceed in the design of other compounds with thiazole core, and discard, the idea of producing molecules with triazole nucleus.

1.3.2. Novel molecules with thiazole scaffold

Considering that the positions 4 and 5 of the thiazole ring have been extensively studied, the focus was now turned to the modification of the biphenyl group, which connects the thiazole 2-position (Figure 1.12) and the fundamental carboxylic group. It was decided to replace one of the two phenyls with a triazole ring, which can be easily inserted by a click-chemistry reaction, namely the azide-alkyne Huisgen cycloaddition. Four different series, with the triazole portion in a different position, have been designed and *in silico* evaluated (Figure

1.13).

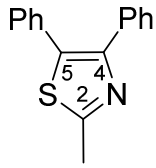


Figure 1.12. Studied position of thiazole core.

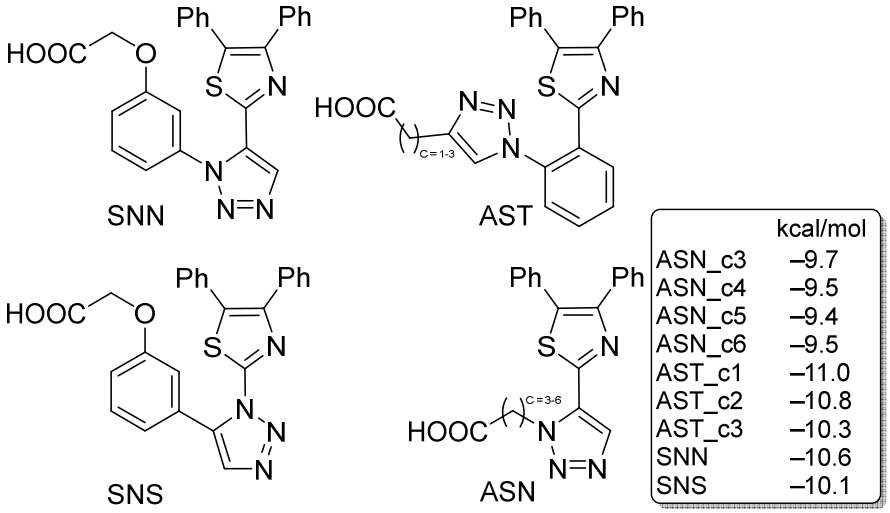
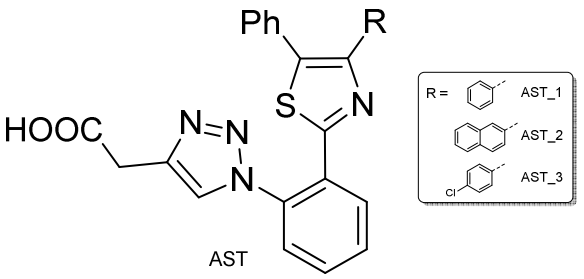


Figure 1.13. Structures of novel designed thiazoles.

The molecules are evaluated, of course, using the same software and the same crystal structure of the previous series. In Figure 1.13 the scoring values of the newly designed molecules are shown. The most powerful molecules, according to these data, seem to be the molecule named AST_c1 (AST_1), with only one carbon atom as a linker between the triazole core and the carboxylic acid moiety.

Considering the molecule promising from the synthetic point of view, other two analogs of this compound were evaluated. The phenyl group in position 4 of the thiazole core was substituted with a naphthyl and with a 4-chlorophenyl group, in order to enhance the steric hindrance, as suggested by the first series of designed compounds (BMS); the results are shown in Figure 1.14.



	Score values (kcal/mol) from AutoDock Vina	AutoDock4
BMS309403	–10.5	–11.2
AST_1	–11.4	–11.6
AST_2	–12.2	–13.4
AST_3	–11.8	–12.4

Figure 1.14. Structures and scores of AST_1 and AST_2.

Compounds AST_1–3 were evaluated using AutoDock Vina and AutoDock4, as shown in Figure 1.14, and both methodologies gave excellent results. The predicted binding poses and the interactions are shown in Figures 1.15 and 1.16, respectively. Given the excellent results of the docking of these molecules in the FABP4 compared with the overall designed compounds (Figure 1.17), it was decided to focus the attention into these three compounds, after a secondary evaluation of the molecules by the prediction capabilities of a 3D-QSAR model (Section 1.4.6).

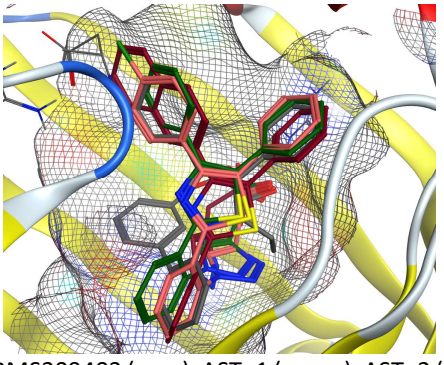


Figure 1.15. Poses of BMS309403 (grey), AST_1 (green), AST_2 (burgundy) and AST_3 (pink).

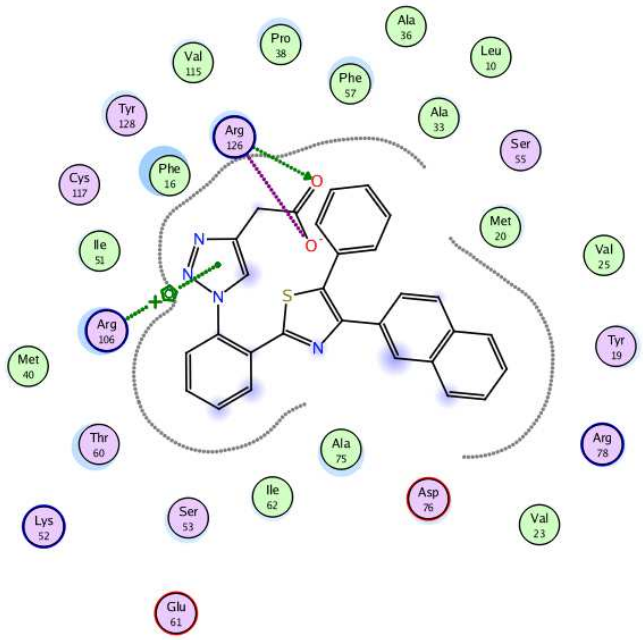


Figure 1.16. Calculated interaction of AST_2 with FABP4.

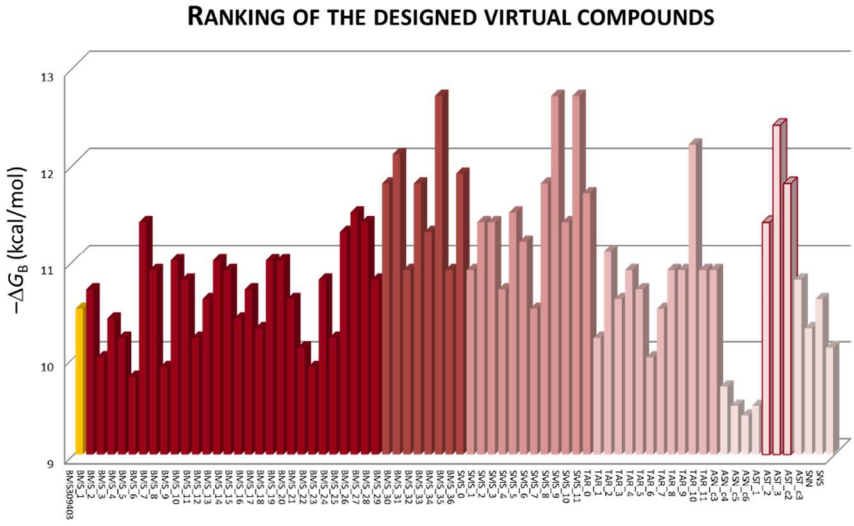


Figure 1.17. Calculated binding energies with AutoDock Vina of the molecules designed in the structure-based approach compared to the binding energy of BMS309403 (yellow-left). AST 1–3 bars are localized on the right of the graph surrounded by a red line.

1.4. Ligand-based approach for FABP4 inhibitor design

Once finished the design, by means of the structure-based approach, of the three compounds (AST_1–3), a 3D-QSAR model for the most potent and selective FABP4 inhibitors was developed. After an accurate selection of compounds, that have already been synthesized and tested against the FABP4 by other researchers, which we published as a review,[33] an internal dataset of molecules was used to build a 3D-QSAR model. The model was built using a software of the Cresset group, Forge.[52]

As we already said, the computer-aided molecular design is an important aspect of drug design and, generally, these methods can be divided into two categories: structure-based and ligand-based drug design.[44] In section 1.3 was used the structure-based approach whereas in this section the attention is focused on the development of a 3D-QSAR model for the already reported FABP4 inhibitors molecules.

Many 3D-QSAR processes determine descriptors by calculating structure properties at the interception points of a 3D grid, which surround the entire space of the aligned molecules. This is due because these methods have no way of knowing which region of space around the molecules is likely to be relevant to the molecular activity. However, Cresset's field point description of molecules provides details about the area of space about a molecule relevant to the molecular description. The 3D-QSAR in Forge is characterized by the fact that uses probe positions that are definite directly from the field points of the aligned molecules training set, and only these positions are used to describe the electrostatic potential and the volume of each molecule. This 3D-QSAR model can help to unfold SAR data and to predict an activity value for new molecules and, on the other hand, will be used to further investigate the activity of AST_1–3.

1.4.1. Biological data and Molecular modeling

As mentioned earlier, several potent small molecules have been identified as FABP4 inhibitors. The 120 structures used to build and

evaluate the 3D-QSAR model were chosen among the structures previously published;[34, 36, 37, 39, 40, 53, 54] all molecules possess an excellent selectivity toward the FABP4 and the range of activities is broad enough to be able to build a good model. The 2D chemical structures were constructed by Marvin Sketch, and all structures were subjected to molecular mechanics energy minimization using the MMFF94 force field present in the same software.[55] Once obtained the 3D structures of all compounds the geometry was also optimized at semi-empirical level of theory using the PM3 Hamiltonian,[56] as implemented in MOPAC 2016 package.[57]

1.4.2. Compound alignment

With the aim to generate a plausible and consistent set of alignment molecules, before running the regression analysis, we evaluated two different types of alignment (Figure 1.18).

First, we evaluated a structure-based alignment, based on the docking of the different ligands on the active site of the protein. All 120 structures, optimized at the PM3 level of theory, have been converted into pdbqt format using Babel,[58] and subsequently docked in the active site of FABP4. Molecular docking was performed using the three-dimensional crystal structures of substrate-free fatty acid binding protein 4 in complex with BMS309403 (PDB ID: 2NNQ) obtained from the Protein Data Bank (PDB, http://www.rcsb.org/pdb). AutoDock Vina (version 1.1.2), a molecular docking program (downloadable at http://vina.scripps.edu) developed in the Molecular Graphics Lab at The Scripps Research Institute,[59] was used for all docking experiments. The default values of the docking parameters in AutoDock Vina were maintained except for the exhaustiveness that was set to 15. A grid box of 18 Å × 18 Å × 18 Å encompassing the inhibitor binding cavity of FABP4 and centered on ligand was used. The binding modes were clustered through the root mean square deviation among the Cartesian coordinates of the ligand atoms. The docking results were ranked based on the binding free energy. Finished the calculation of AutoDockVina, all the generated structures

were manually checked, to ensure a correct positioning within the binding site. Then the generated structures were imported to Forge to build the Structure-based 3D-QSAR model.

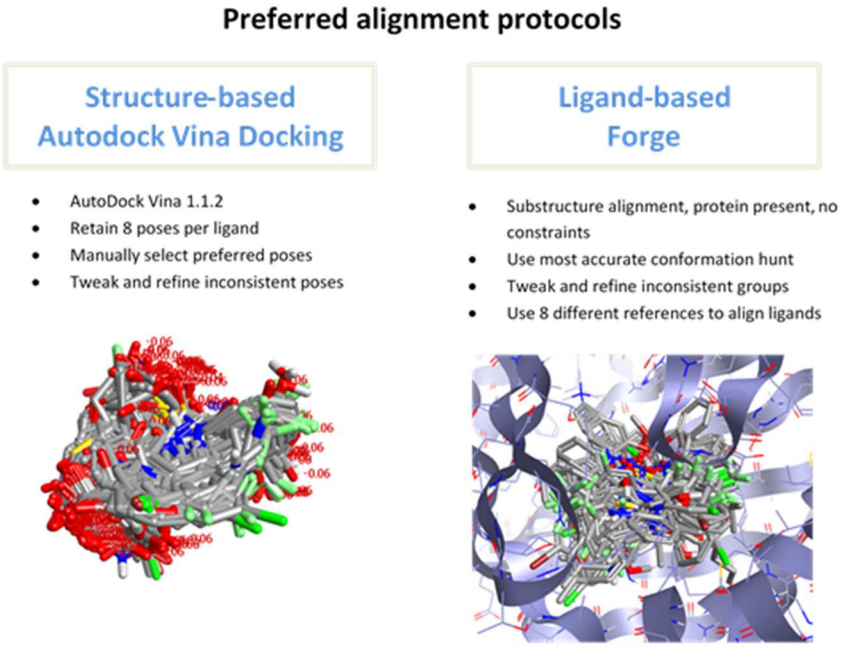


Figure 1.18. Comparison of Alignment Methods.

The second type of alignment that was evaluated is a classic Ligand-based alignment, carried out with the same software used for the building of the model. All the optimized structures, accompanied by their respected IC_{50} values, were imported into Forge (10.4.2, Cresset, Litlington, Cambridgeshire, UK, http://www.cresset-group.com/forge) [52, 60-63] for setting-up the field-based 3D-QSAR model. Eight different molecules were chosen as a template for the calculation of field points and as a template for the alignment. These eight molecules were selected because they are present in crystallized form with FABP4 (PDB IDs: 2NNQ, 3FR2, 3FR4, 3FR5, 4NNS, 4NNT, 1TOU and 1TOW, Table 1.6).[36, 37, 40, 50] The structures, small protein, and inhibitor, were first downloaded from the Protein Data Bank (PDB), the amino acid sequence was then superposed and aligned with YASARA (version 17.8.15) to get even the ligands in the binding site aligned and

superposed, thus the eight molecules were imported on Forge (Figures 1.19 and 1.20).

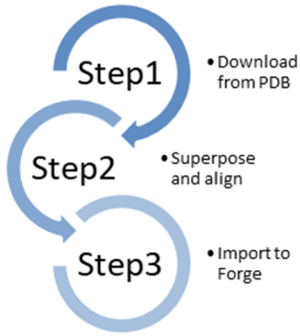
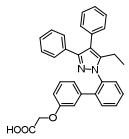
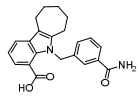
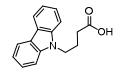
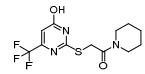


Figure 1.19. Schematic representation of the process used to obtain the template compounds for the ligand-based alignment.

Table 1.6. PDB codes and molecules used as reference compounds for the ligand-based alignment.

PDB code	2D structure	PDB code	2D structure
2NNQ		4NNT	
3FR5		4NNS	
3FR4		1TOW	
3FR2		1TOU	

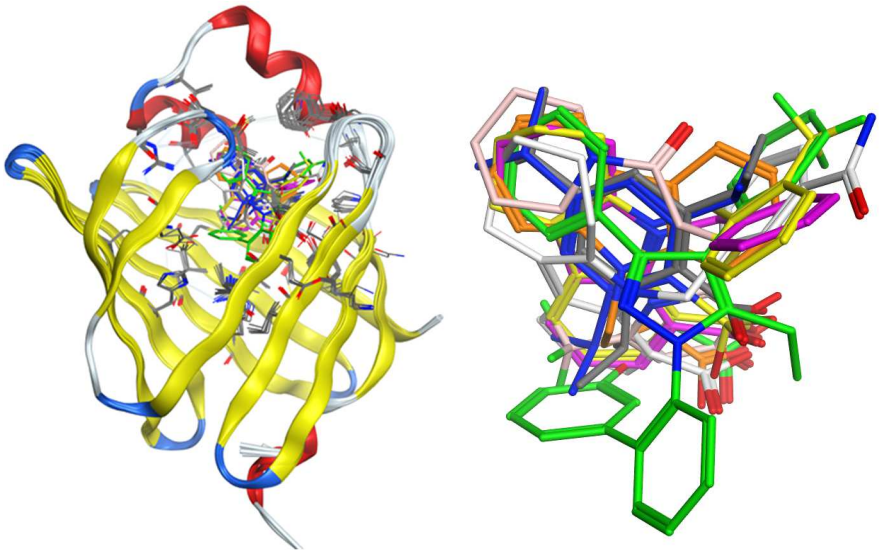


Figure 1.20. A) Protein and inhibitors aligned. B) Aligned inhibitors imported to Forge for the ligand-based alignment.

The XED (eXtended Electron Distribution) force field was used to generate the field point. The force field calculates four different molecular field point: negative and positive electrostatic, van der Waals shape, and hydrophobic. The compounds in the training set were aligned to the reference compound by Maximum Common Substructure using a customized set-up for the conformation hunt:

- Max number of conformations: 500
- RMS cut-off for duplicate conformers: 0.5 Å
- Gradient cut-off for conformer minimization: 0.1 kcal/mol
- Energy window: 2.5 kcal/mol

The RMS cut-off for duplicate conformers parameter controls the similarity threshold below which two conformers are assumed identical. Conformations that gave a minimized energy outside the energy window were discarded.

All the alignments were manually checked to ensure the best possible model. All the field points of the training set were used to derive a gauge invariant set of sampling points, which reduced the number of descriptors that needed to be considered, with a distance of 1 Å between the sample points; sample values were calculated, ensuring that all areas around the molecule that might contribute to the activity are properly described.

1.4.3. Statistical analysis and results

The regression method used in Forge is PLS.[64-68] Specifically, the SIMPLS algorithm was used. The following conditions were used to calculate the field QSAR models:

- Cross-validation type: Leave-one-out
- Number of times to run the leave-many-out operation: 1000
- Percent of data to leave out in leave-many-out mode: 20
- Maximum number of PLS components: 20
- Number of Y scrambles: 50
- Sample point minimum distance threshold: 1 Å

For the validation of the QSAR model, the leave-one-out method was used. 20 was the maximum number of components to extract

from the PLS regression. 50 was the number of Y scrambles to use, this means that in each scramble the activity values are randomly assigned to molecules and the model building is repeated. The higher the number of the scramble sets the stronger is the confirmation of statistical significance. The threshold of the sample point minimum distance was set to 1 Å. This option checks the sphere exclusion algorithm used to reduce the initial number of field sample positions down to a smaller set. Decreasing this value (Sample point minimum distance threshold) increases the number of sample points to consider, which on one hand may improve the model, but on the other hand, increase the probability of over-fitting. A value of 1 Å means that sample points must be at least 1 Å aside from each other. The Leave-one-out method was used during the validation of the QSAR model. This means that the model is built different times, and every single time a single molecule left out of the process, this is then repeated leaving out each training set molecule in turn.

The predictive ability of the generated QSAR models was confirmed by several statistical tests. The cross-validation regression coefficient (Q^2) was calculated based on the PRESS (Prediction error sum of squares) and SSY (Sum of squares of deviation of the experimental values from their mean):

$$Q^2 = 1 - \frac{PRESS}{SSY} = 1 - \frac{\sum_{i=1}^{n}(Yexp - Ypred)^2}{\sum_{i=1}^{n}(Yexp - Ymean)^2}$$

Y*exp* = experimental activity of training set compound
Y*pred* = predicted activity of training set compound
Y*mean* = mean values of the activity of training set compound

The statistical results of both models are reported in Table 1.7 and Figure 1.21.

Table 1.7. 3D-QSAR models statistics.

	Leave-one-out q^2	Training set r^2
3D-QSAR Based on Docked Poses	0.38	0.90
3D-QSAR Based on Forge Alignments	0.64	0.92

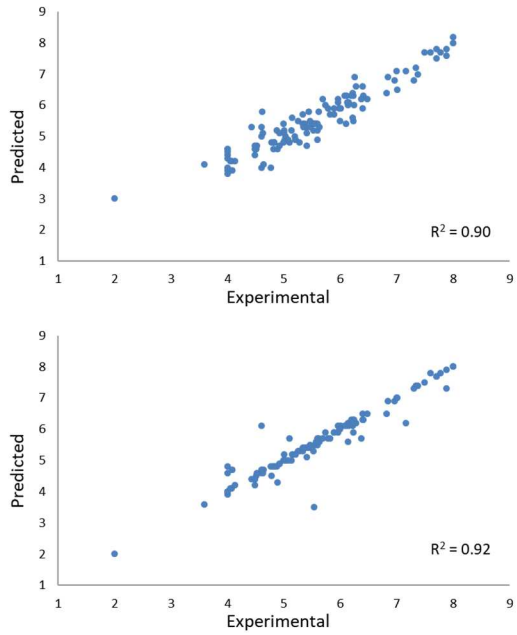


Figure 1.21. 3D-QSAR models statistics: up) 3D-QSAR based on docked poses; down) 3D-QSAR based on Forge alignments.

As is demonstrated by the previous data, the model built with the alignment of Forge is statistically more reliable than the model derived from the docked molecules, which is noisier.

At this point the best model, ligand-based 3D-QSAR align model, was taken as reference and further validated with a set of external compounds. In fact, more reliable estimation of robustness comes from separate training and test sets. Out of 120 molecules, we random choose 96 molecules as a training set to build the model while the remaining 24 compounds served as a test set to evaluate the model. In both the training set and the test set, the selected molecules cover

the whole range of activities of the compounds.

The performance of this model was also validated by the determination coefficient in prediction, R^2_{test}, using the following equation:

$$R^2test = 1 - \frac{\sum_{i=1}^{n}(Ypredtest - Ytest)^2}{\sum_{i=1}^{n}(Ytest - Ymean)^2}$$

Y*predtest* = predicted activity of test set compound by QSAR equation

Y*test* = experimental activity of test set compound

Y*mean* = mean values of the activity of training set compound

The results are shown in Figures 1.22 and 1.23. The 11-components model shows both good predictive and descriptive capability as it is shown by the good R^2 (0.99) and Q^2 (0.69)[69] values for the training and the cross-validated training set (Figure 1.22). The plot of experimental *vs*. predicted activity for the compounds, in the training set and in the cross-validated training set (Q^2 = 0.69), shows a reasonable distribution of the values. The plot of experimental *vs*. predicted activity for the compounds in the test set (Figure 1.23) is still reasonably good with only few outliers and an excellent cross-validated R^2 = 0.73.

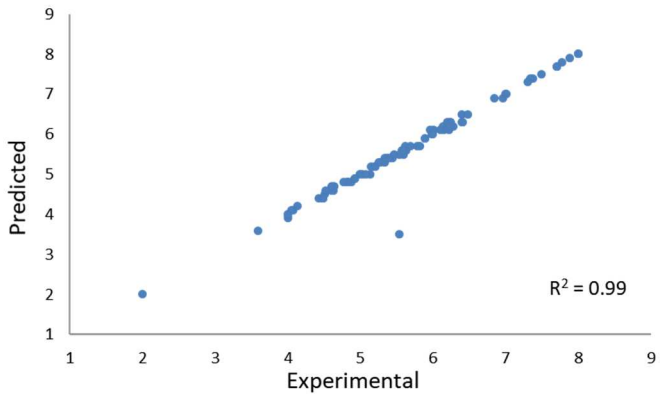


Figure 1.22. The 11-components 3D-QSAR model - Experimental *vs*. Predicted activity of the compounds in the training set.

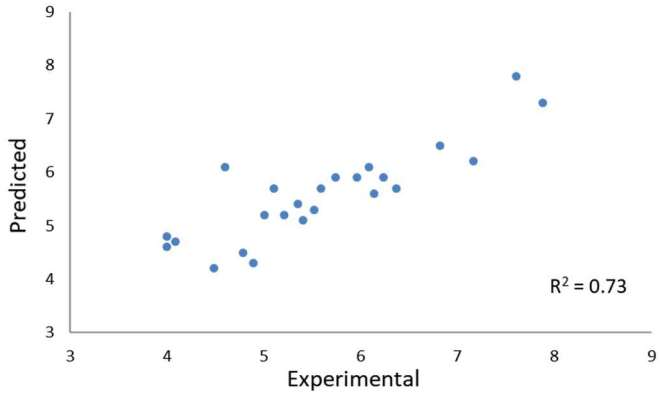


Figure 1.23. The 11-components 3D-QSAR model - Experimental *vs*. Predicted activity of the compounds in the test set.

1.4.4. Discussion

Figure 1.24 shows the 3D visualization of the QSAR final model aligned with BMS309403. The model interpretation view shows the areas where the equation suggests that the local fields have a strong impact on activity. The larger the points the stronger is the correlation between the electrostatic/steric field in that position. In Figure 1.24, the 3D-QSAR model coefficients are superposed to the structure of BMS309403; this 3D-QSAR model is clearly described by both steric and electrostatic effects and well describes the comportment of the different classes of inhibitors inside the binding pocket of FABP4. Forge was used to building a statistically robust 3D-QSAR model for a set of 120 FABP4 inhibitors, covering a wide range of different classes of standard inhibitors of the receptor. As far as it concerns, this is the first 3D-QSAR model of this class of compounds that well describes the individual activities of each molecule.

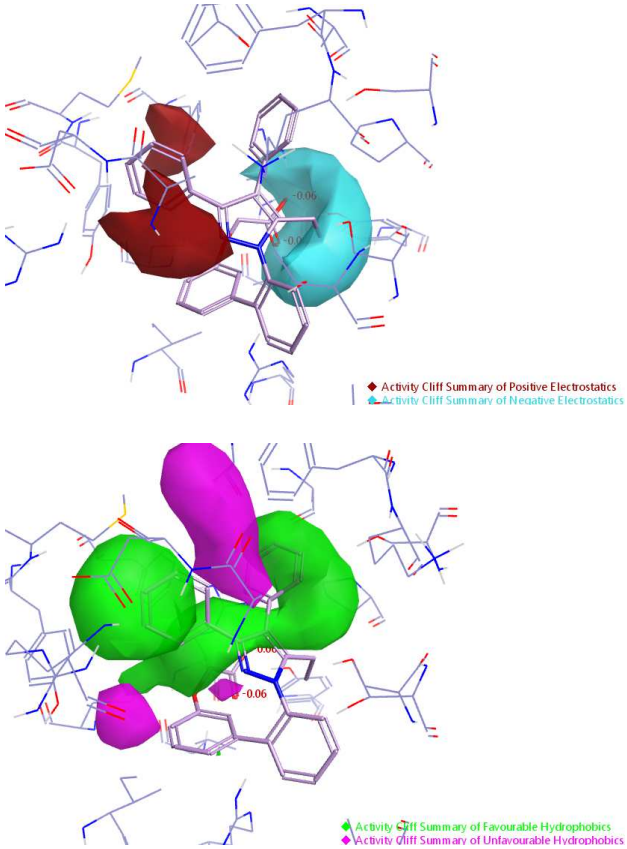


Figure 1.24. Average of positive (red) and negative (blue) fields of actives. Average hydrophobics of actives: favorable (green) and unfavorable (purple).

1.4.5. Finding Bioisosteres and FDA approved drugs

In order to design novel hit compounds with FABP4 inhibitory activity and using the above-reported information by the QSAR model, we decided to use a bioisostere and fragment replacement software tool (Spark; Cresset group), to produce a scaffold hopping analysis and generate a virtual library of FABP4 ligands.[70] On the other hand, a library of FDA approved drugs was also evaluated. We studied the bioisosteric replacement in six different portions of BMS309403. All the different aromatic moieties of the molecules were studied by means of scaffold hopping as represented in Figure 1.25.

Once created, the new virtual compounds were scored assuming that if the fields of the newly designed molecules are very similar to that of the original compounds, the resulting compounds will have

similar biological properties.[71, 72] The bioisosteric replacement was performed using the same 178,558 fragments for each part; in particular, the fragments derive from ChEMBL and Zinc databases (See Experimental section).[73, 74] Five hundred compounds were generated for each substitution for a total of 3,000 compounds, which were then scored in the 3D-QSAR model, together with the FDA drugs (1314 molecules, downloaded from https://www.drugbank.ca/). For the bioisosteric replacement, the first top-scored compounds according to FABP4 predicted pIC_{50} are reported in Figure 1.26, while the top 25 compounds are reported as SMILE in the Tables 1.8–1.13. Furthermore, the 50 top scored FDA approved drugs are reported in the Table 1.14.

The predicted pIC_{50} of the top 25 compounds derived from the bioisosteric replacement of series 1 are ranked between 6.1 and 5.8, those derived from the replacement of series 2 between 6.0 and 5.7, those derived from the replacements of series 3 between 6.5 and 5.9, those derived from the replacements of series 4 between 6.3 and 6, those derived from the replacements of series 5 between 6.7 and 6.0 and those derived from the replacements of series 6 between 7.3 and 6.2. Regarding the predicted activity of the FDA approved drugs, the 50 top scored compounds are ranked between 8.7 and 6.3.

Overall results indicate that the bioisosteric replacement and the following 3D-QSAR model evaluation give compounds with the appropriate chemical structure for the inhibition of FABP4. Interestingly, even from the FDA library, some compounds were predicted to have FABP4 inhibitor activity.

The most interesting results were the ones derived from the series 6 of the bioisosteric replacement and the ones derived from the FDA database. Interestingly, series 6 derived compounds have a structure partially different from the precursor compound BMS309403 but, despite these differences, the compounds retained the activity and the top scored have a better activity than the precursor. This information can be useful for the design of novel compounds with totally different central core and high FABP4 inhibitor activity. Regarding the results

derived from the screening of the FDA database is interesting to notice that different compounds were predicted to have a high affinity for FABP4; this means that part of their activity/side effects can be due to the interaction with FABP4. Actually, some FDA approved drugs were already identified as FABP4 ligands, among them: pimozide, levofloxacin, and trovafloxacin. [42, 75] The pimozide has an IC_{50} of 11.38 µM in FABP4; however, the levofloxacin and trovafloxacin inhibit 70.01 and 57.46% the FABP4 at 10 µM. These results are in total agreement with the ones predicted by the 3D-QSAR scoring, where the pimozide, levofloxacin, and trovafloxacin have a predicted pIC_{50} of 5.1, 5.3 and 4.7, respectively.

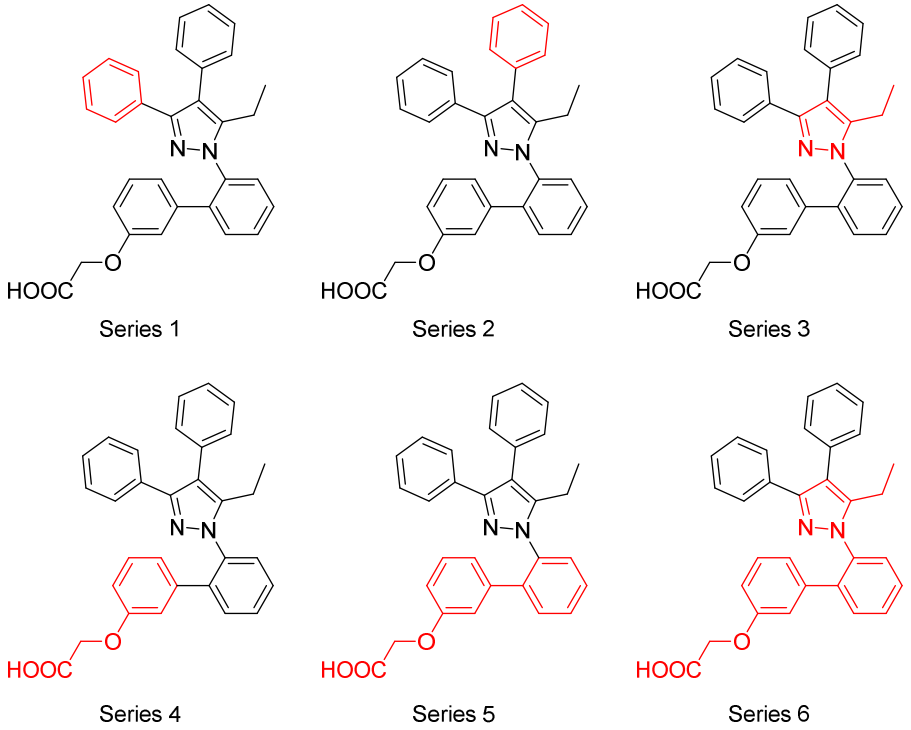


Figure 1.25. Bioisosteric replacement of the selected compound (BMS309403).

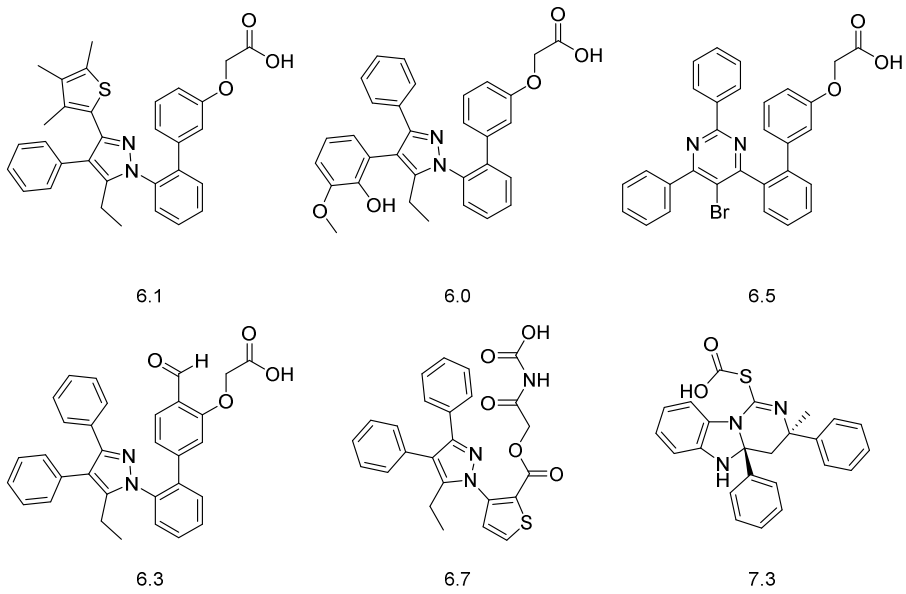


Figure 1.26. First top-scored compounds derived from series 1–6 according to FABP4 predicted pIC_{50}.

Table 1.8. 25 top scored compounds (predicted pIC_{50}) derived from the series 1 of the bioisosteric replacement.

1	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4c(c(c(s4)C)C)C)-c5ccccc5)CC)c1	6.1
2	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)C4CCCCCC4)-c5ccccc5)CC)c1	6.0
3	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4cn(c(c4)C)C)-c5ccccc5)CC)c1	6.0
4	Clc1ccc(O)cc1-c2c(c(n(n2)-c3ccccc3-c4cccc(OCC([O-])=O)c4)CC)-c5ccccc5	6.0
5	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4cc5c(s4)ccs5)-c6ccccc6)CC)c1	5.9
6	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccc5c(occ5C)c4)-c6ccccc6)CC)c1	5.9
7	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccc(c(CC)c4)C)-c5ccccc5)CC)c1	5.9
8	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccc(NC)cc4)-c5ccccc5)CC)c1	5.9
9	Clc1ccc(-c2c(c(n(n2)-c3ccccc3-c4cccc(OCC([O-])=O)c4)CC)-c5ccccc5)cc1O	5.9
10	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4c(cc(s4)C)C)-c5ccccc5)CC)c1	5.9
11	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4cc5cc[nH]c5s4)-	5.9

	c6ccccc6)CC)c1	
12	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4cc5COCCc5s4)-c6ccccc6)CC)c1	5.9
13	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccc5c(c(c[nH]5)C)c4)-c6ccccc6)CC)c1	5.8
14	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccc(s4)NC)-c5ccccc5)CC)c1	5.8
15	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4cccc(N)c4N)-c5ccccc5)CC)c1	5.8
16	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccc([nH]4)C)-c5ccccc5)CC)c1	5.8
17	Oc1c(cc(cc1-c2c(c(n(n2)-c3ccccc3-c4cccc(OCC([O-])=O)c4)CC)-c5ccccc5)C)C	5.8
18	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4cnc(OC)c(c4)C)-c5ccccc5)CC)c1	5.8
19	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4cccc(N)c4)-c5ccccc5)CC)c1	5.8
20	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4cc(oc4)C)-c5ccccc5)CC)c1	5.8
21	Clc1ccc(-c2c(c(n(n2)-c3ccccc3-c4cccc(OCC([O-])=O)c4)CC)-c5ccccc5)cc1N	5.8
22	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4cccc(c4N)C)-c5ccccc5)CC)c1	5.8
23	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4cccc(C5CC5)c4)-c6ccccc6)CC)c1	5.8
24	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccc(N)c(N)c4)-c5ccccc5)CC)c1	5.8
25	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4cc([nH]c4)C)-c5ccccc5)CC)c1	5.8

Table 1.9. 25 top scored compounds (predicted pIC_{50}) derived from the series 2 of the bioisosteric replacement.

1	Oc1c(OC)cccc1-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5	6.0
2	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5c[nH]cc5)CC)c1	5.9
3	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccc[nH]5)CC)c1	5.9
4	Oc1c(N)ccc(-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5)c1	5.9
5	Oc1c(sc(n1)N)-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5	5.9
6	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-	5.8

	c5ccc(c(NC)c5)C)CC)c1	
7	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(-c4cccc(NCC)c4)c(n3)-c5ccccc5)CC)c1	5.8
8	Fc1cccc(O)c1-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5	5.8
9	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(-c4cnc5c(c4)cc[nH]5)c(n3)-c6ccccc6)CC)c1	5.8
10	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5coc(N)c5)CC)c1	5.8
11	Clc1c(N)cccc1-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5	5.8
12	Oc1cc(O)c(O)cc1-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5	5.8
13	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccc(s5)N)CC)c1	5.8
14	Sc1c(N)ccc(-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5)c1	5.8
15	Oc1c(O)cc(cc1-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5)C	5.7
16	Oc1ccc(-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5)cc1OC	5.7
17	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(-c4ccc5c(c4)cc[nH]5)c(n3)-c6ccccc6)CC)c1	5.7
18	FC(F)c1ccccc1-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5	5.7
19	Oc1cc(ncc1-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5)C	5.7
20	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5NN)CC)c1	5.7
21	Brc1cc(-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5)co1	5.7
22	Brc1ccc([nH]1)-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5	5.7
23	Oc1c(O)cccc1-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5	5.7
24	Clc1cccc(-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5)c1O	5.7
25	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5c(OC)ccs5)CC)c1	5.7

Table 1.10. 25 top scored compounds (predicted pIC_{50}) derived from the series 3 of the bioisosteric replacement.

1	Brc1c(nc(nc1-c2ccccc2-c3cccc(OCC([O-])=O)c3)-c4ccccc4)-c5ccccc5	6.5
2	[O-]C(=O)COc1cccc(-c2ccccc2-n3c(N)c(c(c3C)-c4ccccc4)-c5ccccc5)c1	6.3
3	Clc1c(nc(nc1-c2ccccc2-c3cccc(OCC([O-])=O)c3)-c4ccccc4)-c5ccccc5	6.3
4	[O-]C(=O)COc1cccc(-c2ccccc2-c3c(c(nc(n3)-c4ccccc4)-c5ccccc5)CC)c1	6.3
5	Fc1c(nc(nc1-c2ccccc2-c3cccc(OCC([O-])=O)c3)-c4ccccc4)-c5ccccc5	6.2
6	[O-]C(=O)COc1cccc(-c2ccccc2/C(=N/Oc3ccccc3)c4ccccc4Oc5ccccc5)c1	6.1
7	OCCOc1c(c(nc(n1)-c2ccccc2-c3cccc(OCC([O-])=O)c3)-c4ccccc4)-c5ccccc5	6.0
8	[O-]C(=O)COc1cccc(-c2ccccc2-n3c4c(c(n3)-c5ccccc5)c(n[nH]4)-c6ccccc6)c1	6.0
9	[O-]C(=O)COc1cccc(-c2ccccc2OC[C@@H](Nc3ccccc3)COc4ccccc4)c1	6.0
10	[O-]C(=O)COc1cccc(-c2ccccc2-c3c(CCC)c(c([nH]3)-c4ccccc4)-c5ccccc5)c1	6.0
11	[O-]C(=O)COc1cccc(-c2ccccc2-c3csc4c(c(nn34)-c5ccccc5)-c6ccccc6)c1	6.0
12	[O-]C(=O)COc1cccc(-c2ccccc2-c3c(c(c([nH]3)-c4ccccc4)-c5ccccc5)CC)c1	6.0
13	[O-]C(=O)COc1cccc(-c2ccccc2-n3cc(c(c3CC)-c4ccccc4)-c5ccccc5)c1	6.0
14	O=C(O[C@@H]([C@H](OC(=O)C)c1ccccc1)c2ccccc2)c3ccccc3-c4cccc(OCC([O-])=O)c4	5.9
15	Clc1c(OC(=O)C)c(c(nc1-c2ccccc2-c3cccc(OCC([O-])=O)c3)-c4ccccc4)-c5ccccc5	5.9
16	[O-]C(=O)CCc1c(nn(c1-c2ccccc2)-c3ccccc3)-c4ccccc4-c5cccc(OCC([O-])=O)c5	5.9
17	[O-]C(=O)COc1cccc(-c2ccccc2-n3c4c(c(c(o4)-c5ccccc5)-c6ccccc6)c(n3)C)c1	5.9
18	[O-]C(=O)COc1cccc(-c2ccccc2-c3cc(c(n3C)-c4ccccc4)-c5ccccc5)c1	5.9
19	O=S(=O)(Nc1cc(c(cc1-c2ccccc2-c3cccc(OCC([O-])=O)c3)-c4ccccc4)-c5ccccc5)C	5.9
20	Clc1c(-c2ccccc2)cc(nc1-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5	5.9
21	O=C(OC)c1c(c(nn1-c2ccccc2-c3cccc(OCC([O-])=O)c3)-c4ccccc4)-c5ccccc5	5.9
22	[O-]C(=O)COc1cccc(-c2ccccc2-c3cc(c(n3CC)-c4ccccc4)-c5ccccc5)c1	5.9
23	[O-]C(=O)COc1cccc(-c2ccccc2-c3c(c(c([nH]3)-c4ccccc4)-c5ccccc5)C)c1	5.9
24	[O-]C(=O)COc1cccc(-c2ccccc2[C@@H]3CN([C@](O3)(c4ccccc4)C)c5ccccc5)c1	5.9
25	[O-]C(=O)COc1cccc(-c2ccccc2-c3cc(c([nH]3)-c4ccccc4)-c5ccccc5)c1	5.9

Table 1.11. 25 top scored compounds (predicted pIC_{50}) derived from the series 4 of the bioisosteric replacement.

1	O=Cc1ccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)cc1OCC([O-])=O	6.3
2	[O-]C(=O)Nc1c2c(sn1)ccc(-c3ccccc3-n4c(c(c(n4)-c5ccccc5)-c6ccccc6)CC)c2	6.3
3	Fc1ccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)cc1[C@H](O)C([O-])=O	6.3
4	O=C(N1CCc2ccc(-c3ccccc3-n4c(c(c(n4)-c5ccccc5)-c6ccccc6)CC)cc21)CC([O-])=O	6.3
5	O=C(OC)c1ccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)cc1OCC([O-])=O	6.3
6	O=C(OCC([O-])=O)N1CCC[C@H]1Cc2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC	6.3
7	O=C(N1[C@H](CC[C@@H](C1)c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)C)CC([O-])=O	6.2
8	O=C(c1c(cc(c(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)c1)C)C)CC([O-])=O	6.2
9	Fc1ccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)cc1CC([O-])=O	6.1
10	O=S(=O)(NC([O-])=O)c1ccc(s1)-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC	6.1
11	Oc1cc(c(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)cc1CC([O-])=O)C	6.1
12	[O-]C(=O)Cn1c2cc(-c3ccccc3-n4c(c(c(n4)-c5ccccc5)-c6ccccc6)CC)ccc2cn1	6.1
13	O=S(=O)(CC([O-])=O)c1nc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)cs1	6.1
14	Oc1ccc(S(=O)(=O)NC([O-])=O)cc1-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC	6.1
15	O=C(N)c1ccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)cc1NCC([O-])=O	6.1
16	Clc1ccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)cc1OCC([O-])=O	6.1
17	O=S(=O)(NC([O-])=O)c1c(N)ccc(c1)-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC	6.1
18	O=C(NCC([O-])=O)CCNc1ccccc1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC	6.1
19	O=C(N1C[C@H](S[C@@H](C1)C)c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)CC([O-])=O	6.0
20	[O-]C(SSCCCc1ccccc1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)=O	6.0
21	Oc1ccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)cc1CC([O-])=O	6.0
22	O=C(SC([O-])=O)c1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-	6.0

	c5ccccc5)CC)c1	
23	Fc1ccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)cc1/C=C\C([O-])=O	6.0
24	Fc1cc(F)c(cc1-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)CC([O-])=O	6.0
25	O=C1COc2ccc(-c3ccccc3-n4c(c(c(n4)-c5ccccc5)-c6ccccc6)CC)cc2N1CC([O-])=O	6.0

Table 1.12. 25 top scored compounds (predicted pIC_{50}) derived from the series 5 of the bioisosteric replacement.

1	O=C(NC([O-])=O)COC(=O)c1c(-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)ccs1	6.7
2	[O-]C(=O)c1nnc2n1N=C(c3ccccc3-n4c(c(c(n4)-c5ccccc5)-c6ccccc6)CC)CS2	6.5
3	O=C(O[C@H](C)C([O-])=O)c1ccccc1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC	6.4
4	O=C(N[C@H]1CCC[C@@H]([C@@H]1n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)C)CSC([O-])=O	6.4
5	O=C1C(CCC([O-])=O)=CCN1CCn2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC	6.3
6	O=C(N[C@H](CC([O-])=O)C)c1c(-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)ccs1	6.3
7	S/C(NCc1ccccc1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)=N/C([O-])=O	6.3
8	[O-]C(=O)/N=C(\NCCc1ccccc1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)N	6.3
9	O=C(OCC([O-])=O)CS[C@H](n1c(c(c(n1)-c2ccccc2)-c3ccccc3)CC)C	6.2
10	O=C(Nc1ccc(N)cc1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)CCC([O-])=O	6.2
11	O=C(NCCC([O-])=O)c1cccn1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC	6.2
12	O=C(NCc1ccccc1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)/C=C\C([O-])=O	6.2
13	O=C(OCC([O-])=O)CSCn1c(c(c(n1)-c2ccccc2)-c3ccccc3)CC	6.2
14	O=C(NCCC([O-])=O)c1ccc(N)cc1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC	6.1
15	[O-]C(=O)Cc1cn(nn1)C[C@H]2CC[C@H](O2)n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC	6.1
16	O=C(O[C@H](C)C([O-])=O)c1c(-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)c5ccccc5o1	6.1
17	Clc1cccc(NC(=O)CSC([O-])=O)c1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC	6.1
18	O=C(OCC([O-])=O)Cc1ccccc1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC	6.1
19	O=C(NCC([O-])=O)CSCCn1c(c(c(n1)-c2ccccc2)-c3ccccc3)CC	6.1
20	[O-]C(=O)[C@@H]1c2c(CCO1)cn(n2)CCn3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC	6.1
21	Clc1cccc(-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)c1CC(=O)NCC([O-])=O	6.1

22	O=C(Nc1ccccc1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)CCC([O-])=O	6.1
23	O=C(Oc1ccc(cc1-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)C)CCC([O-])=O	6.0
24	O=C(NCCC([O-])=O)c1c(-n2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC)nc(s1)N	6.0
25	[O-]C(=O)Cc1ccccc1OCCn2c(c(c(n2)-c3ccccc3)-c4ccccc4)CC	6.0

Table 1.13. 25 top scored compounds (predicted pIC_{50}) derived from the series 6 of the bioisosteric replacement.

1	[O-]C(SC1=[N+][C@](C[C@@]2(Nc3ccccc3N12)c4ccccc4)(c5ccccc5)C)=O	7.3
2	O=C(NC(C([O-])=O)C([O-])=O)[C@H](Cc1ccccc1)c2ccccc2	7.2
3	S=C(N(Nc1ccccc1)c2ccccc2)CCCC([O-])=O	6.6
4	S=C(NN=C(Cc1ccccc1)Cc2ccccc2)C([O-])=O	6.5
5	O=C([C@@H]1[C@H](O[C@](O1)(c2ccccc2)C)c3ccccc3)[C@H](OC)C([O-])=O	6.5
6	O=C(OC([O-])=O)CS[C@H]1[C@@H](O[C@@H](C1)c2ccccc2)c3ccccc3	6.5
7	[O-]C(=O)NCC[C@@H](c1ccccc1)COc2ccccc2	6.4
8	[O-]C(=O)CC([C@@H]([C@H]([C@H](c1ccccc1)C)c2ccccc2)C)=C	6.4
9	[O-]C(=O)[C@H]1[C@H](OC[C@@H](O1)[C@@H](Cc2ccccc2)c3ccccc3)OC	6.4
10	O=C(N(Cc1ccccc1)Cc2ccccc2)[C@@H](CC([O-])=O)C	6.4
11	O=C1[C@@H](N2C(=O)[C@](O[C@@]2([C@@H]3CCCN13)c4ccccc4)(c5ccccc5)C)CC([O-])=O	6.3
12	[O-]C(=O)/C=C\C[C@H](Cc1ccccc1)c2ccccc2	6.3
13	O=C1C[C@@H]([C@@H]([C@@H]1c2ccccc2)/C=C/[C@H](O)C([O-])=O)c3ccccc3	6.3
14	[O-]C(=O)NC[C@H]1C[C@@H]2CC[C@]1([C@@]2(c3ccccc3)C)c4ccccc4	6.3
15	[O-][N+](=O)c1c(N)cc2c(C[C@H]([C@](O2)(c3ccccc3)C)c4ccccc4)c1C([O-])=O	6.2
16	[O-]C(=O)C(OC[C@H](Cc1ccccc1)c2ccccc2)C([O-])=O	6.2
17	[O-]C(O[C@H](OC)CC/C(c1ccccc1)=C/c2ccccc2)=O	6.2
18	O[C@]1(CCC[C@]2(CC[C@@H](C[C@@]12c3ccccc3)C(C([O-])=O)=C)C)c4ccccc4	6.2
19	Clc1ccc(N(Cc2ccccc2)Cc3ccccc3)cc1C([O-])=O	6.2
20	O=C(OC)[C@@H]1[C@@]2(O1)[C@@H](O[C@@H]([C@@H]2c3ccccc3)c4ccccc4)CC([O-])=O	6.2
21	O[C@@H]1CC(=C[C@H]([C@H]1c2ccccc2)c3ccccc3)C(=O)C([O-])=O	6.2

22	O[C@@H](C[C@H]1[C@@]([C@@H]1c2ccccc2)(c3ccccc3)C)C(=O)C([O-])=O	6.2
23	O[C@H]1[C@@H](C[C@@H](C[C@H]1CC([O-])=O)c2ccccc2)c3ccccc3	6.2
24	[O-]C(=O)[C@H]1C[C@H]1CC[C@@H](Nc2ccccc2)c3ccccc3	6.2
25	O=S(=O)(NC1[C@@]([C@@]1(c2ccccc2)C)(c3ccccc3)C)N(C)C([O-])=O	6.2

Table 1.14. 50 top scored FDA approved drugs with their predicted pIC_{50}.

Cefditoren	8.7
Febuxostat	8.4
Alatrofloxacin	8.2
Eprosartan	8.2
Quinapril	7.9
Nedocromil	7.7
Pentosan Polysulfate	7.7
Sacubitril	7.6
Enalaprilat	7.3
Ceftriaxone	7.1
Riboflavin Phosphate	7.1
Penicillin G	7.0
Tizoxanide	7.0
Amfenac	6.9
Apixaban	6.9
Argatroban	6.9
Cabergoline	6.9
Cefixime	6.9
Gemifloxacin	6.9
Apremilast	6.8
Axitinib	6.8
Bromfenac	6.8
Dolutegravir	6.8
Azilsartan Kamedoxomil	6.8
Nalidixic Acid	6.8
Sulfathiazole	6.8
Aztreonam	6.7
Lodoxamide	6.7
Lomefloxacin	6.7
Cefotaxime	6.6
Cidofovir	6.6
Cinoxacin	6.6

Ketoprofen	6.6
Lisinopril	6.6
Quinaprilat	6.6
Benazeprilat	6.5
Eluxadoline	6.5
Finafloxacin	6.5
Moxalactam	6.5
Sulfasalazine	6.5
Belinostat	6.4
Besifloxacin	6.4
Dinoprost	6.4
Doxycycline	6.4
Ethacrynate	6.4
Meloxicam	6.4
Valdecoxib	6.4
Carglumic Acid	6.3
Felbamate	6.3
Ketorolac	6.3

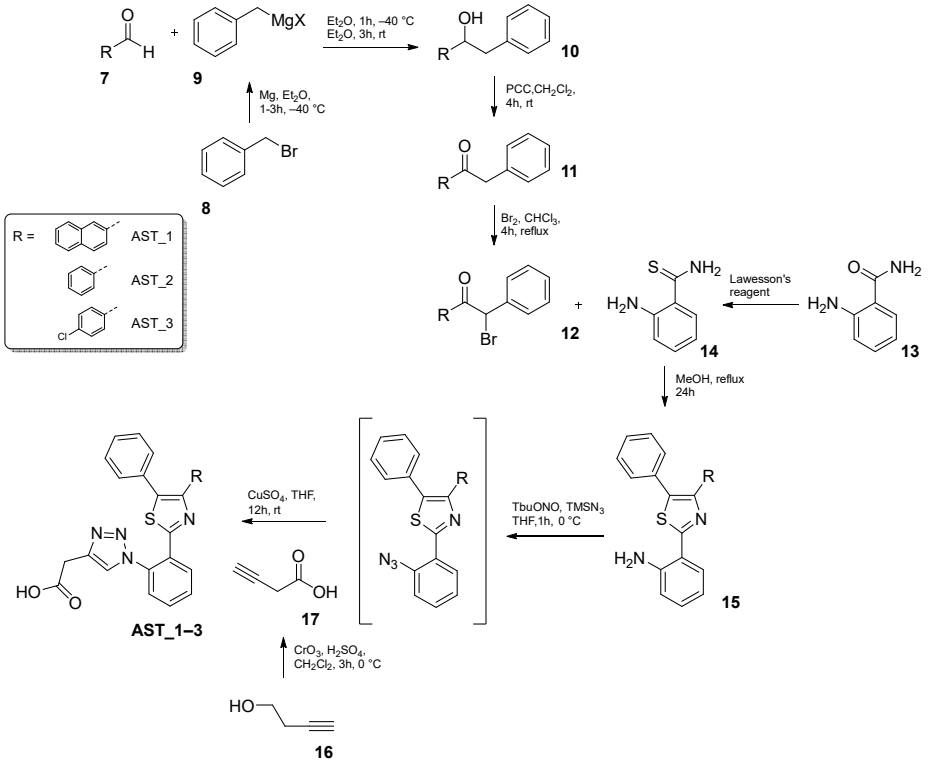
1.4.6 Evaluation of AST_1–3 in the 3D-QSAR model

Considering the designed thiazoles AST_1–3 promising from the synthetic point of view, and promising accordingly with the docking results, the three molecules were also evaluated in the 3D-QSAR model. The molecules were imported into Forge and aligned with the 3D-QSAR model. Once aligned the molecules were scored assuming that if the fields of AST_1–3 are very similar to that of the original compounds, the resulting compounds will have similar biological properties.[71, 72] Interestingly, even in the ligand-based approach, the selected molecules gave positive results. The three molecules AST_1–3 were predicted to have a predicted pIC_{50} of 5.4, 5.9 and 5.6, respectively, which correspond to an IC_{50} of 3.98, 1.25 and 2.51 μM.

1.5. Synthesis of the leads compounds

Considering that the three thiazole-based compounds AST_1–3 gave promising results in both the structure- and the ligand-based approaches, a synthetic pathway for the three compounds was designed. The synthesis (Scheme 1.1) starts from commercially

available aldehydes: benzaldehyde, naphthaldehyde, and 4-chloro benzaldehyde, and the Grignard reagent generated *in situ* by benzyl bromide and magnesium. After obtaining the secondary alcohol, this was first oxidized with PCC and then monohalogenated, at the alpha position, with bromine under acidic conditions. The 2-aminobenzene-1-carbothioamide was generated from the respective amide and the Lawesson's reagent. Then the thioamide was reacted with the compound **12** to give the thiazole derivatives. After obtaining the butynoic acid by the oxidation of the 3-butyn-1-ol, and the azide, from the amine portion of compound **15** treated with *t*-BuONO and azidotrimethylsilane, the final structure was assembled by the azide-alkyne Huisgen cycloaddition in the copper(I) catalyzed variant, in which organic azides and terminal alkynes are united to afford 1,4-regioisomers of 1,2,3-triazoles as sole products.



Scheme 1.1. Synthetic scheme for the molecules AST_1–3.

1.6. FABP4 inhibition evaluation

Inhibition activity assay for FABP4 was performed by using a commercially available FABP4 inhibitor screening assay kit. FABP4 inhibition activity was determined by measuring the decreasing in fluorescence of a detection reagent present in the kit when displaced by an inhibitor of FABP4. In fact, the detection reagent exhibits increased fluorescence when bound to FABP4, the binding of the detection reagent can be monitored by exciting at 370 nm and measuring the emission at 470 nm. Therefore, any powerful inhibitor of such a protein can displace the detection reagent, which binds to the same binding pocket and thereby reducing the fluorescence. Inhibition activity data are expressed as IC_{50} (µM) and obtained results are outlined in Table 1.14. Arachidonic acid, a known powerful ligand of FABP4, was used as a positive control and revealed an IC_{50} of 3.06 µM. Our set of compounds, AST_1–3, revealed IC_{50} values of 5.59, 3.70, and 4.31, respectively.

Table 1.14. measured IC_{50} values for Arachidonic acid and AST_1–3 in FABP4.

Compounds	IC_{50} (µM)
Arachidonic acid	3.06 ± 0.22
AST_1	5.59 ± 0.79
AST_2	3.70 ± 0.34
AST_3	4.31 ± 0.67

1.7. Conclusion and perspective

FABP4 provides an attractive therapeutic and diagnostic target for a variety of diseases. FABP4 plays a relevant role in the development of insulin resistance, atherosclerosis, and metabolic diseases; therefore, the inhibition of such a transporter can be useful for the treatment of these diseases, as demonstrated by the animal models. [76-79] FABP4 inhibitors could also be effective for the treatment of cancer patients via the inhibition or reduction of early-stage tumors and metastasis, and can also be effective as a biomarker for the diagnosis of different types of cancer.[80-86]

Additionally, FABPs are transporters for endocannabinoids.[87-90] It

is possible to identify an agent for its effect on the treatment of a neurological disorder by testing the agent for its ability to modulate the interaction of an endocannabinoid with an intracellular FABP. Therefore, FABP inhibitors could also be used for the treatment of neurological disorders.

In this project, different molecules were theorized as possible inhibitors of FABP4 by means of CADD. Particularly, using a bimodal approach; structure-based followed by ligand-based, three new potent ligands of FABP4 were individuated and synthesized. Their effective binding properties toward FABP4 were tested by means of a displacement assay and the three molecules results as FABP4 binders with an IC_{50} between 3.70 and 5.59 μM. Moreover, a huge number of different other molecules were theorized to be as effective as the three synthesized, and interesting results also came out from the scaffold hopping of the BMS309403 and from the evaluation of an FDA approved drugs database in the QSAR model. Accordingly, with these results, a simple modification of different part of the gold standard BMS309403 can retain the activity and in different cases can also improve the activity of the molecule, and different already approved drugs can be effective FABP4 binders. These findings, of course, should be verified but can be a good starting point for the researchers in the field, particularly considering that the same model that is predicting these results was effective in predicting the ability of the synthesized and tested compounds AST_1–3. Surely, all of these findings will lead our future research in the field of the development of FABP4 inhibitors considering this transporter as a very promising target the ultimate goal is to finally bring optimal molecules into clinical research phase.

Chapter 2. Design and synthesis of gallium-68 chelators for positron emission tomography

2.1. Introduction

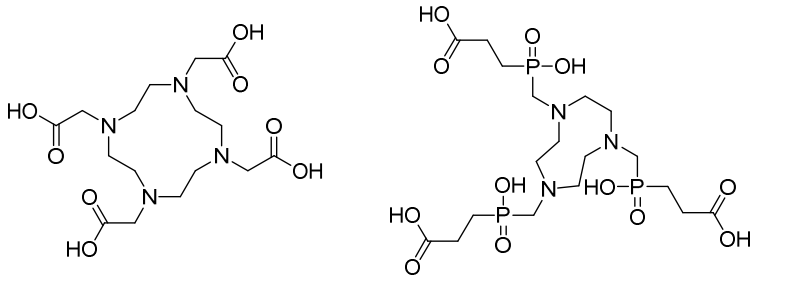
Radiometals are radioactive isotopes that can be exploited for applications in medical diagnosis, as well as for cancer therapy, and gallium has been recently employed as a radiometal. To use effectively these isotopes for the designed specific biological applications, the unbound ion of the radiometal must be sequestered from aqueous solution using chelators agents. A biologically targeting molecule is generally covalently bonded to the chelating molecules used for this application, making an active radiopharmaceutical agent. When injected into a patient, the targeting molecule, which tightly binds a radiometal ion thanks to the chelating moieties, can deliver the isotope without radiometal loss and effectively supply an *in vivo* site-specific radioactive source for imaging or therapy. Every day, a rapidly expanding number of radiometals, with a broad variety of half-lives, emission types, energies, and branching ratios, are generated. The availability of a wide range of radiometal ions makes it possible to precisely pick the specific nuclear properties that are needed for a different number of applications. ^{68}Ga, ^{64}Cu, ^{86}Y, ^{89}Zr, and ^{44}Sc are some examples of radiometals that can be used for positron emission tomography (PET) imaging, providing a sensitive, quantitative, and non-invasive image of a variety of molecular processes and targets into the human body. Differently to the PET, the single photon emission computed tomography (SPECT) is an older (since the 1960s) and more ubiquitous imaging technique and the ^{99}Tc has been the most used isotope of SPECT for a long time. More recently, other radiometals such as ^{67}Ga, ^{111}In, and ^{177}Lu have been started to be used for SPECT imaging, particularly in chelator-based radiopharmaceuticals. On the other hand, different particle emitters such as ^{111}In (Auger electron emitter), ^{90}Y and ^{177}Lu (β^-), and ^{225}Ac, ^{212}Pb, and ^{213}Bi (α), are being heavily investigated for therapy applications, typically in conjunction with antibody vectors or peptides. Each radiometal ion has unique

aqueous coordination chemistry properties; these different properties must be properly exploited if these isotopes have to be safely used for medical applications and *in vivo* use. The major difference between the hot-radioactive and the cold-nonradioactive metal ion chemistry is that radiochelation is normally performed under extremely dilute conditions, with radiometal ions being used at nM to pM concentrations. This means that the chelating moiety of the radiopharmaceutical must be coordinated efficaciously by the metal, in order to work at such low concentrations. Interesting to notice that several of the elements being discussed have multiple radioactive isotopes that are useful for diagnostic or therapeutic purposes (*e.g.* $^{86/90}$Y, $^{67/68}$Ga, $^{44/47}$Sc, $^{60/61/62/64}$Cu), and all isotopes of a given element have identical chemistry. Moreover, is also thanks to this that a single radiopharmaceutical agent can easily be labeled with different radioisotopes. Infrastructures in hospitals and radiopharmacies have been developed around Technetium-99 labeling protocols, where a generator is eluted multiple times per day to produce diverse tracers by reconstituting commercially available "cold kits" compatible with good manufacturing practice.[91] In these environments, speed, simplicity, and reproducibility of radiolabeling are of fundamental importance. With the growth of ^{18}F and ^{11}C PET tracers for imaging, more diverse, complex and costly infrastructure were generated to support PET because a kit model is not compatible for that tracers and to the need for an on-site cyclotron and more complex synthetic chemistry. Nevertheless, ^{68}Ga generators are compatible with good manufacturing practice and can be available to kit production if a simple chelation step can be achieved.[92, 93] This would make ^{68}Ga tracers widely available without the costly infrastructure associated with ^{18}F and ^{11}C tracer production. This concept was suggested more than twenty years ago,[94, 95] but despite several recent attempts of kit-based ^{68}Ga tracer production,[96] an ideal one-step procedure for ^{68}Ga radiolabeling, with the same simplicity of the long-established ^{99m}Tc-labeling procedures requiring only the addition of generator eluate to the kit vial, has not yet been achieved. To reach this end, the chelator

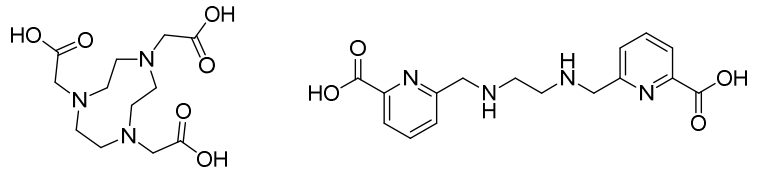
must have several characteristic: its radiolabeling should reach completion (more than 95%) quickly (in less than 5 min) at room temperature, must not be affected by common trace metals, and must be ready to use (without additional steps to concentrate, buffer or purify). Its complex should also resist to the *in vivo* transchelation (*e*.*g*. by transferrin and other iron-binding molecules), and the conjugation and radiolabeling should not produce mixtures of diastereomers, enantiomers or geometric isomers. Of course, no adverse pharmacokinetics e.g. delayed clearance or non-specific binding, should also be produced.

The current generation of ^{68}Ga chelators does not satisfy all these criteria. For example, the widely adopted macrocycle DOTA (Figure 2.1),[97] while complexing Ga^{3+} with extraordinarily high kinetic stability, has very slow complexing kinetics that requires heat, a large amount of the biomolecule (DOTA conjugate), and low pH. Moreover, the low yields (less than 95%) necessitate a purification step. These factors increase the process complexity, limits the specific activity and may damage the biomolecule. Recently, different new molecules have been published with improved characteristics, but they have not eliminated all the previously reported problems. NOTA, TRAP, and DEDPA (Figure 2.1) are promising but, like DOTA, require acidic conditions, and are particularly vulnerable to competition from contaminating trace metals.[98-101] The DATA series of chelators show rapid, room temperature labeling at pH 5; the DATAPPh variant can be labeled in 15 min at pH 7 but requires preprocessed eluate.[102]

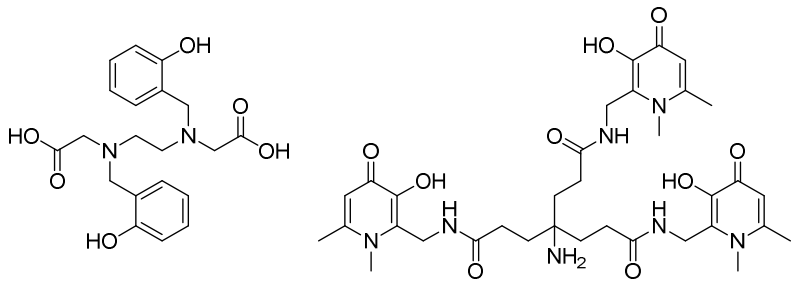
A new class of radiochelators that promises to have the requirements for a kit-based labeling is the THP system. The THP (trishydroxypyridinone) can complex ^{68}Ga rapidly at room temperature and pH = 7, with high yield and purity. Its performance has previously been evaluated against a range of common chelators, including HBED, resulting e better candidate for radiolabeling under mild conditions. THP has also been functionalized for conjugation to peptides and proteins while retaining the required mild radiolabeling and *in vivo* targeting properties.[103-105]



DOTA TRAP



NOTA DEDPA



HBED THP

Figure 2.1. Structures of some ^{68}Ga chelators of current generation.

2.1.1 ^{68}Ga uses and production

^{68}Ga is normally utilized in radiopharmaceuticals for cancer diagnostics; however, it is also used for imaging of myocardial perfusion, pulmonary perfusion, and ventilation as well as inflammation and infection. The monitoring of the transplantation and survival of beta cells in diabetes mellitus is one more growing application area of ^{68}Ga. [106, 107] The characteristics of ^{68}Ga, like its availability from a generator system and the possibility for kit type radiopharmaceutical preparation, make this radionuclide as useful as Technetium-99, but with some additional advantages such as higher sensitivity, resolution, quantification and dynamic scanning.

Moreover, some therapeutic radionuclides resemble the coordination chemistry of Ga^{3+}, making possible the chelation of the radioisotope easily. The most relevant advantages of the Gallium in PET imaging are the high positron emission fraction (89%, E_{max}: 1899 keV, E_{mean}: 890 keV) and the half-life of 68 min. In this way, this radioisotope provides sufficient levels of radioactivity for high-quality images while minimizing radiation dose to the patient and personnel. It also needs short scanning time and enables repetitive examinations. In modern generators, ^{68}Ga is obtained in ionic (Ga^{3+}) form. Ga^{3+} is the only oxidation state stable at physiological pH; moreover, it can be easily chelated in an octahedral fashion by different chelating agents.[108] The long shelf-life generator ($t_{½}(^{68}Ge)$ = 270.95 d) is affordable, easy to use and a steady source of the radionuclide for medical centers compared with a cyclotron. However, it does not require radiation-shielding constructions, consumption of energy and qualified personnel for the running and maintenance. A generator is a contained system containing an equilibrium of parent/daughter radionuclide mixture. Modern commercial generators are formed by a small chromatographic column situated in a shielding container (Figure 2.2). ^{68}Ge is produced in a cyclotron from stable Ga-69 isotope. Then, ^{68}Ge is immobilized on a column inside the generator where it spontaneously decays to ^{68}Ga Eq. (1), which can then be eluted by an eluent. ^{68}Ga decays in its turn to stable Zn(II) Eq.(2). Thus Ge, Ga, and Zn elements coexist in the generator all together and can be found in the eluate solution.

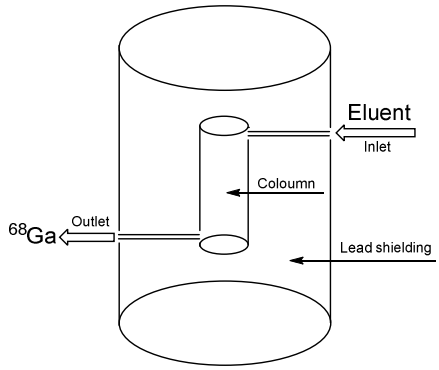


Figure 2.2. Schematic presentation of the cross-section of a column-based gallium generator.

$$^{68}_{32}Ge + {}^{0}_{-1}e \rightarrow {}^{68}_{31}Ga + v \quad (1)$$

$$^{68}_{31}Ga \rightarrow {}^{68}_{30}Zn + \beta^{+} + v\,; p \rightarrow n + \beta^{+} + v \quad (2)$$

Historically, there have been two different Germanium-Gallium separation methods. The liquid/liquid extraction and the chromatography with various eluents. The column technique is most widely used with various sorbents made of inorganic, organic and also mixed materials (Table 2.1). Most of the commercially available generators use acidic eluent since it provides cationic Ga^{3+} for the further direct chelation chemistry. Inorganic column sorbents are used more widely as they are less sensitive to radiolysis.

Table 2.1. Examples of column based ^{68}Ge/^{68}Ga generators.

^{68}Ge/^{68}Ga Generator Column Matrix	
Inorganic (Eluent)	**Organic (Eluent)**
SnO_2 (1 M HCl)	Pyrogallol-formaldehyde (0.3 M HCl)
TiO_2 (0.1 M HCl)	Nanoceria-polyacrylonitrile (0.1 M HCl)
CeO_2 (0.02 M HCl)	

The modern generators are based on chromatography and provide different advantages such as long shelf-life of 1–2 year, stable column matrixes, cationic chemical form of ^{68}Ga(III) allowing subsequent versatile and direct labeling chemistry as well as reproducible performance. There are several available generators, with variation in the molarity of HCl eluent and metal cation content. The first generator of pharmaceutical grade appeared on the market in 2014.

2.1.2 Targeting peptide bioconjugates against cancer

Natural peptides account for up to seven thousand different entities with central roles in mammalian physiology. They can have a different function in relation to their structure (*e.g*. hormones, growth factors, channel protein ligands, neurotransmitters, and antimicrobial agents).[109-113] Peptides are normally characterized by selectivity and efficiency in the activation of their targets, and peptides are usually considered safe and well tolerated by patients. In fact, the ones we call targeting peptides and the bioconjugate-targeting peptides are nowadays make up a weapon of choice for the selective targeting of

different and complex diseases, with therapeutic and imaging applications.[114]

Cancer-related deaths are estimated to reach 11.5 million in 2030.[115] For different cancer types, the cause of the death of the patient is due to the absence of effective treatments, systemic toxicity caused by the chemotherapy and/or development of resistance against the chosen treatment. Indeed, many of the available drugs which don't target the cancer cell only cause unwanted adverse effects. For example, it is worth to remember that cytotoxic chemotherapies normally have as main target the high division rate of cancer cells, damaging the exposed DNA or blocking the cell proliferation. Unfortunately, normal non-cancerogenic cells are also affected by such drugs, causing serious side-effects to the patients. One possible solution to this challenge is the one to use targeted therapies.[116] Most of the recently proposed targeted therapies consist of small molecules and antibodies that specifically target essential molecular mechanisms of the tumor growth. For instance, erlotinib was approved by the U.S. Food and Drug Administration as a tyrosine kinase inhibitor, its mechanism of action include a blocking of the activity of the key growth factor receptors responsible for tumor cell proliferation and metastasis, it blocks the epidermal growth factor receptor (EGFR) overexpressed in pancreatic or lung cancers and has shown to benefit the patients.[117-119] On the other hand, different targeting antibodies have been studied and developed. Bevacizumab is a recombinant humanized monoclonal antibody and, in 2004, it became the first clinically used angiogenesis inhibitor. This antibody binds to the vascular endothelial growth factor (VEGF) and prevents its binding to the receptor and induction of the angiogenic process. In the absence of the signal induced by the VEFG the tumor mass collapse due to an inadequate nurturing via the blood flow.[120] However, the productions of antibodies and proteins are associated with high cost and remain the major limitations of the protein-and the antibody-based biopharmaceuticals. Fortunately, peptides appear at the interface of small molecules and proteins in regard to the production

complexity and costs. Thus, they have been intensively investigated in pre-clinical studies for the development of new imaging moieties and drug conjugates to visualize and destroy tumor cells.[121]

Due to their unique properties, selective receptor-targeting peptide-based agents have also attracted attention in molecular imaging of tumors cells that overexpress corresponding peptide receptors. Among them, many radiolabeled peptides have already been translated into the clinic with impressive diagnostic accuracy and sensitivity.[122] Peptide-based imaging agents in nuclear medicine have tremendous utility in diagnosis, prognosis, and selection of therapeutic regimes for patients. Radiometals can be incorporated into clinically relevant peptides via a bifunctional chelator, providing effective and sensitive radiotracers that can be prepared conveniently in a radiopharmacy.[123-128]

2.1.3 c-Met

c-Met is a receptor tyrosine kinase that, after activated by its ligand, hepatocyte growth factor (HGF), mediates a different range of intracellular signaling pathways. Particularly, some of the signaling pathways are related to proliferation, motility, migration, and invasion of cancer cells.[129] Its natural ligand is the HGF and, of course, they have, under non-tumorigenic physiological conditions, different function (*e.g*. during early embryogenesis).[130, 131]

Aberrant HGF/c-Met signaling is involved in the development and metastatic progression of several tumor types, including colorectal cancer, gastric cancer, and gastro-esophageal cancer, renal, ovarian, lung, breast, cervical, pancreatic, as well as melanoma are others examples.[132-139] Aberrant c-Met signaling activation occurs frequently in gastrointestinal tumors and can result from multiple mechanisms, including protein overexpression, MET (the gene that encoded for c-Met) amplification or enhanced transcription and/or aberrant autocrine or paracrine secretion of HGF. Once activated, the c-Met signaling results in enhanced cancer cell proliferation, survival and invasion, and the elevated c-Met expression/amplification has been

associated with a poor clinical outcome, an elevated occurrence of metastasis, and an increased drug resistance of cancer.

All of these characteristics indicate that this protein receptor is an active participant in cancer initiation and progression, and thus the monitoring of c-Met expression in real time could potentially assist in the diagnosis and in the monitoring of response to therapy.[140-142]

Numerous studies also suggest that c-Met is an interesting target for small inhibitor molecules in cancer therapy because blocking of this protein pathway can inhibit tumor growth and metastasis.[143, 144] Multiple clinical trials have adopted small-molecules inhibitors and monoclonal antibodies against c-Met pathways as cancer therapeutic agents.[145-147]

Improved diagnostic methods for the identification of suitable patient population for c-Met targeted therapy are of fundamental importance to improve clinical outcome of c-Met aberrant cancers. Nowadays, the patient selection is normally centered on immunohistochemistry or fluorescent in situ hybridization. Even if these methods can provide quantitative information about c-Met abundance, they have important limitations in at least two different scenarios. They are not able to reflect the c-Met expression fluctuation over time, and they cannot deal with the c-Met heterogeneity in different tumor sites. Considering this scenario, PET imaging can overcome these limitations because of its high sensitivity in the real-time detection of molecular events.

Several radiolabeled antibodies against c-Met have already been used for *in vivo* tumor targeting. PRS-110, an anticalin (Pieris Pharmaceuticals, Inc.) with monovalent specificity for c-Met, was labeled with ^{89}Zr for imaging U87MG glioblastoma.[148] In another study, the authors successfully developed and characterized a modified HGF, that was then labeled with Copper-64 (^{64}Cu-NOTA-rh-HGF) for PET imaging of tumor c-Met expression *in vivo*. c-Met specificity of ^{64}Cu-NOTA-rh-HGF was demonstrated by several control experiments *in vitro*, *in vivo*, and *ex vivo*, making it a promising PET tracer with broad potential applications in the clinical diagnosis and

treatment monitoring of cancer.[149]

2.1.4 GLP-1

The glucagon-like peptide type-1 (GLP-1) is a 30-amino acids long peptide hormone released in the gut following the ingestion of nutrients. The peptide is involved in the stimulation of insulin release from the β-cells of the pancreas to maintain glucose homeostasis. GLP-1 is an incretin that aside from the increase in the glucose-dependent secretion of insulin, it also inhibits glucagon production, hepatic gluconeogenesis, gastric mobility, and suppresses appetite.[150] Because of this stimulation of insulin release, it is not surprising to find out that the target receptor for GLP-1, which is called GLP-1R, is highly expressed in the *β*-cells of the pancreas. However, the action of GLP-1, as every incretin, must be strictly controlled to prevent the development of hypoglycemia as the blood glucose concentration is reduced by the actions of insulin. To perform this task the human body rapidly degrades the peptide by the action of an enzyme called dipeptidyl-peptidase-IV (DPP-4).[151-153] Consequently, the native peptide is rapidly degraded and inactivated. GLP-1R agonists and inhibitors of DPP-4 that prolong the biological half-life of GLP-1, are used as drugs in the treatment of type 2 diabetes.

Furthermore, GLP-1R has been shown to be expressed in high density and high incidence in some types of cancers derived from endocrine, neuroendocrine, and embryonic origins.[154, 155] Among them, there is the insulinoma, which is a cancer of pancreatic *β*-cells of neuroendocrine origins that show high levels of both GLP-1R and somatostatin receptors. It is a rare form of cancer, and its lesions are difficult to detect by standard clinical imaging methods due to their small size and the anatomical location in the pancreas.[156]

Due to the rapid metabolization and inactivation of the GLP-1 is quite impossible to think of a targeting peptide with the same sequence.[157] In fact, some drugs that are normally used for the therapy of type 2 diabetes can be a peptide, but they obviously have a different sequence from the native GLP-1, and despite they maintain

the activity toward the GLP-1R, they are not metabolized by the DPP-4.[158] One example is Exendin-4. Exendin-4 is a subcutaneously administered peptide drug that is used in the treatment of type 2 diabetes.[159] Originally discovered in the saliva of a lizard, exendin-4 was the first of a class of incretin mimetics that showed potent glucoregulatory activity. There are different already published examples of a modified version of the exendin-4 adapted for the use in PET imaging. Some example are listed in Figure 2.3 ([Lys^{40}(Ahx-DOTA)]exendin-4,[160] [Lys^{40}(Ahx-DTPA)]exendin-4,[161] [Lys^{40}(AhxHYNIC)]exendin-4,[162, 163] [Cys^{40}(FBEM)]exendin-4,[164] [Cys^{0}(FBEM)]exendin-4[164]). In these examples, the primary sequence of the exendin-4 was in all cases extended to the carboxylic or the amino terminal groups and the added amino acid has the right characteristic for making it a radiopharmaceutical agent.[165]

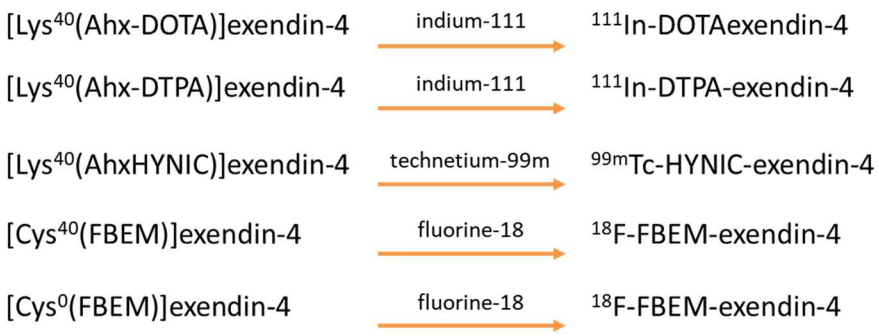


Figure 2.3. Example of GLP-1-like radioligands retaining high binding affinity to GLP-1R.

2.2. Design and synthesis of c-Met-peptide–THP

Considering that the noninvasive PET imaging with radiolabeled c-Met binders may support the selection of patients for c-Met–targeting drugs and can support the identification and the responding and nonresponding patients for such therapeutics, in this study it was decided to select a targeting peptide for c-Met and to conjugate to the peptide a THP, which can be able to chelate free gallium(III) and therefore to be used in PET imaging.

As already reported, the hepatocyte growth factor/c-Met signaling

axis has been described as a promoter of cancer cell growth, angiogenesis, invasion, and metastasis.[166] Moreover, c-Met is overexpressed in various solid tumors, including breast cancer and colorectal cancer.[167-172] In addition, its expression on the cell membrane makes extracellular epitopes accessible for targeting with fluorescent imaging agents as recently reported.[173]

GE-137 is a water-soluble 26-amino acid cyclic peptide labeled with a cyanine dye with a high affinity (K_d = 2 nM) for human c-Met (Figure 2.4).[173] The water-soluble 26-amino acid cyclic peptide component of the probe was selected in a recent study, from a library of peptides, on the basis of its high affinity for the extracellular domain of human c-Met, its lack of competition with hepatocyte growth factor and its lack of effect on HGF-stimulated proliferation *in vitro*.[173] Additionally, the lyophilized (dry) peptide has a documented shelf life of 12 months when stored at 2–8 °C, and the chemical and physical stability of the reconstituted product (in solution) has been demonstrated for 24 h at 2–8 °C, and is safe in humans.

For this study, it was decided to conjugate the same peptide (AGSCYCSGPPRFECWCYETEGT) with the same linker peptide (GGGK) and then to a hexadentate 3,4-hydroxypyridinone ligand (THP) and evaluate the applicability as gallium chelating agent and its use in PET imaging (Figure 2.4). The peptide was manually assembled using Fmoc chemistry starting with 0.1 mmol Rink Amide Novagel resin (Figure 2.5, step 1). An excess of pre-activated amino acids (1 mmol, using HATU) was applied in the coupling steps. The final synthesized sequence on the resin was:

Ac-Ala-Gly-Ser(tBu)-Cys(Trt)-Tyr(tBu)-Cys(StBu)-Ser(tBu)-Gly-Pro-Pro-Arg(Pbf)-Phe-Glu(OtBu)-Cys(Mmt)-Trp(Boc)-Cys(Trt)-Tyr(tBu)-Glu(OtBu)-Thr(tBu)-Glu(OtBu)-Gly-Thr(tBu)-Gly-Gly-Gly-Lys(Dde)–Rink Amide Novagel resin.

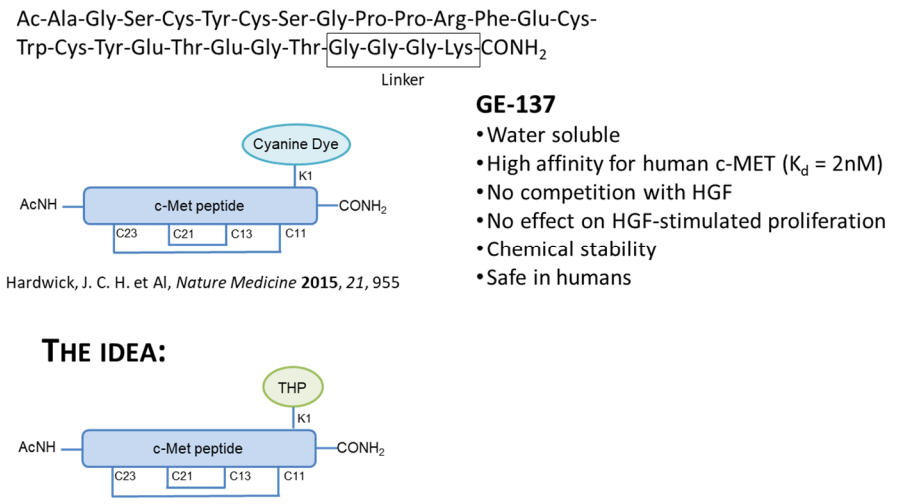


Figure 2.4. Structure and sequence of GE-137. Characteristics of GE-137. Design of c-Met-peptide-THP.

An accurate selection of the appropriates protecting groups on the different cysteines was studied for the selective formation of the different disulfide bonds between the C11–C23 and C13–C21. In the area of peptide science, a variety of chemical transformations have been accomplished on a solid support, including the formation of intramolecular disulfide bridges.[174] Conventional solution-phase techniques for intramolecular disulfide bond formation typically involve oxidation of free thiol precursors at high dilution. Similarly, the oxidation on solid support utilizes pseudo dilution, a kinetic phenomenon expected to favor facile intramolecular interactions, in the microporous resins.[175]

In the case of c-Met-peptide, two trityls-based protecting groups for the cysteines 11 and 23, Cys-(Trt) and Cys-(Mmt) were used and the other two cysteines were protected with *t*-butylthio (S-*t*-Bu) and 4-methoxytrityl. In the first step (Figure 2.6), *S*-*t*-Bu was removed by reduction to liberate free thiol, treating the resin with 20% mercaptoethanol in DMF for 3 h. The resin was then reacted with a 10-fold excess of DTNP (2,2'-Dithiobis(5-nitropyridine)) in DCM for 1 h, and the free thiol was thus protected and activated with the 5-nitropyridin-sulfenyl (5-Npys) group. This was followed by the

cyclization step in which the resin was treated with 1% TFA in DCM in the presence of TIPS (Triisopropylsilane) as the scavenger.

The disulfide bond formation is based on the sequence of two disulfide exchange reactions, both of which do not require an oxidizing agent and can be carried out in an inert (N_2) atmosphere. The first reaction involves a nucleophilic attack of free thiol, obtained following the deprotection of Cys(*S*-*t*-Bu), on the activated disulfide of DTNP. Because of the polymer-bound nature of the thiol, the byproduct 5-nitropyridine-thione can be easily washed along with the excess of DTNP. The cyclization is the second disulfide exchange reaction that involves another nucleophilic attack of the free thiol, obtained following the deprotection of Cys-(Mmt), on the polymer-bound 5-Npys-activated disulfide. Solution-phase counterparts of such disulfide exchange reactions require acidic conditions and inert atmosphere to avoid cross-linking and dimeric products. Moreover, it is well-known that trityl-based functional groups can be selectively cleaved under different acidic conditions. In fact, it has been reported that Mmt can be selectively cleaved in the presence of the Trt group under mildly acidic conditions of 0.5–1% TFA in DCM.[176, 177] The coupling of the peptide with a THP derivative was made after the first disulfide bond. The deprotected THP (**19**) was synthesized and used (Figure 2.7). After the formation of the C13–C21 disulfide bond (Figure 2.5, step 2), the lysine in position 1 was selectively deprotected (Dde removal) in presence of a 2% solution of hydrazine in DMF (Figure 2.5, step 3). After the removal of the Dde, the THP (**19**) was coupled to the peptide using HATU as a coupling agent (Figure 2.5, step 4). THP (**19**) was synthesized by the reaction with Glutaric anhydride and THPd (**18**). After the coupling of the hexadentate tris(3,4-hydroxypyridinone), the peptide was removed from the resin (Figure 2.5, step 5), the Trt groups were removed and the second disulfide bond was made in air-oxidation condition (Figure 2.5, step 6). Particularly, the air-oxidation reaction was performed on the crude products dissolved in a 0.1 M NH_4HCO_3 solution.

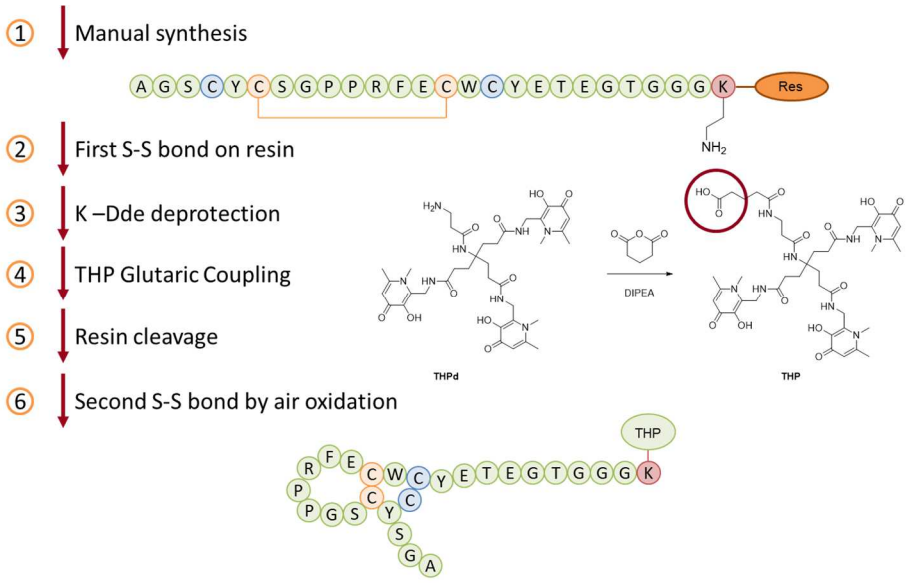


Figure 2.5. Schematic synthesis of c-Met-peptide-THP.

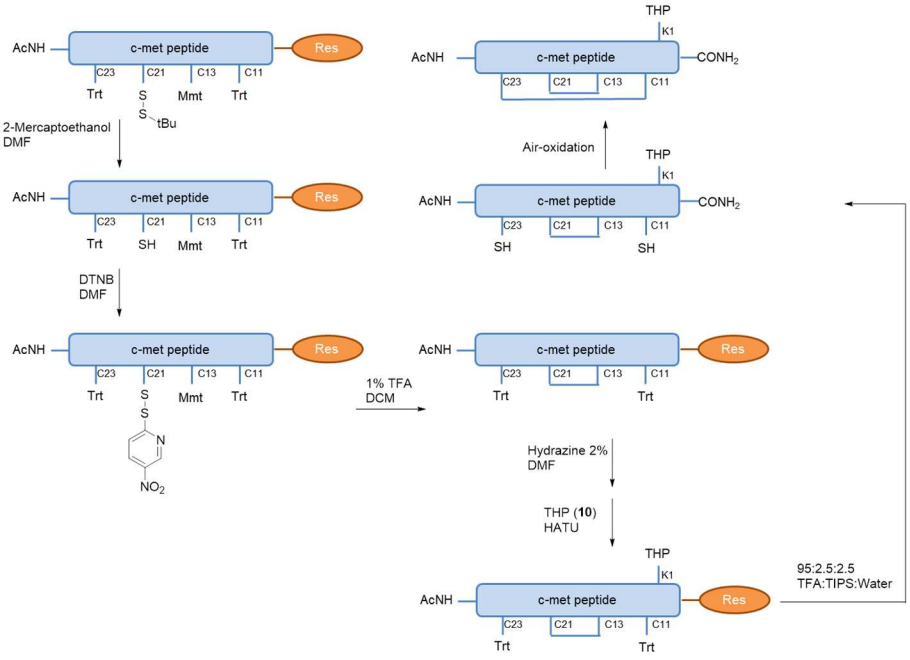


Figure 2.6. Schematic synthesis of the disulfide bonds formation in the c-Met-peptide-THP.

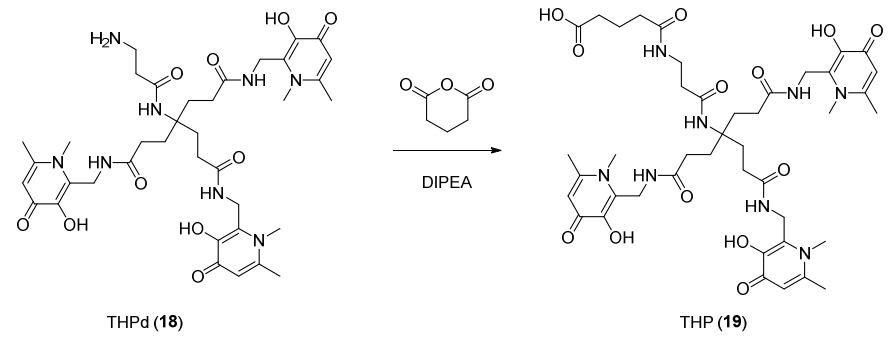


Figure 2.7. Synthesis of THP (**19**).

Once finished the synthesis of the c-Met-peptide-THP (**20**, Figure 2.8), the compound was chromatographed *via* preparative HPLC/UV. The isolated fractions were further analyzed by analytical HPLC-DAD, and the pure fractions analyzed by mass spectrometry (HRMS, ESI+). The effective presence of the final product was confirmed by the mass spectrometry experiments as highlighted in Figure 2.9, where the $(M + 4H)^{4+}$ of c-Met-peptide-THP is showed.

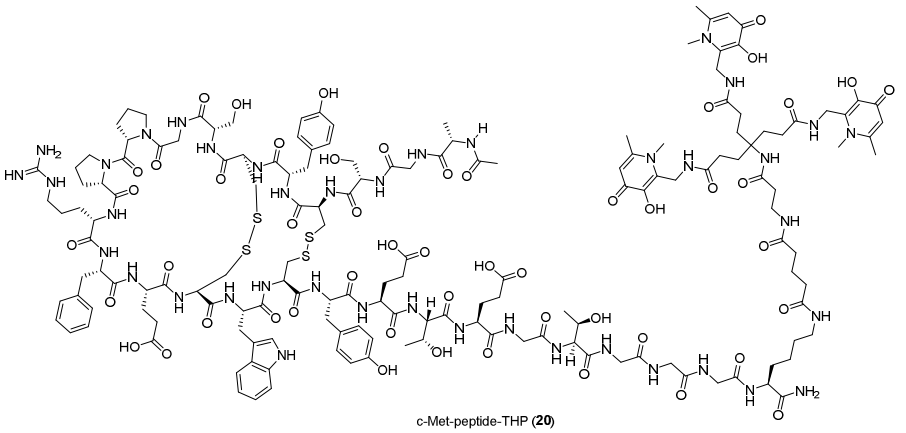


Figure 2.8. Structure of c-Met-peptide-THP.

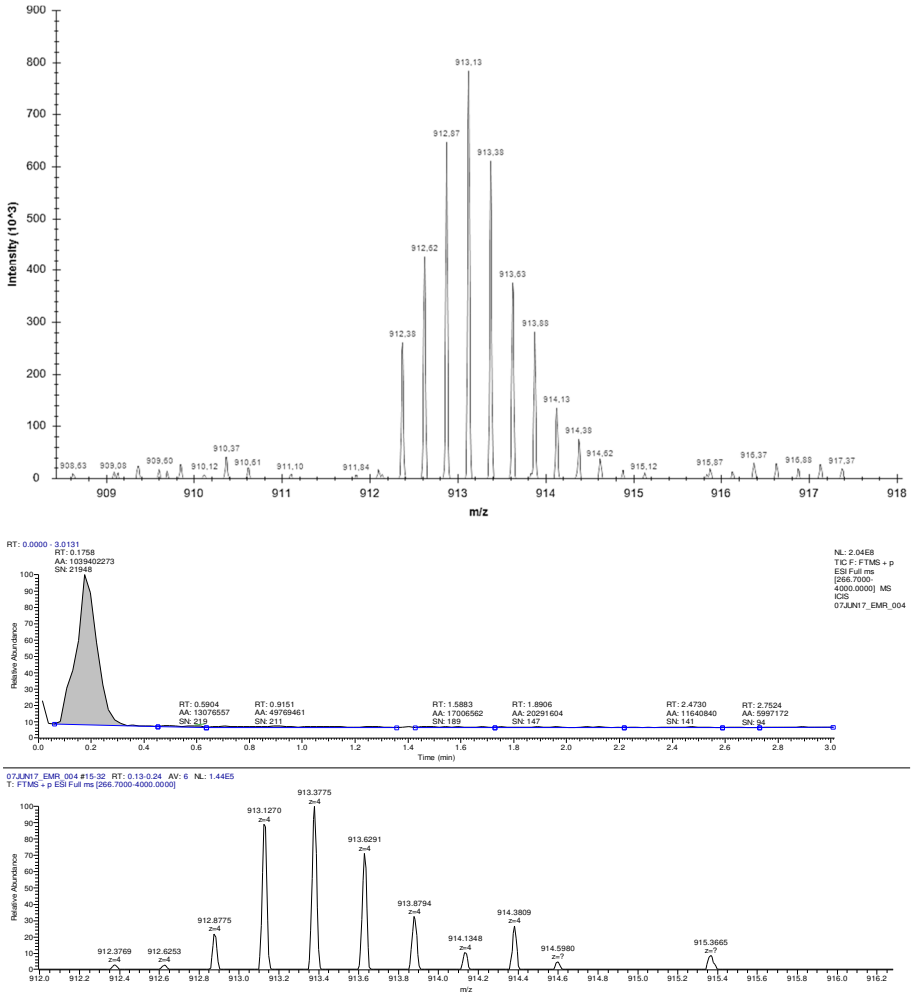


Figure 2.9. Mass spectrum showing $(M + 4H)^{4+}$ where M = c-Met-peptide-THP.

2.3. Design and synthesis of GLP-1-peptide–THP

Considering the PET imaging with radiolabeled GLP-1R binders may support the selection and the diagnosis of patients for GLP-1R aberrant cancer, in this study it was decided to select a targeting peptide for GLP-1R and to conjugate to the peptide a THP, which can be able to chelate the gallium(III) and therefore to be used in PET imaging.

As already reported the GLP-1R has been shown to be expressed in high density and high incidence in some types of cancers derived from endocrine, neuroendocrine, and embryonic origins (e.g. insulinoma). A PET imaging tracer can be particularly useful for the detection of the

small lesions related with the insulinoma and other types of GLP-1R related cancers.

Exendin-4 is a peptide already used in the therapy of type 2 diabetes, and it was already reported that the peptide can be opportunely modified for the use in PET imaging making it an active radiopharmaceutical agent. As shown in Figure 2.3 the peptide can be modified by the insertion of a terminal linker amino acid (e.g. lysine or cysteine) without losing its activity/affinity toward the GLP-1 receptor. For this study, it was decided to conjugate the same peptide (HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS) with a lysine as a linker peptide to the carboxylic terminal position and then to a hexadentate 3,4-hydroxypyridinone ligand (THP) (Figure 2.10) and evaluate the applicability as gallium chelating agent and its use in PET imaging. The peptide was manually assembled using Fmoc chemistry starting with 0.1 mmol 2-Chlorotrityl chloride resin (Figure 2.11, step 1). An excess of pre-activated amino acids (1 mmol, using HATU) was applied in the coupling steps.

The final synthesized sequence on the resin was:
Ac-Ser-Pro-Pro-Pro-Ala-Gly-Ser-Ser-Pro-Gly-Gly-Asn-Lys-Leu-Trp-Glu-Ile-Phe-Leu-Arg-Val-Ala-Glu-Glu-Glu-Met-Gln-Lys-Ser-Leu-Asp-Ser-Thr-Phe-Thr-Gly-Glu-Gly-His-Lys(Dde)–2-Chlorotrityl chloride resin.

Figure 2.10. Structure and sequence of GLP-1-peptide (modified Exendin-4).

Differently to the c-Met targeting peptide, that required the formation of two selective disulfide bonds, after the synthesis of the linear sequence of the GLP-1-peptide, it only required the coupling with a THP to the linker terminal amino acid. After the removal of the Dde (Figure 2.11, step 2), the same THP (**19**, Figure 2.7) used for the c-Met targeting peptide was coupled to the peptide using HATU as a coupling agent (Figure 2.11, step 3). After the coupling of the hexadentate

tris(3,4-hydroxypyridinone) (**19**), the peptide was removed from the resin (Figure 2.11, step 4).

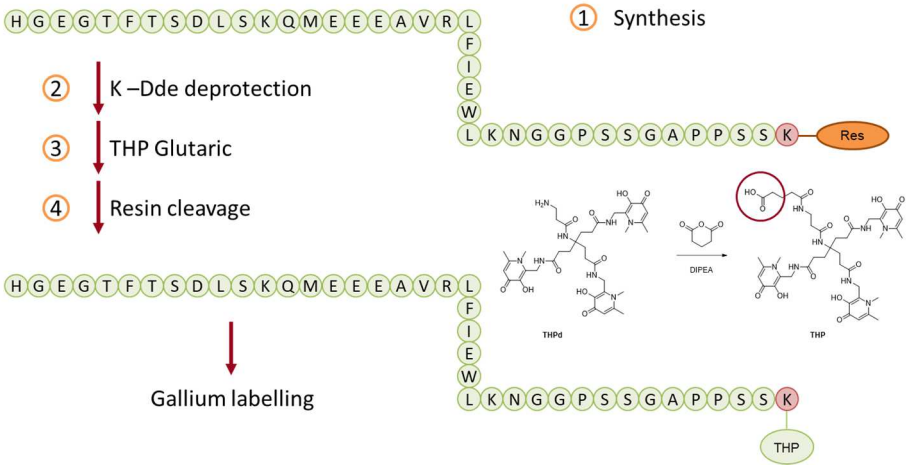


Figure 2.11. Schematic synthesis of GLP-1-peptide-THP.

Once finished the synthesis of the GLP-1-peptide-THP (**21**), the compound was chromatographed *via* preparative HPLC/UV. The isolated fractions were further analyzed by analytical HPLC-DAD, and the pure fractions analyzed by mass spectrometry (HRMS, ESI^+). The effective presence of the final product was confirmed by the mass spectrometry experiments as highlighted in Figure 2.12, where the $(M + 4H)^{4+}$, $(M + 5H)^{5+}$, $(M + 6H)^{6+}$, $(M + 7H)^{7+}$, GLP-1-peptide-THP are showed.

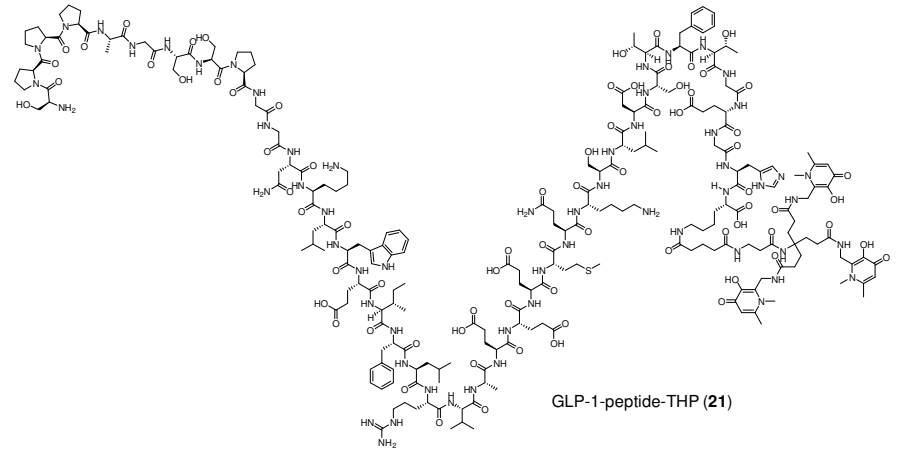


Figure 2.12. Structure of GLP-1-peptide-THP.

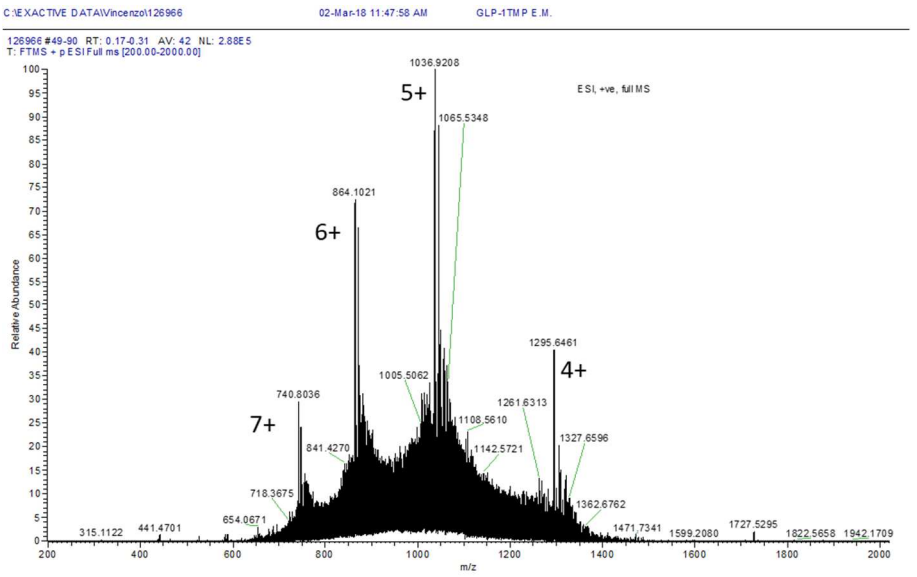


Figure 2.13. Mass spectrum showing GLP-1-peptide-THP.

2.4. Radio HPLC and ITLC of C-met-peptide-THP and GLP-1-peptide-THP

To evaluate the chelating ability of the synthesized compounds, C-met-peptide-THP and GLP-1-peptide-THP, for ^{68}Ga, instant thin layer chromatography (ITLC) analyses and radio HPLC analyses were used. Two different methods used in the ITLC analyses. The different methods were called “acetate method” and “citrate method” and the results are reported in Table 2.2. Unfortunately, the acetate method was unable to identify the chelating products for both of the peptides, in fact in these conditions both C-met-peptide-THP and GLP-1-peptide-THP have a retention factor of 0 which is the same of the free gallium. Differently, the citrate method clearly describes differences between the free gallium and the gallium in the presence of the analyzed peptides. In these cases, both of the peptides have a retention factor of 0, but the free gallium in the presence of the citrate has a retention factor of 0.8–1. This result means that in both cases the peptides are chelating the gallium and constrain it to maintain the position of the peptide (that is not moving in the TLC) because the strong chelation.

Similar results were obtained with the HPLC analyses. In this case, two different column were used: reverse phase (C18) and size exclusion. Even in this case, the results tell that C-met-peptide-THP and GLP-1-peptide-THP are both able to chelate the ^{68}Ga (Table 2.3). In the HPLC analyses, the reverse phase column shows retention times of 1.9, 13–14 and 13.7–13.9 min for ^{68}Ga unbound, C-met-peptide-THP and GLP-1-peptide-THP, respectively. In the size exclusion column, the retention times were 8–13, 14–15 and 12–15 min for ^{68}Ga unbound, C-met-peptide-THP and GLP-1-peptide-THP, respectively. Each analysis gave a single major peak in the HPLC radiochromatogram (See Experimental part), and under these conditions, the radiochemical yield for all of these products was >95% (determined by ITLC).

Table 2.2. Results of the ITLC analyses.

	Acetate method R_f	**Citrate method R_f**
^{68}Ga unbound	0	0.8-1
[^{68}Ga]C-met-peptide-THP	0	0
[^{68}Ga]GLP-1-peptide-THP	0	0

Table 2.3. Results of the HPLC analyses.

	Reverse phase R_t (min)	**Size Exclusion R_t (min)**
^{68}Ga unbound	1.9	8–13
[^{68}Ga]C-met-peptide-THP	13–14	14–15
[^{68}Ga]GLP-1-peptide-THP	13.7–13.9	12–15

2.5 Conclusion and perspectives

The combination of targeted imaging agents with PET allows a non-invasive diagnosis of specific cancers with precise delineation of tumors and metastases and thus disease staging. Moreover, the quantification of the target receptor expression, the uptake kinetics, and the pre-therapeutic dosimetry may allow a more effective selection of the treatments and planning as well as monitoring response to the therapy and early detection of recurrent disease resulting in personalized medicine and, in particular, radiotheranostic.

The main aims of the individualized patient management are to

optimize the treatments and so the therapeutic and to minimize risks and toxicity as well as reduce cost and patient distress. Clinical studies with different ^{68}Ga-based imaging agents demonstrated the significance of individualized patient management.[178-180]

We truly believe that, in future clinics, noninvasive PET imaging with radiolabeled tracers such as the synthesized c-Met-peptide-THP and GLP-1-peptide-THP may support the diagnosis of particular types of cancers with overexpression of HGFr or GLP-1R and the selection of patients for c-Met–targeting drugs and identify responding and nonresponding patients for such therapeutics.

In conclusion, considering the excellent results in the synthesis of C-met-peptide-THP and GLP-1-peptide-THP and the optimal preliminary ^{68}Ga chelating efficiency for both of them, other assays are planned for going further in the projects. The serum stability of the different peptides will be evaluated and then two opportune animal models will be used to evaluate the *ex vivo* and the *in vivo* biodistribution of the C-met-peptide-THP and GLP-1-peptide-THP and the efficiency as PET imaging agents will be of course also evaluated.

Chapter 3. Experimental part

Molecular modeling and QSAR model

All of the compounds were drawn using Marvin Sketch[181] and subjected to a first molecular mechanics energy minimization by Merck molecular force field (MMFF94) optimization using the Marvin Sketch geometrical descriptors plugin.[181] The protonation states of the molecules were calculated assuming a pH of 7. After having obtained the 3D structures for all compounds, the geometry was also optimized at semi-empirical level using the parameterized model number 3 (PM3) semi-empirical Hamiltonian as implemented in MOPAC package (vMOPAC2016).[56, 182, 183] Docking was performed using AutodockVina[59] and AutoDock4[184] using the default docking parameters, the point charges were initially assigned according to the AMBER14 force field,[185] and then damped to mimic the less polar Gasteiger charges used to optimize the AutoDock scoring function. The setup was done with the YASARA molecular modeling program.[186, 187] For Autodock4 Docking was performed by applying the Lamarckian genetic algorithm (LGA) implemented in AutoDock. The ligand-centered maps were generated by the program AutoGrid with a spacing of 0.375 Å and dimensions that encompass all atoms extending 5 Å from the surface of the ligand. All of the parameters were inserted at their default settings. In the docking tab, the macromolecule and ligand are selected, and GA parameters are set as ga_runs = 100, ga_pop_size = 150, ga_num_evals = 20000000, ga_num_generations = 27000, ga_elitism = 1, ga_mutation_rate = 0.02, ga_crossover_rate = 0.8, ga_crossover_mode = two points, ga_cauchy_alpha = 0.0, ga_cauchy_beta = 1.0, number of generations for picking worst individual = 10. The binding modes were clustered through the root mean square deviation among the Cartesian coordinates of the ligand atoms. The docking results were ranked based on the calculated binding free energy. The binding modes with the lowest binding free energy and the most cluster members were selected for optimum docking conformation. PyMOL (downloadable at

https://www.pymol.org/) is an open-source visualization system that produces high-quality 3D-images of small molecules and biological macromolecules (Schrödinger, LLC). The binding results in our study were illustrated using PyMOL Molecular Graphics System (version 1.8).

Forge 10.4.2 (Cresset BioMolecular Discovery Ltd., https://www.cresset-group.com/) was used for the building of the 3D-QSAR model. Spark 10.4.0 was used for the scaffold-hopping analysis (Cresset BioMolecular Discovery Ltd., https://www.cresset-group.com/). Forge's parameters used for the conformation hunt, alignment and used for the build of the QSAR model are reported in Figures 3.1–3.3. The model statistics for the 3D-QSAR model are reported in Figure 3.4. Spark's parameters used for the bioisosteric replacement are reported in Figure 3.5. in Tables 3.1 and 3.2 are reported the compounds used as a training set and as a test set for the set-up of the QSAR model.

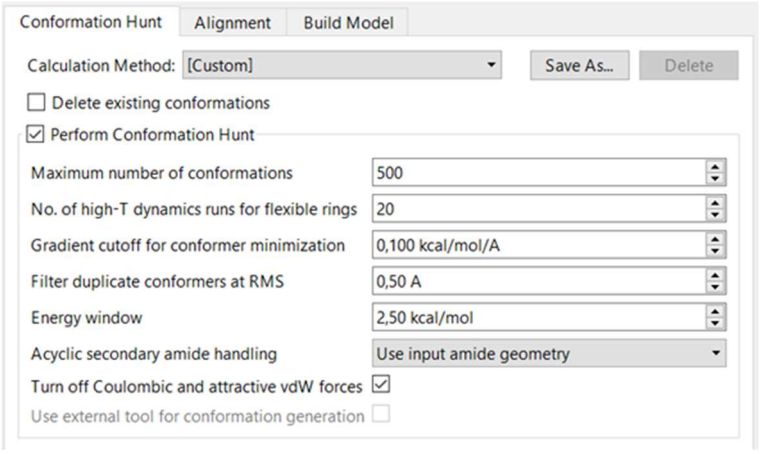


Figure. 3.1. Forge's parameters used for the conformation hunt.

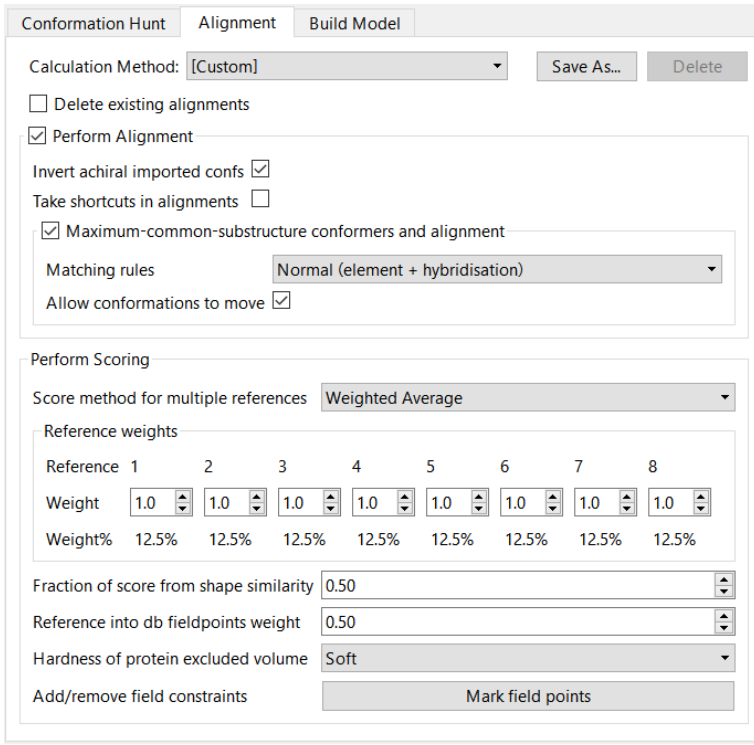


Figure. 3.2. Forge's parameters used for the alignment.

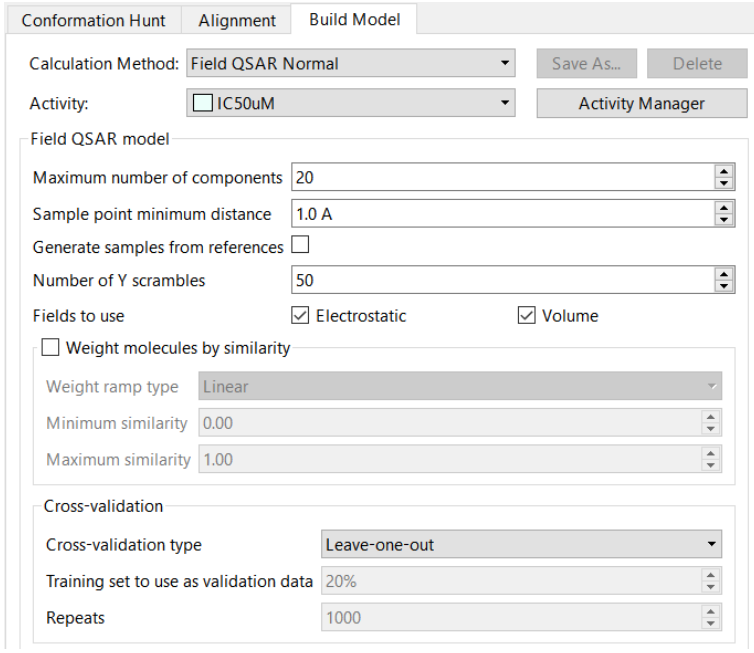


Figure. 3.3. Forge's parameters used for the build of the QSAR model.

Model statistics:
==========

Comps	R^2	Q^2	Test R^2	RMSE	RMSEpred	Tau	Tau-pred
0	-0.044	-0.023	-0.001	1.093	1.096	0.494	-0.910
1	0.620	0.508	0.641	0.666	0.767	0.587	0.492
2	0.781	0.541	0.625	0.501	0.740	0.675	0.514
3	0.897	0.585	0.718	0.347	0.711	0.787	0.573
4	0.948	0.618	0.745	0.244	0.681	0.855	0.568
5	0.968	0.636	0.743	0.192	0.663	0.895	0.598
6	0.983	0.651	0.772	0.142	0.651	0.920	0.594
7	0.988	0.671	0.746	0.118	0.631	0.929	0.621
8	0.993	0.673	0.757	0.088	0.629	0.947	0.621
9	0.996	0.683	0.744	0.068	0.619	0.958	0.633
10	0.997	0.687	0.733	0.058	0.616	0.957	0.626
11*	0.998	0.688	0.731	0.044	0.615	0.969	0.626
12	0.999	0.683	0.729	0.037	0.619	0.969	0.618
13	0.999	0.680	0.726	0.030	0.622	0.976	0.616
14	1.000	0.679	0.731	0.022	0.623	0.983	0.613
15	1.000	0.679	0.730	0.018	0.623	0.985	0.615
16	1.000	0.678	0.728	0.015	0.624	0.988	0.616
17	1.000	0.678	0.728	0.011	0.625	0.988	0.615
18	1.000	0.677	0.730	0.008	0.625	0.991	0.615
19	1.000	0.678	0.728	0.006	0.625	0.993	0.616
20	1.000	0.677	0.728	0.005	0.625	0.995	0.616

Preferred number of components: 11

Figure. 3.4. Model statistics for FABP4 model.

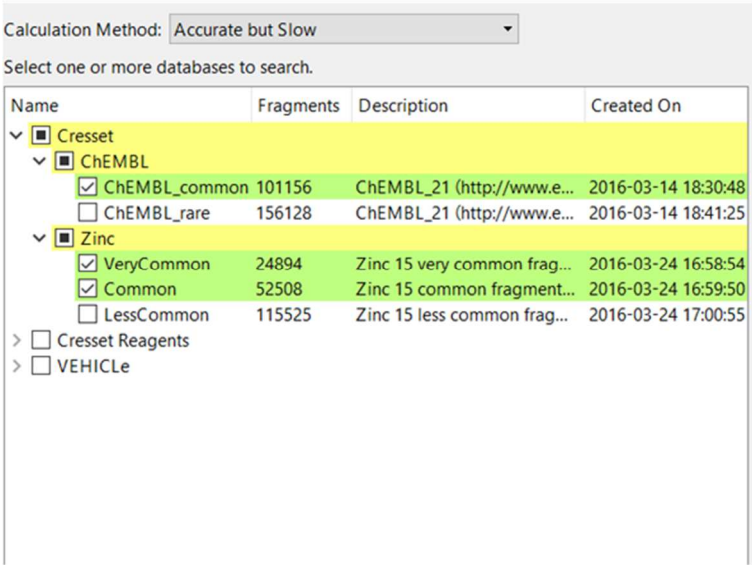


Figure. 3.5. Spark's parameters used for the bioisosteric replacement.

Table 3.1. SMILES, experimental and predicted pIC_{50} values of the molecules in the training set.

		pIC_{50}	
N°	**SMILE**	**Exp.**	**Pred.**
1	FC(F)(F)[C@H]1CCc2c(C1)c(c(c(n2)C3CCCC3)C=4[N-]N=NN4)-c5ccnc(c5)C	8	8
2	CC1(CCCC1)c2c(c(c3c(n2)CCCCC3)-c4ccnc(c4)C)C=5[N-]N=NN5	8	8
3	Clc1c(F)cc2c(c(c(c(N(CC)CC)n2)C=3[N-]N=NN3)-c4ccccc4)c1	7.88	7.9
4	Clc1c(F)cc2c(c(c(c(n2)C(CC)CC)C=3[N-]N=NN3)-c4ccccc4)c1	7.78	7.8
5	OCC1(CCCC1)c2c(c(c3c(n2)CCCCC3)-c4ccnc(c4)C)C=5[N-]N=NN5	7.7	7.7
6	CCCCC[C@H]1CCc2c(C1)c(c(c(n2)C3(CCCC3)COC)C=4[N-]N=NN4)-c5ccccc5	7.7	7.7
7	FC(F)(F)c1ccc2c(c(c(c(N3CCCCC3)n2)C=4[N-]N=NN4)-c5ccccc5)c1	7.49	7.5
8	Clc1ccc2c(c(c(c(n2)C3CC3)C([O-])=O)-c4ccccc4)c1	7.37	7.4
9	Clc1ccc2c(c(c(c(N(CC)C)n2)C=3[N-]N=NN3)-c4ccccc4)c1	7.34	7.4
10	Clc1cc(Cl)cc(NC(=O)NC2(CCCC2)C([O-])=O)c1-c3ccccc3	7.3	7.3
11	Clc1c(F)cc(c(NC(=O)NC2(CCCC2)C([O-])=O)c1)-c3ccccc3	7	7
12	O=C(N)c1ccccc1Cn2c3c(cccc3c4CCCCCc42)C([O-])=O	7.01	7
13	n1c2c(CCCCC2)c(c(c1C3CCCCC3)C=4[N-]N=NN4)-c5ccncc5	6.96	6.9
14	Clc1ccc(c(NC(=O)NC2(CCCC2)C([O-])=O)c1)-c3ccc(F)cc3	6.85	6.9
15	FC(F)(F)c1ccccc1Cn2c3c(cccc3c4CCCCc42)C([O-])=O	6.4	6.5
16	Fc1ccc(-c2c(c(n(n2)-c3ccccc3-c4cccc(OCC([O-])=O)c4)CC)-c5ccccc5)cc1	6.48	6.5
17	[O-]C(=O)c1cccc2c3CCCCCc3n(c12)Cc4ccccc4	6.23	6.3
18	Fc1ccccc1Cn2c3c(cccc3c4CCCCc42)C([O-])=O	6.4	6.3
19	Fc1cccc(Cn2c3c(cccc3c4CCCCc42)C([O-])=O)c1	6.41	6.3
20	FC(F)(F)c1ccccc1Cn2c3c(cccc3c4CCCCCc42)C([O-])=O	6.19	6.3
21	[O-]C(=O)CCCn1c2ccccc2c3ccccc31	6.24	6.3
22	FC(F)(F)c1ccc(c(NC(=O)NC2(CCCC2)C([O-])=O)c1)-c3ccccc3	6.28	6.2
23	[O-]C(=O)c1cccc2c3CCCCc3n(c12)Cc4cccc(OC)c4	6.26	6.2
24	Fc1cccc(Cn2c3c(cccc3c4CCCCCc42)C([O-])=O)c1	6.14	6.2
25	FC(F)(F)c1cc(O)nc(SCc2ccc(OC)cc2)n1	6.22	6.2
26	[O-]C(=O)c1ccc2c(n(c3CCCCc23)Cc4ccccc4)c1	6.12	6.1
27	[O-]C(=O)c1cccc2c3CCCc3n(c12)Cc4ccccc4	6.01	6.1
28	[O-]C(=O)c1cccc2c3CCCCc3n(c12)Cc4ccccc4OC	6.22	6.1
29	[O-]C(=O)c1cccc2c3CCCCc3n(c12)Cc4ccc(C)cc4	5.96	6.1
30	Fc1ccccc1Cn2c3c(cccc3c4CCCCCc42)C([O-])=O	6.15	6.1
31	Fc1ccc(Cn2c3c(cccc3c4CCCCCc42)C([O-])=O)cc1	6.11	6.1
32	[O-]C(=O)CCCCn1c2ccccc2c3ccccc31	6.1	6.1

33	FC(F)(F)c1cccc(Cn2c3c(cccc3c4CCCCCc42)C([O-])=O)c1	5.98	6
34	FC(F)(F)c1cc(O)nc(SCC(=O)N2CCCCC2)n1	6	6
35	O=S(=O)(n1ccc2ccc(cc21)C)c3ccsc3C([O-])=O	5.89	5.9
36	Brc1ccc2c(ccn2S(=O)(=O)c3ccsc3C([O-])=O)c1	5.89	5.9
37	FC(F)(F)c1cccc(Cn2c3c(cccc3c4CCCCc42)C([O-])=O)c1	5.78	5.7
38	FC(F)(F)c1ccc(Cn2c3c(cccc3c4CCCCc42)C([O-])=O)cc1	5.62	5.7
39	FC(F)(F)c1ccc(Cn2c3c(cccc3c4CCCCCc42)C([O-])=O)cc1	5.69	5.7
40	O=S(=O)(n1cc(c2ccccc21)C)c3ccsc3C([O-])=O	5.82	5.7
41	[O-]C(=O)c1cccc2c3CCCCc3n(c12)Cc4ccc(OC)cc4	5.57	5.6
42	[O-]C(=O)[C@H](Oc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)c1)C	5.63	5.6
43	O=S(=O)(n1ccc2cccc(OC)c21)c3ccsc3C([O-])=O	5.59	5.6
44	O/N=C/1CCCc2c1c3cccc(c3n2Cc4ccccc4)C([O-])=O	5.46	5.5
45	Clc1cccc(-n2c(-c3ccccc3)cc(n2)-c4ccccc4OCCCC([O-])=O)c1	5.6	5.5
46	[O-]C(=O)[C@H](Oc1cccc(-c2ccccc2-n3c(c(c(n3)-c4ccccc4)-c5ccccc5)CC)c1)CC	5.54	5.5
47	Fc1ccc2c(ccn2S(=O)(=O)c3ccsc3C([O-])=O)c1	5.46	5.5
48	[O-]C(=O)c1cccc2c(c(n(c12)Cc3ccccc3)C)C	5.44	5.4
49	Clc1ccc(-n2c(-c3ccccc3)cc(n2)-c4ccccc4OCCCC([O-])=O)cc1	5.36	5.4
50	Clc1ccccc1-n2c(-c3ccccc3)cc(n2)-c4ccccc4OCCCC([O-])=O	5.37	5.4
51	[O-]C(=O)c1c(C(C)C)cc(C(C)C)cc1C(C)C	5.4	5.4
52	O=S(=O)(n1c2ccccc2c3ccccc31)c4ccccc4C([O-])=O	5.37	5.4
53	Fc1ccc2ccn(S(=O)(=O)c3ccsc3C([O-])=O)c2c1	5.33	5.4
54	FC(F)(F)c1cc(O)nc(NCc2ccc(OC)cc2)n1	5.41	5.4
55	[O-]C(=O)CCCOc1ccccc1-c2cc(n(n2)-c3ccccc3)-c4ccc(cc4)C	5.27	5.3
56	Brc1ccc(-n2c(-c3ccccc3)cc(n2)-c4ccccc4OCCCC([O-])=O)cc1	5.25	5.3
57	Fc1ccc(-c2c(nn(c2CC)-c3ccccc3-c4cccc(OCC([O-])=O)c4)-c5ccccc5)cc1	5.33	5.3
58	[O-]C(=O)CCCCOc1ccccc1-c2cc(n(n2)-c3ccccc3)-c4ccccc4	5.19	5.2
59	O=S(=O)(n1ccc2cc(ccc21)C)c3ccsc3C([O-])=O	5.15	5.2
60	O=S(=O)(n1ccc2ccc(OC)cc21)c3ccccc3C([O-])=O	5.21	5.2
61	Brc1ccc(-c2cc(nn2-c3ccccc3)-c4ccccc4OCCCC([O-])=O)cc1	5.02	5
62	Fc1ccc(-n2c(-c3ccccc3)cc(n2)-c4ccccc4OCCCC([O-])=O)cc1	4.99	5
63	[O-]C(=O)CCCOc1ccccc1-c2cc(n(n2)-c3ccc(C(C)C)cc3)-c4ccccc4	4.99	5
64	[O-]C(=O)CCn1c2ccccc2c3ccccc31	5.03	5
65	O=S(=O)(n1ccc2c(cccc21)C)c3ccsc3C([O-])=O	5.13	5
66	O=S(=O)(n1ccc2cc(OC)ccc21)c3ccsc3C([O-])=O	5	5
67	O=S(=O)(n1cc(c2ccccc21)C)c3ccccc3C([O-])=O	5.07	5
68	O=S(=O)(n1ccc2c(cccc21)C)c3ccccc3C([O-])=O	4.92	4.9
69	Brc1ccc2c(ccn2S(=O)(=O)c3ccccc3C([O-])=O)c1	4.92	4.9

70	[O-]C(=O)CCCOc1ccccc1-c2cc(n(n2)-c3ccc(OC)cc3)-c4ccccc4	4.87	4.8
71	[O-]C(=O)CCCOc1ccccc1-c2cc(n(n2)C3CCCCCC3)-c4ccccc4	4.81	4.8
72	Brc1ccc2c(n(S(=O)(=O)c3c(C(C)C)cc(C(C)C)cc3C(C)C)cn2)c1	4.83	4.8
73	Clc1ccc2c(nc(n2S(=O)(=O)c3c(C(C)C)cc(C(C)C)cc3C(C)C)C)c1	4.83	4.8
74	O=S(=O)(n1cncc1)c2c(C(C)C)cc(C(C)C)cc2C(C)C	4.77	4.8
75	Clc1ccccc1CNc2nc(O)cc(n2)C(F)(F)F	4.64	4.7
76	FC(F)(F)c1cc(O)nc(n1)CCc2ccc(OC)cc2	4.6	4.7
77	O=C1CCCc2c1c3cccc(c3n2Cc4ccccc4)C([O-])=O	4.6	4.6
78	[O-]C(=O)CCCOc1ccccc1-c2cc(n(n2)C3CCCCC3)-c4ccccc4	4.63	4.6
79	O=S(=O)(n1ccc2cc(ccc21)C)c3ccccc3C([O-])=O	4.52	4.6
80	FC(F)(F)c1cc(O)nc(n1)N(Cc2ccccc2)C	4.62	4.6
81	Clc1ccc(-c2cc(nn2-c3ccccc3)-c4ccccc4OCCCCCCCC([O-])=O)cc1	4.51	4.5
82	FC(F)(F)c1cc(O)nc(NCC(=O)N2CCCCC2)n1	4.43	4.4
83	Clc1cccc(CNc2nc(O)cc(n2)C(F)(F)F)c1	4.48	4.4
84	FC(F)(F)c1cc(O)nc(NCc2ccc(C)cc2)n1	4.48	4.4
85	Clc1ccc(-c2cc(nn2-c3ccccc3)-c4ccccc4OCCCCCC([O-])=O)cc1	4.13	4.2
86	Brc1ccc(-c2cc(nn2-c3ccccc3)-c4ccccc4OCCCCCC([O-])=O)cc1	4.07	4.1
87	O=S(=O)(n1ccc2c(OC)cccc21)c3ccccc3C([O-])=O	4.05	4.1
88	O=S(=O)(N)c1c(C(C)C)cc(C(C)C)cc1C(C)C	4	4
89	[O-]C(=O)Cn1c2ccccc2c3ccccc31	4	4
90	FC(F)(F)c1cc(O)nc(n1)NCc2ccc(-c3ccccc3)cc2	4	4
91	FC(F)(F)c1cc(O)nc(NCc2ccncc2)n1	4	4
92	FC(F)(F)c1cc(O)nc(n1)CCc2ccccc2	4	4
93	FC(F)(F)c1cc(O)nc(NCCc2ccccc2)n1	4	3.9
94	[O-]C(=O)CCCCOc1ccccc1-c2cc(n(n2)-c3ccccc3)-c4ccc(cc4)C	3.59	3.6
95	Clc1ccc(CNc2nc(O)cc(n2)C(F)(F)F)cc1	5.54	3.5
96	Clc1ccc(-c2cc(nn2-c3ccccc3)-c4ccccc4OCC([O-])=O)cc1	2	2

Table 3.2. SMILES, experimental, and predicted pIC_{50} values of the molecules in the test set.

		pIC_{50}	
N°	**SMILE**	**Exp.**	**Pred.**
1	FC(F)(F)c1ccc2c(c(c(c(N(CC)CC)n2)C=3[N-]N=NN3)-c4ccccc4)c1	7.6	7.8
2	Clc1c(F)cc2c(c(c(c(N3CCCCC3)n2)C=4[N-]N=NN4)-c5ccccc5)c1	7.88	7.3
3	Clc1ccc(c(NC(=O)NC2(CCCC2)C([O-])=O)c1)-c3ccccc3	6.82	6.5
4	O=C(N)c1cccc(Cn2c3c(cccc3c4CCCCCc42)C([O-])=O)c1	7.16	6.2
5	[O-]C(=O)c1ccc2c(c3CCCCc3n2Cc4ccccc4)c1	4.6	6.1
6	Fc1ccc(Cn2c3c(cccc3c4CCCCc42)C([O-])=O)cc1	6.08	6.1
7	[O-]C(=O)c1cccc2c3CCCCc3n(c12)Cc4ccccc4	6.23	5.9
8	Fc1cccc(c1Cn2c3c(cccc3c4CCCCc42)C([O-])=O)C(F)(F)F	5.74	5.9

9	O=S(=O)(n1c2ccccc2c3ccccc31)c4ccsc4C([O-])=O	5.96	5.9
10	[O-]C(=O)c1cccc2c3CCCCCc3n(CCC)c12	6.37	5.7
11	[O-]S(=O)(=O)c1c(C(C)C)cc(C(C)C)cc1C(C)C	5.1	5.7
12	O=S(=O)(n1ccc2ccc(OC)cc21)c3ccsc3C([O-])=O	5.59	5.7
13	[O-]C(=O)c1cccc2c3CCCCc3n(CCC)c12	6.14	5.6
14	Fc1cccc2ccn(S(=O)(=O)c3ccsc3C([O-])=O)c12	5.35	5.4
15	[O-]C(=O)CCCOc1ccccc1-c2cc(n(n2)-c3ccccc3)-c4ccccc4	5.52	5.3
16	Clc1ccc(-c2cc(nn2-c3ccccc3)-c4ccccc4OCCCC([O-])=O)cc1	5.21	5.2
17	Fc1cccc2c1ccn2S(=O)(=O)c3ccccc3C([O-])=O	5	5.2
18	Clc1ccc(CN(c2nc(O)cc(n2)C(F)(F)F)C)cc1	5.4	5.1
19	FC(F)(F)c1cc(O)nc(Nc2ccccc2)n1	4	4.8
20	Brc1ccc2c(n(S(=O)(=O)c3c(C(C)C)cc(C(C)C)cc3C(C)C)c(n2)C)c1	4.08	4.7
21	O=S(=O)(n1c(nc2ccccc21)C)c3c(C(C)C)cc(C(C)C)cc3C(C)C	4	4.6
22	[O-]C(=O)CCCOc1ccccc1-c2cc(n(n2)C3CCCC3)-c4ccccc4	4.78	4.5
23	O=S(=O)(n1ccc2c(OC)cccc21)c3ccsc3C([O-])=O	4.89	4.3
24	FC(F)(F)c1cc(O)nc(n1)NCc2ccccc2	4.48	4.2

FABP Inhibitory Activity Assays.

To analyze the inhibitory activity of the compounds FABP4, a displacement assay was utilized as described by the Cayman's instruction, FABP4 Inhibitor/Ligand Screening Assay Kit, Item № 10010231. The compounds (AST_1–3) for activity determination were prepared as a stock solution (1 mM) in DMSO (dimethyl sulfoxide). On the day of activity assay, the compounds were all diluted in phosphate buffered solution (PBS, pH 7.4) to different concentrations (50, 25, 12.5, 6.25, 3.12, 1.56 μM). Appropriate concentrations of DMSO in PBS was used as control. For activity assay, the detection reagent (FABP Assay Detection Reagent, Item № 10010376) was diluted in PBS to a final concentration of 10 μM. The diluted Detection Reagent probe was mixed with 10 μM FABP protein and incubated for 10 min at room temperature. Compounds were then added and equilibrated for another 10 min. The fluorescence signal was finally read at 370 (excitation)/470 (emission) with a CytoFluor® Series 4000 Fluorescence Multi-Well Plate Reader. The plotted displacement curves for Arachidonic acid (positive control) and AST_1–3 are reported in Figures 3.6–3.9.

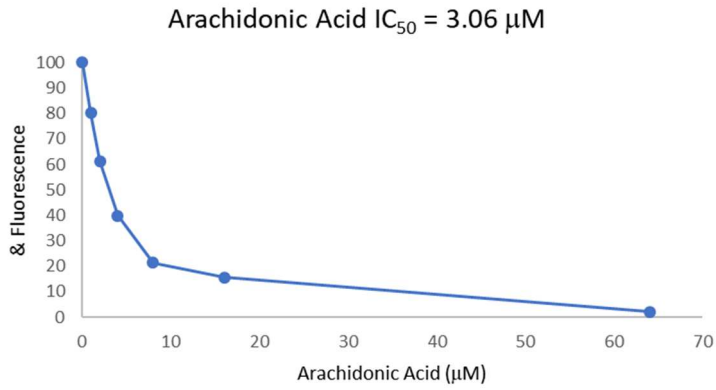


Figure. 3.6. Displacement curve for Arachidonic Acid.

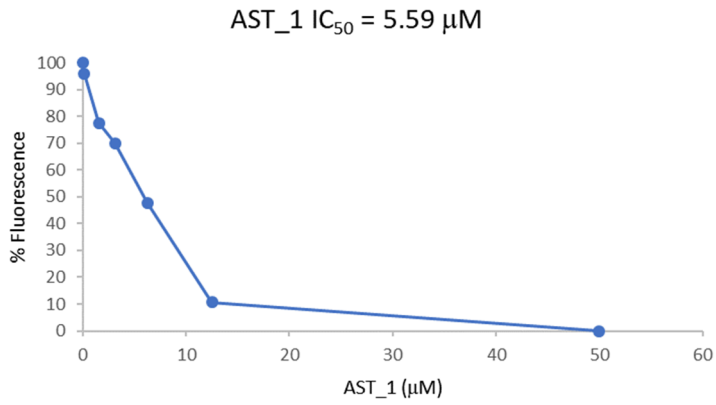


Figure. 3.7. Displacement curve for AST_1.

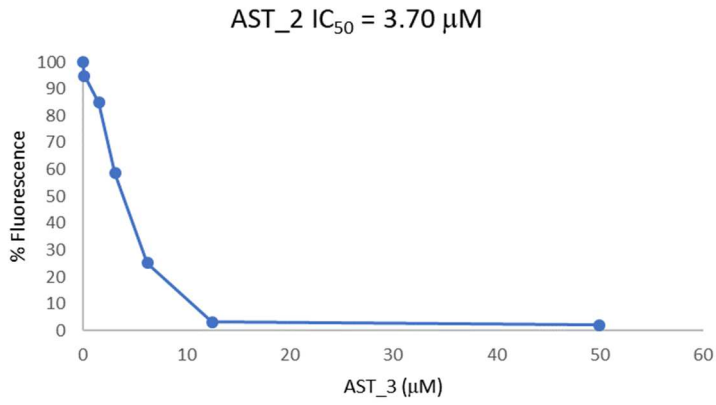


Figure. 3.8. Displacement curve for AST_2.

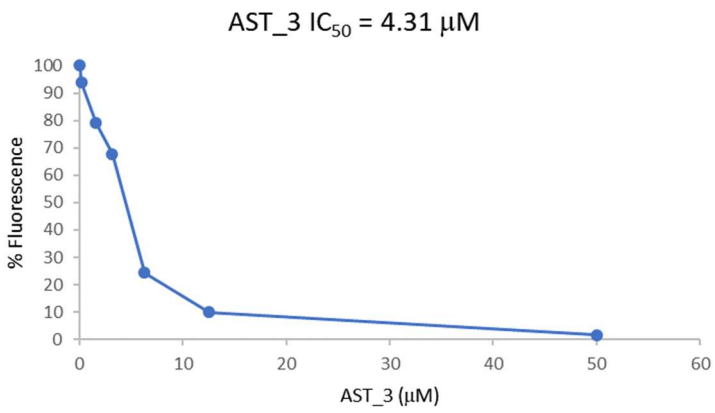


Figure. 3.9. Displacement curve for AST_3.

General information for the synthesis

All chemicals were purchased from Merk, Sigma Aldrich, and Acros Organics and were reagent grade or better. Solvents were purchased from Fisher Scientific, Sigma Aldrich, and VWR. NMR solvents were purchased from GOSS Scientific and NMR tubes were manufactured by Wilmad. Silica gel for column chromatography was purchased from Merck. Samples were dried in a vacuum oven (Gallenkamp, rated to ≤ 250 °C) connected to a vacuum pump (BOC-Edwards.). Pre-coated aluminum sheets (silica gel 60 F254, Merck) were used for thin-layer chromatography (TLC) and spots were visualized under UV light. ^{1}H NMR and ^{13}C NMR spectra were recorded on Varian UNITY Inova spectrometer using $CDCl_3$ or DMSO-d_6 as a solvent and tetramethylsilane (TMS) as an internal standard, at 200 or 500 MHz for ^{1}H NMR and 125 MHz for ^{13}C NMR. ^{13}C spectra were ^{1}H decoupled and multiplicities were determined by the APT pulse sequence. Chemical shift (δ) values are given in ppm. All of the NMR experiments were analyzed using the General NMR Analysis Toolbox (GNAT) [188]. Mass spectra were run at the King's College (Institute of Pharmaceutical Science, London, Stamford Street, London SE1 9NH, UK) on a Thermofisher LCQ DECA XP ion trap mass spectrometer or a Waters - Micromass ZQ - Single quadrupole mass spectrometer. High-resolution LC-MS was performed at the Division of Imaging Sciences, King's College, London on an Agilent 6520 Accurate-Mass Q-TOF LC/MS

connected to an Agilent 1200 HPLC system with a vacuum degasser, quaternary pump, variable wavelength detector, and autosampler. Isotopic distributions were calculated using Molecular Weight Calculator version 6.46. The resulting graph data were extracted and processed in Microsoft Excel. HPLC analyses were performed on Agilent 1100 series HPLC with quadrupole pump, vacuum degasser, variable wavelength detector (set to 254 or 281 nm). The HPLC column was an Agilent Eclipse XDB-C18 (4.6 mm × 150 mm) with a guard column. All the syntheses and experiments for the compounds of the first chapter were performed at the Department of Drug Sciences, University of Catania. The measurement of the IC_{50} of AST_1–3 were performed at the Institute of Pharmaceutical Science, King's College London. All the syntheses and experiments for the compounds of the second chapter were performed at the Institute of Pharmaceutical Science, King's College London.

General procedure for the synthesis of the alcohols 10

A solution of benzyl bromide **8** (0.02 mol) in Et_2O (50 mL) was added dropwise over a period of 2h to a stirred suspension of Mg (0.025 mol) in (50 mL) at room temperature. The mixture was stirred for 3 h more at room temperature to produce benzyl Grignard **9** and cooled to –40 °C. To it, a solution of aldehyde **7** (0.01mol) in Et_2O (50 mL) was added over a period of 1 h. The mixture was stirred for 2 h at –40 °C, gradually brought to room temperature and stirred for 3 h more. Saturated aqueous NH_4Cl (10 mL) was added to it. The mixture was extracted with EtOAc. The combined organic extract was washed with water, brine and dried. Solvent removal under reduced pressure and column chromatography of the residue (silica gel, 93/7, Cyclohexane/EtOAc) afforded the pure alcohol **10**. Yields 92–99%.

General procedure for the synthesis of ketones 11

To a stirred suspension of pyridinium chlorochromate (0.012 mol) in CH_2Cl_2 (60 mL) was added a solution of **10** (0.008 mol). The mixture was stirred at room temperature for nearly 3 h until the disappearance of the starting material (monitored by TLC), diluted with diethyl ether

(50 mL) and filtered through a column of celite. Solvent removal of the filtrate under reduced pressure and column chromatography of the residue (silica gel, 95/5, Cyclohexane/EtOAc) afforded pure **11**. Yields 95–99%.

General procedure for the synthesis of α-Bromoketone 12

Bromine (3.8 mmol) was added dropwise to a solution of **11** (2.52 mmol) in freshly distilled $CHCl_3$ (10 mL). The mixture was heated under reflux until conversion was complete (monitored by TLC). After the mixture had cooled, the solvent was evaporated under reduced pressure to provide a residue that was chromatographed (silica gel, Cyclohexane) to yield pure α-Bromoketone **12**. Yields 70–85%.

General procedure for the synthesis of 2-Aminothiobenzamide 14

To a solution of **13** (14.7 mmol) in THF (75 mL) was added the Lawesson's reagent (8 mmol). The mixture was stirred under N_2 at room temperature for 24h. The solvent was evaporated, and the residue partitioned between EtOAc (50 mL), and 1 N HCl (30 mL). To the aqueous layer was added aq sat. $NaHCO_3$ until pH 8–9. The basic solution extracted with EtOAc (30 mL). The combined organic layers were dried (Na_2SO_4) and filtered. The solvent was evaporated and the resulting solid residue recrystallized from toluene to afford pure **14** as a yellow solid; yield 72%.

General procedure for the synthesis of thiazole 15

The 2-aminothiobenzamide **14** (1.5 mmol) was added to a solution of α-Bromoketone **12** (1 mmol) in methanol. The mixture was refluxed overnight. After the mixture had cooled, the solvent was evaporated under reduced pressure to provide a residue that was column-chromatographed (silica gel, 99/1, Cyclohexane/EtOAc) to yield pure thiazole **15**. 15_1: yield 75%, yellow solid, m.p.: 168–170 °C. ^{1}H NMR (200 MHz, Chloroform-*d*) δ 6.22 (bs, 2H), 6.58 – 6.88 (m, 2H), 7.04 – 7.47 (m, 9H), 7.57 (dddd, *J* = 8.7, 7.8, 7.1, 3.3 Hz, 3H). ^{13}C NMR (50 MHz, Chloroform-*d*) δ 113.63, 116.77, 116.82, 127.60, 127.68, 127.92, 128.16, 128.55, 128.63, 128.87, 129.58, 129.83, 130.59, 145.64,

149.16, 166.96. 15_2: yield 76%, yellow solid, m.p.: 178–180 °C. ^{1}H NMR (200 MHz, Chloroform-*d*) δ 6.42 – 6.85 (m, 2H), 7.17 – 7.54 (m, 9H), 7.54 – 7.89 (m, 6H), 7.99 – 8.22 (m, 1H), 8.68 (s, 1H). ^{13}C NMR (50 MHz, Chloroform-*d*) δ 110.92, 115.13, 126.18, 126.27, 126.89, 127.69, 127.79, 128.19, 128.34, 128.78, 129.26, 129.72, 130.87, 131.32, 131.94, 132.56, 132.86, 133.37, 147.61, 148.92.
15_3: yield 81%, yellow solid, m.p.: 172–173 °C.

General procedure for the synthesis of 3-butynoic acid 17

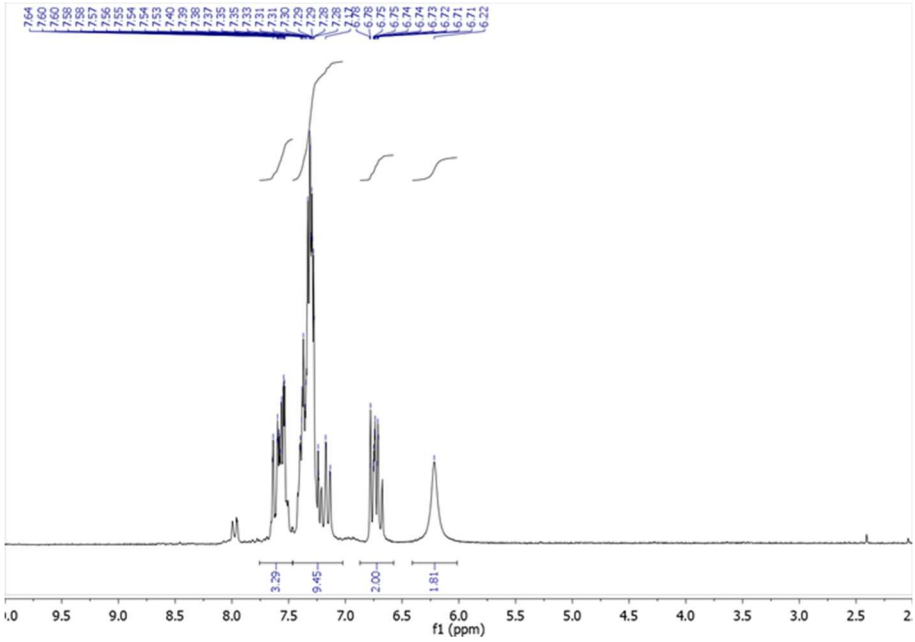
A solution of 21 g (0.3 mol) of 3-butyn-1-ol **16** in acetone (300 mL) was added dropwise to a mechanically stirred solution of 60 g of chromium trioxide in 750 mL of 10 N H_2S0_4 at 0 °C. The reaction mixture was allowed to warm to room temperature and stirring was continued for 1.5 h. The liquid was decanted into a mixture of EtOAc and H_2O. The aqueous layer was extracted five times with EtOAc. The extracts were combined, washed twice with saturated NaCl solution, dried (Na_2SO_4), and evaporated. The solvent was evaporated and the resulting solid residue recrystallized from hexane to afford pure **17** as a white solid, yield 93%.

General procedure for the synthesis of final compound AST_1–3

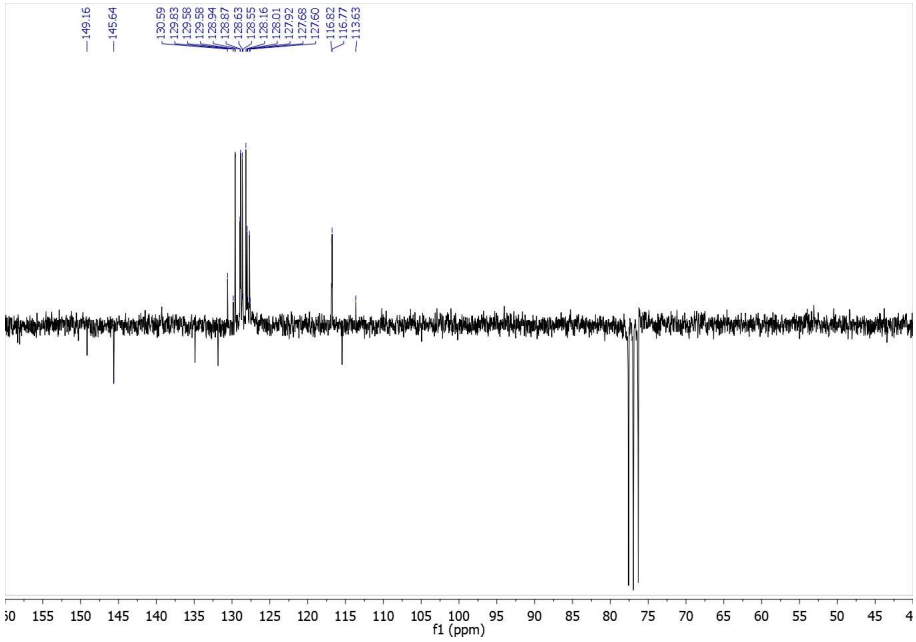
Thiazole **15** (0.54 mmol) was dissolved in THF (2 mL) in a 25 mL round-bottomed flask and cooled to 0 °C in an ice bath. To this stirred mixture was added *t*-BuONO (0.81 mmol) followed by $TMSN_3$ (0.65 mmol) dropwise. The resulting solution was stirred at room temperature for 2 h. 3-Butinoic acid **17** (0.54 mmol), an aq. solution (0.2 mL) of $CuSO_4$ (0.027 mmol), and sodium ascorbate (0.108 mmol) were then added and the reaction was stirred overnight at room temperature. Most of the solvent was evaporated and the product was then precipitated with water to give the final compounds **AST_1**–**3**. AST_1: white solid, m.p.: 238–240 °C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 3.62 (s, 2H), 7.11 – 7.47 (m, 10H), 7.56 (d, *J* = 7.5 Hz, 1H), 7.73 (dt, *J* = 25.3, 7.5 Hz, 2H), 8.08 – 8.47 (m, 2H). AST_2: white solid, m.p.: 241–243 °C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 3.57 (s, 2H), 7.28 (d, *J* = 7.9 Hz, 2H), 7.36 (dd, *J* = 19.7, 7.0 Hz, 3H), 7.43 – 7.52 (m, 2H), 7.55 (d, *J* = 7.9

Hz, 1H), 7.64 – 7.74 (m, 1H), 7.74 – 7.83 (m, 3H), 7.83 – 7.91 (m, 1H), 7.95 (s, 1H), 8.26 (d, *J* = 6.7 Hz, 2H). AST_3: white solid, m.p.: 241–242 °C. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 3.98 (s, 2H), 7.13 (t, *J* = 7.8 Hz, 1H), 7.17 – 7.24 (m, 2H), 7.25 – 7.37 (m, 5H), 7.44 (t, *J* = 1.9 Hz, 1H), 7.50 (dd, *J* = 7.8, 1.3 Hz, 1H), 7.59 (td, *J* = 7.7, 1.5 Hz, 1H), 7.66 (td, *J* = 7.7, 1.5 Hz, 1H), 7.89 (s, 1H), 8.15 (dd, *J* = 7.7, 1.5 Hz, 1H). ^{13}C NMR (50 MHz, Chloroform-*d*) δ 29.68, 126.94, 127.93, 128.11, 128.72, 128.86, 128.96, 129.35, 129.53, 130.35, 130.70, 134.07, 136.05, 160.12.

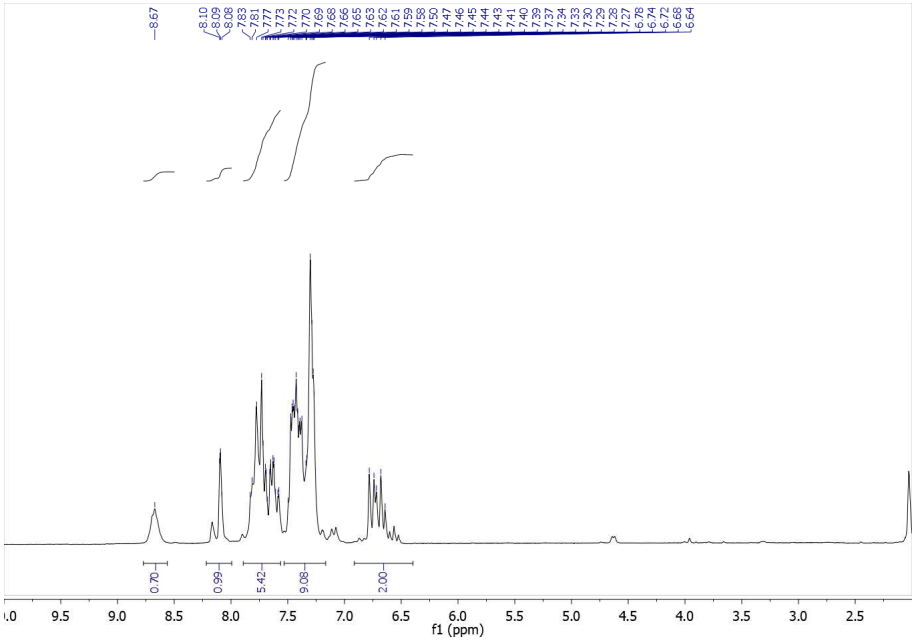
^{1}H NMR of molecule 15_1



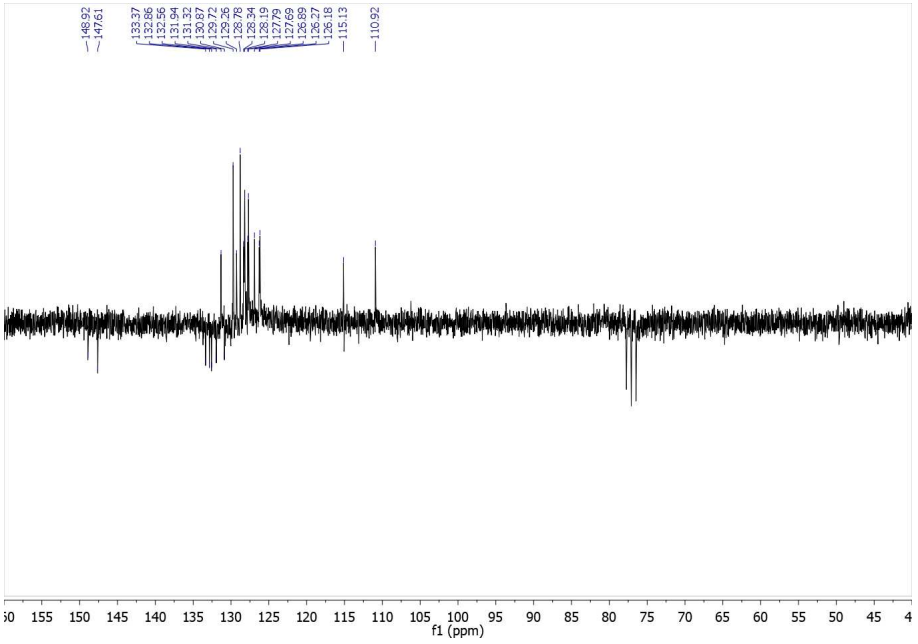
^{13}C NMR of molecule 15_1



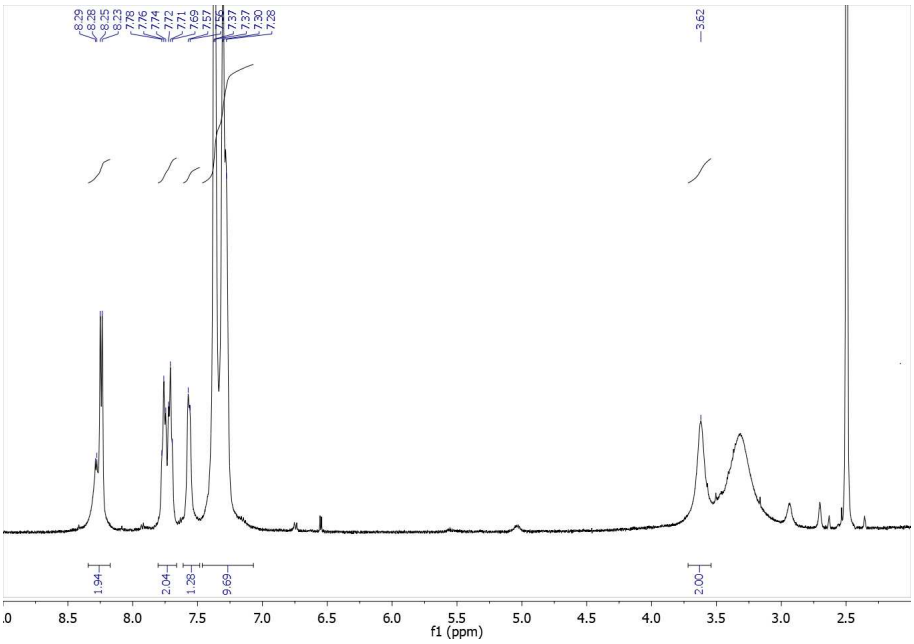
^{1}H NMR of molecule 15_2



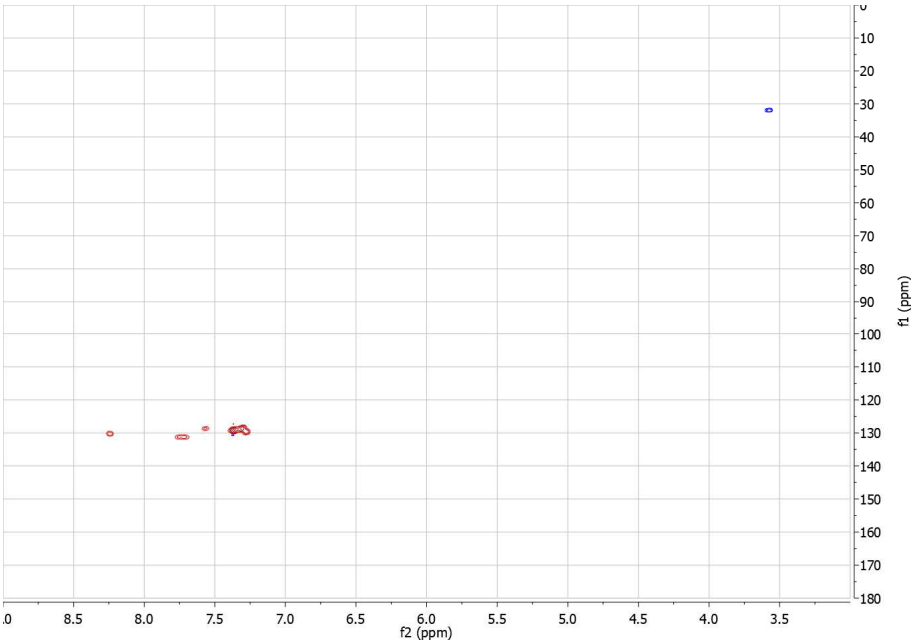
^{13}C NMR of molecule 15_2



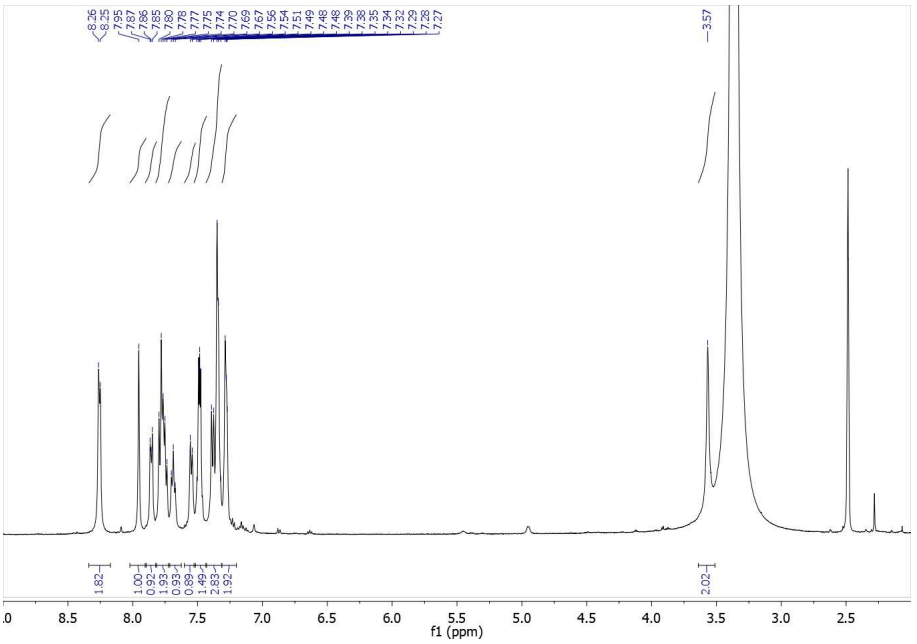
^{1}H NMR of AST_1



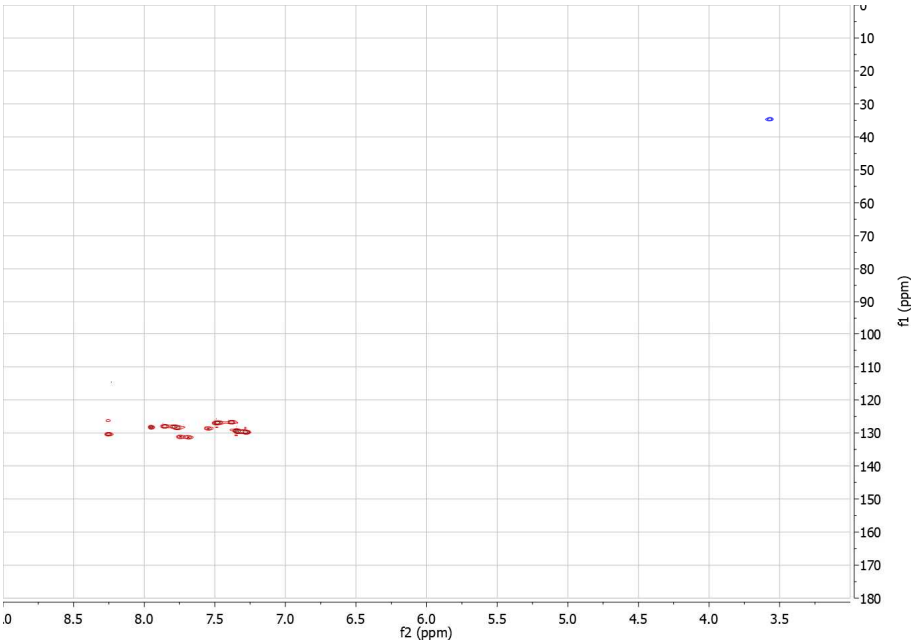
HSQCAD NMR of AST_1



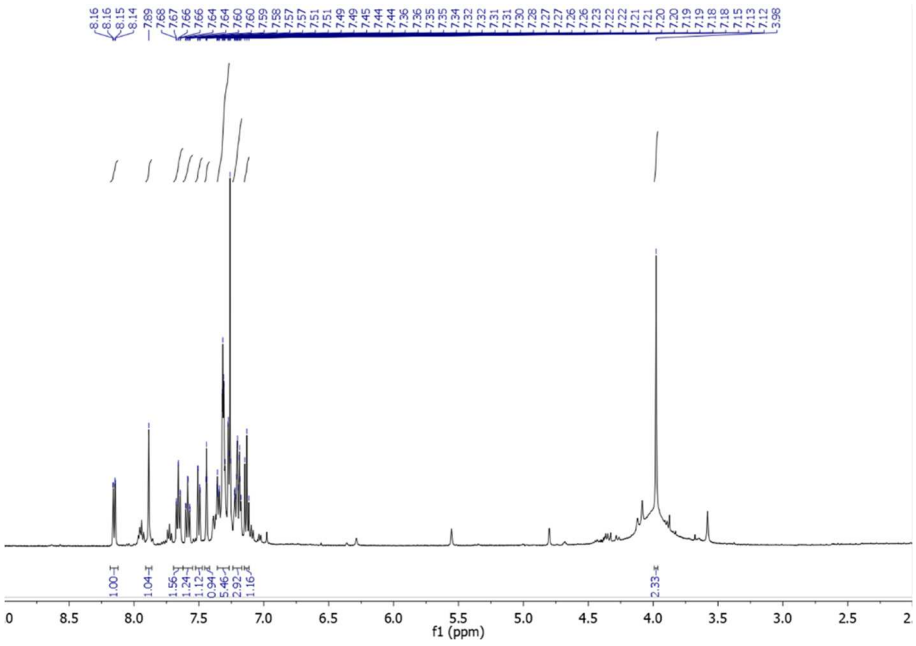
^{1}H NMR of AST_2



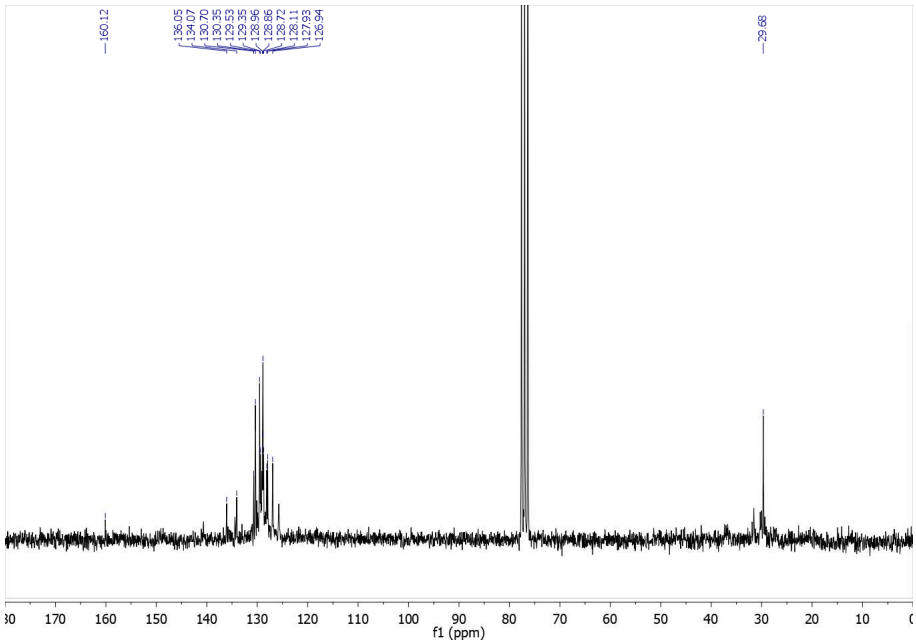
HSQCAD NMR of AST_2



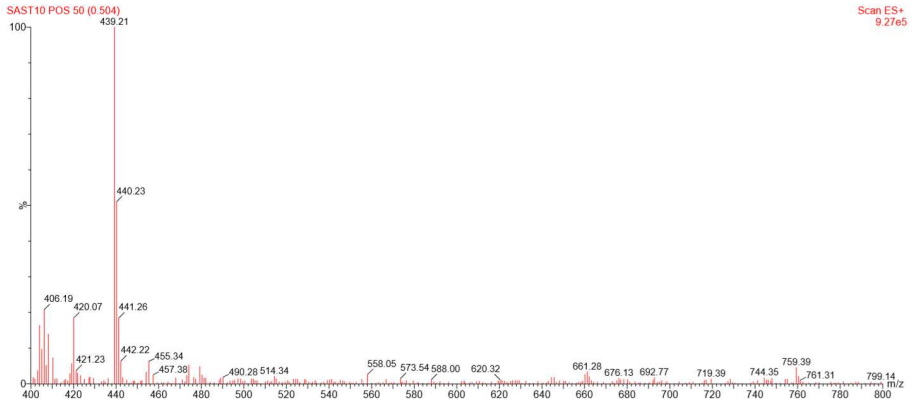
^{1}H NMR of AST_3



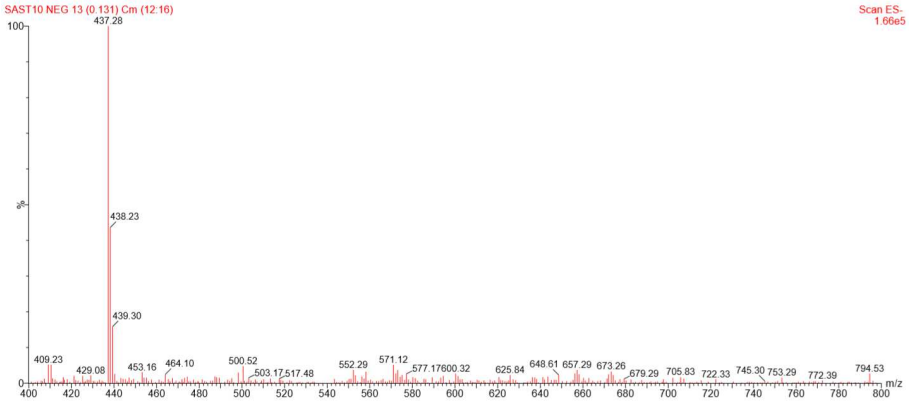
^{13}C NMR of AST_3



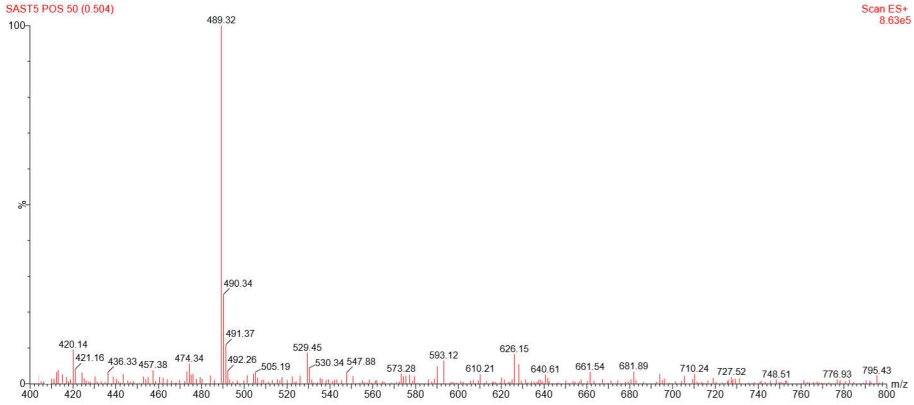
Mass spectrum (ESI+) of AST_1



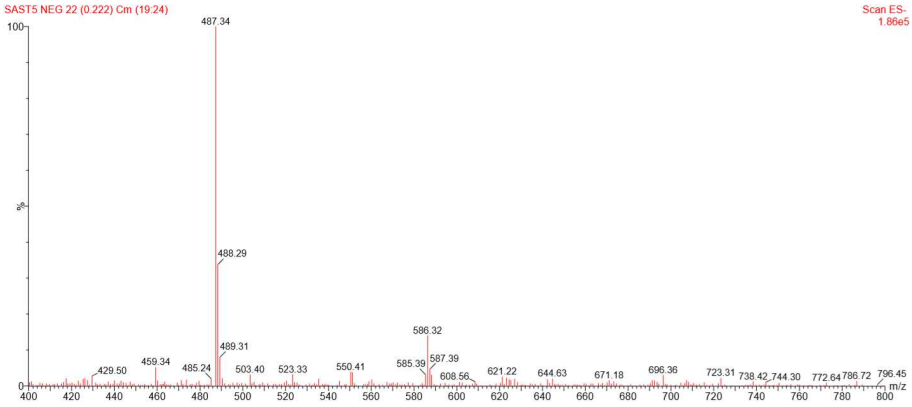
Mass spectrum (ESI–) of AST_1



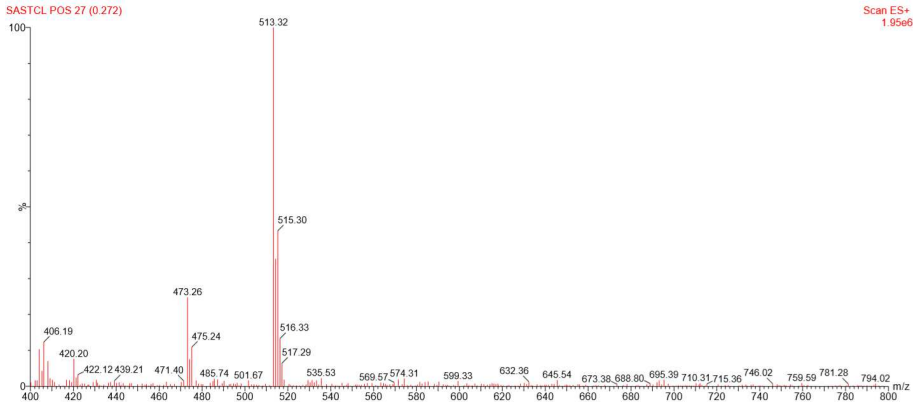
Mass spectrum (ESI+) of AST_2



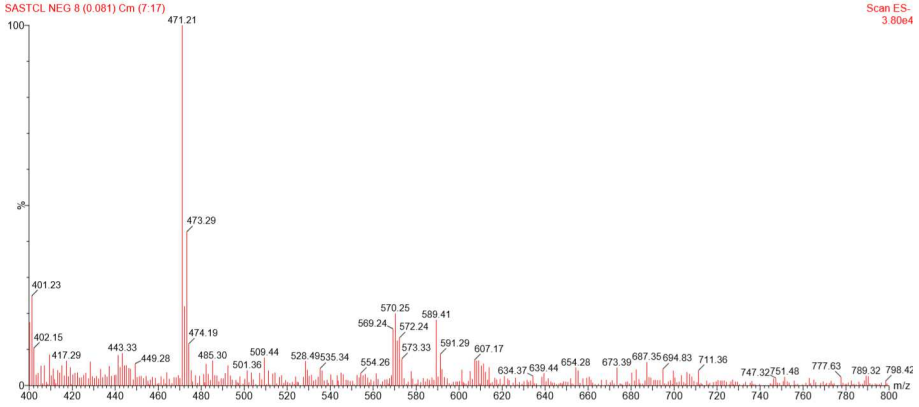
Mass spectrum (ESI–) of AST_2



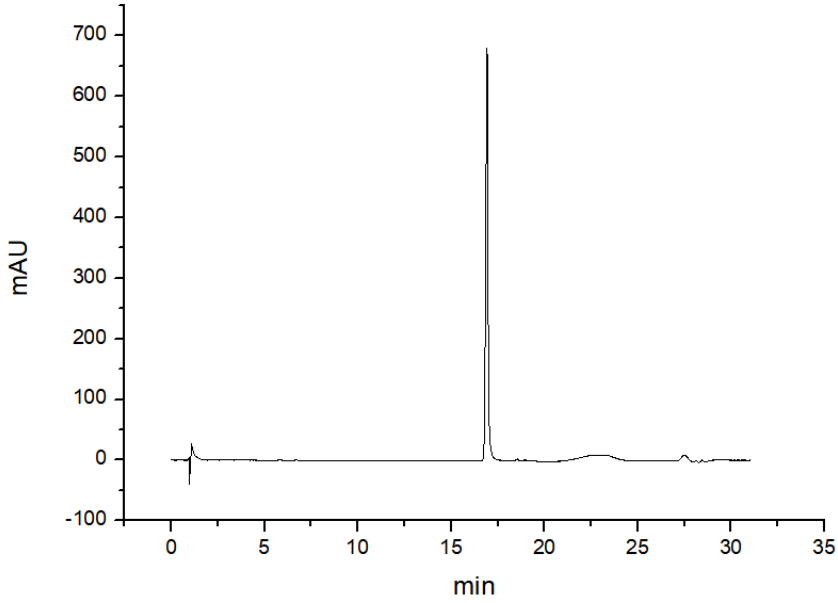
Mass spectrum (ESI+) of AST_3



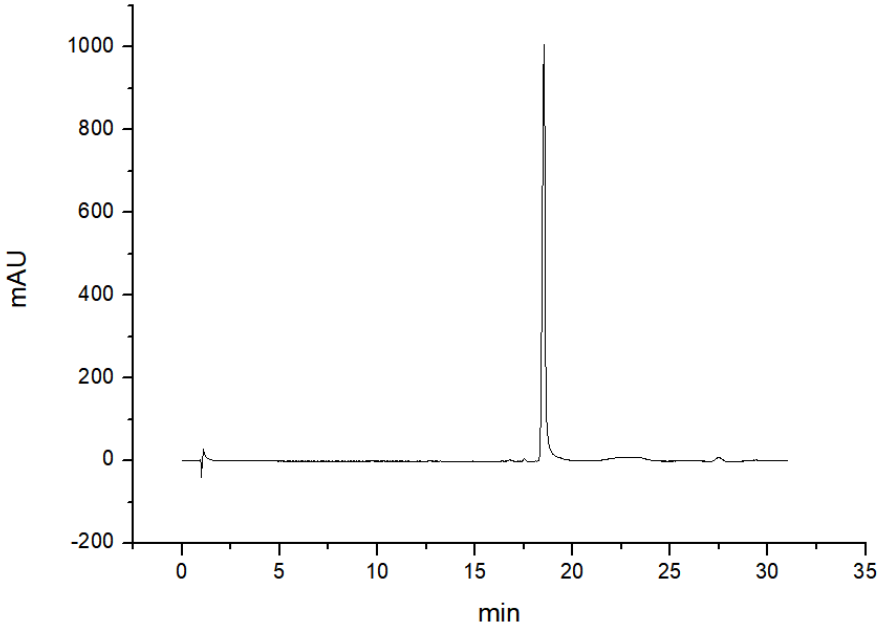
Mass spectrum (ESI–) of AST_3



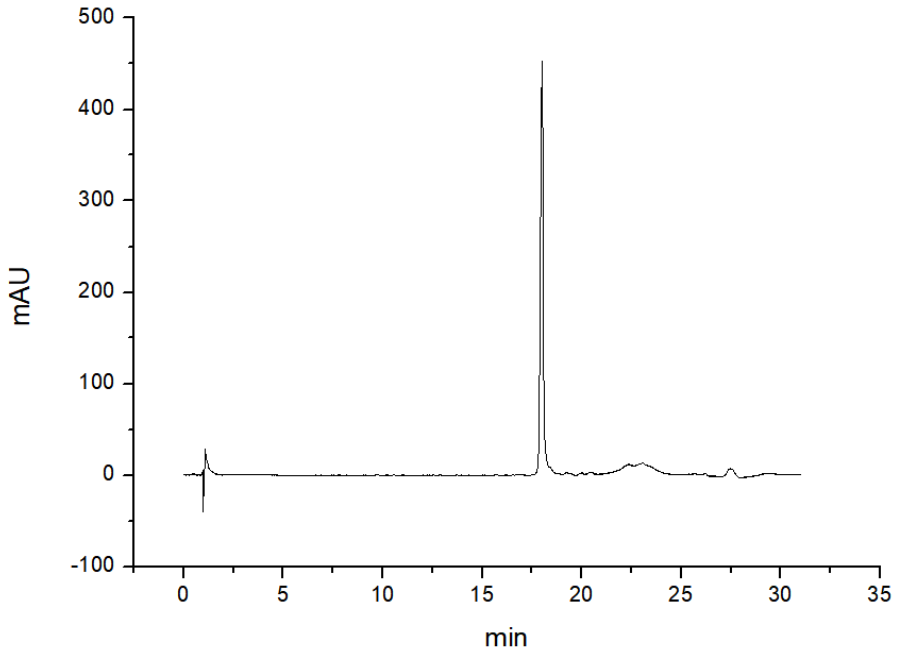
HPLC chromatogram of AST_1



HPLC chromatogram of AST_2



HPLC chromatogram of AST_3



Synthesis of THP (19)

To a cooled (0 °C) solution of THPd (**18**) (1 mmol), synthesized as previously described,[189] in DMF (1 mL) was added DIPEA (6 mmol) and glutaric anhydride (1.2 mmol). The resulting mixture was stirred at RT overnight. The resulting solution was evaporated under reduced pressure, further dried under high vacuum for 3 h, dissolved in a minimal amount of methanol and the product precipitated in ice-cold diethyl ether. Following centrifugation, the pellet was washed several times with diethyl ether. The freeze-dried compound **19** was used without further purification. Yield: 98 %.

Synthesis of c-Met-peptide-THP

The linear sequence of c-met-peptide-THP was synthesized using standard solid-phase peptide synthesis:
Fmoc-Ala-Gly-Ser(tBu)-Cys(Trt)-Tyr(tBu)-Cys(StBu)-Ser(tBu)-Gly-Pro-Pro-Arg(Pbf)-Phe-Glu(OtBu)-Cys(Mmt)-Trp(Boc)-Cys(Trt)-Tyr(tBu)-Glu(OtBu)-Thr(tBu)-Glu(OtBu)-Gly-Thr(tBu)-Gly-Gly-Gly-Lys(Dde)-

Resin.

The NovaPEG Rink Amide resin (Merck Millipore, loading 1.4 mmol/g, 0.3 g) was used as solid support. For loading of the resin 3 equiv. of Fmoc-protected amino acid Lys(Dde), 3 equiv. of oxyme and 3 equiv. of DIC were dissolved in DMF (2 mL) and the reaction mixture was added to the resin. The reaction mixture was shaken overnight at room temperature. The resin was washed three times with DCM (20 mL) and DMF (20 mL). For Fmoc-deprotection the resin was treated two times for 15 min. with 20% piperidine/DMF (10 mL). A standard protocol was used for solid phase peptide synthesis: 4 equiv. Fmoc-protected amino acid, 4 equiv. oxime, 4 equiv. DIC was dissolved in DMF (2 mL). The reaction mixture was shaken for 3 h at room temperature. At the end of the sequence, the last amino acid (Ala) was deprotected and acetylated. For the acetylation 10 equiv. of acetic anhydride and 20 equiv. of DIPEA were added to a solution of DMF (3 mL) and then added to the resin and shaken for 30 min.

Formation of the disulfide bonds of c-Met-peptide-THP and THP coupling

S-*t*-Bu was removed by treating the resin with 20% mercaptoethanol in DMF (10 mL) and shaken for 3 h. Finished the reaction the resin was treated with 10 equiv. of DTNP in DCM (10 mL) for 1 h. Then the resin was treated with 1% TFA in DCM (10 mL) in the presence of TIPS as the scavenger. The reaction was repeated 3 times for 5 minutes, then the resin was washed three times with DCM (20 mL) and DMF (20 mL). After the formation of the C13–C21 disulfide bond, the lysine in 1 position was selectively deprotected (Dde removal) in presence of a solution of hydrazine 2% in DMF (10 mL). The reaction was repeated 2 times for 30 minutes, then the resin was washed three times with DCM (20 mL) and DMF (20 mL). After the removal of the Dde the THP (**19**) was coupled using HATU as coupling agent. 5 equiv. of HATU and 5 equiv. of DIPEA were added to a solution of 5 equiv. of THP (**19**) in DMF (3 mL), then the solution was added to the resin and shaken for 24 h. After the coupling of the hexadentate

tris(3,4-hydroxypyridinone), the peptide was removed from the resin, Trt groups were removed and the second disulfide bond was made in air-oxidation condition. For the cleavage of the peptide from the solid-phase, the resin was treated two times for 15 min. with a solution of TFA/TIPS/thioanisole/Water (97/1/1/1, 10 mL). The solution was then concentrated and the crude product isolated by precipitation into cold diethyl ether. The precipitate was collected by centrifugation and dried in vacuum. The crude products were analyzed by RP-HPLC and mass spectrometry analysis. Afterward, the obtained linear peptide was dissolved in a solution of NH_4HCO_3 (0.1 M; pH 7–8, 3 mL). The reaction mixtures were stirred at room temperature for 24 h and monitored by LC-MS analysis. The final compound was chromatographed via preparative HPLC/UV using H_2O/0.1% TFA as eluent A and CH_3CN/0.1% TFA as eluent B. The elution program used a linear gradient of 0%–40% of eluent B in 60 min. The detection wavelength was 281 nm and the flow rate was 15 mL/min. The isolated fractions were further analyzed by analytical HPLC-DAD, using the same eluent A and B of the preparative purification. The elution program used a linear gradient of 0%–95% of eluent B in 20 min using a C18 column at a 0.3 mL/min flow rate.

Synthesis of GLP-1-peptide-THP

The linear sequence of GLP-1-peptide-THP was synthesized using standard solid-phase peptide synthesis:
Ac-Ser-Pro-Pro-Pro-Ala-Gly-Ser-Ser-Pro-Gly-Gly-Asn-Lys-Leu-Trp-Glu-Ile-Phe-Leu-Arg-Val-Ala-Glu-Glu-Glu-Met-Gln-Lys-Ser-Leu-Asp-Ser-Thr-Phe-Thr-Gly-Glu-Gly-His-Lys(Dde)–resin.

The 2-Chlorotrityl chloride resin (Merck Millipore, loading 1.0–1.8 mmol/g, 0.4 g) was used as solid support. For loading of the resin 3 equiv. of Fmoc-protected amino acid Lys(Dde), 3 equiv. of oxime and 3 equiv. of DIC were dissolved in DMF (2 mL) and the reaction mixture was added to the resin. The reaction mixture was shaken overnight at room temperature. The resin was washed three times with DCM (20 mL) and DMF (20 mL). For Fmoc-deprotection the resin was treated

two times for 15 min. with 20% piperidine/DMF (10 mL). A standard protocol was used for solid phase peptide synthesis: 4 equiv. Fmoc-protected amino acid, 4 equiv. oxime, 4 equiv. DIC was dissolved in DMF (2 mL). The reaction mixture was shaken for 3 h at room temperature. At the end of the sequence, the last amino acid (Ser) was deprotected and acetylated. For the acetylation 10 equiv. of acetic anhydride and 20 equiv. of DIPEA were added to a solution of DMF (3 mL) and then added to the resin and shaken for 30 min. After the removal of the Dde in presence of a solution of hydrazine 2% in DMF (10 mL), the THP (**19**) was coupled using HATU as coupling agent. 5 equiv. of HATU and 5 equiv. of DIPEA were added to a solution of 5 equiv. of THP (**19**) in DMF (3 mL), then the solution was added to the resin and shaken for 24 h. After the coupling of the hexadentate tris(3,4-hydroxypyridinone), the peptide was removed from the resin. For the cleavage of the peptide from the solid-phase, the resin was treated two times for 15 min. with a solution of TFA/TIPS/thioanisole/Water (97/1/1/1, 10 mL). The solution was then concentrated and the crude product isolated by precipitation into cold diethyl ether. The precipitate was collected by centrifugation and dried in vacuum. The crude products were analyzed by RP-HPLC and mass spectrometry analysis. The final compound was chromatographed via preparative HPLC/UV using H2O/0.1% TFA as eluent A and CH3CN/0.1% TFA as eluent B. The elution program used a linear gradient of 0%–40% of eluent B in 60 min. The detection wavelength was 281 nm and the flow rate was 15 mL/min. The isolated fractions were further analyzed by analytical HPLC-DAD, using the same eluent A and B of the preparative purification. The elution program used a linear gradient of 0%–95% of eluent B in 20 min using a C18 column at a 0.3 mL/min flow rate.

Materials for Gallium radiolabeling

Gallium-68 was eluted from an Eckert & Ziegler (E&Z Radiopharma GmbH) $^{68}Ge/^{68}Ga$ generator producing 200-400 MBq of ^{68}Ga, using 5 mL of 0.1 M hydrochloric acid and collected in five 1 mL fractions.

Radiation was counted using a Capintec CRC-25R. High-performance liquid chromatography (HPLC) was performed on an Agilent Technologies 1260 Infinity with inbuilt degasser and quaternary pump, ultraviolet (UV) detection at 220 nm and radioactive detection was done by a Bioscan Inc. B-FC-3200 photomultiplier tube (PMT) detector. HPLC data were collected and analyzed on Laura software. The reverse-phase method used is shown in Method 1.
The size exclusion method used is 100% PBS, flow rate 0.8 mL/min, no gradient, 20-minute run time.

Time / min	Solvent %	
	A	B
0	95	5
5	95	5
20	5	95
25	5	95
25.1	95	5
30	95	5

Method 1: A = water + 0.05% trifluoroacetic acid. B = acetonitrile + 0.05% trifluoroacetic acid. Flow rate: 1 mL/min.

Analytical reverse-phase HPLC was performed on an Agilent Eclipse XDB-C18 column with 5 µm particle size and column dimensions of 4.6 × 150 mm. Size exclusion HPLC was performed on a Phenomenex BioSep 5µm SEC-s2000 145 Å column with dimensions of 200 × 7.8 mm. Radio instant thin layer chromatography (ITLC) was developed on Agilent Technologies glass microfibre chromatography paper impregnated with silica gel and analyzed using a Lablogic Flow-count TLC scanner and a BioScan B-FC-3200 PMT detector using Laura software. ITLC methods were the acetate method (1 M ammonium acetate in water/methanol (1:1)) and the citrate method (0.175 M citric acid, 0.325 M trisodium citrate in water).

Sample preparation for Gallium radiolabeling

Stock solutions of four c-Met-peptide-THP in water were prepared

at a concentration of 1 mg mL-1 (244 µM) using 0.60, 0.27, 0.52 and 0.14 mg of sample respectively. Stock solutions of two GLP-1-peptide-THP samples in water were prepared at a concentration of 5 mg mL 1 (836 µM) using 0.8 and 0.45 mg of sample respectively. Stock solutions were stored at 20 °C when not in use.

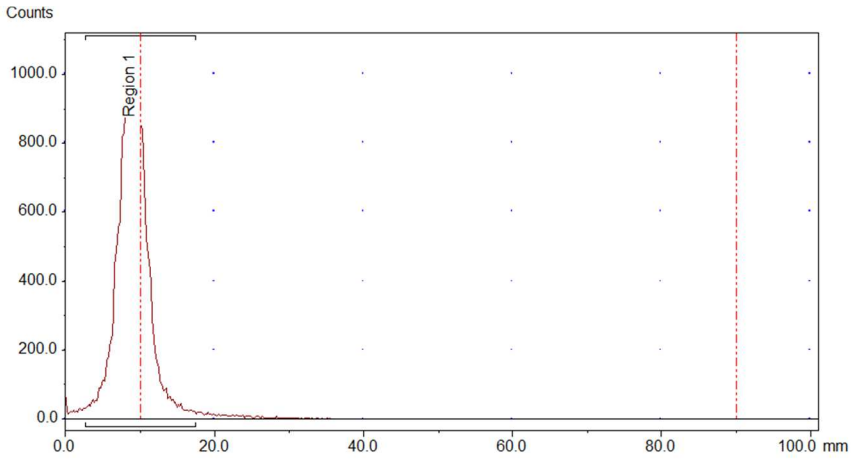
c-Met-peptide-THP radiolabeling

Stock solution of c-Met-peptide-THP (17.5 µL, 4.26 nmol) was added to $^{68}GaCl_3$ (250 µL, 5-90 MBq). Sodium bicarbonate solution in water (26 µL, 1 M) was added immediately. The mixture was agitated and the pH was checked to ensure it was in the range 6.5–7.5. Radiochemical yield and purity was evaluated after 5 min using ITLC by both the acetate method (^{68}Ga R_f = 0, [^{68}Ga]GaTHP R_f = 0.8-1, [^{68}Ga]GaTHP cMet R_f = 0) and citrate method (^{68}Ga R_f = 0.8-1, [68Ga]GaTHP Rf = 0, [68Ga]GaTHP-cMet R_f = 0) and after 10 min by HPLC (Reverse phase: unbound ^{68}Ga Rt = 1.9 min, [^{68}Ga]GaTHP-cMET R_t = 13–14 min. Size Exclusion: unbound ^{68}Ga Rt = 8-13 min, [^{68}Ga]GaTHP-cMet = 14–15 min).

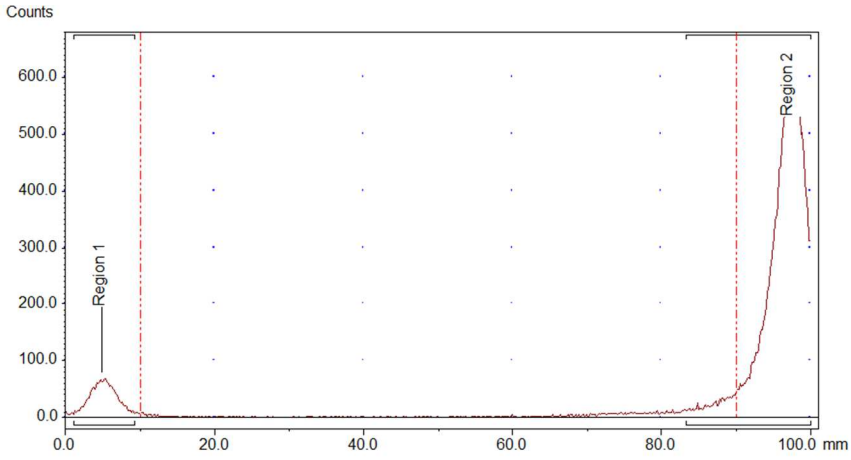
GLP-1-peptide-THP radiolabeling

Stock solution of GLP-1-peptide-THP (5 µL, 4.26 nmol) was added to 68GaCl3 (250 µL, 5-90 MBq). Sodium bicarbonate solution in water (26 µL, 1 M) was added immediately. The mixture was agitated and the pH was checked to ensure it was in the range 6.5–7.5. Radiochemical yield and purity was evaluated after 5 min using ITLC by both the acetate method (^{68}Ga R_f = 0, [^{68}Ga]GaTHP R_f = 0.8-1, [^{68}Ga]GaTHP GLP-1 R_f = 0) and citrate method (^{68}Ga R_f = 0.8-1, [^{68}Ga]GaTHP R_f = 0, [^{68}Ga]GaTHP-GLP-1 R_f = 0) and after 10 min by HPLC (Reverse phase: unbound ^{68}Ga Rt = 1.9 min, [^{68}Ga]GaTHP-GLP-1 R_t = 13.7-13.9 min. Size Exclusion: unbound ^{68}Ga R_t = 8-13 min, [^{68}Ga]GaTHP-GLP-1 = 12-15 min).

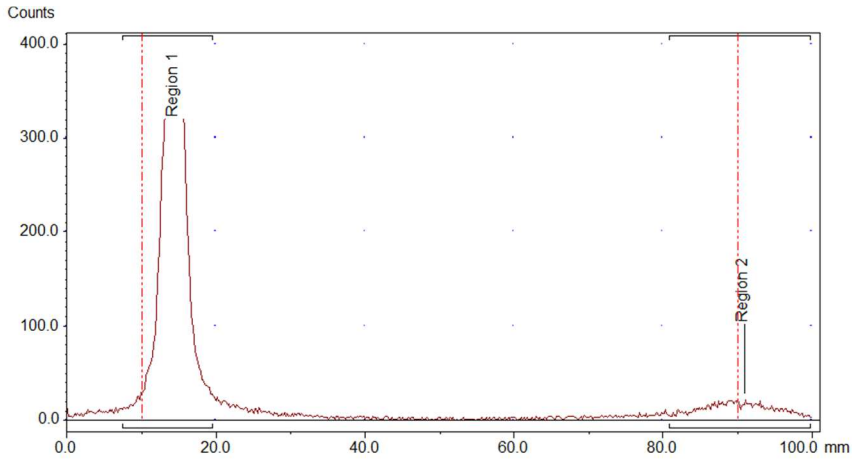
ITLC acetate unbound ^{68}Ga



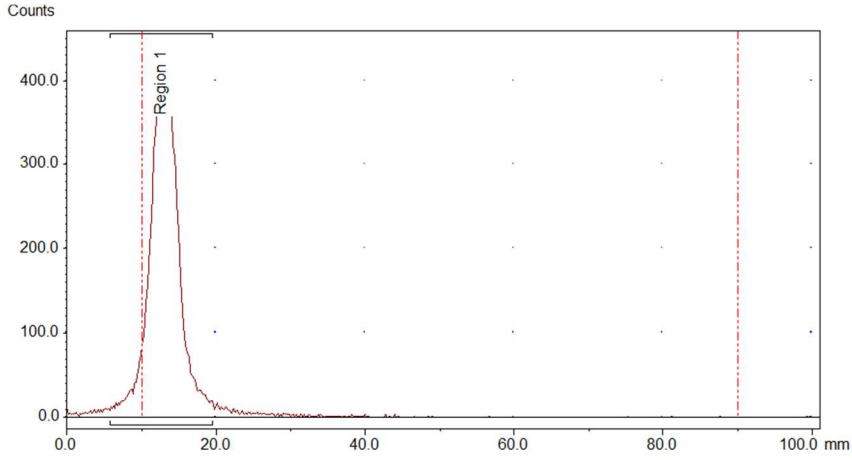
ITLC citrate unbound ^{68}Ga



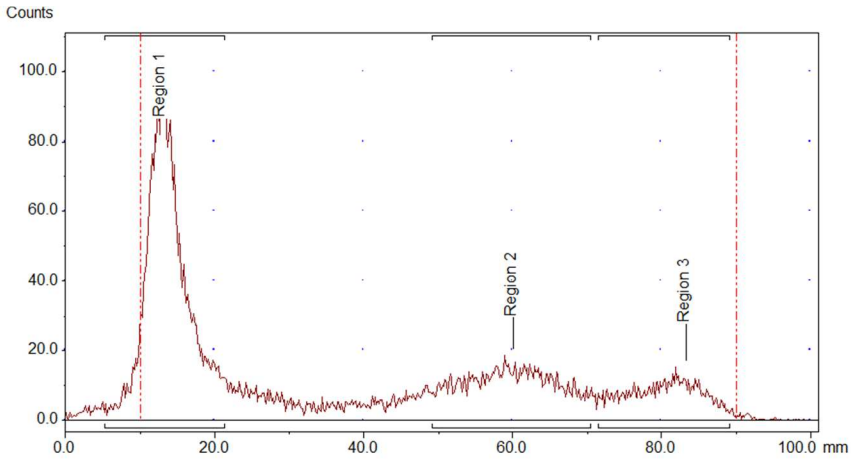
ITLC acetate c-Met-peptide-THP



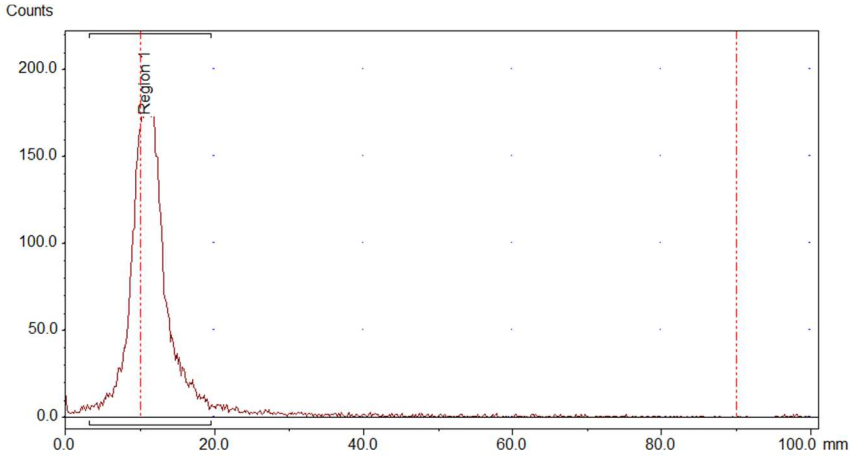
ITLC citrate c-Met-peptide-THP



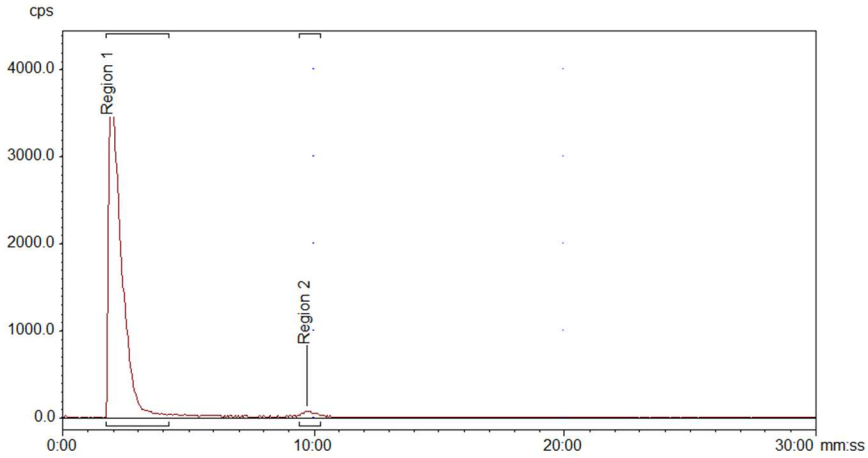
ITLC acetate GLP-1-peptide-THP



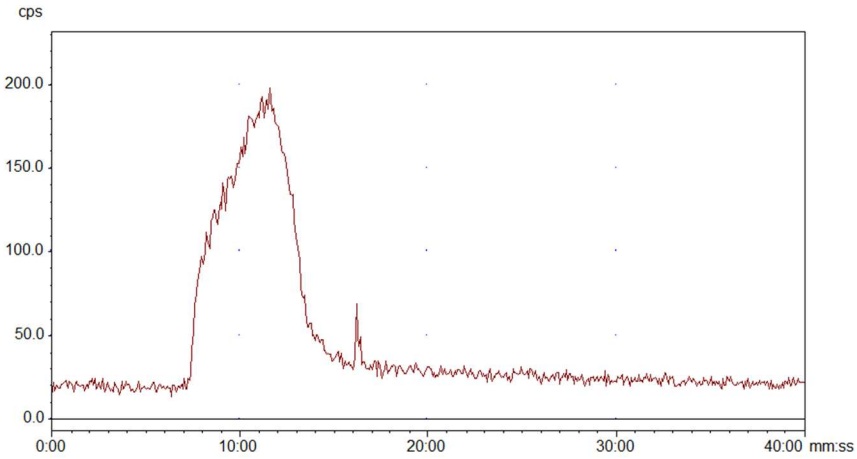
ITLC citrate GLP-1-peptide-THP



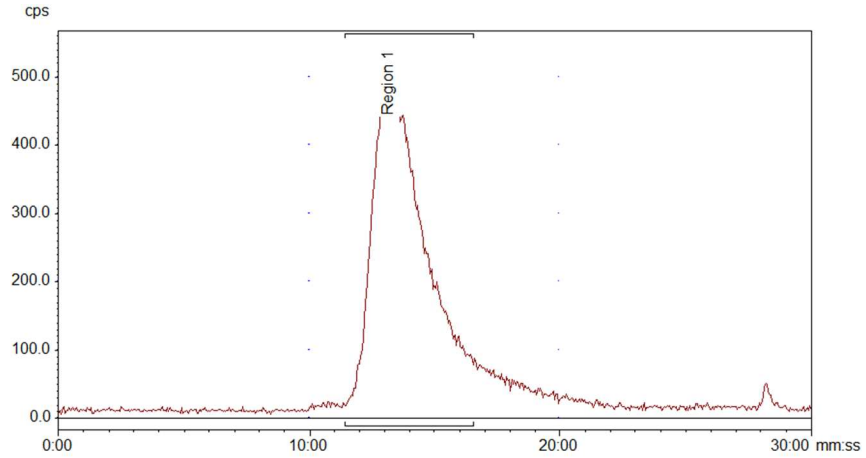
Reverse phase HPLC unbound gallium



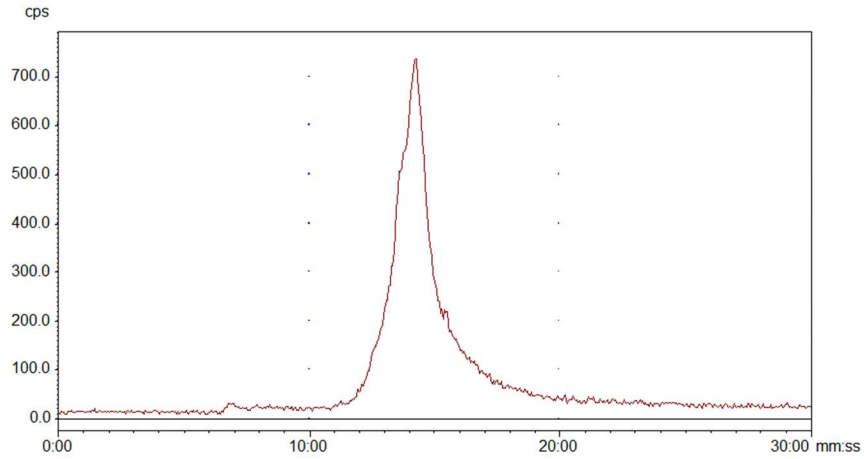
Size exclusion HPLC unbound gallium



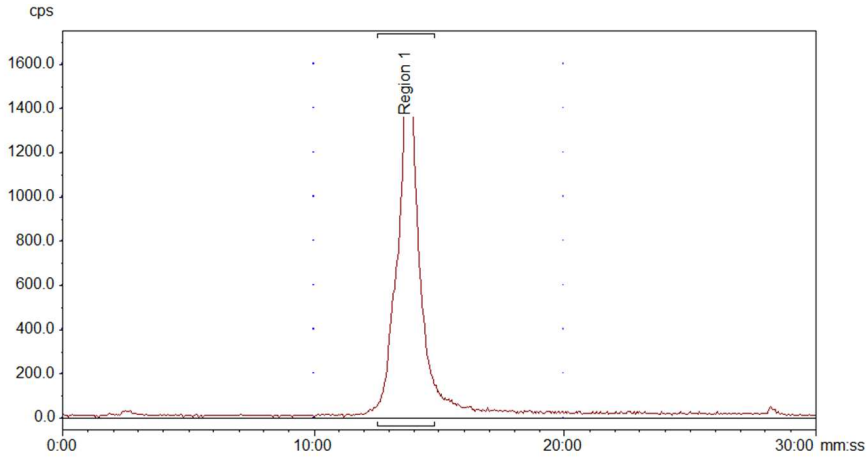
Reverse phase c-Met-peptide-THP



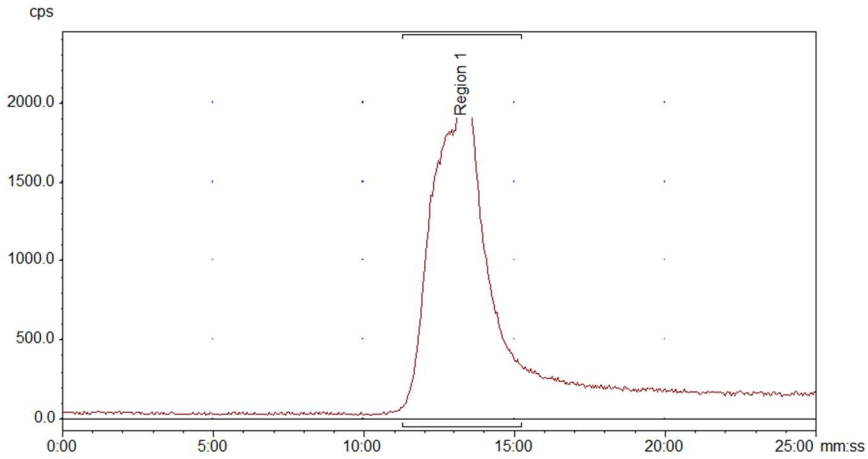
Size exclusion c-Met-peptide-THP



Reverse phase HPLC GLP-1-peptide-THP



Size exclusion HPLC GLP-1-peptide-THP



References

[1] U. N. Das, *Curr Pharm Biotechnol* **2006**, *7*, 467-482.

[2] G. Boden, *Front. Biosci.* **1998**, *3*, d169-175.

[3] R. A. DeFronzo, *Int. J. Clin. Pract. Suppl.* **2004**, 9-21.

[4] S. G. Sheth, F. D. Gordon, S. Chopra, *Ann. Intern. Med.* **1997**, *126*, 137-145.

[5] G. Boden, *Endocrinol Metab Clin North Am* **2008**, *37*, 635-646, viii-ix.

[6] J. Storch, A. E. Thumser, *Biochim. Biophys. Acta* **2000**, *1486*, 28-44.

[7] F. Schroeder, C. A. Jolly, T. H. Cho, A. Frolov, *Chem Phys Lipids* **1998**, *92*, 1-25.

[8] M. I. Queipo-Ortuno, X. Escote, V. Ceperuelo-Mallafre, L. Garrido-Sanchez, M. Miranda, M. Clemente-Postigo, R. Perez-Perez, B. Peral, F. Cardona, J. M. Fernandez-Real, F. J. Tinahones, J. Vendrell, *PLoS One* **2012**, *7*, e48605.

[9] M. R. Syamsunarno, T. Iso, H. Hanaoka, A. Yamaguchi, M. Obokata, N. Koitabashi, K. Goto, T. Hishiki, Y. Nagahata, H. Matsui, M. Sano, M. Kobayashi, O. Kikuchi, T. Sasaki, K. Maeda, M. Murakami, T. Kitamura, M. Suematsu, Y. Tsushima, K. Endo, G. S. Hotamisligil, M. Kurabayashi, *PLoS One* **2013**, *8*, e79386.

[10] B. R. Thompson, A. M. Mazurkiewicz-Munoz, J. Suttles, C. Carter-Su, D. A. Bernlohr, *Journal of Biological Chemistry* **2009**, *284*, 13473-13480.

[11] A. Adida, F. Spener, *Bba-Mol Cell Biol L* **2006**, *1761*, 172-181.

[12] Y. Fu, N. Luo, M. F. Lopes-Virella, W. T. Garvey, *Atherosclerosis* **2002**, *165*, 259-269.

[13] Y. Fu, L. Luo, N. Luo, W. T. Garvey, *Atherosclerosis* **2006**, *188*, 102-111.

[14] K. M. Nieman, H. A. Kenny, C. V. Penicka, A. Ladanyi, R. Buell-Gutbrod, M. R. Zillhardt, I. L. Romero, M. S. Carey, G. B. Mills, G. S. Hotamisligil, S. D. Yamada, M. E. Peter, K. Gwin, E. Lengyel, *Nat Med* **2011**, *17*, 1498-1503.

[15] A. Tolle, S. Suhail, M. Jung, K. Jung, C. Stephan, *BMC Cancer* **2011**, *11*, 302.

[16] H. Uehara, T. Takahashi, M. Oha, H. Ogawa, K. Izumi, *Int J Cancer* **2014**, *135*, 2558-2568.

[17] A. Yang, H. Zhang, Y. Sun, Y. Wang, X. Yang, X. Yang, H. Zhang, W.

Guo, G. Zhu, J. Tian, Y. Jia, Y. Jiang, *Placenta* **2016**, *46*, 49-62.

[18] D. Lee, K. Wada, Y. Taniguchi, H. Al-Shareef, T. Masuda, Y. Usami, T. Aikawa, M. Okura, Y. Kamisaki, M. Kogo, *Oncol. Rep.* **2014**, *31*, 1116-1120.

[19] D. C. Clarke, D. Miskovic, X. X. Han, J. Calles-Escandon, J. F. Glatz, J. J. Luiken, J. J. Heikkila, A. Bonen, *Physiol Genomics* **2004**, *17*, 31-37.

[20] H. M. Goodman, *Physiologist* **1970**, *13*, 75-88.

[21] K. Asanuma, N. Tsuji, T. Endoh, A. Yagihashi, N. Watanabe, *J Immunol* **2004**, *172*, 3922-3929.

[22] B. Schmid, J. F. Rippmann, M. Tadayyon, B. S. Hamilton, *Biochem Biophys Res Commun* **2005**, *328*, 1073-1082.

[23] J. C. Mackall, M. D. Lane, *Biochem Biophys Res Commun* **1977**, *79*, 720-725.

[24] C. Bocos, M. Gottlicher, K. Gearing, C. Banner, E. Enmark, M. Teboul, A. Crickmore, J. A. Gustafsson, *J Steroid Biochem Mol Biol* **1995**, *53*, 467-473.

[25] M. J. Barbera, A. Schluter, N. Pedraza, R. Iglesias, F. Villarroya, M. Giralt, *J Biol Chem* **2001**, *276*, 1486-1493.

[26] Y. Zhang, Y. Sun, E. Rao, F. Yan, Q. Li, Y. Zhang, K. A. Silverstein, S. Liu, E. Sauter, M. P. Cleary, B. Li, *Cancer Res* **2014**, *74*, 2986-2998.

[27] A. Chmurzynska, *J. Appl. Genet.* **2006**, *47*, 39-48.

[28] D. A. Bernlohr, M. A. Simpson, A. V. Hertzel, L. J. Banaszak, *Annual Review of Nutrition* **1997**, *17*, 277-303.

[29] J. H. Veerkamp, *P Nutr Soc* **1995**, *54*, 23-37.

[30] A. Reese-Wagoner, J. Thompson, L. Banaszak, *Biochim Biophys Acta* **1999**, *1441*, 106-116.

[31] Z. H. Xu, D. A. Bernlohr, L. J. Banaszak, *Biochemistry* **1992**, *31*, 3484-3492.

[32] R. S. Sha, C. D. Kane, Z. Xu, L. J. Banaszak, D. A. Bernlohr, *J Biol Chem* **1993**, *268*, 7885-7892.

[33] G. Floresta, V. Pistara, E. Amata, M. Dichiara, A. Marrazzo, O. Prezzavento, A. Rescifina, *Eur J Med Chem* **2017**, *138*, 854-873.

[34] T. Okada, M. Hiromura, M. Otsuka, S. Enomoto, H. Miyachi, *Chem Pharm Bull (Tokyo)* **2012**, *60*, 164-168.

[35] A. V. Hertzel, K. Hellberg, J. M. Reynolds, A. C. Kruse, B. E. Juhlmann, A. J. Smith, M. A. Sanders, D. H. Ohlendorf, J. Suttles, D. A. Bernlohr, *J. Med. Chem.* **2009**, *52*, 6024-6031.

[36] T. Barf, F. Lehmann, K. Hammer, S. Haile, E. Axen, C. Medina, J. Uppenberg, S. Svensson, L. Rondahl, T. Lundback, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 1745-1748.

[37] R. Ringom, E. Axen, J. Uppenberg, T. Lundback, L. Rondahl, T. Barf, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4449-4452.

[38] X. Liu, X. Huang, W. Lin, D. Wang, Y. Diao, H. Li, X. Hui, Y. Wang, A. Xu, D. Wu, D. Ke, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2949-2952.

[39] H. Cai, Q. Liu, D. Gao, T. Wang, T. Chen, G. Yan, K. Chen, Y. Xu, H. Wang, Y. Li, W. Zhu, *Eur. J. Med. Chem.* **2015**, *90*, 241-250.

[40] F. Lehmann, S. Haile, E. Axen, C. Medina, J. Uppenberg, S. Svensson, T. Lundback, L. Rondahl, T. Barf, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4445-4448.

[41] Q. Xu, L. Huang, J. Liu, L. Ma, T. Chen, J. Chen, F. Peng, D. Cao, Z. Yang, N. Qiu, J. Qiu, G. Wang, X. Liang, A. Peng, M. Xiang, Y. Wei, L. Chen, *Eur. J. Med. Chem.* **2012**, *52*, 70-81.

[42] Y. Wang, W. K. Law, J. S. Hu, H. Q. Lin, T. M. Ip, D. C. Wan, *J. Chem. Inf. Model.* **2014**, *54*, 3046-3050.

[43] Y. Zhou, T. Nie, Y. Zhang, M. Song, K. Li, M. Ding, K. Ding, D. Wu, Y. Xu, *Bioorg Med Chem* **2016**.

[44] A. Tropsha, J. Bajorath, *J Med Chem* **2016**, *59*, 1.

[45] T. J. Martinez, *Acc Chem Res* **2017**, *50*, 652-656.

[46] S. C. Basak, G. Restrepo, *Curr Comput Aided Drug Des* **2013**, *9*, 447-448.

[47] S. C. Basak, G. Restrepo, *Curr Comput Aided Drug Des* **2012**, *8*, 83-84.

[48] R. D. Clark, M. Waldman, *J Comput Aided Mol Des* **2012**, *26*, 29-34.

[49] J. Verma, V. M. Khedkar, E. C. Coutinho, *Curr Top Med Chem* **2010**, *10*, 95-115.

[50] R. Sulsky, D. R. Magnin, Y. Huang, L. Simpkins, P. Taunk, M. Patel, Y. Zhu, T. R. Stouch, D. Bassolino-Klimas, R. Parker, T. Harrity, R. Stoffel, D. S. Taylor, T. B. Lavoie, K. Kish, B. L. Jacobson, S. Sheriff, L. P. Adam, W. R. Ewing, J. A. Robl, *Bioorg Med Chem Lett* **2007**, *17*, 3511-3515.

[51] E. I. Howard, B. Guillot, M. P. Blakeley, M. Haertlein, M. Moulin, A. Mitschler, A. Cousido-Siah, F. Fadel, W. M. Valsecchi, T. Tomizaki, T. Petrova, J. Claudot, A. Podjarny, *IUCrJ* **2016**, *3*, 115-126.

[52] T. Cheeseright, M. Mackey, S. Rose, A. Vinter, *J Chem Inf Model* **2006**, *46*, 665-676.

[53] Y. T. Wang, C. H. Liu, H. L. Zhu, *Expert Opin Ther Pat* **2016**, *26*, 767-

776.

[54] Y. Beniyama, K. Matsuno, H. Miyachi, *Bioorg Med Chem Lett* **2013**, *23*, 1662-1666.

[55] A. Cheng, S. A. Best, K. M. Merz, Jr., C. H. Reynolds, *J Mol Graph Model* **2000**, *18*, 273-282.

[56] J. J. Stewart, *J Mol Model* **2004**, *10*, 155-164.

[57] J. J. Stewart, *J Comput Aided Mol Des* **1990**, *4*, 1-105.

[58] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, G. R. Hutchison, *J Cheminform* **2011**, *3*, 33.

[59] O. Trott, A. J. Olson, *J Comput Chem* **2010**, *31*, 455-461.

[60] T. J. Cheeseright, M. D. Mackey, R. A. Scoffin, *Curr Comput Aided Drug Des* **2011**, *7*, 190-205.

[61] T. J. Cheeseright, M. Holm, F. Lehmann, S. Luik, M. Gottert, J. L. Melville, S. Laufer, *J Med Chem* **2009**, *52*, 4200-4209.

[62] T. J. Cheeseright, M. D. Mackey, J. L. Melville, J. G. Vinter, *J Chem Inf Model* **2008**, *48*, 2108-2117.

[63] T. Cheeseright, M. Mackey Phd, S. Rose Phd, A. Vinter Phd, *Expert Opin Drug Discov* **2007**, *2*, 131-144.

[64] J. M. Luco, F. H. Ferretti, *J Chem Inf Comput Sci* **1997**, *37*, 392-401.

[65] A. Surribas, J. M. Amigo, J. Coello, J. L. Montesinos, F. Valero, S. Maspoch, *Anal Bioanal Chem* **2006**, *385*, 1281-1288.

[66] K. Tuppurainen, S. P. Korhonen, J. Ruuskanen, *SAR QSAR Environ Res* **2006**, *17*, 549-561.

[67] G. Palermo, P. Piraino, H. D. Zucht, *Adv Appl Bioinform Chem* **2009**, *2*, 57-70.

[68] T. Asadollahi, S. Dadfarnia, A. M. Shabani, J. B. Ghasemi, M. Sarkhosh, *Molecules* **2011**, *16*, 1928-1955.

[69] A. Golbraikh, A. Tropsha, *J Mol Graph Model* **2002**, *20*, 269-276.

[70] P. H. Olesen, *Current opinion in drug discovery & development* **2001**, *4*, 471-478.

[71] A. Burger, *Prog Drug Res* **1991**, *37*, 287-371.

[72] G. A. Patani, E. J. LaVoie, *Chem Rev* **1996**, *96*, 3147-3176.

[73] J. J. Irwin, B. K. Shoichet, *J Chem Inf Model* **2005**, *45*, 177-182.

[74] A. P. Bento, A. Gaulton, A. Hersey, L. J. Bellis, J. Chambers, M. Davies, F. A. Kruger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos, J. P. Overington, *Nucleic Acids Res* **2014**, *42*, D1083-1090.

[75] Y. Wang, H. Q. Lin, W. K. Law, W. C. Liang, J. F. Zhang, J. S. Hu, T. M. Ip, M. M. Waye, D. C. Wan, *ACS Chem Neurosci* **2015**, *6*, 211-218.

[76] R. Abali, I. Temel Yuksel, M. A. Yuksel, B. Bulut, M. Imamoglu, V. Emirdar, F. Unal, S. Guzel, C. Celik, *J Obstet Gynaecol* **2016**, 1-5.

[77] S. A. Abdelwahab, Y. Owada, N. Kitanaka, A. Adida, H. Sakagami, M. Ono, M. Watanabe, F. Spener, H. Kondo, *Mol Cell Biochem* **2007**, *299*, 99-107.

[78] S. Bag, S. Ramaiah, A. Anbarasu, *J Theor Biol* **2015**, *364*, 344-354.

[79] A. E. Thumser, J. B. Moore, N. J. Plant, *Curr Opin Clin Nutr Metab Care* **2014**, *17*, 124-129.

[80] S. Guaita-Esteruelas, J. Guma, L. Masana, J. Borras, *Mol Cell Endocrinol* **2017**.

[81] K. Kawaguchi, A. Kinameri, S. Suzuki, S. Senga, Y. Ke, H. Fujii, *Biochem J* **2016**, *473*, 449-461.

[82] S. Guaita-Esteruelas, A. Bosquet, P. Saavedra, J. Guma, J. Girona, E. W. Lam, K. Amillano, J. Borras, L. Masana, *Mol Carcinog* **2017**, *56*, 208-217.

[83] M. Furuhashi, S. Saitoh, K. Shimamoto, T. Miura, *Clin Med Insights Cardiol* **2014**, *8*, 23-33.

[84] K. Kawaguchi, S. Senga, C. Kubota, Y. Kawamura, Y. Ke, H. Fujii, *FEBS Open Bio* **2016**, *6*, 190-199.

[85] T. Yamamoto, M. Furuhashi, T. Sugaya, T. Oikawa, M. Matsumoto, Y. Funahashi, Y. Matsukawa, M. Gotoh, T. Miura, *PLoS One* **2016**, *11*, e0167825.

[86] J. Hao, F. Yan, Y. Zhang, A. Triplett, Y. Zhang, D. A. Schultz, Y. Sun, J. Zeng, K. A. T. Silverstein, Q. Zheng, D. A. Bernlohr, M. P. Cleary, N. K. Egilmez, E. Sauter, S. Liu, J. Suttles, B. Li, *Cancer Res* **2018**, *78*, 2343-2355.

[87] X. Peng, K. Studholme, M. P. Kanjiya, J. Luk, D. Bogdan, M. W. Elmes, G. Carbonetti, S. Tong, Y. H. Gary Teng, R. C. Rizzo, H. Li, D. G. Deutsch, I. Ojima, M. J. Rebecchi, M. Puopolo, M. Kaczocha, *Mol Pain* **2017**, *13*, 1744806917697007.

[88] G. G. Martin, D. Landrock, S. Chung, L. J. Dangott, D. R. Seeger, E. J. Murphy, M. Y. Golovko, A. B. Kier, F. Schroeder, *J Neurochem* **2017**, *140*, 294-306.

[89] H. Huang, A. L. McIntosh, G. G. Martin, D. Landrock, S. Chung, K. K. Landrock, L. J. Dangott, S. Li, A. B. Kier, F. Schroeder, *Biochemistry*

2016, *55*, 5243-5255.

[90] M. W. Elmes, M. Kaczocha, W. T. Berger, K. Leung, B. P. Ralph, L. Wang, J. M. Sweeney, J. T. Miyauchi, S. E. Tsirka, I. Ojima, D. G. Deutsch, *J. Biol. Chem.* **2015**, *290*, 8711-8721.

[91] R. N. Dansereau, B. R. Line, *J Nucl Med* **1996**, *37*, 631-631.

[92] M. Chomet, C. Provost, V. Vega, A. Prignon, J. Talbot, V. Nataf, *Eur J Nucl Med Mol I* **2016**, *43*, S471-S472.

[93] S. R. Banerjee, M. G. Pomper, *Appl Radiat Isotopes* **2013**, *76*, 2-13.

[94] E. Deutsch, *J Nucl Med* **1993**, *34*, 1132-1133.

[95] H. N. Wagner, *J Nucl Med* **1991**, *32*, 561-564.

[96] T. Ebenhan, M. Vorster, B. Marjanovic-Painter, J. Wagener, J. Suthiram, M. Modiselle, B. Mokaleng, J. R. Zeevaart, M. Sathekge, *Molecules* **2015**, *20*, 14860-14878.

[97] M. Benesova, M. Schafer, U. Bauder-Wust, A. Afshar-Oromieh, C. Kratochwil, W. Mier, U. Haberkorn, K. Kopka, M. Eder, *J Nucl Med* **2015**, *56*, 914-920.

[98] J. Notni, K. Pohle, H. J. Wester, *Ejnmmi Research* **2012**, *2*.

[99] E. Farkas, J. Nagel, B. P. Waldron, D. Parker, I. Toth, E. Brucher, F. Rosch, Z. Baranyai, *Chem-Eur J* **2017**, *23*, 10358-10371.

[100] C. F. Ramogida, D. Schindler, C. Schneider, Y. L. K. Tan, S. Huh, C. L. Ferreira, M. J. Adam, C. Orvig, *Rsc Adv* **2016**, *6*, 103763-103773.

[101] C. F. Ramogida, J. H. Pan, C. L. Ferreira, B. O. Patrick, K. Rebullar, D. T. T. Yapp, K. S. Lin, M. J. Adam, C. Orvig, *Inorganic Chemistry* **2015**, *54*, 4953-4965.

[102] J. Seemann, B. P. Waldron, F. Roesch, D. Parker, *Chemmedchem* **2015**, *10*, 1019-1026.

[103] D. J. Berry, Y. Ma, J. R. Ballinger, R. Tavare, A. Koers, K. Sunassee, T. Zhou, S. Nawaz, G. E. Mullen, R. C. Hider, P. J. Blower, *Chem Commun (Camb)* **2011**, *47*, 7068-7070.

[104] M. T. Ma, C. Cullinane, C. Imberti, J. Baguna Torres, S. Y. Terry, P. Roselt, R. J. Hicks, P. J. Blower, *Bioconjug Chem* **2016**, *27*, 309-318.

[105] J. D. Young, V. Abbate, C. Imberti, L. K. Meszaros, M. T. Ma, S. Y. A. Terry, R. C. Hider, G. E. Mullen, P. J. Blower, *J Nucl Med* **2017**, *58*, 1270-1277.

[106] E. Blom, B. Langstrom, I. Velikyan, *Bioconjug Chem* **2009**, *20*, 1146-1151.

[107] R. K. Selvaraju, I. Velikyan, L. Johansson, Z. Wu, I. Todorov, J. Shively,

F. Kandeel, O. Korsgren, O. Eriksson, *J Nucl Med* **2013**, *54*, 1458-1463.
[108] I. Velikyan, *Molecules* **2015**, *20*, 12913-12943.
[109] Y. Kai-Larsen, G. H. Gudmundsson, B. Agerberth, *Acta Paediatr* **2014**, *103*, 1000-1008.
[110] R. N. McLay, W. Pan, A. J. Kastin, *Peptides* **2001**, *22*, 2181-2255.
[111] A. Ationu, S. Boateng, *Int J Mol Med* **1998**, *2*, 235-239.
[112] T. N. Fay, J. G. Grudzinskas, *Hum Reprod* **1991**, *6*, 1311-1326.
[113] J. E. Morley, *Psychoneuroendocrinology* **1983**, *8*, 361-379.
[114] V. Le Joncour, P. Laakkonen, *Bioorg Med Chem* **2018**, *26*, 2797-2806.
[115] C. D. Mathers, D. Loncar, *PLoS Med* **2006**, *3*, e442.
[116] T. A. Baudino, *Curr Drug Discov Technol* **2015**, *12*, 3-20.
[117] M. Hojjat-Farsangi, *Int J Mol Sci* **2014**, *15*, 13768-13801.
[118] B. Zhao, W. Zhang, D. Yu, J. Xu, Y. Wei, *Lung Cancer* **2018**, *122*, 10-21.
[119] J. Ansari, E. Batubara, M. Ali, A. Farrag, F. Bashir, H. R. Farghaly, A. M. Ali, A. Shaukat, *Mol Clin Oncol* **2018**, *9*, 92-95.
[120] R. Ronca, M. Benkheil, S. Mitola, S. Struyf, S. Liekens, *Med Res Rev* **2017**, *37*, 1231-1274.
[121] Y. Gilad, M. Firer, G. Gellerman, *Biomedicines* **2016**, *4*.
[122] X. Sun, Y. Li, T. Liu, Z. Li, X. Zhang, X. Chen, *Adv Drug Deliv Rev* **2017**, *110-111*, 38-51.
[123] L. J. Mu, A. Hohne, R. A. Schubiger, S. M. Ametamey, K. Graham, J. E. Cyr, L. Dinkelborg, T. Stellfeld, A. Srinivasan, U. Voigtmann, U. Klar, *Angew Chem Int Edit* **2008**, *47*, 4922-4925.
[124] S. M. Okarvi, *J Nucl Med* **2002**, *43*, 369p-369p.
[125] R. Mansi, D. Tesauro, C. Pedone, E. Benedetti, G. Morelli, *J Pept Sci* **2006**, *12*, 229-229.
[126] S. M. Okarvi, *Medicinal Research Reviews* **2004**, *24*, 685-686.
[127] S. M. Okarvi, *Medicinal Research Reviews* **2004**, *24*, 357-397.
[128] K. Juhl, A. Christensen, M. Persson, M. Ploug, A. Kjaer, *Plos One* **2016**, *11*.
[129] S. L. Organ, M. S. Tsao, *Ther Adv Med Oncol* **2011**, *3*, S7-S19.
[130] A. J. Latimer, J. R. Jessen, *Dev Dyn* **2008**, *237*, 3904-3915.
[131] E. Andermarcher, M. A. Surani, E. Gherardi, *Dev Genet* **1996**, *18*, 254-266.
[132] H. J. Kim, M. J. Baek, D. H. Kang, S. C. Lee, S. B. Bae, K. T. Lee, N. Lee, H. Kim, D. Jeong, T. S. Ahn, M. S. Lee, D. S. Hong, J. H. Won, *Ann Coloproctol* **2018**, *34*, 88-93.

[133] N. Li, Z. Dou, J. Liu, B. Chai, Y. Li, X. An, P. Chu, X. Zhang, *Ann Hepatol* **2018**, *17*, 501-510.
[134] A. Anestis, I. Zoi, M. V. Karamouzis, *Ann Transl Med* **2018**, *6*, 247.
[135] C. Dai, Y. Xie, X. Zhuang, Z. Yuan, *Biomed Pharmacother* **2018**, *104*, 763-770.
[136] A. Konstorum, J. S. Lowengrub, *J Theor Biol* **2018**, *439*, 86-99.
[137] L. Demkova, L. Kucerova, *Mol Cancer* **2018**, *17*, 26.
[138] C. A. Bradley, M. Salto-Tellez, P. Laurent-Puig, A. Bardelli, C. Rolfo, J. Tabernero, H. A. Khawaja, M. Lawler, P. G. Johnston, S. Van Schaeybroeck, M. E. consortium, *Nat Rev Clin Oncol* **2018**, *15*, 150.
[139] K. Noguchi, M. Konno, H. Eguchi, K. Kawamoto, R. Mukai, N. Nishida, J. Koseki, H. Wada, H. Akita, T. Satoh, S. Marubashi, H. Nagano, Y. Doki, M. Mori, H. Ishii, *Oncol Lett* **2018**, *16*, 1892-1898.
[140] M. Trovato, A. Campenni, S. Giovinazzo, M. Siracusa, R. M. Ruggeri, *Biomark Insights* **2017**, *12*, 1177271917701126.
[141] J. J. Yang, J. H. Yang, J. Kim, S. H. Ma, L. Y. Cho, K. P. Ko, A. Shin, B. Y. Choi, H. J. Kim, D. S. Han, C. S. Eun, K. S. Song, Y. S. Kim, S. H. Chang, H. R. Shin, D. Kang, K. Y. Yoo, S. K. Park, *Int J Cancer* **2013**, *132*, 2148-2156.
[142] J. R. Sierra, M. S. Tsao, *Ther Adv Med Oncol* **2011**, *3*, S21-35.
[143] K. Zhu, X. Kong, D. Zhao, Z. Liang, C. Luo, *Expert Opin Ther Pat* **2014**, *24*, 217-230.
[144] J. Porter, *Expert Opin Ther Pat* **2010**, *20*, 159-177.
[145] V. S. Hughes, D. W. Siemann, *Trends Cancer* **2018**, *4*, 94-97.
[146] N. Sharma, A. A. Adjei, *Ther Adv Med Oncol* **2011**, *3*, S37-50.
[147] A. Varkaris, P. G. Corn, S. Gaur, F. Dayyani, C. J. Logothetis, G. E. Gallick, *Expert Opin Investig Drugs* **2011**, *20*, 1677-1684.
[148] A. G. Terwisscha van Scheltinga, M. N. Lub-de Hooge, M. J. Hinner, R. B. Verheijen, A. Allersdorfer, M. Hulsmeyer, W. B. Nagengast, C. P. Schroder, J. G. Kosterink, E. G. de Vries, L. Audoly, S. A. Olwill, *J Nucl Med* **2014**, *55*, 665-671.
[149] H. Luo, H. Hong, M. R. Slater, S. A. Graves, S. Shi, Y. Yang, R. J. Nickles, F. Fan, W. Cai, *J Nucl Med* **2015**, *56*, 758-763.
[150] V. R. Aroda, *Diabetes Obes Metab* **2018**, *20 Suppl 1*, 22-33.
[151] B. Gallwitz, *Handb Exp Pharmacol* **2011**, 53-74.
[152] R. Lugari, A. Dei Cas, D. Ugolotti, A. L. Barilli, C. Camellini, G. C. Ganzerla, A. Luciani, B. Salerni, F. Mittenperger, S. Nodari, A. Gnudi,

R. Zandomeneghi, *Horm Metab Res* **2004**, *36*, 111-115.
[153] B. Gallwitz, T. Ropeter, C. Morys-Wortmann, R. Mentlein, E. G. Siegel, W. E. Schmidt, *Regul Pept* **2000**, *86*, 103-111.
[154] T. Nomiyama, T. Kawanami, S. Irie, Y. Hamaguchi, Y. Terawaki, K. Murase, Y. Tsutsumi, R. Nagaishi, M. Tanabe, H. Morinaga, T. Tanaka, M. Mizoguchi, K. Nabeshima, M. Tanaka, T. Yanase, *Diabetes* **2014**, *63*, 3891-3905.
[155] R. E. Ryder, *Diabet Med* **2013**, *30*, 1148-1155.
[156] A. Mehrabi, L. Fischer, M. Hafezi, A. Dirlewanger, L. Grenacher, M. K. Diener, H. Fonouni, M. Golriz, C. Garoussi, N. Fard, N. N. Rahbari, J. Werner, M. W. Buchler, *Pancreas* **2014**, *43*, 675-686.
[157] J. M. Trujillo, W. Nuffer, S. L. Ellis, *Ther Adv Endocrinol Metab* **2015**, *6*, 19-28.
[158] M. C. Wong, H. H. Wang, M. W. Kwan, D. D. Zhang, K. Q. Liu, S. W. Chan, C. K. Fan, B. C. Fong, S. T. Li, S. M. Griffiths, *PLoS One* **2014**, *9*, e90963.
[159] G. Derosa, P. Maffioli, *Curr Clin Pharmacol* **2012**, *7*, 214-228.
[160] D. Wild, A. Wicki, R. Mansi, M. Behe, B. Keil, P. Bernhardt, G. Christofori, P. J. Ell, H. R. Macke, *J Nucl Med* **2010**, *51*, 1059-1067.
[161] M. Brom, L. Joosten, W. J. Oyen, M. Gotthardt, O. C. Boerman, *Contrast Media Mol Imaging* **2012**, *7*, 160-166.
[162] D. Pach, A. Sowa-Staszczak, A. Jabrocka-Hybel, A. Stefanska, M. Tomaszuk, R. Mikolajczak, B. Janota, M. Trofimiuk-Muldner, E. Przybylik-Mazurek, A. Hubalewska-Dydejczyk, *Int J Endocrinol* **2013**, *2013*, 384508.
[163] A. Sowa-Staszczak, D. Pach, R. Mikolajczak, H. Macke, A. Jabrocka-Hybel, A. Stefanska, M. Tomaszuk, B. Janota, A. Gilis-Januszewska, M. Malecki, G. Kaminski, A. Kowalska, J. Kulig, A. Matyja, C. Osuch, A. Hubalewska-Dydejczyk, *Eur J Nucl Med Mol Imaging* **2013**, *40*, 524-531.
[164] D. O. Kiesewetter, H. Gao, Y. Ma, G. Niu, Q. Quan, N. Guo, X. Chen, *Eur J Nucl Med Mol Imaging* **2012**, *39*, 463-473.
[165] A. Hubalewska-Dydejczyk, A. Sowa-Staszczak, M. Tomaszuk, A. Stefanska, *Q J Nucl Med Mol Imaging* **2015**, *59*, 152-160.
[166] C. Birchmeier, W. Birchmeier, E. Gherardi, G. F. Vande Woude, *Nat Rev Mol Cell Bio* **2003**, *4*, 915-925.
[167] S. Bandla, A. Pennathur, J. D. Luketich, D. G. Beer, L. Lin, A. J. Bass, T.

E. Godfrey, V. R. Little, *Ann Thorac Surg* **2012**, *93*, 1101-1106.

[168] Y. Liu, X. F. Yu, J. Zou, Z. H. Luo, *World J Gastroentero* **2015**, *21*, 3706-3710.

[169] E. S. Qamsari, S. S. Ghaderi, B. Zarei, R. Dorostkar, S. Bagheri, F. Jadidi-Niaragh, M. H. Somi, M. Yousefi, *Tumor Biol* **2017**, *39*.

[170] H. J. Kim, T. S. Ahn, S. C. Lee, S. B. Bae, K. T. Lee, M. S. Lee, M. J. Baek, *Eur J Cancer* **2013**, *49*, S159-S159.

[171] Z. S. Zeng, M. R. Weiser, E. Kuntz, C. T. Chen, S. A. Khan, A. Forslund, G. M. Nash, M. Gimbel, Y. Yamaguchi, A. T. Culliford, M. D'Alessio, F. Barany, P. B. Paty, *Cancer Letters* **2008**, *265*, 258-269.

[172] S. Osada, H. Imai, Y. Sasaki, K. Yoshida, *Annals of Oncology* **2011**, *22*, v95-v95.

[173] J. Burggraaf, I. M. C. Kamerling, P. B. Gordon, L. Schrier, M. L. de Kam, A. J. Kales, R. Bendiksen, B. Indrevoll, R. M. Bjerke, S. A. Moestue, S. Yazdanfar, A. M. J. Langers, M. Swaerd-Nordmo, G. Torheim, M. V. Warren, H. Morreau, P. W. Voorneveld, T. Buckle, F. W. B. van Leeuwen, L. I. Odegardstuen, G. T. Dalsgaard, A. Healey, J. C. H. Hardwick, *Nature Medicine* **2015**, *21*, 955-961.

[174] A. K. Galande, R. Weissleder, C. H. Tung, *Journal of Combinatorial Chemistry* **2005**, *7*, 174-177.

[175] I. Annis, L. Chen, G. Barany, *Journal of the American Chemical Society* **1998**, *120*, 7226-7238.

[176] K. Barlos, D. Gatos, O. Hatzi, N. Koch, S. Koutsogianni, *Int J Pept Prot Res* **1996**, *47*, 148-153.

[177] A. Aletras, K. Barlos, D. Gatos, S. Koutsogianni, P. Mamos, *Int J Pept Prot Res* **1995**, *45*, 488-496.

[178] I. Velikyan, A. Sundin, B. Eriksson, H. Lundqvist, J. Sorensen, M. Bergstrom, B. Langstrom, *Nucl Med Biol* **2010**, *37*, 265-275.

[179] I. Velikyan, U. Rosenstrom, S. Estrada, I. Ljungvall, J. Haggstrom, O. Eriksson, G. Antoni, *Nucl Med Biol* **2014**, *41*, 728-736.

[180] R. K. Selvaraju, I. Velikyan, V. Asplund, L. Johansson, Z. Wu, I. Todorov, J. Shively, F. Kandeel, B. Eriksson, O. Korsgren, O. Eriksson, *Nucl Med Biol* **2014**, *41*, 471-476.

[181] M. Y. Lee, H. Li, Y. Xiao, Z. Zhou, A. Xu, P. M. Vanhoutte, *Br J Pharmacol* **2011**, *162*, 1564-1576.

[182] C. Aleman, F. J. Luque, M. Orozco, *Journal of computational chemistry* **1993**, *14*, 799-808.

[183] S. H. Olesen, J. Y. Zhu, M. P. Martin, E. Schonbrunn, *ChemMedChem* **2016**, *11*, 1137-1144.

[184] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J Comput Chem* **2009**, *30*, 2785-2791.

[185] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, *Journal of computational chemistry* **2003**, *24*, 1999-2012.

[186] E. Krieger, G. Vriend, *Bioinformatics* **2014**, *30*, 2981-2982.

[187] E. Krieger, G. Vriend, *J Comput Chem* **2015**, *36*, 996-1007.

[188] L. Castanar, G. D. Poggetto, A. A. Colbourne, G. A. Morris, M. Nilsson, *Magn Reson Chem* **2018**.

[189] T. Zhou, H. Neubert, D. Y. Liu, Z. D. Liu, Y. M. Ma, X. L. Kong, W. Luo, S. Mark, R. C. Hider, *J Med Chem* **2006**, *49*, 4171-4182.