

# Exploring Allosteric Modulation of the Nuclear Receptor ROR $\gamma$ t from a Drug Discovery Perspective

## ***Citation for published version (APA):***

Meijer, F. A. (2021). *Exploring Allosteric Modulation of the Nuclear Receptor ROR $\gamma$ t from a Drug Discovery Perspective*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Biomedical Engineering]. Eindhoven University of Technology.

## ***Document status and date:***

Published: 23/06/2021

## ***Document Version:***

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

## ***Please check the document version of this publication:***

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## ***General rights***

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

## ***Take down policy***

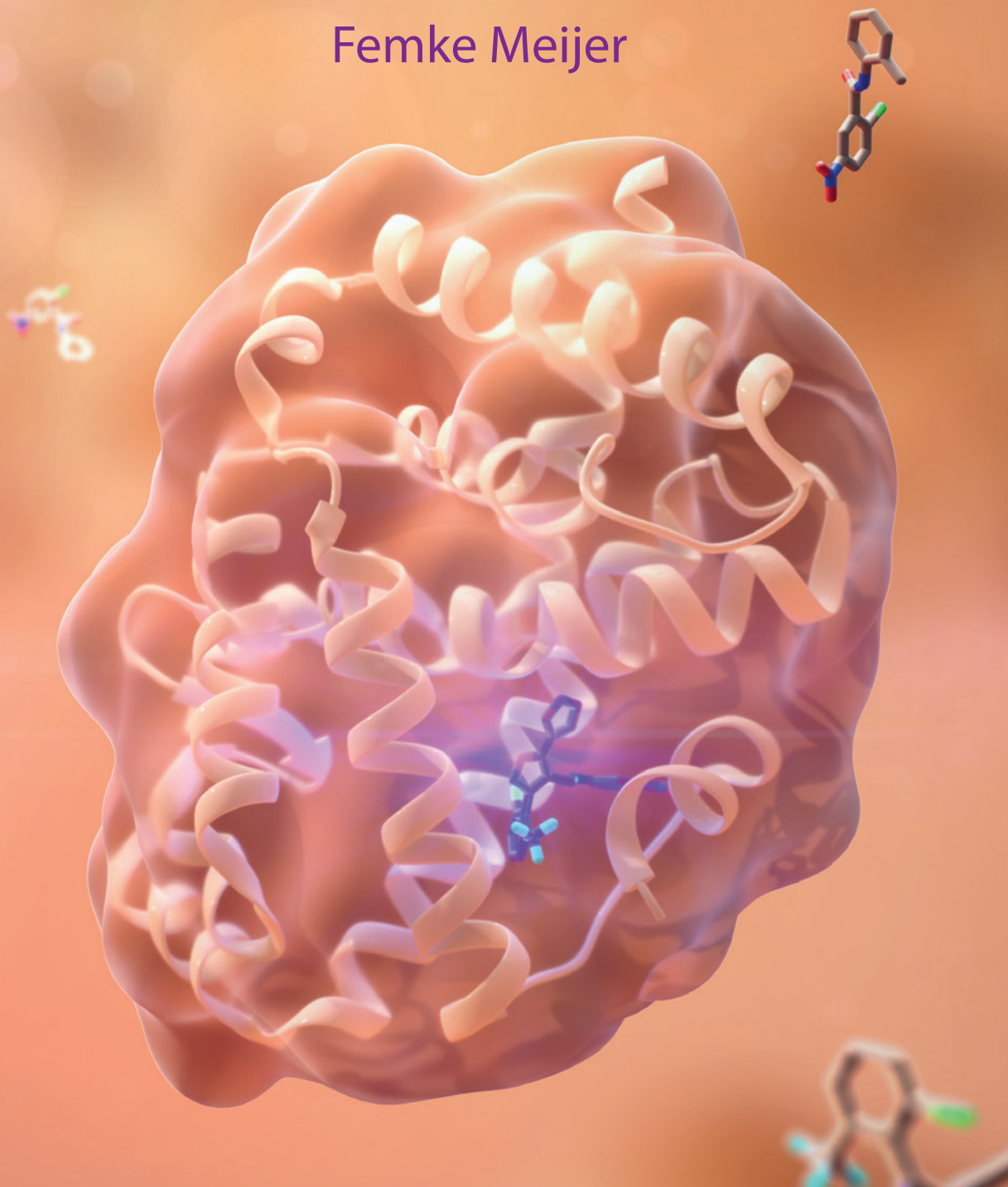
If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# Exploring Allosteric Modulation of the Nuclear Receptor ROR $\gamma$ t from a Drug Discovery Perspective

Femke Meijer





# **Exploring Allosteric Modulation of the Nuclear Receptor ROR $\gamma$ t from a Drug Discovery Perspective**

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit  
Eindhoven, op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens,  
voor een commissie aangewezen door het College voor Promoties, in het  
openbaar te verdedigen op woensdag 23 juni 2021 om 11:00 uur

door

Femke Anouk Meijer

geboren te Breda

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof. dr. M. Merkx
1 <sup>e</sup> promotor:	prof. dr. ir. L. Brunsveld
co-promotor:	dr. R.G. Doveston (University of Leicester)
leden:	prof. dr. E.W. Meijer
	prof. dr. M. van der Stelt (Universiteit Leiden)
	prof. dr. U. Holzgrave (Universität Würzburg)
	dr. C. Ottmann
	dr. A. Barnard (Imperial College London)

*Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.*

*“The important thing is to never stop questioning.  
Curiosity has its own reason for existing.”*

Albert Einstein

F.A. Meijer ©

Cover design: ICMS Animation Studio, TU/e

Printed by: Gildeprint – The Netherlands

A catalogue record is available from the Eindhoven University of Technology Library  
ISBN: 978-90-386-5279-5

This research has been financially supported by the Netherlands Organization for Scientific Research (NWO) through Gravity program 024.001.035 and VICI grant 016.150.366.



# Table of Contents

<b>Chapter 1</b>	<b>I</b>
Allosteric Modulation of Nuclear Receptors	
<b>Chapter 2</b>	<b>29</b>
Ligand-Based Design of Allosteric ROR $\gamma$ t Inverse Agonists	
<b>Chapter 3</b>	<b>73</b>
Structure Activity Relationship Studies of Trisubstituted Isoxazoles as Selective Allosteric Ligands for ROR $\gamma$ t	
<b>Chapter 4</b>	<b>113</b>
Cooperativity between the Orthosteric and Allosteric Ligand Binding Sites of ROR $\gamma$ t	
<b>Chapter 5</b>	<b>141</b>
Covalent Occlusion of the ROR $\gamma$ t Ligand Binding Pocket Allows Unambiguous Targeting of an Allosteric Site	
<b>Chapter 6</b>	<b>169</b>
Dual Targeting of the Orthosteric and Allosteric Binding Sites of ROR $\gamma$ t with a Bitopic Ligand	
<b>Chapter 7</b>	<b>201</b>
Epilogue	
<b>Summary</b>	<b>223</b>
<b>Curriculum Vitae</b>	<b>227</b>
<b>List of Publications</b>	<b>229</b>
<b>Acknowledgements</b>	<b>231</b>



# Chapter 1

## Allosteric Modulation of Nuclear Receptors

### Abstract

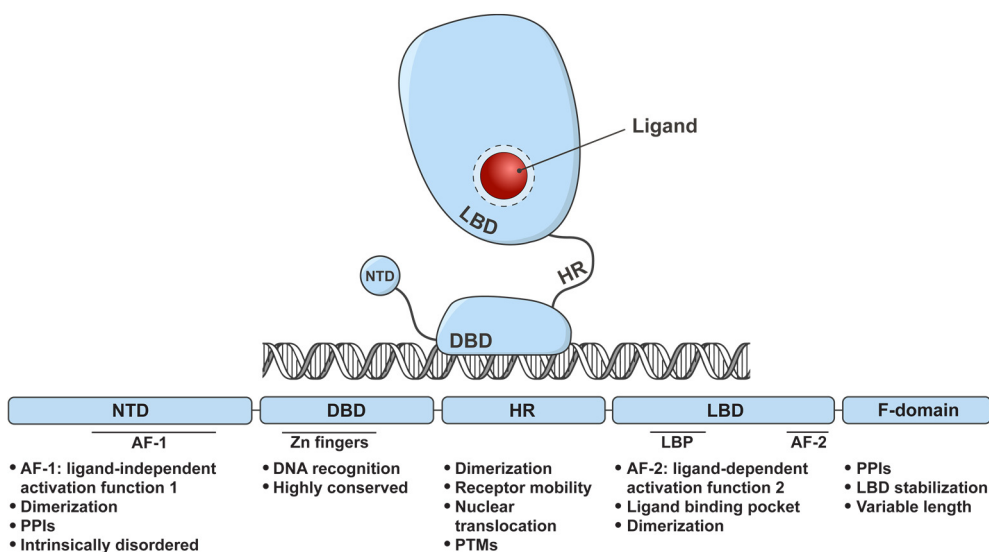
Nuclear receptors (NRs) are multi-domain proteins, whose natural regulation typically occurs via the binding of ligands to a classical, orthosteric binding pocket and via intra- and inter-domain allosteric mechanisms. Allosteric modulation of NRs via synthetic small molecules has recently emerged as an interesting methodology for addressing the need for small molecules targeting NRs in pathology, via novel modes of action and with beneficial profiles. This chapter provides an overview of the possibilities for NR allosteric modulation, via the binding of small molecules to non-canonical, allosteric binding pockets. The main focus will be on the nuclear receptor ROR $\gamma$ t, which is an important target in the treatment of autoimmune diseases. Although the majority of ROR $\gamma$ t modulators target the canonical, orthosteric ligand binding pocket, the discovery of an allosteric and highly unique binding pocket in the ligand binding domain provides an interesting alternative for the modulation of ROR $\gamma$ t.

**Parts of this chapter have been published as:** F.A. Meijer<sup>#</sup>, I.A. Leijten-van de Gevel<sup>#</sup>, R.M.J.M. de Vries<sup>#</sup> & L. Brunsveld. Allosteric Small Molecule Modulators of Nuclear Receptors. *Mol. Cell. Endocrinol.* 485, 20-32 (2019).

<sup>#</sup> These authors contributed equally

## Structural organization of nuclear receptors

Nuclear receptors (NRs) are a family of ligand-regulated transcription factors, consisting of 48 structurally related members.<sup>1,2</sup> NRs play an essential role in various fundamental biological processes, including cell proliferation, reproduction and metabolism.<sup>3</sup> Dysfunction of NRs can therefore result in many pathophysiological processes that propagate diseases such as cancer, diabetes and autoimmune disorders.<sup>4</sup> This has resulted in NRs becoming highly relevant targets for therapeutic intervention, with 16% of all FDA approved drugs targeting this protein class.<sup>5-7</sup>



**Figure 1.1 | Schematic overview of the nuclear receptor domains.** A variable N-terminal domain (NTD), a highly conserved DNA binding domain (DBD), a flexible hinge region (HR), the ligand binding domain (LBD) that typically consists of twelve  $\alpha$ -helices, and the F-domain which is present in only a subset of the nuclear receptors. Abbreviations: PPIs, protein-protein interactions; PTMs, post-translational modifications.

NRs have a conserved domain organization (Figure 1.1), consisting of five different subunits, starting at the N-terminus with the highly variable N-terminal domain (NTD).<sup>1,2,8,9</sup> For most NRs, this domain contains a ligand-independent activation function, termed AF-1.<sup>10</sup> The NTD is intrinsically disordered, but the interaction with binding partners (e.g. cofactor proteins) can induce folding, which enhances transcriptional activity.<sup>10,11</sup> The NTD is followed by the highly conserved DNA binding domain (DBD), containing two zinc fingers that recognize specific hormone response elements (HREs).<sup>12</sup> The affinity for specific HREs is dependent on the NR subtype but also on NR homo- or heterodimerization, or monomeric

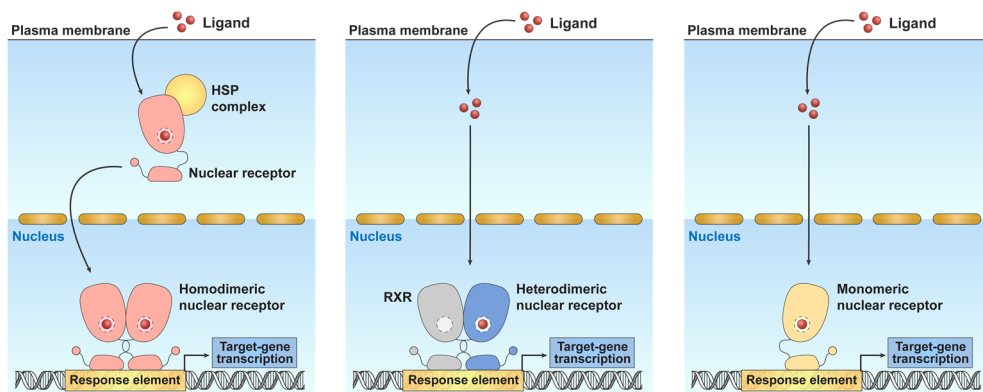


binding to extended HREs.<sup>12</sup> The DBD is connected to the ligand binding domain (LBD) via a hinge region (HR) that plays a key role in dimerization, receptor mobility and nuclear translocation, and is highly decorated with posttranslational modifications (PTMs).<sup>13</sup> The LBD is a highly conserved domain, typically consisting of twelve  $\alpha$ -helices which are organized in an antiparallel, globular arrangement.<sup>14</sup> The domain plays a central role in the function of the receptor, by binding to endogenous ligands or small molecule drugs that regulate the receptor's activity.<sup>8</sup> Furthermore, the LBD contains motifs for NR dimerization and the recruitment of cofactors (AF-2 site).<sup>8</sup> C-terminal to the LBD, some NRs contain an additional F-domain that has a variable length and exerts functions ranging from interacting with other proteins to stabilizing the ligand-bound conformations of the LBD.<sup>15</sup>

### Nuclear receptor mode of action and regulation

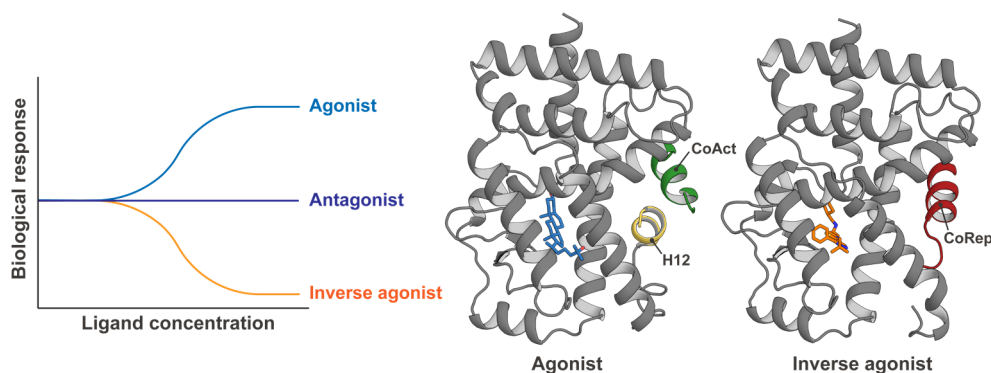
NRs are typically activated by small lipophilic ligands, such as steroid hormones like estradiol and cholesterol, that are able to pass through the cellular membrane into the cell (Figure 1.2).<sup>16</sup> After diffusion through the membrane, the ligand binds to the NR, which can be located either in the cytoplasm or in the nucleus of the cell. Ligand binding triggers the dissociation of inhibiting complexes (i.e. heat shock proteins (HSPs) or corepressors) and the recruitment of coactivator proteins.<sup>17</sup> This allows binding of the NR to a specific response element on the DNA, activating gene transcription, which controls a wide range of physiological processes in the cell.

NRs can be subclassified into four different types, according to their mode of action (Figure 1.2).<sup>17,18</sup> Type I and III NRs (e.g. estrogen receptor  $\alpha$  (ER $\alpha$ ) and androgen receptor (AR)) are in general sequestered in the cytoplasm by HSPs (Figure 1.2, left).<sup>19,20</sup> Ligand binding induces the dissociation of the HSP complex and results in translocation of the NR to the nucleus. These NR subtypes both bind to the DNA as homodimers, where type I NRs recognize inverted repeats and type III direct repeats.<sup>18</sup> Type II NRs (e.g. peroxisome proliferator-activated receptor (PPAR)) (Figure 1.2, middle) are located in the nucleus in complex with corepressor proteins.<sup>18</sup> Upon ligand binding, corepressors dissociate from the NR and coactivator proteins are recruited, resulting in DNA binding (direct repeats).<sup>21</sup> They generally bind to the DNA as heterodimers with the retinoid X receptor (RXR).<sup>18</sup> Finally, type IV NRs (Figure 1.2, right) are orphan receptors, for which the endogenous ligands have not yet been identified or agreed upon.<sup>17</sup> They reside in the nucleus and can bind as monomers to half-site response elements on the DNA.<sup>18</sup> An example of a class IV receptor is the retinoic acid receptor-related orphan receptor (ROR), which has a central role in this thesis.



**Figure 1.2 | Overview of the NR signaling pathway for different NR subtypes. Left)** Type I and III NRs are typically present in the cytoplasm and, upon ligand binding, translocate to the nucleus, where the NR homodimer can bind to the DNA and induce transcription. **Middle)** Type II NRs already reside in the nucleus and bind to the DNA as heterodimer with RXR. Upon ligand binding, they activate gene transcription. **Right)** Orphan NRs (type IV) already reside in the nucleus and bind to the DNA as monomers. Activation of transcription occurs by ligand binding to the NR, PTMs of the NR or changes in NR expression levels. Figure adapted from Glass *et al.*<sup>17</sup>

The regulation of NR activity (i.e. the recruitment of coactivators and corepressors) is controlled by the binding of endogenous or synthetic ligands to the ligand binding pocket of the LBD (Figure 1.3). More specifically, these ligands regulate the NR activity via modulation of the conformation of Helix 12 (H12, also termed AF-2 domain), which is a short  $\alpha$ -helical region on the C-terminus of the LBD.<sup>17</sup> When the LBD is in its apo form, it can adopt a wide range of conformations, as stated by the ‘dynamic stabilization’ model.<sup>8,22</sup> Binding of an agonist to the ligand binding pocket stabilizes H12 in an ‘active’ conformation, creating a hydrophobic cleft to which coactivator proteins can bind via an LXXLL motif (where L is a leucine and X is any amino acid), which results in an overall increase of the transcriptional activity.<sup>8</sup> In contrast, inverse agonists destabilize the ‘active’ fold of H12, disrupting coactivator binding. In some cases (especially in the case of constitutively active NRs), inverse agonists can induce corepressor recruitment (corepressors bind via an LXXXLXXXL motif), repressing the transcriptional activity of the receptor.<sup>8,23</sup> Lastly, antagonists are ligands that occupy the ligand binding pocket and in that way prevent the binding of other ligands (i.e. agonists and inverse agonists), which neutralizes gene transcription.<sup>23</sup> However, many antagonists described for NRs display inverse agonist activity (especially for NRs with a basal transcriptional activity), so it can be hard to define the difference between inverse agonists and antagonists.<sup>23</sup>

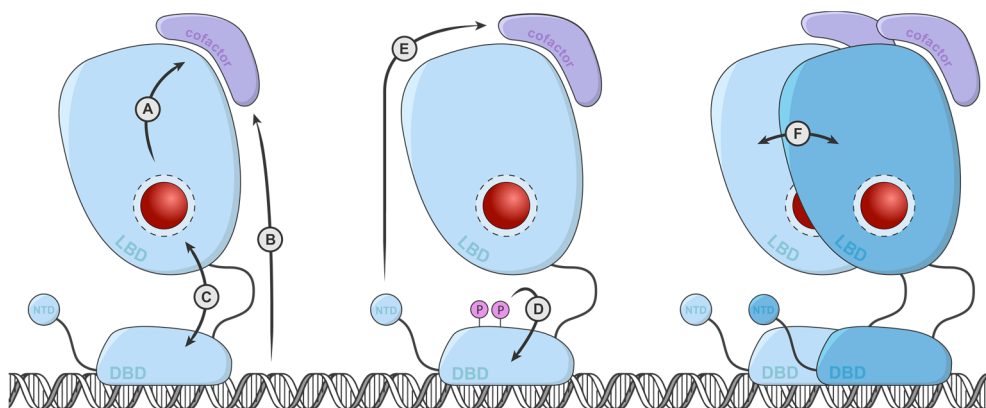


**Figure 1.3 | Schematic representation of the effect of different ligands on the NR biological response and tertiary structures of the NR ligand binding domain (LBD) in complex with an agonist and inverse agonist.** Agonist binding (blue line, blue sticks) to the ligand binding pocket stabilizes helix 12 (shown in yellow) in an ‘active’ conformation, inducing coactivator binding (shown in green) and increasing the biological response (PDB: 3LoL).<sup>24</sup> Inverse agonists (orange line, orange sticks) destabilize the active fold of helix 12 (not visible), disrupting coactivator binding, but inducing corepressor binding (shown in red), decreasing the biological response (PDB: 6A22).<sup>25</sup> Antagonists (purple line) prevent the binding of other ligands, neutralizing gene transcription.

### Inter-domain allosteric regulation of nuclear receptor activity

Up until a few years ago, NR structural information was only available for the separate domains. This provided relevant information on ligand and DNA binding as well as on dimerization of separate LBDs or DBDs, but left a demand for information on the implications of inter-domain communication.<sup>8</sup> Recently, multi-domain structures were elucidated, which provided more insight into the interplay between the different NR domains.<sup>26–29</sup> In particular, it became clear that the different domains within a NR act and communicate via modulatory, allosteric mechanisms.<sup>30–32</sup> Allosterism, in general, is the process where binding of an interaction partner, e.g. ligand or protein, at one site of a protein results in a functional change at another, topographically distinct site.<sup>33,34</sup> This means that there is communication over a distance between the binding site and the site of the biological response, via a conformational change of the protein.<sup>30</sup> Figure 1.4 shows the different conceptual types of allosteric regulatory mechanisms that have been described for NRs.

First, the majority of NR ligands bind to the ligand binding pocket, which causes a conformational change in the LBD, resulting in the regulation of cofactor recruitment at another location on the LBD (Figure 1.4A).<sup>31</sup> Besides this prime example of an intramolecular NR allosteric mechanism, many other allosteric mechanisms can be found in endogenous NR regulation. One of these mechanisms is the allosteric communication between the DBD



**Figure 1.4 | Conceptual modes of endogenous NR allosteric regulation.** A) From the ligand binding pocket to the cofactor binding site. B) From the DNA to the cofactor binding site. C) Between the DNA binding domain (DBD) and ligand binding domain (LBD). D) Via post-translational modifications within and over NR domains. E) From the N-terminal domain (NTD) to the cofactor binding site. F) Between NR dimerization partners.

and the cofactor binding site (AF-2 site) (Figure 1.4B), which has for example been studied in the estrogen receptor (ER).<sup>35</sup> This study showed that the structure of the coactivator pocket is influenced by the type of estrogen response element, pointing to the response element as a regulator of biological activity.<sup>35</sup> Another allosteric NR mechanism is the interplay between the LBD and DBD, which can allosterically influence each other's ligand affinity (Figure 1.4C), typified by an example on the androgen receptor (AR) where mutations in the LBD led to reduced DNA binding without affecting ligand binding, and also the other way around.<sup>36</sup> Post-translational modifications (PTMs), including phosphorylation, acetylation and ubiquitination, are another example of an allosteric NR regulation mechanism (Figure 1.4D). PTMs can occur at a specific site of the protein and can influence, for example, cofactor interactions, cellular localization and protein stability.<sup>37–40</sup> The NTD also shows some exemplary examples of allosteric communication within the NR (Figure 1.4E), since for GR (as an example), interactions between the NTD and certain binding partners resulted in an enhancement of coactivator recruitment and thus transcriptional activity.<sup>41</sup> Lastly, dimerization with RXR is a common phenomenon for NRs, which could lead to a change in structural plasticity and thereby a change in the behavior of a NR. This, in turn, can allosterically affect the recruitment and binding of ligands (Figure 1.4F).<sup>42</sup>

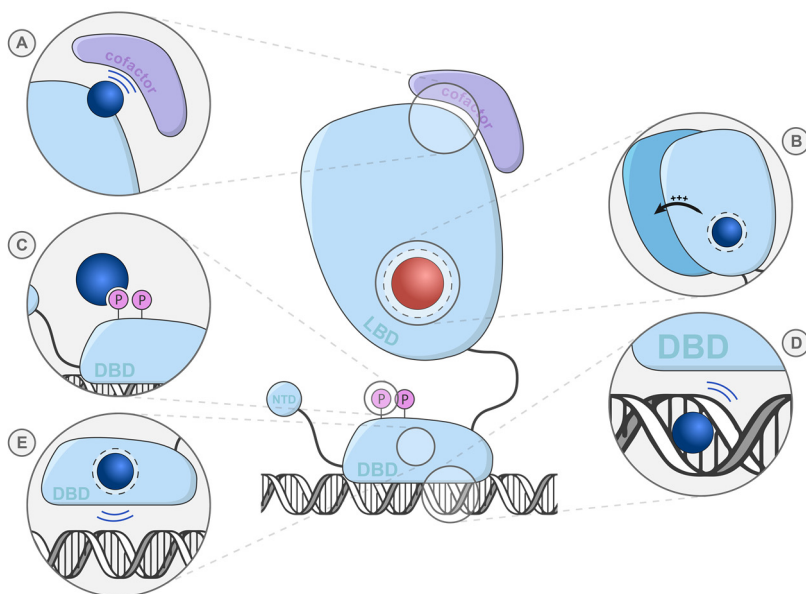


These examples of endogenous NR allosteric regulation illustrate the relevance of these mechanisms on the biological NR function. Allosteric NR modulation therefore poses a promising approach for NR drug discovery.<sup>43</sup>

### **Allosteric modulation of nuclear receptors with small molecules**

The majority of NRs possess endogenous ligands that bind to the canonical ligand binding pocket in the LBD, typically termed the orthosteric binding site. NR modulation via this canonical binding site has proved to be a successful approach, exemplified by the high number of orthosteric NR drugs on the market.<sup>5-7</sup> In contrast, allosteric ligands bind to sites on the protein that do not overlap with the orthosteric pocket.<sup>44</sup> Although allosteric NR modulation with synthetic small molecules is a relatively new and challenging area of research, it could provide a promising alternative to orthosteric targeting, with some potential advantages.<sup>34,45-48</sup> First, the ligand binding pocket shows a high degree of similarity between NRs as noted previously. Thus, whereas orthosteric compounds can be associated with selectivity issues and consequential side effects in some cases, allosteric ligands could result in a higher degree of selectivity due to the greater structural diversity of allosteric sites.<sup>47</sup> Second, allosteric compounds are typically not in competition with the receptor's endogenous ligands.<sup>47</sup> Therefore, allosteric modulators could potentially be used at lower concentrations resulting in reduced toxicity and side effects. Furthermore, in case of drug-resistant mutations in the orthosteric binding site, allosteric ligands can provide a promising alternative entry.<sup>48</sup>

Despite these favorable characteristics, the challenge of discovering allosteric drugs is an important current limitation, since allosteric sites are often not identified or lack specific assay formats for target ligand screening.<sup>34</sup> Although overcoming these challenges could be a difficult task, a significant number of allosteric modulators have already been identified for multiple receptor classes, including GPCRs and ligand-gated ion channels, but also for enzymes.<sup>30,49-53</sup> Some of these allosteric ligands have already been approved by the FDA, such as Sensipar and Gleevec, and multiple others are in clinical trials.<sup>54,55</sup> NRs have similarly seen an increase in attention regarding the identification of allosteric ligands. Although the number of publications is still lagging behind that of other receptor classes, a significant number of examples of allosteric NR modulation are known. These examples involve allosteric sites beyond the LBD (Figure 1.5), as well as sites within the LBD itself (Figure 1.6).



**Figure 1.5 | Different modes of action for small molecule-based allosteric NR modulation outside of the ligand binding domain (LBD).** **A)** Via binding to the cofactor binding site. **B)** Via compounds modulating the NR dimerization. **C)** Via the recognition of post-translational modifications. **D)** Via modulation of the DNA response element. **E)** Via binding to the DNA binding domain.

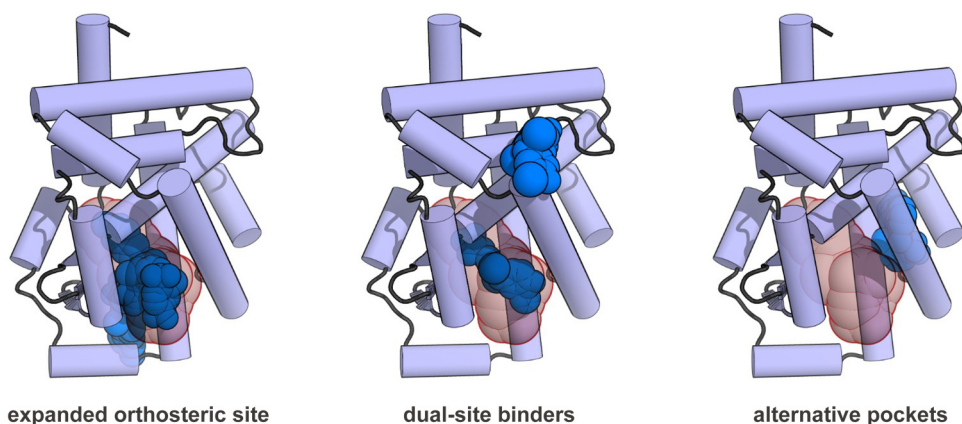
A first type of alternative NR modulation is the direct inhibition of cofactor binding at the AF-2 site (Figure 1.5A).<sup>56</sup> Many types of AF-2 site inhibitors have been described, such as constrained peptides, peptide mimetics and small molecule inhibitors, all typically mimicking the helical LXXLL cofactor binding motif.<sup>45,46,48</sup>

NR homo- and heterodimerization also provide entries for allosteric modulation via small molecules (Figure 1.5B). Especially for the RXR NR, many examples of ligands that affect its dimerization behavior, and thus the transactivation of its dimer partner, have been identified, in some cases even selectively activating RXR heterodimerization with one specific NR over others.<sup>57–59</sup>

Additionally, several ligands are known to allosterically modulate NRs by interacting with specific PTMs (Figure 1.5C). For PPAR $\gamma$ , for example, S273 phosphorylation (which results in the upregulation of certain target genes, decreasing insulin sensitivity) could be inhibited by ligands that bind to the ligand binding pocket and have a minimal agonistic activity.<sup>60,61</sup> A different approach to target a PTM using small molecules to alter NR activity has been shown for ER $\alpha$ , for which a phosphorylation site in the F-domain has been identified which interacts with the hub-protein 14-3-3. This interaction can be stabilized by the small molecule

fusicoccin, which reduces ER $\alpha$  homodimerization, resulting in inhibition of downstream gene regulation, which could provide a promising strategy in the treatment of breast cancer.<sup>62</sup>

Lastly, interactions between the NR DBD and the DNA are crucial for the expression of target genes. Targeting this interaction is therefore another potential strategy for NR allosteric modulation.<sup>45,46</sup> A first strategy is the blocking of specific HREs at the DNA (Figure 1.5D), e.g. via the binding of hairpin polyamides to the minor groove of the DNA, thereby disrupting the interaction with the DBD of for example ER $\alpha$  and AR.<sup>63,64</sup> A second approach is to target the zinc fingers on the DBD of the NR (Figure 1.5E). By using oxidizing agents that cause disulfide bond formation of the cysteine residues, the zinc ions are displaced from the zinc fingers by disrupting the Zn-S interactions, leading to inhibition of DNA binding.<sup>46,65</sup> Because of the high sequence homology of the DBD within the NR class, a challenging issue is the design of selective modulators that target a specific NR DBD.



**Figure 1.6 | Three conceptual modes of allosteric binding sites in the NR ligand binding domain (LBD).** The orthosteric ligand binding pocket is displayed in red and the ligands are shown in blue. The expanded orthosteric site ligands address the classical ligand binding pocket, but also extend into novel pockets, unique for a specific NR. The dual-site binders bind twice to the LBD, typically one time to the ligand binding pocket and a second time to an allosteric site. The alternative pocket modulators bind exclusively to an allosteric pocket.

Apart from these examples on allosterically modulating NRs outside of the LBD, a number of compounds are known that bind to allosteric sites within the LBD, which are categorized into three subgroups, depending on their binding location (Figure 1.6).

Ligands in the first group target expanded orthosteric sites, binding (part of) the orthosteric ligand binding pocket but also extending into neighboring pockets. These ligands are often found in NRs that contain a larger orthosteric pocket, like PPAR and the farnesoid

X receptor (FXR).<sup>66–70</sup> A typical example for FXR is Ivermectin, which displays antidiabetic activities by reducing blood glucose levels and improving insulin sensitivity in an FXR-dependent manner.<sup>70</sup> Interestingly, it shows a unique binding mode in the FXR ligand binding pocket by shifting H2 and H6 outward and thus expanding the pocket, to create extra space to accommodate the larger size of the molecule.<sup>70</sup> By binding to an expanded binding site, the compound allosterically enhances the flexibility of the cofactor binding site resulting in an alteration of the cofactor binding pattern for FXR.<sup>70</sup>

The compounds in the second group, the dual-site binders, bind to the LBD in a two to one stoichiometry where typically one of the two ligands targets the ligand binding pocket and the second ligand binds to a surface exposed site elsewhere on the LBD.<sup>71–73</sup> A number of crystal structures have been reported with one ligand binding to the ligand binding pocket and an additional ligand present in the hydrophobic groove of the coactivator binding site, e.g. hydroxytamoxifen in complex with ER $\beta$ , directly preventing coactivator binding from the AF-2 site.<sup>56</sup> Another example in this group, which instead functions via a real allosteric mechanism, is the toxic bile acid lithocholic acid (LCA) which is an agonist for the vitamin D receptor (VDR).<sup>71</sup> Binding of the ligand to VDR, in turn, facilitates the clearance of LCA metabolites. A recent crystal structure of the VDR LBD co-crystallized with LCA, showed binding of the ligand to the orthosteric pocket as well as to a surface-exposed site near the cofactor binding site.<sup>74</sup> The presence of two LCA molecules leads to a stabilization of H12, the loop between H11 and H12 and the cofactor. Binding of a ligand to this second binding site might therefore reveal a mechanism to selectively induce cofactor binding to VDR.<sup>74</sup>

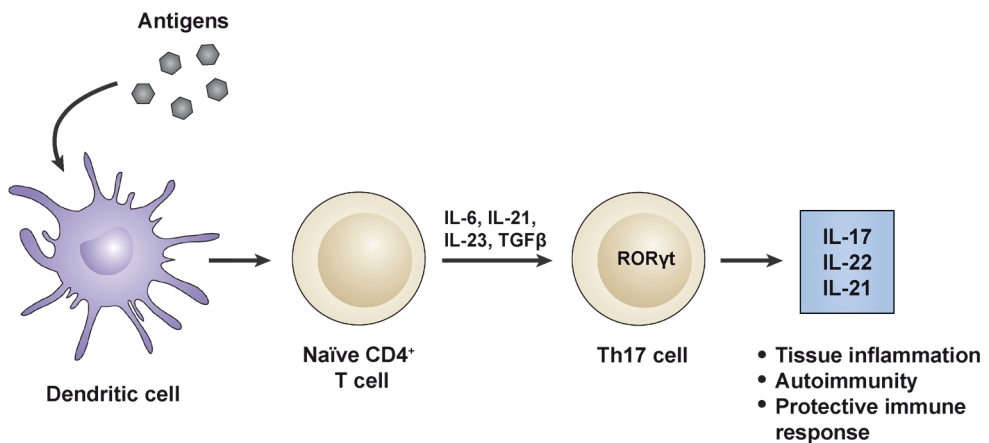
The third class deals with allosteric sites fully outside of the orthosteric ligand binding pocket, which are termed alternative binding pockets. An important note is that binding of ligands to these allosteric pockets could potentially still influence orthosteric ligand binding, because of allosteric effects. First, the Sulindac-derived analogs K-8008 and K-8012 were identified as novel RXR $\alpha$  allosteric antagonists, inhibiting coactivator recruitment with IC<sub>50</sub> values around 10  $\mu$ M.<sup>75</sup> The ligands stabilize RXR $\alpha$  dimer formation by binding to a pocket near the dimer-dimer interface, in that way diminishing the interaction between RXR $\alpha$  and the p85 $\alpha$  subunit of PI3K.<sup>75</sup> Inhibition of this protein-protein interaction leads to apoptosis and inhibition of tumor growth without apparent toxicity.<sup>75</sup> A second example in this class is the binding function-3 (BF-3) allosteric site on AR.<sup>76–79</sup> Ligands targeting this allosteric site regulate AR activity via the inhibition of co-chaperone binding, that normally interact with the AR LBD via the BF-3 site.<sup>80</sup> The recently discovered ligand VPC-13566 was found as a potent and selective ligand, inducing a significant inhibition of cancer cell growth, showing



potential for prostate cancer therapies.<sup>80</sup> Apart from these two examples, another very unique and promising alternative NR binding site is the allosteric site of ROR $\gamma$ t, which will be the focus of this thesis. The next part of this chapter will describe the function and modulation of ROR $\gamma$ t in more detail.

## The physiological function of ROR $\gamma$ t

The retinoic acid receptor-related orphan receptor (ROR) family of NRs has received its name from the sequence homology to the retinoic acid receptor (RAR).<sup>81</sup> The ROR subfamily contains three different subtypes: ROR $\alpha$  (NR1F1), ROR $\beta$  (NR1F2) and ROR $\gamma$  (NR1F3), which share about 50% sequence identity in their LBD.<sup>82–85</sup> Two different isoforms of ROR $\gamma$  exist; ROR $\gamma$ 1 and ROR $\gamma$ 2, with the latter also known as ROR $\gamma$ t, only differing from each other in the length of their NTD.<sup>82,83</sup> The RORs are type IV NRs, binding as monomers to their specific ROR response elements.<sup>17,83</sup> All ROR isoforms have distinct, tissue-specific expression patterns, regulating different target genes and thus different biological processes.<sup>82–84</sup> Interestingly, whereas ROR $\gamma$  is widely expressed in many tissues including the kidney, liver, adipose tissue and skeletal muscle, ROR $\gamma$ t expression is limited to the lymphoid tissues, especially the thymus.<sup>83,84</sup>



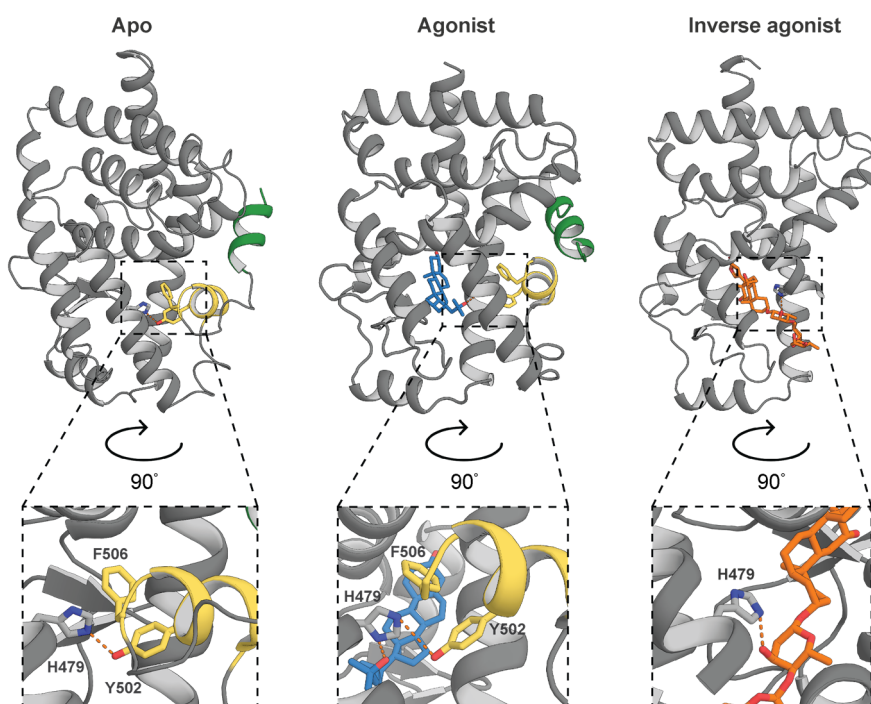
**Figure 1.7 | ROR $\gamma$ t plays a key role in Th17 cell differentiation.** The interaction between antigen-presenting dendritic cells and naïve CD4<sup>+</sup> T cells, in addition to the presence of several exogenous factors (IL-6, IL-21, IL-23 and TGF $\beta$ ), drives the differentiation into Th17 cells. The expression of ROR $\gamma$ t is necessary for Th17 cell differentiation and for the production of several cytokines, including IL-17, IL-22 and IL-21, that induce tissue inflammation and protective immune responses. Figure adapted from Kojetin *et al.*<sup>86</sup>

RORyt has many different biological functions, but most importantly, it plays an essential regulatory role in the immune system. The protein is a key regulator in the differentiation of naïve CD4<sup>+</sup> T cells into T helper 17 (Th17) cells (upon activation by an antigen-presenting dendritic cell) and the production of the pro-inflammatory cytokine IL-17a, needed for tissue inflammation and protective immune responses.<sup>84,87–89</sup> Elevated IL-17a levels are highly associated with the pathogenesis of autoimmune diseases, including psoriasis, multiple sclerosis, rheumatoid arthritis and inflammatory bowel disease.<sup>90–95</sup> Disrupting the Th17/IL-17a pathway could therefore potentially be an effective strategy for the treatment of these diseases.<sup>90</sup> The clinical successes of FDA-approved monoclonal antibodies targeting IL-17a or Th17 cell development (e.g. secukinumab<sup>96</sup> and ixekizumab<sup>97</sup> for the treatment of plaque psoriasis) have already validated the potential of Th17 pathway inhibition as a successful therapeutic strategy.<sup>98,99</sup> However, inhibition of RORyt with small, orally available drugs might be an attractive alternative strategy to decrease IL-17 production in the treatment of these autoimmune diseases, with broader therapeutic outcomes and potential benefits over expensive, intravenously administered antibody-based treatments. The inhibition of RORyt with small molecules has therefore been the focus of many research efforts over the past decades, with several synthetic RORyt inverse agonists progressed to clinical trials, as will be described further in the next paragraph.<sup>100–110</sup>

Apart from its essential role in the development of autoimmune diseases, RORyt has also emerged as an important therapeutic target in cancer. The inhibition of RORyt could for example be a promising strategy for the treatment of prostate cancer, since it stimulates AR gene transcription.<sup>111,112</sup> Conversely, activation of RORyt with agonists has been implicated in enhancing anti-tumor immunity.<sup>113–116</sup> The programmed cell death 1 receptor (PD-1) is a membrane receptor that is present on T cells and binds to PD-L1 on tumor cells.<sup>115</sup> The interaction between PD-1 and PD-L1 inhibits T cell activation and thus leads to an inactivation of the T cell immune response against the cancer cell.<sup>117</sup> Activation of RORyt decreases T cell PD-1 levels, which could prevent immune resistance in tumor cells and thus could be used in cancer immunotherapy.<sup>114,115,118</sup> Although these examples show potential for anti-cancer treatments, the development of such ligands remains an underexplored area of research, compared to the design of ligands for autoimmune therapies.<sup>116</sup>

## Ligand-dependent modulation of ROR $\gamma$ t

ROR $\gamma$ t appears to be constitutively active, which means that it features some level of background transcriptional activity, even in the absence of an agonist. However, ROR $\gamma$ t is still responsive to ligand binding.<sup>85</sup> The majority of ROR $\gamma$ t ligands bind to the canonical ligand binding pocket, termed the orthosteric binding site, within the LBD of ROR $\gamma$ t (Figure 1.10A). Recently, a novel class of ROR $\gamma$ t inverse agonists has been identified, that bind to a topographically distinct, non-canonical binding site in the LBD (Figure 1.10A).<sup>119</sup> An overview of ROR $\gamma$ t ligands and the structural characteristics of the apo and ligand-bound conformations of the ROR $\gamma$ t LBD will be described here in more detail.



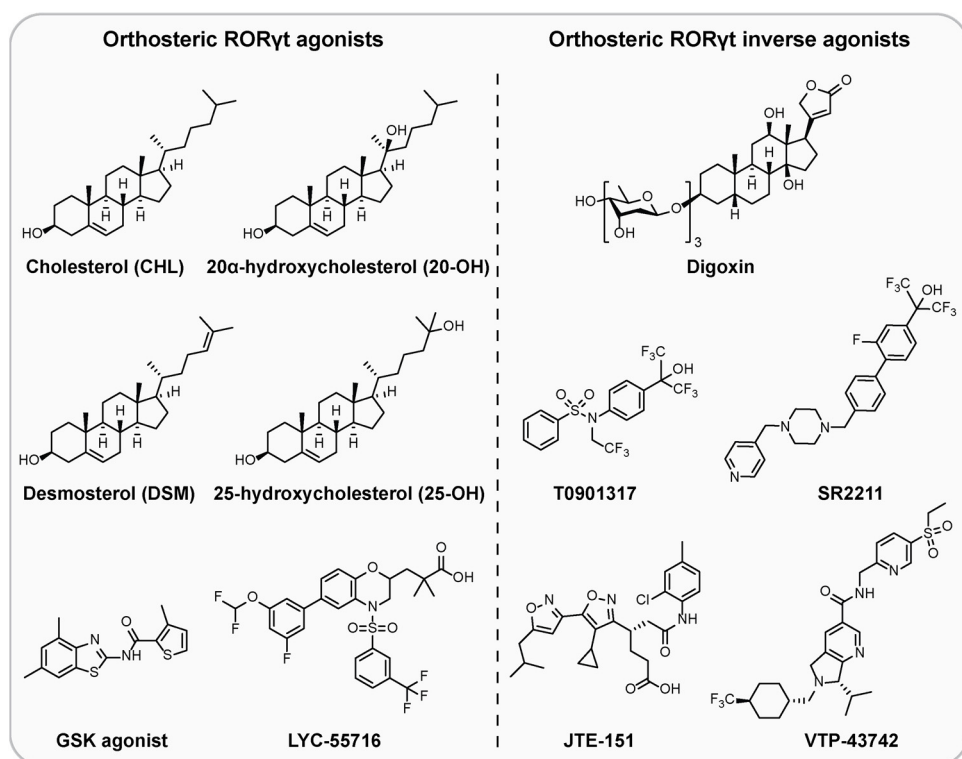
**Figure 1.8 | Crystal structures of the ROR $\gamma$ t LBD in the apo, agonist-bound and inverse agonist-bound state and the effect of ligand binding on the HYF triplet. Left)** Apo ROR $\gamma$ t, covalently linked to the coactivator SRC2 (shown in green) (PDB: 5VB3).<sup>85</sup> Close interactions between His479, Tyr502 and Phe506 ( $\pi$ - $\pi$  stacking and a hydrogen bond interaction between His479 and Tyr502 (orange dashes)) anchor H12 (shown in yellow) in the active conformation. **Middle)** ROR $\gamma$ t co-crystallized with 25-hydroxycholesterol (shown as blue sticks) (PDB: 3LoL).<sup>24</sup> The ligand stabilizes His479 further via a hydrogen bond interaction. **Right)** ROR $\gamma$ t co-crystallized with digoxin (shown as orange sticks) (PDB: 3BoW).<sup>120</sup> H12 is not visible in the crystal structure. The ligand induces a conformational change of His479 via hydrogen bonding, disrupting the His-Tyr lock, which results in destabilization of the active conformation of H12.

In the apo structure, the C-terminal H12 is already positioned in a conformation that enables partial recruitment of coactivator proteins, which causes the constitutively active behavior of ROR $\gamma$ t.<sup>85,121</sup> A key structural element that anchors H12 in the active conformation is the His479-Tyr502-Phe506 (HYF) triplet, which is conserved in the ROR family and not found in any other NR (Figure 1.8, left).<sup>85</sup> Tyr502 and Phe506 are located on the same face of H12, and form close interactions with the His479 residue on H11; Tyr502 makes a hydrogen bond interaction with His479 (termed the His-Tyr lock) and additionally, there is a favorable  $\pi$ - $\pi$  interaction between Tyr502 and Phe506 and edge-to-face  $\pi$ - $\pi$  stacking between His479 and Phe506.<sup>85,121–123</sup>

The apo ROR $\gamma$ t LBD contains a large, mainly hydrophobic ligand binding pocket with a volume of 940 Å<sup>3</sup>, as was revealed by the crystal structure.<sup>85</sup> This orthosteric binding site is highly conserved across the NR family and is located in the core of the globular LBD, mainly embraced by helices 3, 5, 6, 7 and 11.<sup>4,123</sup> Although formally an orphan receptor with no proven endogenous ligands, recent studies have convincingly shown that naturally occurring cholesterol derivatives are physiological ligands for ROR $\gamma$ t.<sup>24,124–126</sup> Cholesterol and its derivatives (i.e. 20 $\alpha$ -hydroxycholesterol (20-OH), 25-hydroxycholesterol (25-OH) and desmosterol (DSM)) (Figure 1.9) have been shown to be effective ROR $\gamma$ t agonists with low nanomolar potency in a coactivator recruitment assay.<sup>24</sup> These agonists, that bind to the orthosteric site, enhance ROR $\gamma$ t transcriptional activity further by stabilizing the active conformation of H12 via (indirect) stabilizing interactions with the HYF triplet, in a way that promotes the recruitment of transcriptional coactivators such as the steroid receptor coactivator (SRC) family.<sup>121–123</sup> As an example, the hydroxy group of agonist 25-OH forms a direct hydrogen bond with the imidazole of His-479 on H11, that in turn stabilizes Tyr502 and Phe506 (Figure 1.8, middle).<sup>122</sup> The mechanism of stabilization can also be indirect, as is the case for cholesterol.<sup>122</sup> It is hypothesized that these ligands stabilize the hydrogen bond between His479 and Tyr502 by occupying the pocket volume near the HYF triplet.<sup>122</sup> Besides these natural ROR $\gamma$ t agonists, some companies, including GSK<sup>102</sup> and Lycera (LYC-55716, phase 2A clinical trials for cancer immunotherapy)<sup>116,127</sup>, have also identified a number of synthetic agonists (Figure 1.9).

In contrast, ROR $\gamma$ t inverse agonists destabilize H12 in a conformation that is unsuitable for coactivator binding, thus leading to diminished gene transcription. Different mechanisms of action have been reported, but most commonly these inverse agonists disturb the His-Tyr lock of the HYF triplet and clash with amino acid residues in the active protein conformation.<sup>85,122,123</sup> A well-known ROR $\gamma$ t inverse agonist is digoxin (Figure 1.9), for which

the trisaccharide extension on its steroidal core induces a conformational change of His479 via hydrogen bonding, disrupting the His-Tyr lock, which results in destabilization of the active conformation of H12 (Figure 1.8, right).<sup>120,128,129</sup> Additionally, the extension of the steroid scaffold would clash when H11 and H12 would be in the active conformation, inducing a displacement of H12 (that is no longer visible in the crystal structure due to its flexible character), which prevents H12 from adopting the necessary conformation for coactivator binding.<sup>120,122</sup> Although the His-catching mechanism is a common mode of action reported for multiple other ROR $\gamma$ t inverse agonists, a number of inverse agonists can also destabilize the His-Tyr lock indirectly by inducing conformational changes in the protein without directly interacting with His479.<sup>122</sup> In addition to the mentioned natural ligand digoxin, numerous synthetic inverse agonists are also known for ROR $\gamma$ t. The first validated synthetic ROR $\gamma$ t ligand was the benzenesulfonamide **T0901317**<sup>130</sup> (Figure 1.9), which was originally identified



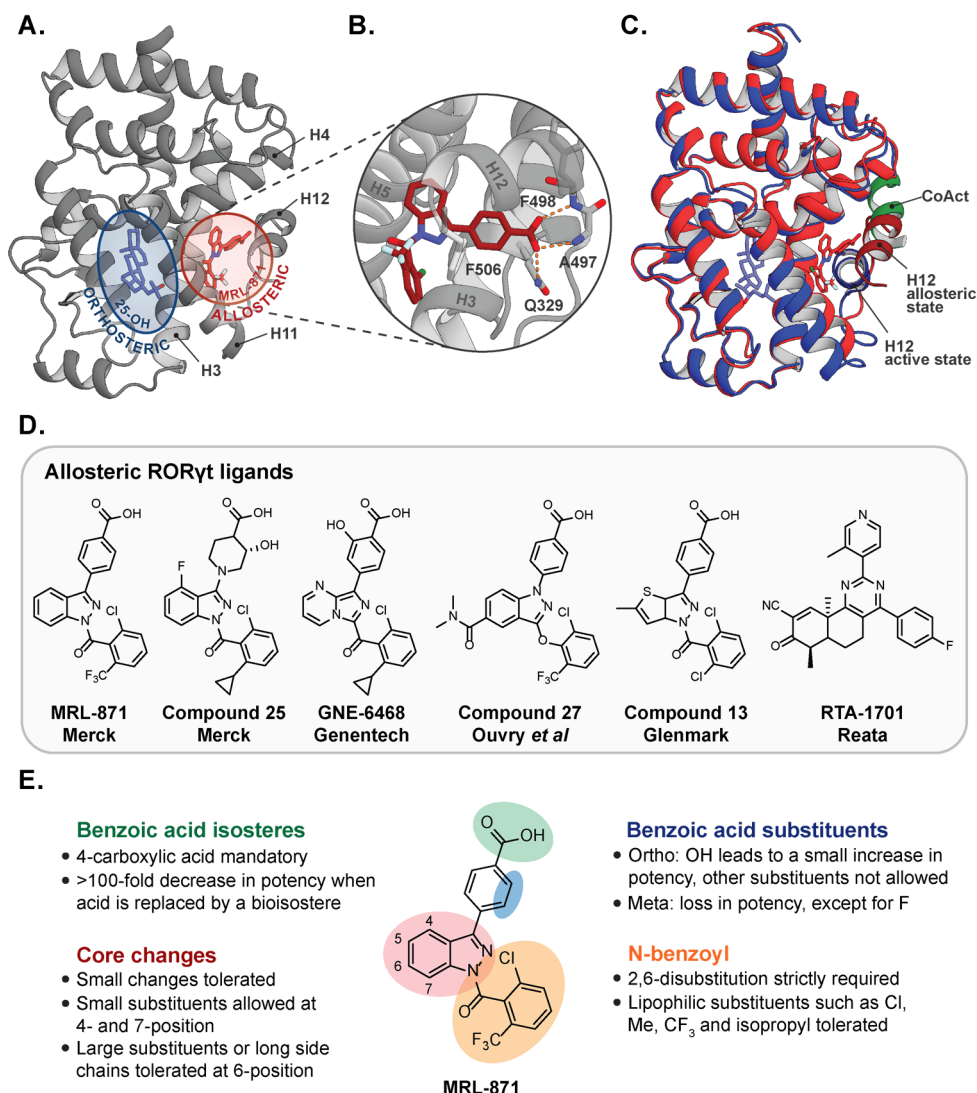
**Figure 1.9 | Overview of ROR $\gamma$ t orthosteric ligands.** Orthosteric agonists: natural ligands cholesterol (CHL), 20 $\alpha$ -hydroxycholesterol (20 $\alpha$ -OH), 25-hydroxycholesterol (25-OH), desmosterol (DSM), and synthetic ligands **LYC-55716** (Lycera) and a GSK agonist. Orthosteric inverse agonists: natural ligand digoxin, synthetic ligands **T0901317** and **SR2211**, and clinical candidates **JTE-151** (Japan Tobacco) and **VTP-43742** (Vitae).

as an agonist for the liver X receptor (LXR) and FXR.<sup>86</sup> A team at the Scripps Florida used this non-selective ligand as a starting point for developing selective ROR $\gamma$ t inverse agonists, such as **SR2211**<sup>131</sup> (Figure 1.9), but also agonists.<sup>115</sup>

Many pharmaceutical companies, including Merck, GSK and AstraZeneca, have made extensive efforts to develop synthetic (orthosteric) ROR $\gamma$ t inverse agonists as clinical candidates for the treatment of autoimmune diseases, with eight inhibitors currently active in clinical studies.<sup>100–104</sup> However, some of these highly active ROR $\gamma$ t ligands in clinical trials were discontinued for further development, including compounds **VTP-43742** (VitaE)<sup>132</sup> and **JTE-151** (Japan Tobacco)<sup>133</sup> (Figure 1.9). The reasons for discontinuation were mainly due to off-target effects, toxicity and poor therapeutic efficacy.<sup>104,123,134</sup> While a number of companies still focus on optimizing chemical structures of orthosteric ROR $\gamma$ t ligands to improve selectivity and physicochemical properties, others have shifted their attention to the investigation of the druggability of an allosteric ROR $\gamma$ t binding site.

### Modulation of ROR $\gamma$ t via a non-canonical, allosteric binding site

Recently, a novel class of ROR $\gamma$ t inverse agonists has been identified, typified by **MRL-871** (disclosed by Merck) (Figure 1.10D), that bind to a previously unreported allosteric binding site.<sup>119,135</sup> This topographically distinct, second binding site on the ROR $\gamma$ t LBD is formed by H3, H4, H11, H11' and reoriented H12, which has a unique conformation by folding back over the ligand (Figure 1.10A).<sup>119,121</sup> The interactions in the allosteric pocket are predominantly hydrophobic and a clear  $\pi$ - $\pi$  stacking interaction between the 2,6-disubstituted phenyl ring and Phe506 can be observed.<sup>119,121</sup> Additionally, hydrogen bonds are formed between the carboxylic acid moiety of **MRL-871** and the main chain amide hydrogen atoms of Ala497 and Phe498 (located on H12) as well as the side chain of residue Gln329 (H3) (Figure 1.10B).<sup>119</sup> The ligands directly interact with the activation function loop between H11 and H12 (AF-2 domain), thus forcing H12 to adopt an unusual conformation that is incompatible with coactivator recruitment (Figure 1.10C).<sup>119,121</sup> Time-resolved fluorescence resonance energy transfer (TR-FRET) assays with the allosteric ligands indeed showed inhibition of coactivator binding with IC<sub>50</sub> values in the low nanomolar range.<sup>119</sup>



**Figure 1.10 | An allosteric binding site is present in the LBD of RORγt.** **A)** The co-crystal structure of RORγt in complex with allosteric ligand **MRL-871** (red sticks) (PDB: 4YPQ)<sup>119</sup>, with orthosteric ligand 25-hydroxycholesterol (25-OH) (blue sticks) shown as an overlay (PDB: 3LoL).<sup>24</sup> **B)** Enlarged view of **MRL-871** (red sticks) in the allosteric pocket, with hydrogen bond interactions shown with orange dashes. **C)** Overlay of the co-crystal structures of RORγt in complex with 20α-hydroxycholesterol (blue sticks) (PDB: 3LoL)<sup>24</sup> and with **MRL-871** (red sticks) (PDB: 4YPQ)<sup>119</sup>, showing **MRL-871** binding at the location where H12 (dark blue) is positioned in the active conformation (when an orthosteric agonist is bound). H12 in the allosteric conformation (dark red) adopts an unusual conformation that is incompatible with coactivator (CoAct, green) recruitment. **D)** Chemical structures of allosteric RORγt ligands **MRL-871** (Merck), **compound 25** (Merck), **GNE-6468** (Genentech), **compound 27** (Ouvry *et al*), **compound 13** (Glenmark) and **RTA-1701** (Reata). **E)** Structure activity relationship (SAR) profile of **MRL-871** and its derivatives.

The allosteric pocket has been shown to be highly unique to ROR $\gamma$ t within the NR family, indicating that ROR $\gamma$ t allosteric ligands could achieve high ROR $\gamma$ t-subtype selectivity.<sup>119,136</sup> Indeed, the MRL-series showed >100-fold selectivity for ROR $\gamma$ t when tested against a panel of NRs.<sup>119</sup> The only notable off-target activity was against the orthosteric binding site of PPAR $\gamma$ , observed at high concentrations of the **MRL-871** derivatives.<sup>119,137</sup> Additionally, although showing the desired biological response in autoimmune models,<sup>138,139</sup> the pharmacokinetic (PK) properties of the MRL-series appeared to be not optimal, since metabolic instability (hydrolytic cleavage of the heterocyclic amide linkage) and toxicity issues have been reported.<sup>137,140</sup>

Extensive follow-up research on this allosteric ligand class has been done by both academia and industry, focusing on additional structure activity relationship (SAR) studies, and on the optimization of the PK properties and selectivity profile, while maintaining the favorable potency for ROR $\gamma$ t.<sup>137,140–144</sup> Merck mainly focused on improving the metabolic stability of a structurally related **MRL-871** analog.<sup>141</sup> Acyl-glucuronidation was found to be one of its major clearance mechanisms, but conversion of the carboxylate into other typical acid bioisosteres or introducing steric hindrance at the ortho position of the carboxylate moiety, resulted in a significant loss of potency.<sup>141</sup> Saturation of the benzoic acid ring moiety led to the discovery of **compound 25** (Figure 1.10D) that showed a highly improved *in vivo* metabolic stability and better physicochemical properties (e.g. reduced lipophilicity and higher solubility), in addition to an even higher potency, a favorable selectivity profile and the expected allosteric binding mode.<sup>141</sup> Two related classes of inhibitors, with an alternative core, were reported by Genentech and Ouvre *et al.*<sup>137,140</sup> The main focus in the design of the new analogs was to improve the selectivity for ROR $\gamma$ t over PPAR $\gamma$  and to remove the potential metabolic liability of the amide bond in **MRL-871**, since related compounds with an *N*-benzoyl moiety had demonstrated *in vivo* instability.<sup>137,140,145</sup> **GNE-6468** and phenoxyindazole **compound 27** (Figure 1.10D) were both identified as highly potent compounds in both biochemical and cellular assays, with an improved selectivity profile (>300-fold selectivity for ROR $\gamma$ t) and physicochemical properties.<sup>137,140</sup> However, because of unfavorable *in vivo* properties, such as high plasma clearance values and high phototoxicity values that could not be improved by additional structural changes, further investigation has been discontinued.<sup>137,140</sup> Finally, Glenmark Pharmaceuticals performed an *in silico* scaffold hopping approach, using the **MRL-871** core as the basis.<sup>143,144</sup> This resulted in the development of a family of thiophenopyrazole inverse agonists demonstrating low nanomolar IC<sub>50</sub> values in biochemical assays.<sup>143,144</sup> **Compound 13** (Figure 1.10D) was identified as the most potent



example<sup>143,144</sup>, which was later proven to bind via an allosteric binding mode.<sup>146</sup> However, the lack of cellular assays or *in vivo* models leaves questions regarding the potential of the (pre)clinical profile.<sup>144</sup>

The combination of the results of the previously described studies elucidated a clear SAR profile for this type of allosteric ROR $\gamma$ t ligands, for which the most important molecular characteristics are shown in Figure 1.10E.<sup>119,135,137,140–144</sup> First, there is a strict requirement for lipophilic 2,6-disubstitution on the *N*-benzoyl moiety (i.e. 2,6-dichloro, 2,6-dimethyl or 2-chloro-6-CF<sub>3</sub>/isopropyl substitution), since these *ortho* substituents impart a specific rotation to the phenyl ring (a perpendicular orientation to the indazole scaffold) and address hydrophobic sites in the allosteric pocket. Removal of one of these substituents, a change in the type of substituent (e.g. fluoro, cyano or methoxy), or substituents elsewhere on the ring lead to a drastic loss of potency. Furthermore, a 4-benzoic acid moiety is mandatory for potency against ROR $\gamma$ t, since the replacement of the carboxylic acid with isosteres like tetrazole, various benzamides and *N*-acylsulfonamides resulted in a >100-fold decrease in potency. Installation of an *ortho*-hydroxy moiety on the benzoic acid group shows small improvements in potency, while other substituents such as chloro, cyano, methyl or methoxy groups are not allowed. *Meta*-substitution generally results in a loss of activity, except for a fluoro residue which leads to a retention in potency. Small differences in the indazole core, such as a flipped central core, or the removal or addition of a nitrogen atom, are in general allowed. Also small substitutions on the indazole scaffold, such as a fluoro or methyl, are allowed (especially at the 4- and 7-position), while at the 6-position, larger substituents such as dimethyl-amides or even amides with long side chains, are tolerated as well.

Apart from **MRL-871** and the mentioned derivatives, one alternative class of ROR $\gamma$ t allosteric ligands has been reported by Reata Pharmaceuticals that contain a different chemical structure (pyrimidine tricyclic enone cores), as exemplified by the highly potent compound **RTA-1701** (Figure 1.10D).<sup>147,148</sup> Interestingly, this series of compounds bind (partly) to the same allosteric site of ROR $\gamma$ t, but via covalent attachment (Michael addition) to Cys476 on helix 11 of the LBD.<sup>147</sup> The compounds show high selectivity towards ROR $\gamma$ t (compared to other NR members), and beneficial PK and biological properties.<sup>147</sup> **RTA-1701** is currently in phase I clinical trials for the treatment of autoimmune disorders, as the first allosteric clinical candidate for ROR $\gamma$ t.<sup>104</sup>

## Aim and outline of this thesis

NRs are a promising target for drug discovery because of their central regulating role within the human genome. The primary focus of NR targeting has classically been via the orthosteric ligand binding pocket. Small molecule modulators binding to this canonical pocket in the LBD have yielded great successes, however issues such as cross-reactivity and competition with endogenous ligands remain a major challenge. In recent years, allosteric modulation of NRs has gained attention as a promising alternative strategy. Allosteric ligands modulate the NR activity via distinct structural mechanisms, by binding to a pocket on the protein that does not overlap with the orthosteric ligand binding pocket. Many successful examples of allosteric NR modulation, especially via allosteric pockets on the LBD, have been shown in this chapter.

ROR $\gamma$ t is an illustrative example of a NR containing both a clearly defined orthosteric binding site, but also a highly unique second binding site, termed allosteric pocket, located near H12 in the LBD. ROR $\gamma$ t plays a critical role in the immune system, regulating the development and differentiation of Th17 cells. Inhibition of ROR $\gamma$ t therefore shows high potential in the treatment of autoimmune diseases. Many orthosteric ROR $\gamma$ t ligands, both natural and synthetic, have been reported in the literature. However, a number of clinical candidates, targeting the orthosteric site, have been discontinued for further development, mainly due to significant off-target activities. Inhibition of ROR $\gamma$ t via the allosteric pocket thus shows a promising alternative to orthosteric modulation. The first example of an allosteric ROR $\gamma$ t ligand was the indazole **MRL-871**, a potent inverse agonist. Although extensive follow-up studies have been performed on this allosteric site, most studies focused on **MRL-871** derivatives, leaving high demand for further exploration of the allosteric pocket besides this known ligand class.

The existence of both an orthosteric and a second, allosteric binding pocket provides plenty of opportunities for the modulation of ROR $\gamma$ t. Therefore, **the aim of the work described in this thesis** is to further explore the modulation of ROR $\gamma$ t via the allosteric binding site, and the eventual interplay between both binding sites, because of three reasons: **1)** ROR $\gamma$ t is a highly interesting target from a drug discovery perspective, and further investigation of ROR $\gamma$ t allosteric inhibition could provide new insights towards novel allosteric, therapeutic ROR $\gamma$ t modulators, **2)** ROR $\gamma$ t inhibitors can be utilized as chemical tools to enhance the understanding of ROR $\gamma$ t-related biology, and **3)** Modulation of ROR $\gamma$ t via the allosteric site could provide a learning platform, for which the insights could also be applied to the allosteric modulation of other NRs, both in drug discovery but also in chemical biology applications.

In **Chapter 2**, *in silico* pharmacophore screening and docking studies were used to discover a novel class of allosteric ROR $\gamma$ t ligands, in order to diversify in the scaffolds targeting the allosteric site. A library of compounds was synthesized and after some rounds of optimization, the trisubstituted isoxazole ligand **FM26** was found as a lead compound. The ligand showed high inhibition of coactivator recruitment in TR-FRET assays in a selective way, and significantly reduced IL-17a mRNA expression levels in RT-PCR assays. The co-crystal structure revealed an allosteric binding mode, similar to **MRL-871**, with an additional hydrogen bond interaction with the backbone of the protein, which could explain the high potency in comparison to compounds lacking a hydrogen bond donor moiety.

**Chapter 3** describes additional structure activity relationship studies around lead compound **FM26**, in order to obtain extra insight into the binding mode of this isoxazole class of allosteric ROR $\gamma$ t ligands. First, the synthesis route for these derivatives was optimized to obtain a Boc-protected core intermediate that could be used for the synthesis of all derivatives, containing different substituents and linkers. The ligands were tested in both TR-FRET and thermal shift assays, showing that especially the compounds with small lipophilic linkers (e.g. an ether or alkene instead of amine) demonstrate an increased potency compared to **FM26**. The co-crystal structures for some of these derivatives showed the expected allosteric binding mode and could aid in elucidating a clear SAR profile for this new type of compounds. Additionally, a high selectivity profile for ROR $\gamma$ t over PPAR $\gamma$  was observed in TR-FRET assays and promising PK properties were measured, which shows high potential for this novel class of allosteric ROR $\gamma$ t ligands.

In **Chapter 4**, TR-FRET studies revealed the simultaneous binding of an orthosteric and allosteric ligand to both binding sites of ROR $\gamma$ t, called dual ligand binding. The binding of both ligands shows a cooperative behavior, as the potency of an allosteric ligand is enhanced in the presence of an orthosteric ligand. In order to explore the mechanism behind this cooperative dual ligand binding, the co-crystal structures of ROR $\gamma$ t, in complex with both an orthosteric and allosteric ligand, were solved for all combinations of ligands. The structural data elucidated a conformational change within the protein, resulting in a clamping motion around the allosteric pocket upon binding of an orthosteric ligand. Molecular dynamics simulations shed light on the mechanism behind this clamping motion, emphasizing the ability of Ala355 to switch between helix 4 and helix 5. The orthosteric ROR $\gamma$ t ligands regulate the conformation of this Ala355 residue, thereby shifting helix 4 towards the allosteric pocket and enhancing the potency of the allosteric inverse agonists.

**Chapter 5** describes the occlusion of the orthosteric site of ROR $\gamma$ t using covalent chemical probes, with the ultimate aim of developing a screening approach to identify novel allosteric ROR $\gamma$ t ligands. When a screening on the native ROR $\gamma$ t LBD would be performed, a combination of orthosteric and allosteric ligands would be identified as they both act as ROR $\gamma$ t inhibitors. Therefore, the orthosteric site is occluded, which will prevent orthosteric ligands from binding, enabling the selective screening of allosteric ligands. A small library of covalent probes was synthesized and they were ligated to a native cysteine residue in the orthosteric binding site of ROR $\gamma$ t. A selection of these probes showed efficient inhibition of orthosteric ligand binding in TR-FRET and thermal shift assays, while the allosteric site was still capable of binding ligands. These covalent probes are thus excellent tool compounds for the development of a screening method that could be used for the unambiguous and rapid identification of novel allosteric ROR $\gamma$ t ligands.

In **Chapter 6**, the dual targeting of the orthosteric and allosteric binding sites of ROR $\gamma$ t was investigated via the development of a bitopic ligand that comprises of a covalently linked orthosteric and allosteric ROR $\gamma$ t pharmacophore. Three candidate bitopic ligands were designed, synthesized and biochemically evaluated. The most promising results were obtained for **Bit-L15**, containing a PEG-15 linker that connects the two pharmacophores. Biochemical assays provided evidence for the concomitant engagement of both binding pockets of the ROR $\gamma$ t LBD, suggesting a bitopic mode of action, while the overall efficacy compared to the allosteric pharmacophore **MRL-871** was maintained in both a biochemical and a cellular context. In addition, selectivity studies revealed that **Bit-L15** displays increased selectivity for ROR $\gamma$ t over other NRs. These results demonstrate that bitopic modulation of ROR $\gamma$ t might enable desirable and advantageous properties over classical, monovalent targeting strategies.

Finally, **Chapter 7** (the epilogue) provides a future perspective on the allosteric modulation of NRs, and specifically ROR $\gamma$ t, based on the work described in this thesis. Preliminary insights from currently ongoing work are presented as well. Additional experimental strategies for the discovery of novel allosteric ligands are described, as well as anticipated challenges regarding ROR $\gamma$ t (allosteric) inhibition.

## References

1. Robinson-Rechavi, M., Escriva Garcia, H. & Laudet, V. The nuclear receptor superfamily. *J. Cell Sci.* **116**, 585–586 (2003).
2. Burris, T. P. *et al.* Nuclear Receptors and Their Selective Pharmacologic Modulators. *Pharmacol. Rev.* **65**, 710–778 (2013).
3. Huang, P., Chandra, V. & Rastinejad, F. Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. *Annu. Rev. Physiol.* **72**, 247–272 (2010).
4. Gronemeyer, H., Gustafsson, J.-Å. & Laudet, V. Principles for modulation of the nuclear receptor superfamily. *Nat. Rev. Drug Discov.* **3**, 950–964 (2004).
5. Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 (2017).
6. Moore, J. T., Collins, J. L. & Pearce, K. H. The Nuclear Receptor Superfamily and Drug Discovery. *ChemMedChem* **1**, 504–523 (2006).
7. Zhao, L., Zhou, S. & Gustafsson, J.-Å. Nuclear Receptors: Recent Drug Discovery for Cancer Therapies. *Endocr. Rev.* **40**, 1207–1249 (2019).
8. Rastinejad, F., Huang, P., Chandra, V. & Khorasanizadeh, S. Understanding nuclear receptor form and function using structural biology. *J. Mol. Endocrinol.* **51**, T1–T21 (2013).
9. Tata, J. R. Signalling through nuclear receptors. *Nat. Rev. Mol. Cell Biol.* **3**, 702–710 (2002).
10. Kumar, R. & Litwack, G. Structural and functional relationships of the steroid hormone receptors' N-terminal transactivation domain. *Steroids* **74**, 877–883 (2009).
11. Helsen, C. & Claessens, F. Looking at nuclear receptors from a new angle. *Mol. Cell. Endocrinol.* **382**, 97–106 (2014).
12. Helsen, C. *et al.* Structural basis for nuclear hormone receptor DNA binding. *Mol. Cell. Endocrinol.* **348**, 411–417 (2012).
13. Haelens, A., Tanner, T., Denayer, S., Callewaert, L. & Claessens, F. The hinge region regulates DNA binding, nuclear translocation, and transactivation of the androgen receptor. *Cancer Res.* **67**, 4514–23 (2007).
14. Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H. & Moras, D. Crystal structure of the ligand-binding domain of the human nuclear receptor RXR- $\alpha$ . *Nature* **375**, 377–382 (1995).
15. Patel, S. R. & Skafar, D. F. Modulation of nuclear receptor activity by the F domain. *Mol. Cell. Endocrinol.* **418**, 298–305 (2015).
16. Mangelsdorf, D. J. *et al.* The nuclear receptor superfamily: the second decade. *Cell* **83**, 835–839 (1995).
17. Glass, C. K. & Ogawa, S. Combinatorial roles of nuclear receptors in inflammation and immunity. *Nat. Rev. Immunol.* **6**, 44–55 (2006).
18. Heldin, C.-H., Lu, B., Evans, R. & Gutkind, J. S. Signals and Receptors. *Cold Spring Harb. Perspect. Biol.* **8**, a005900 (2016).
19. Echeverria, P. C. & Picard, D. Molecular chaperones, essential partners of steroid hormone receptors for activity and mobility. *Biochim. Biophys. Acta - Mol. Cell Res.* **1803**, 641–649 (2010).
20. Sever, R. & Glass, C. K. Signaling by nuclear receptors. *Cold Spring Harb. Perspect. Biol.* **5**, 1–4 (2013).
21. Bulynko, Y. A. & O'Malley, B. W. Nuclear receptor coactivators: structural and functional biochemistry. *Biochemistry* **50**, 313–328 (2011).
22. Pissios, P., Tzamelis, I., Kushner, P. J. & Moore, D. D. Dynamic Stabilization of Nuclear Receptor Ligand Binding Domains by Hormone or Corepressor Binding. *Mol. Cell* **6**, 245–253 (2000).
23. Kojetin, D. J. & Burris, T. P. Small molecule modulation of nuclear receptor conformational dynamics: implications for function and drug discovery. *Mol. Pharmacol.* **83**, 1–8 (2013).
24. Jin, L. *et al.* Structural Basis for Hydroxycholesterols as Natural Ligands of Orphan Nuclear Receptor ROR $\gamma$ . *Mol. Endocrinol.* **24**, 923–929 (2010).
25. Noguchi, M. *et al.* Ternary crystal structure of human ROR $\gamma$  ligand-binding-domain, an inhibitor and corepressor peptide provides a new insight into corepressor interaction. *Sci. Rep.* **8**, 17374 (2018).
26. Chandra, V. *et al.* Structure of the intact PPAR- $\gamma$ -RXR- $\alpha$  nuclear receptor complex on DNA. *Nature* **456**, 350–356 (2008).
27. Ricci, C. G., Silveira, R. L., Rivalta, I., Batista, V. S. & Skaf, M. S. Allosteric Pathways in the PPAR $\gamma$ -RXR $\alpha$  nuclear receptor complex. *Sci. Rep.* **6**, 19940 (2016).
28. Orlov, I., Rochel, N., Moras, D. & Klaholz, B. P. Structure of the full human RXR/VDR nuclear receptor heterodimer complex with its DR3 target DNA. *EMBO J.* **31**, 291–300 (2012).
29. Rochel, N. *et al.* Common architecture of nuclear receptor heterodimers on DNA direct repeat elements with different spacings. *Nat. Struct. Mol. Biol.* **18**, 564–570 (2011).
30. Changeux, J.-P. & Christopoulos, A. Allosteric Modulation as a Unifying Mechanism for Receptor Function and Regulation. *Cell* **166**, 1084–1102 (2016).

31. Hilser, V. J. & Thompson, E. B. Structural dynamics, intrinsic disorder, and allostery in nuclear receptors as transcription factors. *J. Biol. Chem.* **286**, 39675–39682 (2011).
32. Billas, I. & Moras, D. Allosteric controls of nuclear receptor function in the regulation of transcription. *J. Mol. Biol.* **425**, 2317–2329 (2013).
33. Motlagh, H. N., Wrabl, J. O., Li, J. & Hilser, V. J. The ensemble nature of allostery. *Nature* **508**, 331–339 (2014).
34. Nussinov, R. & Tsai, C.-J. Allostery in Disease and in Drug Discovery. *Cell* **153**, 293–305 (2013).
35. Hall, J. M., McDonnell, D. P. & Korach, K. S. Allosteric Regulation of Estrogen Receptor Structure, Function, and Coactivator Recruitment by Different Estrogen Response Elements. *Mol. Endocrinol.* **16**, 469–486 (2002).
36. Helsen, C. *et al.* Evidence for DNA-binding domain–ligand-binding domain communications in the androgen receptor. *Mol. Cell. Biol.* **32**, 3033–3043 (2012).
37. Becares Salles, N., Gage, M. C. & Pineda-Torra, I. Post-translational modifications of lipid-activated nuclear receptors: Focus on metabolism. *Endocrinology* **158**, 213–225 (2016).
38. Faus, H. & Haendler, B. Post-translational modifications of steroid receptors. *Biomed. Pharmacother.* **60**, 520–528 (2006).
39. Anbalagan, M., Huderson, B., Murphy, L. & Rowan, B. G. Post-translational modifications of nuclear receptors and human disease. *Nucl. Recept. Signal.* **10**, e001 (2012).
40. Brunmeir, R., Xu, F., Brunmeir, R. & Xu, F. Functional Regulation of PPARs through Post-Translational Modifications. *Int. J. Mol. Sci.* **19**, 1738 (2018).
41. Khan, S. H., Ling, J. & Kumar, R. TBP Binding-Induced Folding of the Glucocorticoid Receptor AF1 Domain Facilitates Its Interaction with Steroid Receptor Coactivator-1. *PLoS One* **6**, e21939 (2011).
42. Jin, L. & Li, Y. Structural and functional insights into nuclear receptor signaling. *Adv. Drug Deliv. Rev.* **62**, 1218–1226 (2010).
43. Fernandez, E. J. Allosteric pathways in nuclear receptors — Potential targets for drug design. *Pharmacol. Ther.* **183**, 152–159 (2018).
44. van Westen, G. J. P., Gaulton, A. & Overington, J. P. Chemical, Target, and Bioactive Properties of Allosteric Modulation. *PLoS Comput. Biol.* **10**, e1003559 (2014).
45. Moore, T. W., Mayne, C. G. & Katzenellenbogen, J. a. Minireview: Not picking pockets: nuclear receptor alternate-site modulators (NRAMs). *Mol. Endocrinol.* **24**, 683–695 (2010).
46. Caboni, L. & Lloyd, D. G. Beyond the ligand-binding pocket: targeting alternate sites in nuclear receptors. *Med. Res. Rev.* **33**, 1081–1118 (2013).
47. Lu, S., Li, S. & Zhang, J. Harnessing allostery: a novel approach to drug discovery. *Med. Res. Rev.* **34**, 1242–1285 (2014).
48. Tice, C. M. & Zheng, Y.-J. Non-canonical modulators of nuclear receptors. *Bioorg. Med. Chem. Lett.* **26**, 4157–4164 (2016).
49. Wenthur, C. J., Gentry, P. R., Mathews, T. P. & Lindsley, C. W. Drugs for allosteric sites on receptors. *Annu. Rev. Pharmacol. Toxicol.* **54**, 165–184 (2014).
50. Taly, A., Hénin, J., Changeux, J.-P. & Cecchini, M. Allosteric regulation of pentameric ligand-gated ion channels: an emerging mechanistic perspective. *Channels (Austin)* **8**, 350–360 (2014).
51. Thal, D. M., Glukhova, A., Sexton, P. M. & Christopoulos, A. Structural insights into G-protein-coupled receptor allostery. *Nature* **559**, 45–53 (2018).
52. Gentry, P. R., Sexton, P. M. & Christopoulos, A. Novel Allosteric Modulators of G Protein-coupled Receptors. *J. Biol. Chem.* **290**, 19478–19488 (2015).
53. Hardy, J. A. & Wells, J. A. Searching for new allosteric sites in enzymes. *Curr. Opin. Struct. Biol.* **14**, 706–715 (2004).
54. Jensen, A. A. & Bräuner-Osborne, H. Allosteric modulation of the calcium-sensing receptor. *Curr. Neuropharmacol.* **5**, 180–186 (2007).
55. Hassan, A. Q., Sharma, S. V & Warmuth, M. Allosteric inhibition of BCR-ABL. *Cell Cycle* **9**, 3710–3714 (2010).
56. Wang, Y. *et al.* A second binding site for hydroxytamoxifen within the coactivator-binding groove of estrogen receptor beta. *Proc. Natl. Acad. Sci.* **103**, 9908–9911 (2006).
57. Kojetin, D. J. *et al.* Structural mechanism for signal transduction in RXR nuclear receptor heterodimers. *Nat. Commun.* **6**, 8013 (2015).
58. Scheepstra, M. *et al.* Ligand Dependent Switch from RXR Homo- to RXR-NURR1 Heterodimerization. *ACS Chem. Neurosci.* **8**, 2065–2077 (2017).
59. Leibowitz, M. D. *et al.* Biological Characterization of a Heterodimer-Selective Retinoid X Receptor Modulator: Potential Benefits for the Treatment of Type 2 Diabetes. *Endocrinology* **147**, 1044–1053 (2006).
60. Choi, J. H. *et al.* Anti-diabetic drugs inhibit obesity-linked phosphorylation of PPAR $\gamma$  by Cdk5. *Nature* **466**, 451–456 (2010).

61. Choi, J. H. *et al.* Antidiabetic actions of a non-agonist PPAR $\gamma$  ligand blocking Cdk5-mediated phosphorylation. *Nature* **477**, 477–481 (2011).
62. De Vries-van Leeuwen, I. J. *et al.* Interaction of 14-3-3 proteins with the Estrogen Receptor Alpha F domain provides a drug target interface. *Proc. Natl. Acad. Sci.* **110**, 8894–8899 (2013).
63. Gearhart, M. D. *et al.* Inhibition of DNA Binding by Human Estrogen-Related Receptor 2 and Estrogen Receptor  $\alpha$  with Minor Groove Binding Polyamides. *Biochemistry* **44**, 4196–4203 (2005).
64. Nickols, N. G. & Dervan, P. B. Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc. Natl. Acad. Sci.* **104**, 10418–10423 (2007).
65. Whittall, R. M. *et al.* Preferential Oxidation of Zinc Finger 2 in Estrogen Receptor DNA-binding Domain Prevents Dimerization and, Hence, DNA Binding. *Biochemistry* **39**, 8406–8417 (2000).
66. Hughes, T. S. *et al.* An alternate binding site for PPAR $\gamma$  ligands. *Nat. Commun.* **5**, 3571 (2014).
67. Li, Y. *et al.* T2384, a novel antidiabetic agent with unique peroxisome proliferator-activated receptor gamma binding properties. *J. Biol. Chem.* **283**, 9168–9176 (2008).
68. Pellicciari, R. *et al.* Back door modulation of the farnesoid X receptor: Design, synthesis, and biological evaluation of a series of side chain modified chenodeoxycholic acid derivatives. *J. Med. Chem.* **49**, 4208–4215 (2006).
69. Meyer, U., Costantino, G., Macchiarul, A. & Pellicciari, R. Is Antagonism of E/Z-Guggulsterone at the Farnesoid X Receptor Mediated by a Noncanonical Binding Site? A Molecular Modeling Study. *J. Med. Chem.* **48**, 6948–6955 (2005).
70. Jin, L. *et al.* The antiparasitic drug ivermectin is a novel FXR ligand that regulates metabolism. *Nat. Commun.* **4**, 1937 (2013).
71. Makishima, M. *et al.* Vitamin D Receptor As an Intestinal Bile Acid Sensor. *Science* **296**, 1313–1316 (2002).
72. Bernardes, A. *et al.* Molecular mechanism of peroxisome proliferator-activated receptor  $\alpha$  activation by WY14643: a new mode of ligand recognition and receptor stabilization. *J. Mol. Biol.* **425**, 2878–2893 (2013).
73. Souza, P. C. T. *et al.* Identification of a New Hormone-Binding Site on the Surface of Thyroid Hormone Receptor. *Mol. Endocrinol.* **28**, 534–545 (2014).
74. Belorusova, A. Y. *et al.* Structural insights into the molecular mechanism of vitamin D receptor activation by lithocholic acid involving a new mode of ligand recognition. *J. Med. Chem.* **57**, 4710–4719 (2014).
75. Chen, L. *et al.* Sulindac-derived RXR $\alpha$  modulators inhibit cancer cell growth by binding to a novel site. *Chem. Biol.* **21**, 596–607 (2014).
76. Estebanez-Perpina, E. *et al.* A surface on the androgen receptor that allosterically regulates coactivator binding. *Proc. Natl. Acad. Sci.* **104**, 16074–16079 (2007).
77. Lack, N. A. *et al.* Targeting the Binding Function 3 (BF3) Site of the Human Androgen Receptor through Virtual Screening. *J. Med. Chem.* **54**, 8563–8573 (2011).
78. Munuganti, R. S. N. *et al.* Identification of a potent antiandrogen that targets the BF $_3$  site of the androgen receptor and inhibits enzalutamide-resistant prostate cancer. *Chem. Biol.* **21**, 1476–1485 (2014).
79. Ban, F. *et al.* Discovery of 1 H-indole-2-carboxamides as novel inhibitors of the androgen receptor binding function 3 (BF3). *J. Med. Chem.* **57**, 6867–6872 (2014).
80. Lallous, N. *et al.* Targeting Binding Function-3 of the Androgen Receptor Blocks Its Co-Chaperone Interactions, Nuclear Translocation, and Activation. *Mol. Cancer Ther.* **15**, 2936–2945 (2016).
81. Rutz, S., Eidenschenk, C., Kiefer, J. R. & Ouyang, W. Post-translational regulation of ROR $\gamma$  - A therapeutic target for the modulation of interleukin-17-mediated responses in autoimmune diseases. *Cytokine Growth Factor Rev.* **30**, 1–17 (2016).
82. Jetten, A. M. Retinoid-related orphan receptors (RORs): critical roles in development, immunity, circadian rhythm, and cellular metabolism. *Nucl. Recept. Signal.* **7**, e003 (2009).
83. Zhang, Y., Luo, X., Wu, D. & Xu, Y. ROR nuclear receptors: structures, related diseases, and drug discovery. *Acta Pharmacol. Sin.* **36**, 71–87 (2015).
84. Solt, L. A. & Burris, T. P. Action of RORs and their ligands in (patho)physiology. *Trends Endocrinol. Metab.* **23**, 619–627 (2012).
85. Li, X. *et al.* Structural studies unravel the active conformation of apo ROR $\gamma$  NR and a common inverse agonism of two diverse classes of ROR $\gamma$  inhibitors. *J. Biol. Chem.* **292**, 11618–11630 (2017).
86. Kojetin, D. J. & Burris, T. P. REV-ERB and ROR nuclear receptors as drug targets. *Nat. Rev. Drug Discov.* **13**, 197–216 (2014).
87. Ivanov, I. I. *et al.* The Orphan Nuclear Receptor ROR $\gamma$  Directs the Differentiation Program of Proinflammatory IL-17+ T Helper Cells. *Cell* **126**, 1121–1133 (2006).

88. Manel, N., Unutmaz, D. & Littman, D. R. The differentiation of human TH-17 cells requires TGF- $\beta$  and induction of the nuclear receptor ROR $\gamma$ t. *Nat. Immunol.* **9**, 641–649 (2008).
89. Kuwabara, T., Ishikawa, F., Kondo, M. & Kakiuchi, T. The Role of IL-17 and Related Cytokines in Inflammatory Autoimmune Diseases. *Mediators Inflamm.* **2017**, e3908061 (2017).
90. Miossec, P. & Kolls, J. K. Targeting IL-17 and Th17 cells in chronic inflammation. *Nat. Rev. Drug Discov.* **11**, 763–776 (2012).
91. Yang, J., Sundrud, M. S., Skepner, J. & Yamagata, T. Targeting Th17 cells in autoimmune diseases. *Trends Pharmacol. Sci.* **35**, 493–500 (2014).
92. Xue, X. *et al.* Pharmacologic modulation of ROR $\gamma$ t translates to efficacy in preclinical and translational models of psoriasis and inflammatory arthritis. *Sci. Rep.* **6**, 1–17 (2016).
93. Burkett, P. R. & Kuchroo, V. K. IL-17 Blockade in Psoriasis. *Cell* **167**, 1669 (2016).
94. Lock, C. *et al.* Gene-microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis. *Nat. Med.* **8**, 500–508 (2002).
95. Withers, D. R. *et al.* Transient inhibition of ROR- $\gamma$ t therapeutically limits intestinal inflammation by reducing Th17 cells and preserving group 3 innate lymphoid cells. *Nat. Med.* **22**, 319–323 (2016).
96. Hueber, W. *et al.* Effects of AIN457, a fully human antibody to interleukin-17A, on psoriasis, rheumatoid arthritis, and uveitis. *Sci. Transl. Med.* **2**, 52ra72 (2010).
97. Mease, P. J. *et al.* Ixekizumab, an IL-17A specific monoclonal antibody, for the treatment of biologic-naïve patients with active psoriatic arthritis: results from the 24-week randomised, double-blind, placebo-controlled and active (adalimumab)-controlled period of t. *Ann. Rheum. Dis.* **76**, 79–87 (2017).
98. Isono, F., Fujita-Sato, S. & Ito, S. Inhibiting ROR $\gamma$ t/Th17 axis for autoimmune disorders. *Drug Discov. Today* **19**, 1205–1211 (2014).
99. Tabarkiewicz, J., Pogoda, K., Karczmarczyk, A., Pozarowski, P. & Giannopoulos, K. The Role of IL-17 and Th17 Lymphocytes in Autoimmune Diseases. *Arch. Immunol. Ther. Exp. (Warsz)*. **63**, 435–449 (2015).
100. Fauber, B. P. & Magnuson, S. Modulators of the nuclear receptor retinoic acid receptor-related orphan receptor- $\gamma$  (ROR $\gamma$  or ROR $\epsilon$ ). *J. Med. Chem.* **57**, 5871–5792 (2014).
101. Bronner, S. M., Zbieg, J. R. & Crawford, J. J. ROR $\gamma$  antagonists and inverse agonists: a patent review. *Expert Opin. Ther. Pat.* **27**, 101–112 (2017).
102. Cyr, P., Bronner, S. M. & Crawford, J. J. Recent progress on nuclear receptor ROR $\gamma$  modulators. *Bioorganic Med. Chem. Lett.* **26**, 4387–4393 (2016).
103. Pandya, V. B., Kumar, S., Sachchidanand, Sharma, R. & Desai, R. C. Combating Autoimmune Diseases With ROR $\gamma$  Inhibitors: Hits and Misses. *J. Med. Chem.* **61**, 10976–10995 (2018).
104. Sun, N., Guo, H. & Wang, Y. Retinoic acid receptor-related orphan receptor gamma-t (ROR $\gamma$ t) inhibitors in clinical development for the treatment of autoimmune diseases: a patent review (2016-present). *Expert Opin. Ther. Pat.* **29**, 663–674 (2019).
105. Kotoku, M. *et al.* Discovery of Second Generation ROR $\gamma$  Inhibitors Composed of an Azole Scaffold. *J. Med. Chem.* **62**, 2837–2842 (2019).
106. Duan, J. J.-W. *et al.* Structure-based Discovery of Phenyl (3-Phenylpyrrolidin-3-yl)sulfones as Selective, Orally Active ROR $\gamma$ t Inverse Agonists. *ACS Med. Chem. Lett.* **10**, 367–373 (2019).
107. Tanis, V. M. *et al.* 3-Substituted Quinolines as ROR $\gamma$ t Inverse Agonists. *Bioorg. Med. Chem. Lett.* **29**, 1463–1470 (2019).
108. Huh, J. R. *et al.* Identification of Potent and Selective Diphenylpropanamide ROR $\gamma$  Inhibitors. *ACS Med. Chem. Lett.* **4**, 79–84 (2013).
109. René, O. *et al.* Minor structural change to tertiary sulfonamide RORC ligands led to opposite mechanisms of action. *ACS Med. Chem. Lett.* **6**, 276–281 (2015).
110. Harcken, C. *et al.* Discovery of a Series of Pyrazinone ROR $\gamma$  Antagonists and Identification of the Clinical Candidate BI 730357. *ACS Med. Chem. Lett.* **12**, 143–154 (2021).
111. Wang, J. *et al.* ROR- $\gamma$  drives androgen receptor expression and represents a therapeutic target in castration-resistant prostate cancer. *Nat. Med.* **22**, 488–496 (2016).
112. Zhang, Y. *et al.* Discovery and Characterization of XY101, a Potent, Selective, and Orally Bioavailable ROR $\gamma$  Inverse Agonist for Treatment of Castration-Resistant Prostate Cancer. *J. Med. Chem.* **62**, 4716–4730 (2019).
113. Jones, S. A., Sutton, C. E., Cua, D. & Mills, K. H. G. Therapeutic potential of targeting IL-17. *Nature immunology* **13**, 1022–1025 (2012).
114. Strutzenberg, T. S. *et al.* HDX-MS reveals structural determinants for ROR $\gamma$  hyperactivation by synthetic agonists. *Elife* **8**, e47172 (2019).
115. Chang, M. R. *et al.* Synthetic ROR $\gamma$ t Agonists Enhance Protective Immunity. *ACS Chem. Biol.* **11**, 1012–1018 (2016).



116. Qiu, R. & Wang, Y. Retinoic acid receptor-related orphan receptor-gamma-t (RORyt) agonists as potential small molecule therapeutics for cancer immunotherapy. *J. Med. Chem.* **61**, 5794–5804 (2018).
117. Rui, Y., Honjo, T. & Chikuma, S. Programmed cell death 1 inhibits inflammatory helper T-cell development through controlling the innate immune response. *Proc. Natl. Acad. Sci.* **110**, 16073–16078 (2013).
118. Hu, X. *et al.* Novel synthetic RORy agonist compounds as a potential anti-tumor therapeutic approach. *J. Immunother. Cancer* **2**, 194 (2014).
119. Scheepstra, M. *et al.* Identification of an allosteric binding site for RORyt inhibition. *Nat. Commun.* **6**, e8833 (2015).
120. Fujita-Sato, S. *et al.* Structural basis of digoxin that antagonizes RORgamma t receptor activity and suppresses Th17 cell differentiation and interleukin (IL)-17 production. *J. Biol. Chem.* **286**, 31409–31417 (2011).
121. Yuan, C.-M. *et al.* Molecular dynamics simulations on RORyt: insights into its functional agonism and inverse agonism. *Acta Pharmacol. Sin.* **40**, 1480–1489 (2019).
122. Kallen, J. *et al.* Structural States of RORgamma: X-ray Elucidation of Molecular Mechanisms and Binding Interactions for Natural and Synthetic Compounds. *ChemMedChem* **12**, 1014–1021 (2017).
123. Huang, M., Bolin, S., Miller, H. & Ng, H. L. RORy Structural Plasticity and Druggability. *Int. J. Mol. Sci.* **21**, (2020).
124. Hu, X. *et al.* Sterol metabolism controls TH17 differentiation by generating endogenous RORy agonists. *Nat. Chem. Biol.* **11**, 141–147 (2015).
125. Santori, F. R. *et al.* Identification of natural RORy ligands that regulate the development of lymphoid cells. *Cell Metab.* **21**, 286–298 (2015).
126. Soroosh, P. *et al.* Oxysterols are agonist ligands of RORyt and drive Th17 cell differentiation. *Proc. Natl. Acad. Sci.* **111**, 12163–12168 (2014).
127. Weems, G. A. *et al.* Lyc-55716: A novel small-molecule RORy agonist immuno-oncology agent: Rationale for tumor selection and clinical evaluation of gastric and esophageal carcinoma in phase 2a expansion. *J. Clin. Oncol.* **36**, 67 (2018).
128. Huh, J. R. *et al.* Digoxin and its derivatives suppress TH17 cell differentiation by antagonizing RORyt activity. *Nature* **472**, 486–490 (2011).
129. Olsson, R. I. *et al.* Benzoxazepines Achieve Potent Suppression of IL-17 Release in Human T-Helper 17 (TH17) Cells through an Induced-Fit Binding Mode to the Nuclear Receptor RORy. *ChemMedChem* **11**, 207–216 (2016).
130. Kumar, N. *et al.* The benzenesulfoamide T0901317 is a novel retinoic acid receptor-related orphan receptor-alpha/gamma inverse agonist. *Mol. Pharmacol.* **77**, 228–236 (2010).
131. Kumar, N. *et al.* Identification of SR2211: A potent synthetic RORy-selective modulator. *ACS Chem. Biol.* **7**, 672–677 (2012).
132. Gege, C. Retinoid-related orphan receptor gamma t (RORyt) inhibitors from Vitae Pharmaceuticals (WO2015116904) and structure proposal for their Phase I candidate VTP-43742. *Expert Opin. Ther. Pat.* **26**, 737–744 (2016).
133. Japan Tobacco Inc. Amide compound and pharmaceutical application therefor. *PCT Int. Appl.* WO2012/147916A1 (2012).
134. Fauber, B. P. *et al.* Reduction in lipophilicity improved the solubility, plasma-protein binding, and permeability of tertiary sulfonamide RORc inverse agonists. *Bioorg. Med. Chem. Lett.* **24**, 3891–3897 (2014).
135. Karstens, W. F. J. *et al.* RORgammaT Inhibitors. *PCT Int. Appl.* WO2012/106995 (2012).
136. Leijten-van de Gevel, I. A. & Brunsvel, L. Delineation of the molecular determinants of the unique allosteric binding site of the orphan nuclear receptor RORyt. *J. Biol. Chem.* **295**, 9183–9191 (2020).
137. Fauber, B. P. *et al.* Discovery of imidazo[1,5-a]pyridines and -pyrimidines as potent and selective RORc inverse agonists. *Bioorganic Med. Chem. Lett.* **25**, 2907–2912 (2015).
138. Guo, Y. *et al.* Inhibition of RORyT Skews TCRα Gene Rearrangement and Limits T Cell Repertoire Diversity. *Cell Rep.* **17**, 3206–3218 (2016).
139. De Wit, J. *et al.* RORyt inhibitors suppress TH17 responses in inflammatory arthritis and inflammatory bowel disease. *J. Allergy Clin. Immunol.* **137**, 960–963 (2016).
140. Ouvry, G. *et al.* Discovery of phenoxyindazoles and phenylthioindazoles as RORy inverse agonists. *Bioorganic Med. Chem. Lett.* **26**, 5802–5808 (2016).
141. Zhang, H. *et al.* Discovery of N-(Indazol-3-yl)piperidine-4-carboxylic Acids as RORyt Allosteric Inhibitors for Autoimmune Diseases. *ACS Med. Chem. Lett.* **11**, 114–119 (2020).

- 142. Shaikh, N. S. *et al.* Discovery and pharmacological evaluation of indole derivatives as potent and selective ROR $\gamma$ t inverse agonist for multiple autoimmune conditions. *Bioorg. Med. Chem. Lett.* **29**, 2208–2217 (2019).
- 143. Chaudari, S. S. *et al.* Bicyclic Heterocyclic Compounds as ROR gamma Modulators. *PCT Int. Appl.* WO2015/008234 (2015).
- 144. Gege, C. Retinoid-related orphan receptor  $\gamma$  t modulators: comparison of Glenmark's me-too patent application (WO2015008234) with the originator application from Merck Sharp and Dohme (WO2012106995). *Expert Opin. Ther. Pat.* **25**, 1215–1221 (2015).
- 145. Liu, K. *et al.* Selective PPAR $\gamma$  modulators with improved pharmacological profiles. *Bioorg. Med. Chem. Lett.* **15**, 2437–2440 (2005).
- 146. de Vries, R. M. J. M., Meijer, F. A., Doveston, R. G. & Brunsveld, L. Elucidation of an Allosteric Mode of Action for a Thienopyrazole ROR $\gamma$ t Inverse Agonist. *ChemMedChem* **15**, 561–565 (2020).
- 147. Jiang, X. *et al.* A novel series of cysteine-dependent, allosteric inverse agonists of the nuclear receptor ROR $\gamma$ t. *Bioorg. Med. Chem. Lett.* **30**, e126967 (2020).
- 148. Xin, J. *et al.* Pyrimidine tricyclic enone derivatives for inhibition of RORgamma and other uses. *PCT Int. Appl.* WO2018/111315 (2018).

# Chapter 2

## Ligand-Based Design of Allosteric RORyt Inverse Agonists

### Abstract

RORyt is a nuclear receptor protein associated with the pathogenesis of autoimmune diseases. Allosteric inhibition of RORyt is conceptually new, unique for this specific nuclear receptor, and offers advantages over traditional orthosteric inhibition. In this chapter, a highly efficient *in silico*-guided approach is described that led to the discovery of novel allosteric RORyt inverse agonists with a distinct isoxazole chemotype. The most potent compound, **FM26**, displayed sub-micromolar inhibition in a coactivator recruitment assay and effectively reduced IL-17a mRNA production in EL4 cells, a marker of RORyt activity. The projected allosteric mode of action of **FM26** was confirmed by biochemical experiments and co-crystallization with the RORyt ligand binding domain. The isoxazole compounds showed promising pharmacokinetic properties comparable to other allosteric ligands, but with a more diverse chemotype. The efficient ligand-based design approach adopted demonstrates its versatility in generating chemical diversity for allosteric targeting of RORyt.

**This chapter has been published as:** F.A. Meijer<sup>#</sup>, R.G. Doveston<sup>#</sup>, R.M.J.M. de Vries, G.M. Vos, A.A.A. Vos, S. Leysen, M. Scheepstra, C. Ottmann, L.-G. Milroy & L. Brunsveld. Ligand-Based Design of Allosteric RORyt Inverse Agonists. *J. Med. Chem.* 63, 241-259 (2020).

<sup>#</sup> These authors contributed equally

**A patent application has been granted for a selection of ligands shown in this chapter:** L. Brunsveld, R.G. Doveston, S. Leysen, L.-G. Milroy, F.A. Meijer, M. Scheepstra & C. Ottmann. Substituted Heterocyclic Compounds and their Use as RORyt Inhibitors. *PCT Int. Appl.* WO2020/149740 (2020).

## Introduction

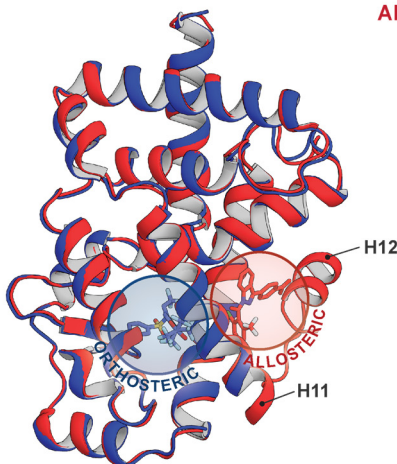
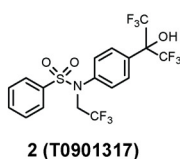
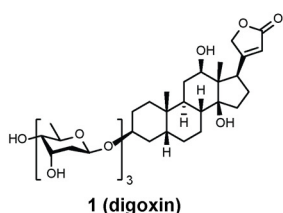
The nuclear receptor (NR) ROR $\gamma$ t has emerged as an important therapeutic target in recent years because of its essential role in both cancer and autoimmune disease. Inhibition of ROR $\gamma$ t is a promising therapeutic strategy for the treatment of prostate cancer because it stimulates androgen receptor (AR) gene transcription.<sup>1,2</sup> However, ROR $\gamma$ t is most prominently targeted for inhibition because of its essential role in promoting T helper 17 (Th17) cell differentiation.<sup>3–5</sup> Th17 cells produce the cytokine IL-17a which is strongly implicated in the pathogenesis of autoimmune diseases<sup>6</sup> such as psoriasis,<sup>7</sup> multiple sclerosis<sup>8</sup> and inflammatory bowel disease.<sup>9</sup> Disrupting the Th17/IL-17a pathway using IL-17a monoclonal antibodies (mAb) is a successful therapeutic strategy, with three mAbs approved for the treatment of plaque psoriasis: secukinumab (Cosentyx),<sup>10</sup> brodalumab (Siliq),<sup>11</sup> and ixekizumab (Taltz).<sup>12</sup> Inhibition of ROR $\gamma$ t with small molecules to disrupt the Th17/IL-17a pathway has been the focus of much research in recent years,<sup>13–20</sup> with several compounds progressed to clinical trials.<sup>14</sup>

ROR $\gamma$ t contains a hydrophobic ligand binding pocket located within the ligand binding domain (LBD) that is highly conserved across the NR family.<sup>21</sup> However, its transcriptional activity is not dependent on ligand binding because the apo protein retains the C-terminal helix 12 (H12) in a conformational state that allows for partial recruitment of coactivator proteins.<sup>22,23</sup> Although formally an orphan receptor with no proven endogenous ligands, ROR $\gamma$ t is responsive to the binding of naturally occurring cholesterol derivatives. Hydroxy-cholesterols have been shown to be effective agonists that stabilize H12 in such a way to further promote coactivator binding.<sup>24</sup> In contrast, digoxin (**1**, Figure 2.1) is an inverse agonist that stabilizes H12 in a conformation that is unsuitable for coactivator binding but promotes corepressor binding, thus leading to diminished gene transcription.<sup>25</sup> Numerous synthetic inverse agonists are also known, including **To901317** (**2**, Figure 2.1).<sup>26</sup> In all these cases, the ligands target the same orthosteric ligand binding pocket (Figure 2.1).

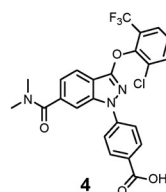
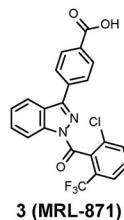
NR orthosteric ligand binding pockets are the target for numerous highly effective drug molecules.<sup>27</sup> Nevertheless, the highly conserved nature of this pocket across the NR family has led to issues associated with selectivity and mutation-induced resistance. Furthermore, dosing levels must be appropriate to compete with endogenous ligands. Molecules that target allosteric binding sites on NRs could circumvent such problems, for example because of the chemical uniqueness of the pocket and the absence of a competitive endogenous ligand.<sup>28–30</sup> Such allosteric compounds are therefore extremely valuable for both drug discovery and chemical biology applications.<sup>28–30</sup> The discovery that the potent ROR $\gamma$ t inverse agonists

**MRL-871** (**3**)<sup>31</sup> and later **4**<sup>32</sup> (Figure 2.1) target a previously unreported allosteric binding site within the RORyt LBD was therefore highly significant. These ligands were observed to directly interact with the activation function loop between H11 and H12 (AF-2 domain), thus forcing H12 to adopt an unusual conformation that prevents coactivator recruitment (Figure 2.1).<sup>31</sup>

#### Orthosteric Inverse Agonists



#### Allosteric Inverse Agonists



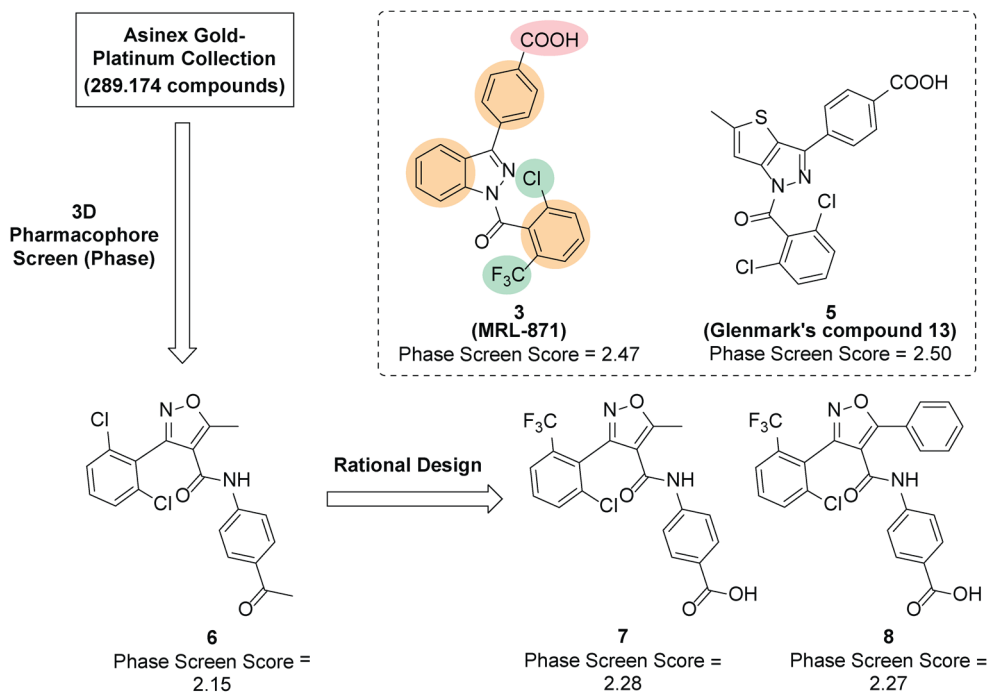
**Figure 2.1 | RORyt ligand binding domain (LBD) and a selection of orthosteric and allosteric ligands.** Orthosteric and allosteric RORyt ligand binding sites are shown as an overlay of the co-crystal structures of RORyt in complex with orthosteric inverse agonist **2** (To901317) (blue, PDB: 4NB6) and with allosteric inverse agonist **3** (MRL-871) (red, PDB: 4YPQ). The structures of the orthosteric inverse agonist **1** (digoxin) and allosteric inverse agonist **4** are also shown.

Allosteric modulation of RORyt has enormous potential as a novel therapeutic strategy; but, the examples of ligands that unambiguously target the allosteric pocket have been limited to compounds based on closely related chemotypes containing indazole or imidazopyridine cores.<sup>28</sup> As an example, indazoles **3** and **4** display promising *in vivo* activity,<sup>33,34</sup> but challenges remain, such as PPAR $\gamma$  cross-activity and pharmacokinetic (PK) profiles, for which novel chemotypes are needed.<sup>15</sup> In order to better exploit the strategy of allosteric modulation for therapeutic purposes, there is thus an urgent need to identify novel chemotypes targeting the allosteric site. This chapter reports on the design, synthesis and evaluation of a novel class of RORyt allosteric inverse agonists. The novel chemotype, discovered by *in silico*-guided pharmacophore screening and optimization, is based on a trisubstituted isoxazole core that, following efficient optimization of two substituents, led to the discovery of a sub-micromolar

inverse agonist. Protein X-ray crystallography and biophysical data unambiguously proved the designed allosteric mode of action. The compounds effectively inhibited cellular IL-17a expression and thus constitute valuable leads in the development of treatments for autoimmune diseases. To the best of our knowledge, our highly efficient *in silico*-guided approach is the first example of a medicinal chemistry program to overtly identify and develop a novel chemotype that targets the ROR $\gamma$ t allosteric site.

### ***In silico* pharmacophore screen**

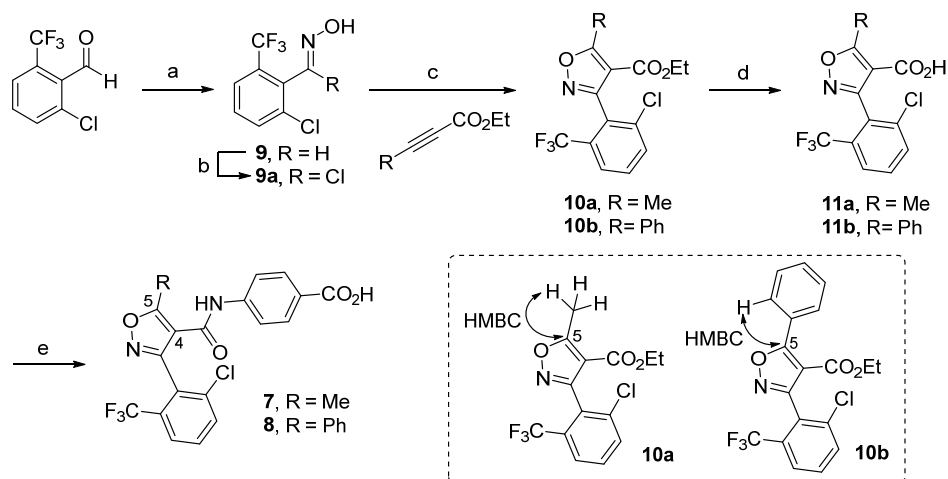
In order to identify novel chemotypes for chemical optimization, we used the crystal structure of the ROR $\gamma$ t LBD in complex with **3** as the basis for an *in silico* 3D pharmacophore screen against virtual compound libraries. An analogous scaffold hopping approach had been used previously by Glenmark Pharmaceuticals to identify similar scaffolds to **3** (MRL-871) such as the potent allosteric inverse agonist thienopyrazole **5** (Glenmark's **compound 13**) (Figure 2.2).<sup>35,36</sup> We created a 3D pharmacophore hypothesis based on the crystal structure of **3** bound to the allosteric pocket, using Phase (Schrödinger 2017-2).<sup>37,38</sup> Six structural features of **3** known to be important for activity were incorporated into the hypothesis: the three six-membered aromatic rings, an anionic group and two hydrophobic substituents (Figure 2.2). This hypothesis was used to interrogate a virtual library of 289,174 compounds from the Asinex Gold-Platinum collection of drug-like molecules.<sup>39</sup> Compounds matching at least four out of the six pharmacophore features were deemed to be a good hit. These were ranked using the 'Phase Screen Score' with higher scores indicating a better alignment with the hypothesis. The Phase Screen Scores for **3** and **5** were used as contextual references. The four highest ranking hit structures were all found to be based around the same trisubstituted isoxazole scaffold with **6** returned as the best match (Figure 2.2). This same scaffold was present in 13 of the top 30 hits. However, in each case we noted that only four out of six pharmacophore features were matched. Therefore, we designed two virtual ligands, **7** and **8**, that incorporated five and six of the features, respectively. As expected this led to improved Phase Screen Scores (Figure 2.2) and these compounds were therefore selected as initial targets for experimental investigation.



**Figure 2.2 | 3D Pharmacophore screening identifies a compound class with a novel isoxazole-based chemotype for experimental evaluation.** The structural features of **3** incorporated into the pharmacophore hypothesis are indicated: orange = aromatic rings, green = hydrophobic groups, red = anionic group.

### Exploratory structure activity relationship study

Isoxazoles **7** and **8** were synthesized via a 1,3-dipolar cycloaddition between a nitrile oxide (generated *in situ* from the oxime chloride **9a**) and a commercially available alkyne (Scheme 2.1).<sup>40</sup> The regiochemistry of the resulting trisubstituted isoxazole esters **10a** and **10b** was confirmed by 2D-NMR experiments (key HMBC correlations are highlighted in Scheme 2.1). Ester hydrolysis followed by amide coupling of *tert*-butyl-4-amino benzoate via the respective acid chloride, and finally deprotection of the *tert*-butyl ester furnished the target compounds in an efficient manner (Scheme 2.1).



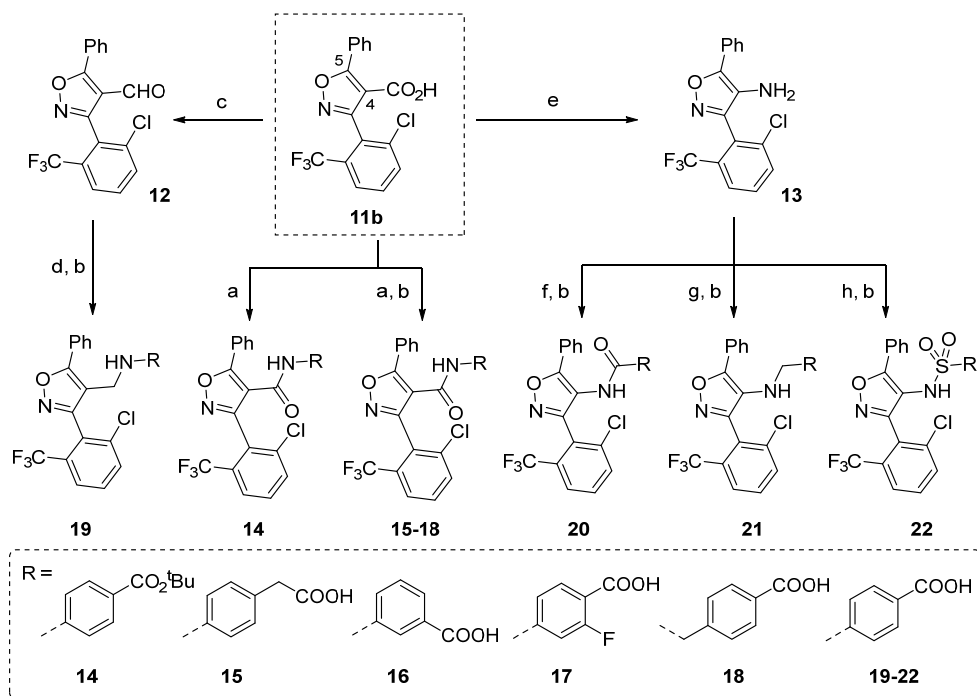
**Scheme 2.1 | Synthesis of trisubstituted isoxazoles **7** and **8**.** Reagents and conditions: (a)  $\text{NH}_2\text{OH}\cdot\text{HCl}$ ,  $\text{NaOH}$  (aq), EtOH, rt, 18 h, 83%; (b) NCS, DMF, 60 °C, 18 h, 86%; (c) alkyne,  $\text{NEt}_3$ , THF, 80 °C, 4 h, 69% (**10a**), 80% (**10b**); (d) LiOH, EtOH,  $\text{H}_2\text{O}$ , 70 °C, 8 h, 84% (**11a**), 95% (**11b**); (e) i)  $\text{SOCl}_2$ , 50 °C, 2 h; ii) *tert*-butyl-4-amino benzoate,  $\text{NEt}_3$ ,  $\text{CH}_2\text{Cl}_2$ , 45 °C, 6 h; iii) TFA,  $\text{CH}_2\text{Cl}_2$ , rt, 18 h, 42% (**7**), 69% (**8**). Key HMBC correlations used to confirm the regiochemistry of **10a** and **10b** are shown (bottom right). The  $^{13}\text{C}$ -NMR signals for the C-5 carbons are distinctively downfield at 175 and 173 ppm, respectively.

To determine if the compounds showed a functional response in terms of RORyt affinity for a coactivator, **7** and **8** were tested in a time-resolved fluorescence resonance energy transfer (TR-FRET) coactivator recruitment assay, based on FRET pairing between an anti-His terbium cryptate donor and a d2-labelled coactivator peptide (see Figure 2.3C, left).<sup>31</sup> Remarkably, both compounds inhibited coactivator recruitment in a dose-dependent manner. The phenyl derivative **8** was found to be significantly more potent than the methyl derivative **7**, with half-maximum inhibitory concentration ( $\text{IC}_{50}$ ) values of  $54 \pm 3 \mu\text{M}$  for **8** compared to  $>100 \mu\text{M}$  for **7**. In line with previous reports, **3** and **5** were determined to be significantly more potent with an  $\text{IC}_{50}$  of  $7.8 \pm 0.5 \text{ nM}$  and  $430 \pm 60 \text{ nM}$ , respectively (Table 2.1).

In view of these highly promising TR-FRET results, with the *in silico*-derived compounds around the trisubstituted isoxazole scaffold already showing activity, phenyl isoxazole **8** was selected as the focus of a subsequent structure activity relationship (SAR) study focusing on the isoxazole C-4 position. As such, a small library of eleven derivatives was synthesized using carboxylic acid **11b** as the cornerstone intermediate (Scheme 2.2) and evaluated using the TR-FRET coactivator recruitment assay (Table 2.1). While limited in size, this SAR study indicated

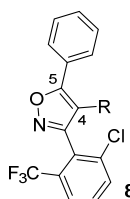


that a benzoic acid-containing substituent at the C-4 position was essential for potency: examples bearing no C-4 substitution (**11b**), a *para*-benzoate (**14**) or a methylene carboxylic acid (**15**) showed much reduced potency compared to the initial hit. Moving the acid moiety to the *meta*-position (**16**) or adding a *meta*-fluoro substituent (**17**) somewhat lowered the activity. However, the insertion of a single methylene unit between the amide and benzoic acid moieties (**18**) led to a six-fold increase in potency compared to the initial hit. The corresponding amine (**19**) displayed similar activity. Finally, reversing the relative positions of carbonyl and nitrogen components of the amide bond (**20-22**) did not result in a corresponding increase in potency.

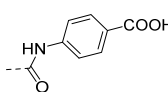

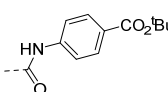
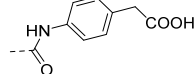
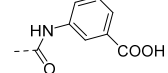
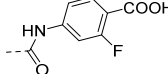
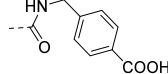
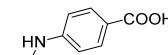
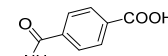
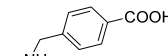
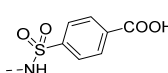


**Scheme 2.2 | Synthesis of C-4 isoxazole derivatives.** Reagents and conditions: (a) i)  $\text{SOCl}_2$ ,  $50^\circ\text{C}$ , 2 h; ii)  $\text{NH}_2\text{R}$ ,  $\text{NEt}_3$ ,  $\text{CH}_2\text{Cl}_2$ ,  $45^\circ\text{C}$ , 6 h, 27-87%; (b)  $\text{LiOH}$ ,  $\text{MeOH}$ ,  $\text{H}_2\text{O}$ ,  $70^\circ\text{C}$ , 8 h, 43-99%; (c) (i)  $\text{SOCl}_2$ ,  $50^\circ\text{C}$ , 2 h; ii)  $\text{MeNH}(\text{OMe})$ ,  $\text{NEt}_3$ ,  $\text{CH}_2\text{Cl}_2$ , rt, 6 h; iii)  $\text{LiAlH}_4$ ,  $\text{THF}$ ,  $0^\circ\text{C}$ , 30 min, 65%; (d) i) Ethyl-4-aminobenzoate,  $\text{AcOH}$ ,  $\text{MeOH}$ , reflux, 24 h; ii)  $\text{NaCNBH}_3$ ,  $\text{MeOH}$ , reflux, 12 h, 31%; (e) i)  $\text{DPPA}$ ,  $t\text{-BuOH}$ ,  $85^\circ\text{C}$ , 18 h; ii)  $\text{TFA}$ ,  $\text{CH}_2\text{Cl}_2$ , rt, 8 h, 59%; (f) i) Mono-methyl terephthalate,  $\text{SOCl}_2$ ,  $50^\circ\text{C}$ , 2 h; ii) **13**,  $\text{NEt}_3$ ,  $\text{CH}_2\text{Cl}_2$ , 76%; (g) i) Methyl-4-formyl benzoate,  $\text{AcOH}$ ,  $\text{MeOH}$ , reflux, 24 h; ii)  $\text{NaCNBH}_3$ ,  $\text{MeOH}$ , reflux, 18 h, 43%; (h) Methyl-4-(chlorosulfonyl)benzoate, pyridine,  $60^\circ\text{C}$ , 24 h, 71%.

**Table 2.1 | Structure activity relationship studies around the C-4 isoxazole position.** TR-FRET IC<sub>50</sub> values (μM) from coactivator recruitment assays and respective Glide docking scores are shown. Abbreviations: n.d., not determined. TR-FRET data are representative of three independent experiments (recorded in triplicate) and are presented as mean ± SD.



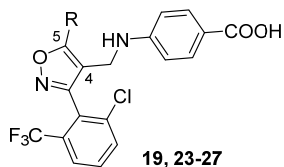
**8, 11b, 14-22**





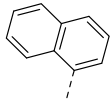
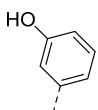
Compound	R (C-4 substituent)	IC <sub>50</sub> (μM)	Glide Score
3	—	0.0078 ± 0.0005	−14.58
5	—	0.43 ± 0.06	−13.11
7	—	> 100	−13.37
8		54 ± 3	−14.18
11b		> 100	−10.13
14		> 100	n.d.
15		> 100	−13.72
16		74 ± 3	−13.00
17		91 ± 5	−14.31
18		8.8 ± 0.5	−12.02
19		9.6 ± 0.6	−14.01
20		> 100	−13.55
21		31 ± 1	−13.52
22		63 ± 4	−13.00

### ***In silico* docking directs secondary SAR study**

In order to further improve the potency of our compounds we next explored the SAR at the isoxazole C-5 position. For this, molecular docking (Glide, Schrödinger 2017-2)<sup>41,42</sup> was used to select - with an attention to synthetic resource - C-5 substituents that were optimal for allosteric binding and therefore activity. For the study, a single C-4 substituent, the amine of compound **19**, was chosen based on its experimental activity and *in silico* docking score (Table 2.1). A virtual library of 84 C-5 analogues was enumerated using the open-source ChemT software.<sup>43</sup> This library was docked against the allosteric site of RORyt as defined by the X-ray crystal structure of **3** in complex with the RORyt LBD.<sup>31</sup> A single docking pose was returned for each virtual ligand and these were ranked using the 'Glide Score', which is an empirical scoring function that approximates the ligand binding free energy (the more negative the values, the higher the expected potency).<sup>44</sup> We contextualized these scores by comparison to those of compounds with known activity. The results (summarized in Table 2.2, see Supporting Table S2.1 for full information) indicated that smaller heteroaromatic moieties at the C-5 position would improve allosteric binding of the isoxazole ligands relative to **19**, with heteroatoms at the 2-position predicted to be optimal e.g. furan **23** and thiophene **24** (Table 2.2). The introduction of a hydrogen bond donor on the ring (specifically at the 3-position) was predicted to be even more beneficial: docking poses indicated that an additional hydrogen bond interaction with the backbone of helix 4 might be possible (e.g. pyrrole **25**, Table 2.2, Supporting Figure S2.2A). Bulkier substituents were predicted to be detrimental for binding (e.g. naphthyl **26**). To explore the predicted effect of a hydrogen bond donating group further, we interrogated a designed subset of ligands in the same docking experiment (see Supporting Table S2.2). None of these ligands showed an improved Glide score compared to pyrrole **25**. However, we noted that 3-hydroxyl substitution of the C-5 phenyl ring (**27**) was predicted to significantly enhance binding relative to **19**. To validate our findings experimentally we selected a cross-section of five derivatives for synthesis (i.e. **23-27**).

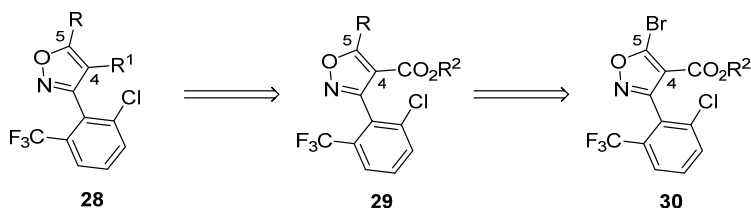
**Table 2.2 | Structure activity relationship studies around the C-5 isoxazole position.** TR-FRET  $IC_{50}$  values ( $\mu M$ ) from coactivator recruitment assays and respective Glide docking scores are shown. TR-FRET data are representative of three independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD.



Compound	R (C-5 substituent)	$IC_{50}$ ( $\mu M$ )	Glide Score
19		$9.6 \pm 0.6$	-14.01
23		$1.1 \pm 0.1$	-14.30
24		$1.8 \pm 0.2$	-14.18
25		$0.26 \pm 0.02$	-15.74
26		>100	-10.84
27		$6.6 \pm 0.5$	-14.60

## Docking-guided C-5 SAR study

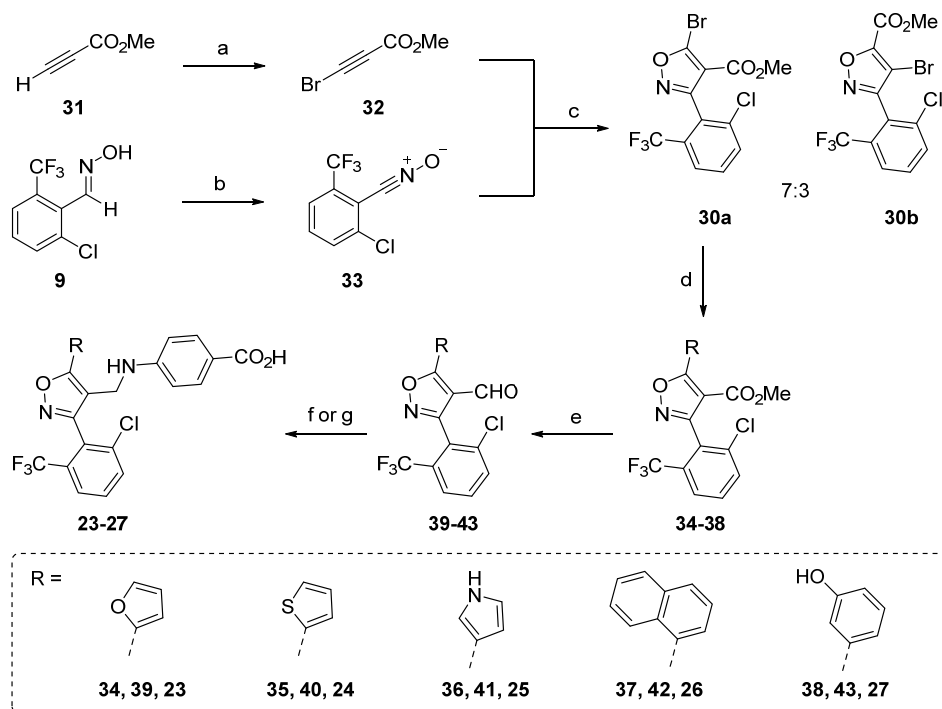
To expedite the synthesis of isoxazole analogues with various C-5 and C-4 substituents we re-designed our synthetic approach. It was envisaged that 5-bromo-4-carboxy isoxazole intermediate **30** (Scheme 2.3) would enable later stage introduction of the desired C-5 substituents via palladium-mediated cross-coupling chemistry. Introduction of C-4 substituents by manipulation of a carbonyl functional group (as developed previously) would then be possible (Scheme 2.3).



**Scheme 2.3 | Retrosynthetic analysis of trisubstituted isoxazole **28**, allowing for late-stage diversification.**

The intermediate **30** was prepared using analogous methodology to that used previously, as shown in Scheme 2.4. In this case it was necessary to isolate nitrile oxide **33** prior to 1,3-dipolar cycloaddition with alkynyl bromide **32**.<sup>45</sup> An efficient cycloaddition reaction led to an essentially quantitative recovery of a 7:3 mixture of 5-bromoisoxazole **30a** and 4-bromoisoxazole **30b** as determined by <sup>1</sup>H-NMR. This result was in close alignment with literature examples that indicated the 5-bromo isomer would predominate.<sup>45</sup> The mixture of regioisomers was purified by recrystallization from hot n-heptane resulting in the isolation of a 97:3 regiomer mixture (43% recovery) that was employed in subsequent steps. Assignment of the 5-bromoisoxazole **30a** as the major regioisomer was confirmed by 2D-NMR analysis of downstream products and by synthesis via an independent route (see Supporting Figure S2.1 and Supporting Scheme S2.1).

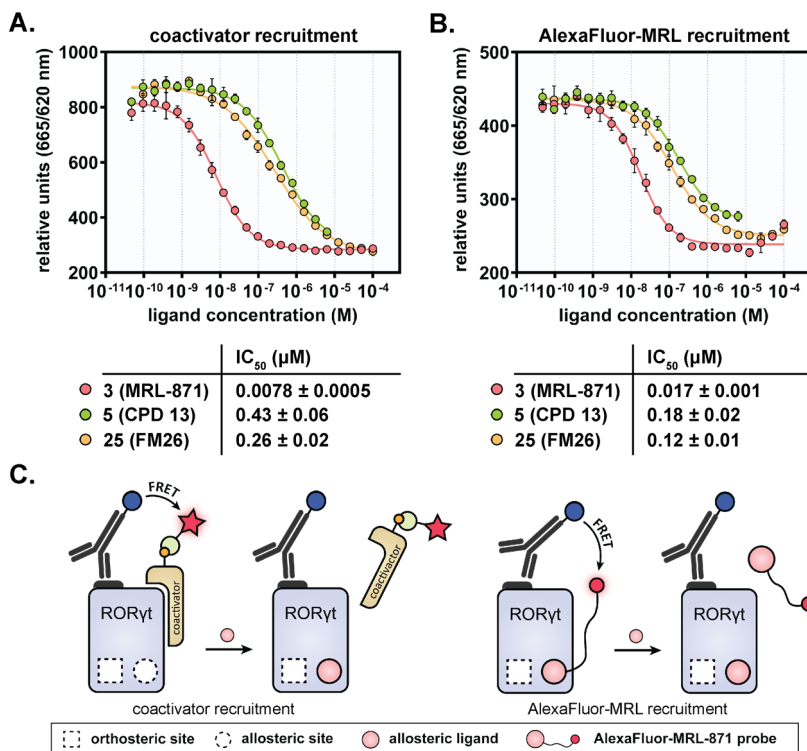
The desired substituents were introduced at the C-5 position by way of a Suzuki cross-coupling reaction with a pinacol boronate<sup>46</sup> (to give intermediates **34-38**) before conversion of the C-4 ester to an aldehyde (**39-43**) and reductive amination (Scheme 2.4). The lability of the 5-bromo group under the conditions for ester reduction dictated the order in which the synthesis steps were performed. Hydrolysis of the benzoic methyl ester to the free acid yielded the desired compounds **23-27**.



**Scheme 2.4 | Synthesis of isoxazole C-5 analogues 23-27.** Reagents and conditions: (a) NBS, AgNO<sub>3</sub>, Me<sub>2</sub>CO, rt, 20 h, 80%; (b) i) NCS, DMF, 60 °C, 18 h; ii) NEt<sub>3</sub>, THF, rt, 30 min, 85%; (c) THF, 80 °C, 4 h, 43% (30a); (d) RB(pin), Pd(dppf)Cl<sub>2</sub>, Cs<sub>2</sub>CO<sub>3</sub>, DME, 85 °C, 8 h, 39-58%; (e) i) LiAlH<sub>4</sub>, THF, 0 °C → rt, 2 h; ii) DMP, CH<sub>2</sub>Cl<sub>2</sub>, rt, 8 h, 51-96%; (f) i) *tert*-Butyl-4-amino benzoate, MeOH, AcOH, reflux, 24 h; ii) NaBH<sub>4</sub>, EtOH, 85 °C, 2-6 h, 16-24%; iii) TFA, CH<sub>2</sub>Cl<sub>2</sub>, rt, 18 h, **23**, **24**, **26**, 48-73%; (g) i) Methyl-4-amino benzoate, MeOH, AcOH, reflux, 24 h; ii) NaBH<sub>4</sub>, MeOH, reflux, 2-4 h, 16-19%; iii) LiOH, MeOH, H<sub>2</sub>O, 70 °C, 8 h, **25**, 57%, **27**, 99%.

In order to explore the SAR around the isoxazole C-5 position, the five analogues prepared in this second synthesis campaign were evaluated using the TR-FRET coactivator recruitment assay (Table 2.2). We were gratified to observe that furan **23** gave a 9-fold improvement in potency compared to phenyl **19**. By comparison, thiophene **24** was slightly less potent. Most significantly, pyrrole **25** (**FM26**), which also showed the most beneficial Glide Score, was 36-fold more potent than **19** and with an IC<sub>50</sub> value lower than the allosteric modulator **5** (IC<sub>50</sub> = 260 ± 20 nM for **25** vs. 430 ± 60 nM for **5**) (Figure 2.3A). These results were in excellent agreement with the *in silico* Glide Scores obtained (Table 2.2) and the improvements in potency are a notable step toward emulating the high potency of indazole **3** (Figure 2.3A). As predicted, the bulky naphthyl group of **26** was detrimental for activity such that no dose-

response curve could be fitted. The phenol derivative **27** showed a small improvement in potency compared to **19**. For this more bulky group at the C-5 position, compared to pyrrole **25**, the potential for additional hydrogen bonding, as indicated in the docking study, is thus not strongly expressed.



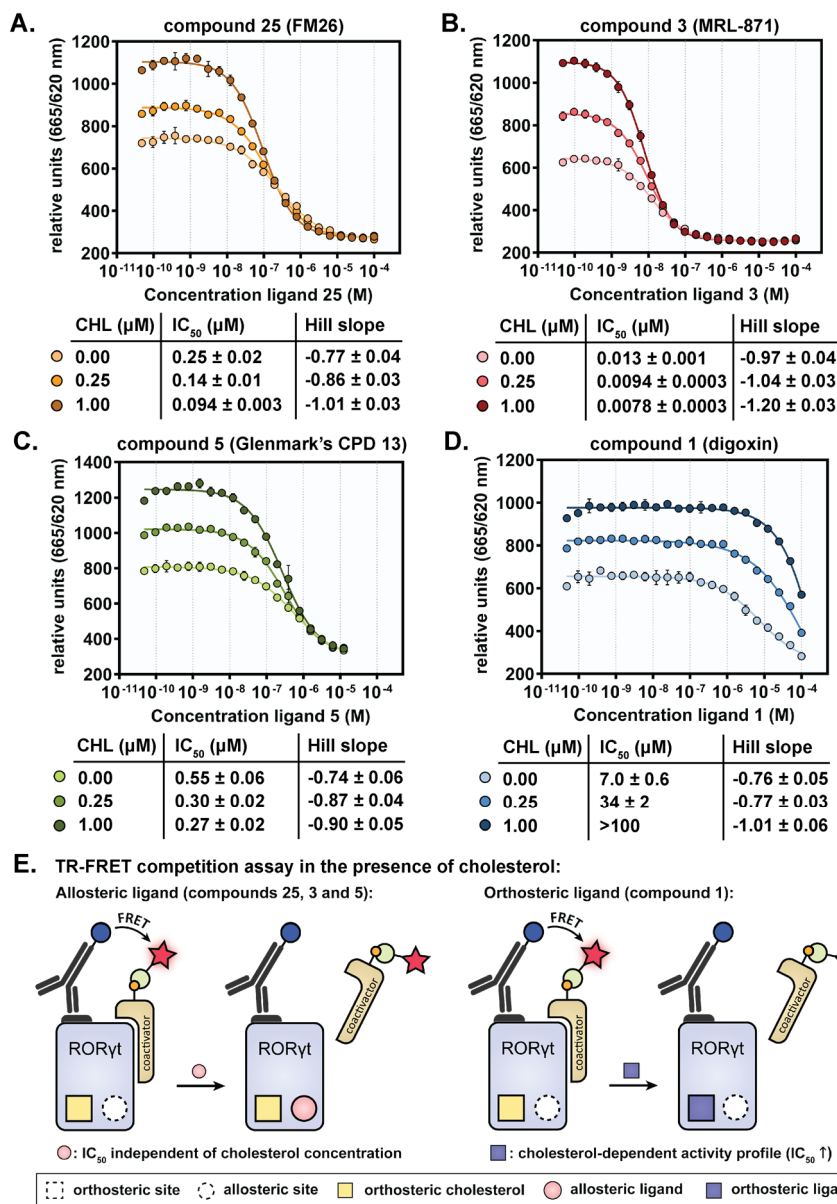
**Figure 2.3 | Biochemical RORyt assay data for ligands **3** (MRL-871), **5** (Glenmark's compound **13**) and **25** (FM26).** A/B) Dose-response curves from the TR-FRET coactivator recruitment assays (A) and AlexaFluor-MRL recruitment assays (B), including an overview of the IC<sub>50</sub> values. (The last data points for **5** are not shown because of solubility issues at high concentrations.) Data are representative of three independent experiments (recorded in triplicate) and are presented as mean ± SD. C) Left: Schematic representation of the TR-FRET coactivator recruitment assay. When RORyt is in its apo or agonist-bound state, the LBD binds to the coactivator, resulting in FRET pairing between an anti-His terbium cryptate donor (blue circle) and a d2-labelled streptavidin acceptor (red star). Inverse agonist binding results in coactivator displacement, thus a lower FRET pairing. Right: Schematic representation of the TR-FRET AlexaFluor-MRL recruitment assay. When the AlexaFluor647-labelled MRL-871 probe binds to the RORyt LBD, there is FRET pairing between the anti-His terbium cryptate donor and the probe. Allosteric inverse agonist binding results in probe displacement thus a lower FRET pairing.

## Mode of action studies

The allosteric mode of action for the novel lead compound **25** was first explored using a competitive TR-FRET coactivator recruitment assay against fixed concentrations of cholesterol (CHL) (an orthosteric agonist). If an allosteric ligand and cholesterol bind in a non-competitive manner at different sites on the ROR $\gamma$ t LBD, then the IC<sub>50</sub> of the allosteric ligand should be independent of cholesterol concentrations (Figure 2.4E, left). By contrast, ligands competing for the same binding site should show a cholesterol-dependent activity profile, whereby increasing cholesterol concentrations should result in a corresponding increase in IC<sub>50</sub> of the competing ligand (Figure 2.4E, right).<sup>31</sup> In our assay, increasing concentrations of **25** perturbed coactivator recruitment in the absence of cholesterol with an IC<sub>50</sub> value of  $250 \pm 20$  nM. Interestingly, increasing concentrations of cholesterol indeed resulted not in an increase but in a further decrease of the IC<sub>50</sub> value for **25** (**FM26**) with a concomitant sharpening of the Hill slope (Figure 2.4A). This result provides strong evidence not only for an allosteric mode of action, but also for a cooperative behavior between orthosteric and allosteric ligand binding. The same profile was observed for **5** (Glenmark's **compound 13**) (Figure 2.4C), providing evidence that this compound also modulates ROR $\gamma$ t activity in an allosteric fashion. Indazole **3** (**MRL-871**) also exhibited this behavior (Figure 2.4B). By comparison, the IC<sub>50</sub> values for the orthosteric inverse agonist **1** (digoxin) increased as the concentration of cholesterol increased, showing that digoxin and cholesterol compete for binding in the orthosteric pocket (Figure 2.4D). Collectively, our competitive assay data provide strong evidence that **25** functions as an allosteric inverse agonist.

To further confirm the allosteric mode of action for **25** on ROR $\gamma$ t we used an orthogonal assay to directly probe for allosteric ligand binding, as opposed to measuring indirect effects on coactivator recruitment. This assay used the previously described AlexaFluor647-labelled MRL-871 derivative (Supporting Figure S2.3),<sup>31</sup> which upon binding to ROR $\gamma$ t shows FRET pairing with an anti-His terbium cryptate donor on the protein (see Figure 2.3C, right).<sup>32</sup> The results of this experiment (Figure 2.3B) indeed corroborated the data obtained from the competitive coactivator recruitment assay: the isoxazole **25** displaced the allosteric probe with an IC<sub>50</sub> value of  $120 \pm 10$  nM, which was lower than that of **5** (IC<sub>50</sub> =  $180 \pm 20$  nM). As expected, indazole **3** was highly potent (IC<sub>50</sub> =  $17 \pm 1$  nM).





**Figure 2.4 | Dose-response curves from the competitive TR-FRET coactivator recruitment assays on ROR $\gamma$ t.** Titration of compounds 25 (FM26) (A), 3 (MRL-871) (B), 5 (Glenmark's compound 13) (C) and 1 (digoxin) (D) to ROR $\gamma$ t in the presence of fixed concentrations of cholesterol (CHL) (0.00  $\mu$ M, 0.25  $\mu$ M and 1.00  $\mu$ M). The IC<sub>50</sub> and Hill slope values for the compounds are shown below. Data are representative of three independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD. E) Schematic representation of the competition TR-FRET assay, titrating an allosteric (left) or orthosteric ligand (right) in the presence of cholesterol. Allosteric ligand binding is independent of cholesterol concentrations, while orthosteric ligand binding shows a cholesterol-dependent activity profile (IC<sub>50</sub> values increase, due to competition for the same binding site).

Indazole **3** had previously been shown to be selective for ROR $\gamma$ t over other NRs (>100-fold), with only minor cross-activity on PPAR $\gamma$ .<sup>31</sup> To give an indication of the cross-reactivity of the isoxazole series on PPAR $\gamma$ , a TR-FRET coactivator recruitment assay was performed with compounds **3**, **5** and isoxazoles **19** and **23-27**, measuring the competition with rosiglitazone<sup>47</sup> (an orthosteric PPAR $\gamma$  ligand). Compounds **3** and **5** showed IC<sub>50</sub> values of respectively 7.2  $\mu$ M and 15  $\mu$ M (Table 2.3), demonstrating some cross-reactivity on PPAR $\gamma$  at high ligand concentrations. Compound **25** and all other compounds of the isoxazole series resulted in only weak to no activity on PPAR $\gamma$  (IC<sub>50</sub> values >50  $\mu$ M), indicating that these compounds lead to favorably low PPAR $\gamma$  cross-reactivity. Thus, these data indicate that the novel class of allosteric isoxazole inverse agonists features potential as efficacious and selective ROR $\gamma$ t inverse agonists.

**Table 2.3 | IC<sub>50</sub> values observed in the competitive TR-FRET coactivator recruitment assays on PPAR $\gamma$ .** Data are representative of three independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD.

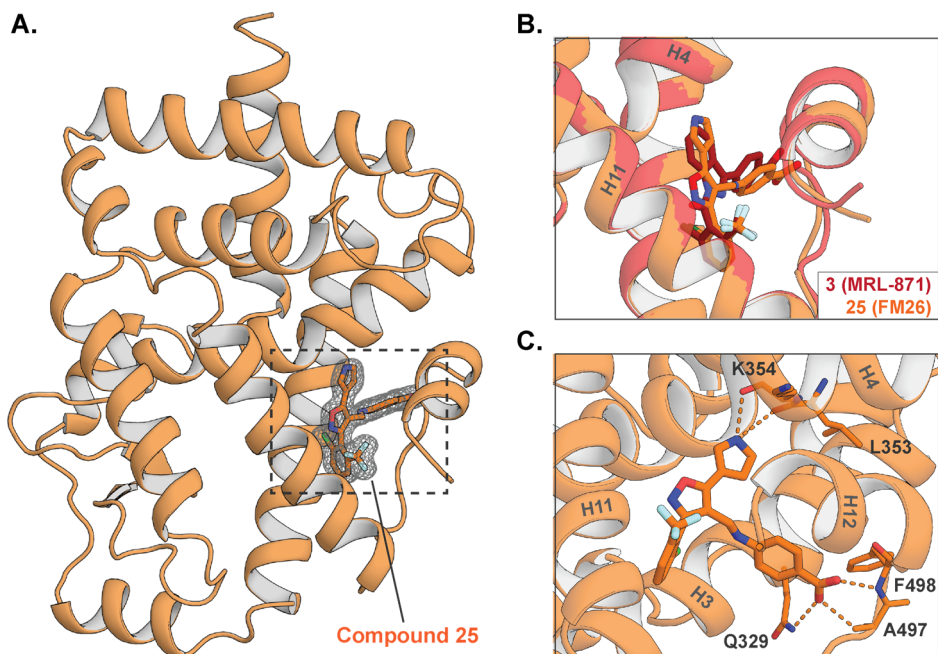
Compound	IC <sub>50</sub> ( $\mu$ M) PPAR $\gamma$
<b>3</b>	7.2 $\pm$ 0.8
<b>5</b>	15 $\pm$ 1
<b>19</b>	79 $\pm$ 6
<b>23</b>	> 100
<b>24</b>	> 100
<b>25</b>	99 $\pm$ 6
<b>26</b>	> 100
<b>27</b>	> 100

## Crystallography

Co-crystallization studies were performed for the most potent isoxazole **25** with the ROR $\gamma$ t LBD, to provide molecular insights in the ligand-receptor interaction. Crystals grew in a *P6<sub>1</sub>22* space group and diffracted to a resolution of 1.61 Å (Supporting Table S2.3). In the experimental electron density map, clear density for compound **25** was observed in the allosteric site, formed by helices 3, 4, 11 and 12 (Figure 2.5A, Supporting Figure S2.4). The compound binds to this allosteric site in a similar orientation as **3** (Figure 2.5B), as was predicted by our docking studies (Supporting Figure S2.2A). The 2,6-disubstituted phenyl ring, common to both **3** and **25**, is located in the exact same part of the binding pocket (Figure 2.5B). Moreover, hydrogen bond interactions between the carboxylic acid group and the

backbone amide hydrogen atoms of Ala497 and Phe498 (located on the loop between helix 11 and 12), as well as with the side chain of residue Gln329, are also evident in both structures (Figure 2.5B,C).

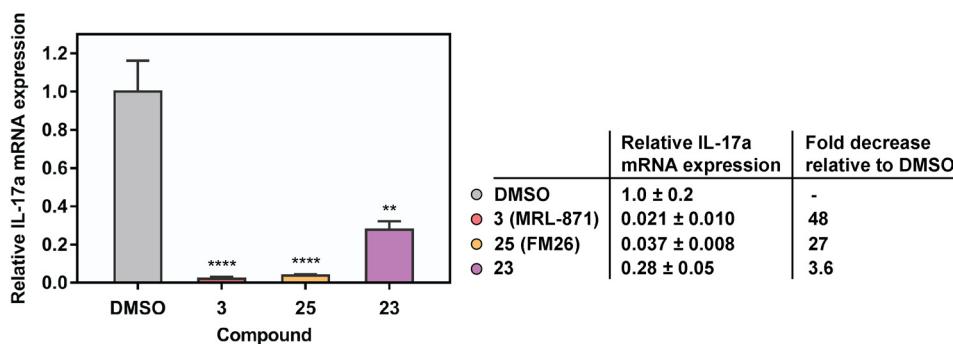
Unique to **25** is the pyrrole ring, which is oriented to allow a hydrogen bond interaction with the backbone carbonyls of residues Leu353 and Lys354 (Figure 2.5C). The isoxazole scaffold also allows a deeper penetration of this compound towards helix 4 of ROR $\gamma$ t, resulting in a modest shift of this helix. Additionally, the benzoic acid moiety of **25** shows a different orientation in the pocket compared to **3**, resulting in a change in the overall fold of the loop between helix 11 and helix 12 (Figure 2.5B). These structural data provide clear evidence for the allosteric binding of **25** to ROR $\gamma$ t in an orientation that was predicted with remarkable accuracy in the docking study (Supporting Figure S2.2B), but with specific additional molecular effects resulting from the novel isoxazole scaffold and pyrrole-based substitution pattern.



**Figure 2.5 | Co-crystal structure of the ROR $\gamma$ t LBD in complex with compound **25** (FM26) (PDB: 6SAL). A)** The tertiary structure of ROR $\gamma$ t bound to **25** (orange stick representation). The final  $2F_o - F_c$  electron density map of **25** is shown as an isomesh contoured at  $1\sigma$ . **B)** Overlay of the co-crystal structures of ROR $\gamma$ t in complex with **25** (orange sticks) and with **3** (MRL-871) (red sticks) (PDB: 5C4O). **C)** Enlarged view of the allosteric pocket of ROR $\gamma$ t showing the interactions between **25** and the protein. Hydrogen bond interactions are shown with orange dashes.

## Isoxazole 25 inhibits IL-17a expression in EL4 cells

EL4 is a murine lymphoblast cell line that constitutively expresses ROR $\gamma$ t. Because ROR $\gamma$ t promotes IL-17a production, an effective approach to determine the cellular activity of ROR $\gamma$ t inverse agonists is to measure the reduction in IL-17a mRNA expression levels by quantitative reverse transcriptase PCR (RT-PCR). To this end, EL4 cells were treated with 10  $\mu$ M of **3**, **25** and **23** for 24 h before IL-17a mRNA levels were measured (Figure 2.6). The most potent isoxazole in biochemical assays, **25**, significantly reduced IL-17a mRNA expression 27-fold whilst the weaker inverse agonist **23** showed a smaller reduction (3.6-fold) compared to the DMSO control. As expected, **3** led to the most significant decrease in IL-17a expression (48-fold), which was in line with previous reports.<sup>31</sup> These results demonstrate that the allosteric modulation of ROR $\gamma$ t by optimized trisubstituted isoxazoles leads to an effective cellular response, correlating with the biochemical protein binding data, and which is known to be beneficial for the treatment of autoimmune diseases.<sup>10–12</sup>



**Figure 2.6 | IL-17a mRNA expression levels in EL4 cells treated with ligands **3** (MRL-871), **25** (FM26) and **23** (10  $\mu$ M, 24 h) or DMSO and fold decrease of IL-17a expression relative to DMSO.** The level of IL-17a expression was normalized to that of GAPDH expression. Data are representative of three independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD. The relative gene expression was calculated by the  $2^{-\Delta\Delta C_t}$  (Livak) method using the DMSO control as a calibrator. Statistical analysis was performed using a one-way analysis of variance (ANOVA) compared against the DMSO control following Dunnett's *post hoc* test; \*\*  $P < 0.01$  and \*\*\*\*  $P < 0.0001$ .

## Absorption, distribution, metabolism and excretion (ADME) profile

To further assess the potential of **23** and isoxazole analogues such as **23** and **8**, we investigated the ADME profile of these compounds, and compared them to indazole **3** (Table 2.4). The isoxazole compounds showed favorable profiles compared to **3** in terms of chemical stability, solubility and permeability through artificial plasma membranes (PAMPA). A metabolic stability study with human liver microsomes indicated that the 4-methylamino isoxazoles **23** and **25** were more liable to phase I metabolism compared to indazole **3**, which showed good stability. Compounds **23** and **25** showed promising phase II stability. In blood plasma, whilst inferior to **3**, the stability of **23** and **25** was acceptable, although all these compounds showed high levels of binding to plasma proteins. Pleasingly, the 5-phenyl-4-amido isoxazole **8** showed a good ADME profile, with comparable microsomal stability to **3** and reduced plasma protein binding. This likely indicates that further optimization of the C-4 and C-5 isoxazole substituents has the potential to produce candidate molecules with desirable *in vivo* efficacy.

Table 2.4 | ADME properties for compounds **3**, **8**, **23** and **25**.

Cmpd	Chemical Stability (% remain)	Solubility ( $\mu$ M)	PAMPA (% Flux)	Microsomal Stability		Plasma stability (% remain)	Plasma protein binding (% bound)
				Phase I (Cl <sub>int</sub> , $\mu$ L/min/mg)	Phase II (% remain)		
<b>3</b>	81.0	390	23.7	– 1.2	47.1	100	99.9
<b>8</b>	95.4	490	60.1	– 0.1	100	99.9	97.8
<b>23</b>	100	392	50.4	43.2	92.8	86.5	100
<b>25</b>	95.3	411	33.6	20.7	69.8	85.9	99.9

## Conclusions

To summarize, we report the design, synthesis and early optimization of a novel class of ROR $\gamma$ t allosteric inverse agonists. The chemotype of the central aromatic ring system differs significantly from all the other fused bicyclic ring systems reported thus far. To identify this novel, more diverse, molecular scaffold we used the crystal structure of **3** bound to the ROR $\gamma$ t allosteric site as the basis for a 3D pharmacophore screen against a virtual compound library. Rational design steps led to the discovery of the *in silico* designed hit **8**, which already featured a modest inhibition of transcriptional coactivator recruitment to the ROR $\gamma$ t LBD and served as starting point for further optimization in a SAR campaign. A second and highly efficient

iteration of lead optimization was guided by *in silico* docking studies. Through the synthesis of just five derivatives (Table 2.2), this process delivered **25 (FM26)**, a sub-micromolar allosteric inverse agonist. It is highly noteworthy that there was a strong correlation between the Glide docking scores and the RORyt biochemical activity within this new class of isoxazole. Whereas screening approaches do not overtly identify allosteric ligands, our rational scaffold hopping approach is much more targeted, with less demand on experimental resource. Overall, the discovery workflow adopted, with a central role for structure-driven *in silico* screening and optimization, showed to be highly effective and might have wider applications in expediting NR allosteric drug discovery.

Competitive coactivator recruitment and allosteric probe displacement assays were used to confirm the allosteric mode of action, with concomitant cooperative RORyt binding with an orthosteric agonist. The co-crystal structure of **25** in complex with the RORyt LBD unequivocally proved the allosteric binding mode, via a similar mechanism to **3** and was impressively similar to the initially docked structure of **25** in RORyt. The crystal structure revealed a number of unique interactions and structural RORyt modifications that bring forward intriguing insights and new lines of exploration regarding RORyt allosteric ligand binding, selectivity, and affinity optimization. Furthermore, compound **25** was shown to significantly reduce IL-17a mRNA expression levels in EL4 cells and to have a promising ADME profile. These factors highlight the potential of this new isoxazole-based ligand class and overt targeting of the RORyt allosteric site to deliver effective treatments for autoimmune diseases.

## Acknowledgements

Richard Doveston is greatly acknowledged for his work on the chemical synthesis and *in silico* experiments, and for the valuable discussions. Gaël Vos and Alex Vos are thanked for the synthesis of some ligands and Rens de Vries is acknowledged for performing X-ray crystallography studies. Furthermore, Matthias Baumann and Bert Klebl from the Lead Discovery Center in Dortmund are kindly acknowledged for performing ADME measurements.

This chapter was reprinted (adapted) with permission from *J. Med. Chem.* 2020, 63, 241-259. Copyright 2020 American Chemical Society. Further permissions related to this material should be directed to the ACS.

## Experimental Section

**Pharmacophore screening.** The receptor-ligand complex structure (PDB: 4YPQ) was prepared using the Protein Preparation Wizard within Maestro (version 2017-2, Schrödinger LLC, New York, NY, USA) (default parameters). A 3D pharmacophore model for **3** bound to the allosteric pocket of RORyt LBD was created using Phase (version 2017-2, Schrödinger LLC, default hypothesis settings). Energy-minimized 3D ligand conformations for each molecule to be investigated were generated using the Ligand Preparation wizard within Maestro (default parameters). These were screened against the hypothesis whereby up to 50 ligand conformations were generated for each molecule. A hit was returned for compounds that matched four out of six pharmacophore features and these were ranked using the Phase Screen Score.

**Molecular docking studies.** The receptor-ligand structure (PDB: 4YPQ) was prepared as described above. A receptor grid centered on the bound ligand was created using Glide (version 2017-2, Schrödinger LLC). All parameters were kept as the default. Ligand libraries were either enumerated in SMILES format using the open-access Chem-T software, or generated manually. Ligands were prepared using the Ligand Preparation wizard as described above. Ligands were docked using Glide (version 2017-2, Schrödinger LLC) in standard precision mode with flexible ligand sampling. The predicted binding modes of all ligands were ranked according to their Glide Score (see Supporting Information for selected examples).

**General chemistry.** All non-aqueous reactions were performed under an argon atmosphere unless otherwise stated. Water-sensitive reactions were performed in oven-dried glassware, cooled under argon before use. Solvents were removed *in vacuo* using a Büchi rotary evaporator and a diaphragm pump. THF and CH<sub>2</sub>Cl<sub>2</sub> were dried and purified by means of a MBRAUN Solvent Purification System (MB-SPS-800). Anhydrous DMF was obtained in SureSeal bottles from Sigma-Aldrich. All other solvents used were of chromatography or analytical grade and supplied by Biosolve or Sigma-Aldrich. Commercially available starting materials were obtained from Sigma-Aldrich, Fluka, Acros, Alfa-Aesar or Fluorochem and were used without further purification unless stated. TLC was carried out on aluminium-backed silica (Merck silica gel 60 F254) plates supplied by Merck. Visualization of the plates was achieved using an ultraviolet lamp ( $\lambda_{\text{max}} = 254 \text{ nm}$ ), KMnO<sub>4</sub>, anisaldehyde or ninhydrin. Column chromatography was either performed manually using silica gel (60-63  $\mu\text{m}$  particle size) or using an automated Grace Reveleris X2 chromatograph with pre-packed silica columns supplied by Buchi/Grace (40  $\mu\text{m}$  particle size). LC-MS analysis was carried out with a system comprising a Thermo Fisher LCQ Fleet Ion Trap Mass Spectrometer and C18 Jupiter SuC4300A 150 x 2.0 mm column using a gradient of 5-100% MeCN in water (+ 0.1% HCOOH) over 15 min. The purity of the samples was assessed using a UV detector at 254 nm. Unless otherwise stated all final compounds were >95% pure as judged by HPLC.

GC-MS analysis was performed on a Phenomenex Zebron ZB-5MS 30 m x 0.25 mm x 0.25  $\mu$ m column (with a gradient of 80 °C for 1 min to 300 °C for 1 min, with a rate of 30 °C/min) in helium gas, connected to a GCMS-QP2010 Plus Quadrupole Mass Spectrometer. High resolution mass spectra (HRMS) were recorded using a Waters ACQUITY UPLC I-Class LC system coupled to a Xevo G2 Quadrupole Time of Flight (Q-TOF) mass spectrometer. Proton ( $^1\text{H}$ ) and carbon ( $^{13}\text{C}$ ) NMR spectral data were collected on a 400 MHz Bruker Cryomagnet or 400 MHz Varian Gemini. Chemical shifts ( $\delta$ ) are quoted in parts per million (ppm) and referenced to the residual solvent peak. Coupling constants ( $J$ ) are quoted in Hertz (Hz) and splitting patterns reported in an abbreviated manner: app. (apparent), s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet). Assignments were made with the aid of 2D COSY, HSQC and HMBC experiments.

### Synthetic procedures

**General Procedure for Ester Hydrolysis.** LiOH.H<sub>2</sub>O (5.0 eq) was added to a suspension of ester (1.0 eq) in a 4:1 mixture of MeOH/H<sub>2</sub>O (0.2 M). The reaction mixture was heated to 70 °C until TLC analysis indicated complete consumption of the starting material. MeOH was removed *in vacuo* and the resulting aqueous mixture was acidified to pH 3 using 10% v/v aqueous HCl and extracted with a 9:1 mixture of CH<sub>2</sub>Cl<sub>2</sub>/MeOH (5 x). The combined organic phase was dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo* to furnish a carboxylic acid which was purified as described.

**General Procedure for Amide Coupling.** Carboxylic acids (1.0 eq) were dissolved in SOCl<sub>2</sub> (50 eq) and heated to 50 °C for 2 h. The excess SOCl<sub>2</sub> was removed *in vacuo* to furnish an acid chloride intermediate that was immediately dissolved in CH<sub>2</sub>Cl<sub>2</sub> (0.1 M). To this was added NEt<sub>3</sub> (3.0 eq), the appropriate amine or aniline (1.5 eq) and DMAP (0.1 eq) and the reaction mixture was stirred at reflux for 18 h. The reaction mixture was diluted with saturated aqueous NH<sub>4</sub>Cl solution and extracted with ethyl acetate (3 x). The combined organic phase was washed with brine, dried over MgSO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO<sub>2</sub>) using the specified eluent.

**General Procedure for *tert*-Butyl Ester Deprotection.** Esters (1.0 eq) were treated with a 20% trifluoroacetic acid solution in CH<sub>2</sub>Cl<sub>2</sub> (0.2 M). The reaction mixture was stirred at the specified temperature for the specified amount of time and then concentrated *in vacuo*. The crude product was purified as indicated.

**General Procedure for Suzuki Coupling.** Under an inert atmosphere the pinacol boronate (2.0 eq), Cs<sub>2</sub>CO<sub>3</sub> (2.0 eq) and Pd(dppf)Cl<sub>2</sub> (0.1 eq) were added to a solution of bromide **30a** (1.0 eq) in de-gassed DME. The reaction mixture was heated at 85 °C for 8 h, cooled to room temperature, diluted with saturated aqueous NH<sub>4</sub>Cl and extracted with EtOAc (3 x). The combined organic phase was dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography using the specified eluent.

**General Procedure for Conversion of Esters to Aldehydes.** LiAlH<sub>4</sub> (1 M in THF, 1.0 eq) was added dropwise to a solution of ester (1.0 eq) in THF (0.2 M) at 0 °C. The reaction mixture was warmed to room temperature and stirred until TLC analysis indicated complete consumption of the starting material. The reaction mixture was cooled to 0 °C, quenched by the addition of saturated aqueous NH<sub>4</sub>Cl solution and extracted with EtOAc (3 x). The combined organic phase was dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo*. The resulting intermediate product was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (0.2 M). To this was added Dess-Martin Periodinane (1.5 eq) and the reaction mixture was stirred at room temperature until TLC analysis indicated complete consumption of the intermediate. The reaction mixture was quenched by the addition of 10% aqueous Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> solution and extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x). The combined organic



phase was washed with saturated aqueous NaHCO<sub>3</sub> and H<sub>2</sub>O, dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo* to furnish the title compound which was purified as described.

**General Procedure for Reductive Amination.** The chosen amine or aniline (1.0 eq) was added to a solution of the appropriate aldehyde (1.0 eq) and AcOH (0.1 eq) in MeOH or EtOH (0.25 M). The reaction mixture was heated at reflux for 24 h and then concentrated *in vacuo*. The intermediate imine was isolated by flash column chromatography using the specified eluent and then dissolved in MeOH or EtOH (0.2 M), cooled to 0 °C (ice) and treated with NaBH<sub>4</sub> (5.0 eq). The reaction mixture was held at the specified temperature until TLC analysis indicated complete consumption of the imine. The solvent was removed *in vacuo* and the mixture was dissolved in CH<sub>2</sub>Cl<sub>2</sub> and washed with water. The aqueous phase was further extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x), dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO<sub>2</sub>) using the specified eluent.

**4-(3-(2-chloro-6-(trifluoromethyl)phenyl)-5-methyl-1,2-oxazole-4-amido)benzoic acid (7).** According to the General Procedure for amide coupling carboxylic acid **11a** (60.0 mg, 0.199 mmol) was coupled with *tert*-butyl-4-amino benzoate. The crude product was purified by flash column chromatography, eluting with 20% EtOAc in *c*-hexane, to furnish the amide (46.0 mg, 48%). The intermediate product (43.0 mg, 0.089 mmol) was subject to *tert*-butyl ester deprotection (see General Procedure for *tert*-Butyl ester deprotection) and purified by trituration with Et<sub>2</sub>O to furnish **7** (33.0 mg, 87%) as a white solid. *R*<sub>f</sub> = 0.52 (9:1 CH<sub>2</sub>Cl<sub>2</sub>/MeOH); <sup>1</sup>H-NMR (400 MHz, MeOD): δ (ppm) 7.97 (2H, d, *J* = 8.8 Hz, benzoate H-2), 7.81 (2H, app. d, *J* = 8.0 Hz, Ar H-3 and Ar H-5), 7.67 (1H, t, *J* = 8.0 Hz, Ar H-4), 7.61 (2H, d, *J* = 8.8 Hz, benzoate H-3), 2.76 (3H, s, CH<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, MeOD): δ (ppm) 171.3 (C-5), 161.2 (CO<sub>2</sub>H), 159.5 (C-3), 143.8 (benzoate C-4), 137.4 (ArC-2), 134.4 (ArC-3), 132.7 (q, *J* = 30.9 Hz, ArC-6), 132.6 (ArC-4), 131.8 (benzoate C-2), 127.6 (ArC-1), 127.5 (benzoate C-1), 126.1 (q, *J* = 5.1 Hz, ArC-5), 120.6 (benzoate C-3), 116.1 (C-4), 12.8 (CH<sub>3</sub>). LC-MS (ESI): calc. for C<sub>19</sub>H<sub>13</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 425.04, observed: 425.17 (*R*<sub>t</sub> = 6.22 min). HRMS (ESI): calc. for C<sub>19</sub>H<sub>13</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 425.0516, observed: 425.0511.

**4-(3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazole-4-amido)benzoic acid (8).** *tert*-Butyl benzoate **14** (30.0 mg, 0.055 mmol) was deprotected according to the General Procedure for *tert*-Butyl ester deprotection. The crude product was purified by trituration with Et<sub>2</sub>O to furnish **8** (21.0 mg, 78%) as a white solid. *R*<sub>f</sub> = 0.55 (9:1 CH<sub>2</sub>Cl<sub>2</sub>/MeOH); <sup>1</sup>H-NMR (400 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) 7.97 (1H, d, *J* = 8.2 Hz, ArH-3 or ArH-5), 7.92 (1 H, d, *J* = 7.9 Hz, ArH-3 or ArH-5), 7.85-7.83 (4H, m, benzoate H-2, PhH-ortho), 7.78 (1H, app. t, *J* = 8.1 Hz, ArH-4), 7.58-7.54 (5H, m, benzoate H-3, PhH-meta, PhH-para); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) 167.0 (C-5), 166.8 (CO<sub>2</sub>H), 158.9 (CO<sub>2</sub>NH), 158.7 (C-3), 142.3 (benzoate C-4), 135.5 (ArC-2), 133.7 (ArC-3), 132.4 (PhC-quart.), 131.6 (ArC-4), 130.6 (ArC-6), 130.3 (benzoate C-2), 129.4 (PhC-ortho), 127.3 (PhC-meta), 126.0 (PhC-para), 125.9 (benzoate C-1), 125.3 (ArC-5 and ArC-1), 119.3 (benzoate C-3), 113.4 (C-4), quartet for CF<sub>3</sub> not observed. LC-MS (ESI): calc. for C<sub>24</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 487.06, observed: 487.17 (*R*<sub>t</sub> = 7.00 min). HRMS (ESI): calc. for C<sub>24</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 487.0672, observed: 487.0662.

**(E)-N-((2-chloro-6-(trifluoromethyl)phenyl)methylidene)hydroxylamine (9).** Hydroxylamine hydrochloride (3.95 g, 57.0 mmol, 1.2 eq) was suspended in EtOH (20 mL) and 10% w/v aqueous solution of NaOH (20 mL) was added such that the final pH of the resulting solution was < pH 9. 2-Chloro-6-(trifluoromethyl)benzaldehyde (9.88 g, 47.4 mmol, 1.0 eq) was then added as a solution in EtOH (20 mL) and the mixture was stirred at room temperature for 18 h. The reaction mixture was diluted with H<sub>2</sub>O and extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x). The combined organic phase was dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo* to furnish **9** (8.77 g, 83%) as a white solid which was used without further purification. *R*<sub>f</sub> = 0.45 (4:1 *c*-hexane/EtOAc); <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) 8.97 (1H, s, N=CH), 8.36 (1H, s, NOH), 7.67-7.64 (2H, m, H-3, H-5), 7.45 (1H, app. t, *J* = 8.0 Hz, H-4); <sup>13</sup>C-NMR (100 MHz,

$\text{CDCl}_3$ ):  $\delta$  (ppm) 145.3 (N=CH), 135.7 (C-2), 133.6 (C-3), 131.3 (q,  $J = 31.2$  Hz, C-6), 130.2 (C-4), 129.0 (C-1), 125.1 (q,  $J = 5.5$  Hz, C-5), 123.22 (q,  $J = 274.2$  Hz,  $\text{F}_3\text{C}$ ). LC-MS (ESI): calc. for  $\text{C}_8\text{H}_6\text{ClF}_3\text{NO}$   $[\text{M}+\text{H}]^+$ : 224.00, observed: 224.00 ( $R_t = 5.82$  min).

**(Z)-2-chloro-N-hydroxy-6-(trifluoromethyl)benzene-1-carbonimidoyl chloride (9a).** N-Chlorosuccinamide (5.22 g, 39.1 mmol, 1.0 eq) was added to a solution of hydroxylamine **9** (8.74 g, 39.1 mmol, 1.0 eq) in DMF (80 mL). The reaction mixture was stirred at 60 °C for 18 h, then diluted with  $\text{H}_2\text{O}$  (150 mL) and extracted with  $\text{Et}_2\text{O}$  (3 x). The combined organic phase was washed with  $\text{H}_2\text{O}$  (3 x) and brine, dried ( $\text{MgSO}_4$ ), filtered and concentrated *in vacuo* to furnish **9a** (9.10 g, 95% purity, 86%) which was used immediately in the next step without further purification.  $R_f = 0.42$  (4:1 c-hexane/ $\text{EtOAc}$ );  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 8.50 (1H, s, NOH), 7.68-7.66 (2H, m, H-3, H-5), 7.53 (1H, app. t,  $J = 8.1$  Hz, H-4).

**Ethyl 3-(2-chloro-6-(trifluoromethyl)phenyl)-5-methyl-1,2-oxazole-4-carboxylate (10a).**  $\text{NEt}_3$  (0.570 mL, 4.07 mmol, 2.5 eq) was added to a solution of imidoyl chloride **9a** (0.421 g, 1.63 mmol, 1.0 eq) in THF (2.5 mL). A white precipitate formed immediately after which was added ethyl 2-butyrate (0.190 mL, 1.63 mmol, 1.0 eq). The reaction mixture was heated to reflux for 4 h, filtered and concentrated *in vacuo*. The crude material was purified by automated flash column chromatography, eluting with a gradient of 2-5%  $\text{EtOAc}$  in c-hexane, to furnish **10a** (0.375 g, 69%) as a colourless oil.  $R_f = 0.22$  (9:1 c-hexane/ $\text{EtOAc}$ );  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 7.69 (1H, d,  $J = 8.0$  Hz, ArH-3 or ArH-5), 7.66 (1H, d,  $J = 8.0$  Hz, ArH-3 or ArH-5), 7.53 (1H, app. t,  $J = 8.0$  Hz, ArH-4); 4.07 (2H, 2x dq (overlapping),  $J = 12.7, 7.1$  Hz,  $\text{CO}_2\text{CH}_2\text{CH}_3$ ), 2.78 (3H, s,  $\text{CH}_3$ ), 0.98 (3H, t,  $J = 7.1$  Hz,  $\text{OCH}_2\text{CH}_3$ );  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 175.4 (C-5), 161.1 ( $\text{CO}_2\text{Et}$ ), 158.1 (C-3), 136.3 (ArC-2), 132.7 (ArC-3), 131.6 (q,  $J = 31.6$  Hz, ArC-6), 130.6 (ArC-4), 127.5 (ArC-1), 124.6 (q,  $J = 5.0$  Hz, ArC-5), 123.0 (q,  $J = 274.5$  Hz,  $\text{F}_3\text{C}$ ), 110.2 (C-4), 60.6 ( $\text{OCH}_2\text{CH}_3$ ), 13.66 ( $\text{CH}_3$ ), 13.3 ( $\text{OCH}_2\text{CH}_3$ ). LC-MS (ESI): calc. for  $\text{C}_{14}\text{H}_{12}\text{ClF}_3\text{NO}_3$   $[\text{M}+\text{H}]^+$ : 334.04, observed: 334.08 ( $R_t = 7.47$  min).

**Ethyl 3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazole-4-carboxylate (10b).**  $\text{NEt}_3$  (3.92 mL, 28.1 mmol, 2.5 eq) was added to a solution of imidoyl chloride **9a** (2.90 g, 11.3 mmol, 1.0 eq) in THF (17 mL). A white precipitate formed immediately after which was added ethyl 3-phenyl propionate (1.85 mL, 11.3 mmol, 1.0 eq). The reaction mixture was heated to reflux for 6 h, filtered and concentrated *in vacuo*. The crude material was purified by automated flash column chromatography, eluting with a gradient of 0-25%  $\text{EtOAc}$  in n-heptane, to furnish **10b** (3.59 g, 80%) as a white solid.  $R_f = 0.47$  (4:1 c-hexane/ $\text{EtOAc}$ );  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 8.16-8.13 (2H, m, PhH-ortho), 7.74 (1H, d,  $J = 7.9$  Hz, ArH-3 or ArH-5), 7.71 (1H, d,  $J = 8.2$  Hz, H-3 or H-5), 7.59-7.51 (4H, m, H-4, ArH), 4.06 (2H, dq (2 overlapping),  $J = 11.8, 7.1$  Hz,  $\text{CO}_2\text{CH}_2\text{CH}_3$ ), 0.88 (3H, t,  $J = 7.1$  Hz,  $\text{OCH}_2\text{CH}_3$ );  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 173.0 (C-5), 160.8 ( $\text{CO}_2\text{Et}$ ), 159.2 (C-3), 136.3 (ArC-2), 132.6 (ArC-3), 131.7 (PhC-quart.), 131.5 (q,  $J = 30.5$  Hz, ArC-6), 130.5 (ArC-4), 129.2 (PhC-ortho), 128.5 (PhC-meta), 127.8 (q,  $J = 1.8$  Hz, ArC-1), 126.3 (PhC-meta), 124.5 (q,  $J = 5.0$  Hz, ArC-5), 122.9 (q,  $J = 274.3$  Hz,  $\text{F}_3\text{C}$ ), 109.3 (C-4), 60.8 ( $\text{OCH}_2\text{CH}_3$ ), 13.3 ( $\text{OCH}_2\text{CH}_3$ ). LC-MS (ESI): calc. for  $\text{C}_{19}\text{H}_{14}\text{ClF}_3\text{NO}_3$   $[\text{M}+\text{H}]^+$ : 396.06, observed: 396.17 ( $R_t = 8.40$  min).

**3-(2-chloro-6-(trifluoromethyl)phenyl)-5-methyl-1,2-oxazole-4-carboxylic acid (11a).** According to the General Procedure for ester hydrolysis, ester **10a** (0.273 g, 0.820 mmol) was hydrolyzed in 8 h to furnish **11a** (0.210 g, 84%) as a white solid which required no further purification.  $R_f = 0.61$  ( $\text{EtOAc}$ );  $^1\text{H-NMR}$  (400 MHz,  $\text{DMSO-d}_6$ ):  $\delta$  (ppm) 13.09 (1H, br. s,  $\text{CO}_2\text{H}$ ), 7.96 (1H, d,  $J = 8.0$  Hz, ArH-3 or ArH-5), 7.90 (1H, d,  $J = 8.0$  Hz, ArH-3 or ArH-5), 7.77 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 2.76 (3H, s,  $\text{CH}_3$ );  $^{13}\text{C-NMR}$  (100 MHz,  $\text{DMSO-d}_6$ ):  $\delta$  (ppm) 175.3 (C-5), 161.9 ( $\text{CO}_2\text{H}$ ), 158.2 (C-3), 135.3 (ArC-2), 133.4 (ArC-3), 131.8 (ArC-4), 130.1 (q,  $J = 30.8$  Hz, ArC-6), 126.8 (ArC-1), 125.1 (q,  $J = 5.0$  Hz, ArC-5), 123.0 (q,  $J = 274.4$  Hz,

F<sub>3</sub>C), 110.5 (C-4), 12.8 (CH<sub>3</sub>). LC-MS (ESI): calc. for C<sub>12</sub>H<sub>6</sub>ClF<sub>3</sub>NO<sub>3</sub> [M-H]<sup>-</sup>: 304.01, observed: 304.17 (R<sub>t</sub> = 5.82 min).

**3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazole-4-carboxylic acid (11b).** According to the General Procedure for ester hydrolysis, ester **10b** (3.66 g, 9.26 mmol) was hydrolyzed in 8 h to furnish **11b** (3.25 g, 95%) as a white solid which required no further purification. *R*<sub>f</sub> = 0.35 (1:1 *c*-hexane/EtOAc); <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) 8.05–8.03 (2H, m, PhH-ortho), 7.71 (1H, d, *J* = 8.0 Hz, ArH-3 or ArH-5), 7.68 (1H, d, *J* = 8.1 Hz, H-3 or H-5), 7.60–7.50 (4H, m, H-4, ArH); <sup>13</sup>C-NMR (100 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) 171.8 (C-5), 161.5 (CO<sub>2</sub>H), 159.3 (C-3), 135.2 (ArC-2), 133.4 (ArC-3), 132.0 (PhC-quart.), 131.8 (ArC-4), 130.0 (q, *J* = 30.5 Hz, ArC-6), 129.0 (PhC-ortho), 128.7 (PhC-meta), 126.7 (ArC-1), 125.7 (PhC-meta), 125.2 (q, *J* = 5.0 Hz, ArC-5), 123.0 (q, *J* = 274.4 Hz, F<sub>3</sub>C), 99.4 (C-4). LC-MS (ESI): calc. for C<sub>17</sub>H<sub>10</sub>ClF<sub>3</sub>NO<sub>3</sub> [M+H]<sup>+</sup>: 368.02, observed: 368.08 (R<sub>t</sub> = 7.03 min). HRMS (ESI): calc. for C<sub>17</sub>H<sub>10</sub>ClF<sub>3</sub>NO<sub>3</sub> [M+H]<sup>+</sup>: 368.0301, observed: 368.0299.

**3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazole-4-carbaldehyde (12).** Carboxylic acid **11b** (2.0 g, 5.44 mmol, 1.0 eq) was dissolved in SOCl<sub>2</sub> (10.0 mL, 138 mmol, 25 eq) and heated to 60 °C for 2 h. Excess SOCl<sub>2</sub> was removed *in vacuo* and the intermediate acid chloride was immediately dissolved in CH<sub>2</sub>Cl<sub>2</sub> (20 mL) and cooled to 0 °C. To this was added NEt<sub>3</sub> (2.27 mL, 16.3 mmol, 3.0 eq) and *N,O*-dimethyl hydroxylamine hydrochloride (0.580 g, 5.98 mmol, 1.1 eq). The reaction mixture was allowed to warm to room temperature with stirring over 16 h before being quenched by the addition of saturated aqueous NH<sub>4</sub>Cl solution and extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x). The combined organic phase was dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo*. The resulting Weinreb amide was dissolved in THF (20 mL) and cooled to 0 °C. To this was added LiAlH<sub>4</sub> (1 M in THF, 2.72 mL, 2.72 mmol, 0.5 eq) and the reaction mixture was stirred at 0 °C for 1 h before being quenched by the addition of saturated aqueous NH<sub>4</sub>Cl solution and extracted with EtOAc (3 x). The combined organic phase was dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo* to furnish a crude product which was purified by automated flash column chromatography, eluting with a gradient of 10–30% EtOAc in *n*-heptane, to furnish **12** (1.25 g, 65%) as a white solid. *R*<sub>f</sub> = 0.65 (1:1 *c*-hexane/EtOAc); <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) 9.93 (1 H, s, CHO), 8.06–8.04 (2 H, m, PhH-ortho), 7.77 (1H, d, *J* = 7.9 Hz, ArH-3 or ArH-5), 7.74 (1H, d, *J* = 7.8 Hz, ArH-3 or ArH-5), 7.68–7.59 (4H, m, H-4, PhH); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) 182.8 (CHO), 174.6 (C-5), 160.8 (CO<sub>2</sub>Et), 158.6 (C-3), 136.3 (ArC-2), 133.2 (ArC-3), 132.7 (PhC-quart.), 132.0 (q, *J* = 30.9 Hz, ArC-6), 131.3 (ArC-4), 129.5 (PhC-ortho), 128.9 (PhC-meta), 125.9 (ArC-1), 125.7 (PhC-para), 125.0 (ArC-5), 123.0 (q, *J* = 274.9 Hz, F<sub>3</sub>C), 116.2 (C-4). LC-MS (ESI): calc. for C<sub>17</sub>H<sub>10</sub>ClF<sub>3</sub>NO<sub>2</sub> [M+H]<sup>+</sup>: 352.03, observed: 352.08 (R<sub>t</sub> = 7.28 min).

**3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazol-4-amine (13).** NEt<sub>3</sub> (0.760 mL, 5.44 mmol, 1.1 eq) and diphenylphosphoryl azide (1.06 mL, 4.95 mmol, 1.0 eq) were added to a pre-warmed solution of acid **11b** (1.82 g, 4.95 mmol, 1.0 eq) in *t*-BuOH (18 mL) at 50 °C. The reaction mixture was then heated to 85 °C for 18 h after which it was cooled to room temperature and diluted with 1 M aqueous HCl (50 mL) and extracted with EtOAc (3 x). The combined organic phase was washed with saturated aqueous NaHCO<sub>3</sub> and brine, dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo*. The crude product was purified by automated flash column chromatography, eluting with a gradient of 0–20% EtOAc in *n*-heptane, to furnish a carbamate (1.41 g, 65%) as a white solid. Trifluoroacetic acid (3.0 mL) was added to a solution of the carbamate (1.13 g, 2.58 mmol, 1.0 eq) in CH<sub>2</sub>Cl<sub>2</sub> (9.0 mL). The reaction mixture was heated at reflux for 4 h then cooled to room temperature and concentrated *in vacuo*. The crude product was dissolved in CH<sub>2</sub>Cl<sub>2</sub> and washed with saturated aqueous NaHCO<sub>3</sub> and water, dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo* to furnish **13** (0.798 g, 91%) as a pale yellow solid that was used without further purification. *R*<sub>f</sub> = 0.56 (1:1 *c*-hexane/EtOAc); <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) 7.86 (1H, d, *J* = 7.0 Hz, ArH-3 or ArH-5), 7.79–7.76 (2H, m, PhH-ortho), 7.61 (1H, app. t, *J* = 8.0 Hz, ArH-3 or ArH-5), 7.54–

7.49 (3H, m, PhH), 7.41 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 2.97 (2H, s,  $\text{NH}_2$ );  $^{13}\text{C}$ -NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 164.5 (C-5), 154.7 (C-3), 137.0 (ArC-2), 133.4 (ArC-3), 131.4 (PhC-quart.), 129.1 (PhC-ortho), 128.9 (ArC-4), 128.3 (PhC-para), 126.5 (ArC-5), 126.0 (q,  $J = 4.0$  Hz, ArC-1), 125.2 (PhC-meta), 122.9 (q,  $J = 274.4$  Hz,  $\text{CF}_3$ ), 110.4 (C-4). LC-MS (ESI): calc. for  $\text{C}_{16}\text{H}_{11}\text{ClF}_3\text{N}_2\text{O}$   $[\text{M}+\text{H}]^+$ : 339.04, observed: 339.08 ( $R_t = 7.12$  min).

***Tert*-butyl 4-(3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazole-4-amido)benzoate (14).** According to the General Procedure for amide coupling, carboxylic acid **11b** (0.200 g, 0.540 mmol) was coupled with *tert*-butyl-4-amino benzoate. The crude product was purified by flash column chromatography, eluting with 15% EtOAc in *c*-hexane, to furnish amide **14** (0.260 g, 88%) as a white solid.  $R_f = 0.55$  (3:2 *c*-hexane/EtOAc);  $^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 7.96 (2H, dd,  $J = 7.9, 1.7$  Hz, PhH-ortho), 7.86 (2H, d,  $J = 8.7$  Hz, benzoate H-2), 7.77 (2H, app. t,  $J = 8.1$  Hz, ArH-3 and ArH-5), 7.63-7.56 (4H, m, PhH-meta, PhH-para, ArH-4), 7.37 (1H, br. s,  $\text{C}(\text{O})\text{NH}$ ), 7.29 (2H, d,  $J = 8.7$  Hz, benzoate H-3), 1.55 (9H, s,  $\text{CO}_2\text{C}(\text{CH}_3)_3$ );  $^{13}\text{C}$ -NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 169.4 (C-5), 165.1 ( $\text{CO}_2\text{H}$ ), 158.4 (C-3 and  $\text{CO}_2\text{NH}$ ), 140.8 (benzoate C-4), 136.6 (ArC-2), 133.3 (ArC-3), 132.1 (PhC-quart.), 131.5 (ArC-4), 130.7 (ArC-6), 129.5 (PhC-ortho and benzoate C-2), 128.7 (PhC-meta), 128.2 (benzoate C-1), 126.2 (ArC-1), 126.0 (PhC-para), 125.1 (ArC-5), 118.7 (benzoate C-3), 113.0 (C-4), 81.1 ( $\text{CO}_2\text{C}(\text{CH}_3)_3$ ), 28.3 ( $\text{CO}_2\text{C}(\text{CH}_3)_3$ ), quartet for  $\text{CF}_3$  not observed. LC-MS (ESI): calc. for  $\text{C}_{28}\text{H}_{23}\text{ClF}_3\text{N}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 543.12, observed: 543.08 ( $R_t = 8.88$  min). HRMS (ESI): calc. for  $\text{C}_{28}\text{H}_{23}\text{ClF}_3\text{N}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 543.1298, observed: 543.1292.

**2-(4-(3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazole-4-amido)phenyl) acetic acid (15).** According to the General Procedure for amide coupling, carboxylic acid **11b** (0.200 g, 0.540 mmol) was coupled with methyl-(4-aminophenyl) acetate. The crude product was purified by flash column chromatography, eluting with a gradient of 15-45% EtOAc in *n*-heptane, to furnish the amide (0.173 g, 62%). The intermediate product (0.117 g, 0.230 mmol) was subject to ester hydrolysis according to the General Procedure for ester hydrolysis and purified by trituration with  $\text{Et}_2\text{O}$  to furnish **15** (0.099 g, 87%) as a white solid.  $R_f = 0.55$  (9:1  $\text{CH}_2\text{Cl}_2/\text{MeOH}$ );  $^1\text{H}$ -NMR (400 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 12.32 (1H, br. s,  $\text{CO}_2\text{H}$ ), 10.49 (1H, s,  $\text{CONH}$ ), 7.98 (1H, d,  $J = 8.2$  Hz, ArH-3 or ArH-5), 7.93 (1H, d,  $J = 8.0$  Hz, ArH-3 or ArH-5), 7.88-7.87 (2H, m, PhH-ortho), 7.79 (1H, app. t,  $J = 8.1$  Hz, ArH-4), 7.61-7.59 (3H, m, PhH-meta, PhH-para), 7.40 (2H, d,  $J = 8.2$  Hz, benzoate H-2), 7.15 (2H, d,  $J = 8.2$  Hz, benzoate H-3), 3.49 (2H, br. s, benzylic  $\text{CH}_2$ );  $^{13}\text{C}$ -NMR (100 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 172.7 ( $\text{CO}_2\text{H}$ ), 166.6 (C-5), 158.7 ( $\text{CO}_2\text{NH}$ ), 158.4 (C-3), 136.9 (benzoate C-4), 135.5 (ArC-2), 133.6 (ArC-3), 132.3 (PhC-quart.), 131.4 (ArC-4), 130.9 (benzoate C-1), 130.5 (q,  $J = 30.9$  Hz, ArC-6), 129.7 (benzoate C-2), 129.4 (PhC-ortho), 127.2 (PhC-meta), 126.0 (PhC-para), 125.5 (ArC-5), 125.3 (ArC-1), 122.9 (q,  $J = 274.5$  Hz,  $\text{CF}_3$ ), 119.9 (benzoate C-3), 113.8 (C-4), 30.7 (benzylic  $\text{CH}_2$ ). LC-MS (ESI): calc. for  $\text{C}_{25}\text{H}_{17}\text{ClF}_3\text{N}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 501.08, observed: 501.17 ( $R_t = 6.67$  min). HRMS (ESI): calc. for  $\text{C}_{25}\text{H}_{17}\text{ClF}_3\text{N}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 501.0829, observed: 501.0829.

**3-(3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazole-4-amido)benzoic acid (16).** According to the General Procedure for amide coupling, carboxylic acid **11b** (0.200 g, 0.540 mmol) was coupled with methyl-3-amino benzoate. The crude product was purified by flash column chromatography, eluting with 15% EtOAc in *c*-hexane, to furnish the amide (0.212 g, 78%). The intermediate product (0.100 g, 0.200 mmol) was subject to ester hydrolysis according to the General Procedure for ester hydrolysis and purified by trituration with  $\text{Et}_2\text{O}$  to furnish **16** (0.070 g, 72%) as a white solid.  $R_f = 0.51$  (9:1  $\text{CH}_2\text{Cl}_2/\text{MeOH}$ );  $^1\text{H}$ -NMR (400 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 13.01 (1H, br. s,  $\text{CO}_2\text{H}$ ), 10.66 (1H, s,  $\text{CONH}$ ), 8.11 (1H, s, benzoate H-2), 7.98 (1H, d,  $J = 8.1$  Hz, ArH-3 or ArH-5), 7.91 (1H, d,  $J = 7.7$  Hz, ArH-3 or ArH-5), 7.87 (2H, m, PhH-ortho), 7.80 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 7.68-7.64 (2H, m, benzoate H-4 and H-6), 7.63-7.59 (3H, m, PhH-meta, PhH-para), 7.40 (1H, app. t,  $J = 7.8$  Hz, benzoate H-5);  $^{13}\text{C}$ -NMR (100 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 166.9 (C-5), 158.7 ( $\text{CO}_2\text{H}$ ), 158.6 ( $\text{CO}_2\text{NH}$ ), 154.5 (C-3), 138.5 (benzoate C-3), 135.4 (ArC-2), 133.6 (ArC-3), 132.4 (PhC-quart.), 131.5 (ArC-4), 129.4 (PhC-ortho), 129.0 (benzoate

C-5), 127.3 (PhC-meta), 125.9 (PhC-para and benzoate C-1), 125.4 (ArC-5 and ArC-1), 124.9 (benzoate C-6), 124.0 (benzoate C-4), 120.6 (benzoate C-2), 113.5 (C-4), (quartet for CF<sub>3</sub> not observed). LC-MS (ESI): calc. for C<sub>24</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 487.06, observed: 487.25 (R<sub>t</sub> = 7.10 min). HRMS (ESI): calc. for C<sub>24</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 487.0672 observed: 487.0667.

**4-(3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazole-4-amido)-2-fluorobenzoic acid (17).** According to the General Procedure for amide coupling, carboxylic acid **11b** (0.200 g, 0.540 mmol) was coupled with methyl-4-amino-2-fluoro benzoate. The crude product was purified by flash column chromatography, eluting with a gradient of 20-25% EtOAc in n-heptane, to furnish the amide (0.075 g, 27%). The intermediate product (0.063 g, 0.120 mmol) was subject to ester hydrolysis according to the General Procedure for ester hydrolysis and purified by trituration with Et<sub>2</sub>O to furnish **17** (0.053 g, 87%) as a white solid. R<sub>f</sub> = 0.27 (1:1 n-heptane/EtOAc); <sup>1</sup>H-NMR (400 MHz, DMSO-d<sub>6</sub>): δ (ppm) 13.05 (1H, br. s, CO<sub>2</sub>H), 10.94 (1H, s, CONH), 7.99 (1H, d, J = 8.2 Hz, ArH-3 or ArH-5), 7.94 (1H, d, J = 7.9 Hz, ArH-3 or ArH-5), 7.87-7.78 (4H, m, PhH-ortho, ArH-4, benzoate H-6), 7.62-7.59 (3H, m, PhH-meta, PhH-para), 7.51 (1H, d, J = 13.1 Hz, benzoate H-3), 7.28 (1H, d, J = 8.7 Hz, benzoate H-5); <sup>13</sup>C-NMR (100 MHz, DMSO-d<sub>6</sub>): δ (ppm) 167.3 (C-5), 164.5 (CO<sub>2</sub>H), 161.5 (d, J = 256.0 Hz, benzoate C-2), 159.1 (CO<sub>2</sub>NH), 158.7 (C-3), 143.6 (d, J = 11.4 Hz, benzoate C-4), 135.4 (ArC-2), 133.7 (ArC-3), 132.8 (benzoate C-6), 132.4 (PhC-quart.), 131.7 (ArC-4), 130.4 (q, J = 30.6 Hz, ArC-6), 129.4 (PhC-ortho), 127.4 (PhC-meta), 125.7 (PhC-para), 125.4 (ArC-5), 125.1 (ArC-1), 122.9 (q, J = 274.6 Hz, CF<sub>3</sub>), 115.0 (benzoate C-5), 114.2 (d, J = 10.1 Hz, benzoate C-1), 113.1 (C-4), 107.2 (d, J = 27.5 Hz, benzoate C-3). LC-MS (ESI): calc. for C<sub>24</sub>H<sub>14</sub>ClF<sub>4</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 505.05, observed: 505.25 (R<sub>t</sub> = 7.10 min). HRMS (ESI): calc. for C<sub>24</sub>H<sub>14</sub>ClF<sub>4</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 505.0578 observed: 505.0569.

**4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazol-4-yl)formamido)methyl)benzoic acid (18).** According to the General Procedure for amide coupling, carboxylic acid **11b** (0.200 g, 0.540 mmol) was coupled with methyl-4-aminomethyl benzoate. The crude product was purified by flash column chromatography, eluting with 15% EtOAc in c-hexane, to furnish the amide (0.173 g, 62%). The intermediate product (0.106 g, 0.200 mmol) was subject to ester hydrolysis according to the General Procedure for ester hydrolysis and purified by trituration with Et<sub>2</sub>O to furnish **18** (0.096 g, 96%) as a white solid. R<sub>f</sub> = 0.51 (9:1 CH<sub>2</sub>Cl<sub>2</sub>/MeOH); <sup>1</sup>H-NMR (400 MHz, MeOD): δ (ppm) 7.91 (2H, d, J = 8.3 Hz, benzoate H-2), 7.84 (1H, d, J = 7.7 Hz, ArH-3 or ArH-5), 7.83 (1H, d, J = 8.3 Hz, ArH-3 or ArH-5), 7.78-7.76 (2H, m, PhH-ortho), 7.72 (1H, app. t, J = 8.0 Hz, ArH-4), 7.57-7.53 (1H, m, PhH-para), 7.49-7.45 (2H, m, PhH-meta), 7.24 (2H, d, J = 8.3 Hz, benzoate H-3), 4.42 (1H, d, J = 15.2 Hz, benzylic CHa), 4.40 (1H, d, J = 15.2 Hz, benzylic CHb); <sup>13</sup>C-NMR (100 MHz, DMSO-d<sub>6</sub>): δ (ppm) 167.2 (C-5), 166.6 (CO<sub>2</sub>H), 160.0 (CO<sub>2</sub>NH), 158.4 (C-3), 143.6 (benzoate C-4), 135.6 (ArC-2), 133.6 (ArC-3), 132.3 (PhC-quart.), 131.3 (ArC-4), 130.6 (q, J = 31.6 Hz, ArC-6), 129.5 (benzoate C-1), 129.2 (benzoate C-2 and PhC-ortho), 127.5 (PhC-meta), 127.4 (PhC-para), 125.9 (ArC-5), 125.4 (ArC-1), 122.9 (q, J = 274.1 Hz, CF<sub>3</sub>), 118.8 (benzoate C-3), 113.6 (C-4), 42.4 (benzylic CH<sub>2</sub>). LC-MS (ESI): calc. for C<sub>25</sub>H<sub>17</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 501.08, observed: 501.25 (R<sub>t</sub> = 6.67 min). HRMS (ESI): calc. for C<sub>25</sub>H<sub>17</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 501.0829 observed: 501.0818.

**4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazol-4-yl)methyl)amino) benzoic acid (19).** Ethyl-4-aminobenzoate (86 mg, 0.52 mmol, 1.0 eq) and AcOH (0.5 mL) were added to a solution of aldehyde **12** (0.183 g, 0.52 mmol, 1.0 eq) in EtOH (10.0 mL). The reaction mixture was heated at reflux for 4 h after which time it was cooled to room temperature and NaCNBH<sub>3</sub> (65.3 mg, 1.04 mmol, 2.0 eq) was added. The reaction mixture was then heated at reflux for a further 12 h then concentrated *in vacuo*, diluted with EtOAc and washed with saturated aqueous NaHCO<sub>3</sub>, water and brine. The combined organic phase was dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography, eluting with 17% EtOAc in n-heptane, to furnish the ester (81.4 mg, 31%). The intermediate product (0.049 g, 0.098 mmol) was subject to ester hydrolysis according to the

General Procedure for ester hydrolysis and purified by trituration with Et<sub>2</sub>O to furnish **19** (0.046 g, 99%) as a white solid. *R*<sub>f</sub> = baseline (4:1 c-hexane/EtOAc); <sup>1</sup>H-NMR (400 MHz, DMSO-d<sub>6</sub>): δ (ppm) 12.03 (1H, br. s, CO<sub>2</sub>H), 7.92-7.85 (4H, m, ArH-3, ArH-5, PhH-ortho), 7.73 (1H, app. t, *J* = 8.0 Hz, ArH-4), 7.64-7.58 (3H, m, PhH-ortho and PhH-meta), 7.53 (2H, d, *J* = 8.8 Hz, benzoate C-2), 6.48 (1H, t, *J* = 5.3 Hz, CH<sub>2</sub>NH), 6.37 (2H, d, *J* = 8.8 Hz, benzoate H-3), 4.21 (1H, dd, *J* = 14.7, 5.3 Hz, CH<sub>2</sub>NH), 4.12 (1H, dd, *J* = 14.7, 5.3 Hz, CH<sub>2</sub>NH); <sup>13</sup>C-NMR (100 MHz, DMSO-d<sub>6</sub>): δ (ppm) 167.4 (C-5), 166.4 (CO<sub>2</sub>H), 159.6 (C-3), 151.6 (benzoate C-4), 135.5 (ArC-2), 133.7 (ArC-3), 132.2 (PhC-quart.), 130.8 (benzoate C-2), 130.7 (ArC-4), 130.6 (q, *J* = 30.4 Hz, ArC-6), 129.3 (PhC-ortho), 127.2 (PhC-meta), 126.6 (PhC-para), 125.6 (ArC-5), 125.4 (q, *J* = 5.0 Hz, ArC-1), 122.9 (q, *J* = 274.6 Hz, CF<sub>3</sub>), 117.6 (benzoate C-1), 113.1 (C-4), 110.7 (benzoate C-3), 35.7 (CH<sub>2</sub>NH). LC-MS (ESI): calc. for C<sub>24</sub>H<sub>17</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 473.08, observed 473.00 (*R*<sub>t</sub> = 7.35 min). HRMS (ESI): calc. for C<sub>24</sub>H<sub>17</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 473.0880 observed: 473.0862.

#### 4-((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazol-4-yl)carbamoyl) benzoic acid (**20**).

According to the General Procedure for amide coupling, mono-methyl terephthalate (23.0 mg, 0.120 mmol) was coupled with amine **13** (43.0 mg, 0.120 mmol). The crude product was purified by flash column chromatography, eluting with 25% EtOAc in c-hexane, to furnish the amide (46.0 mg, 76%). The intermediate product (35.0 mg, 0.070 mmol) was subject to ester hydrolysis according to the General Procedure for ester hydrolysis and purified by trituration with Et<sub>2</sub>O to furnish **20** (22.6 mg, 66%) as a white solid. *R*<sub>f</sub> = 0.56 (9:1 CH<sub>2</sub>Cl<sub>2</sub>/MeOH); <sup>1</sup>H-NMR (400 MHz, DMSO-d<sub>6</sub>): δ (ppm) 13.21 (1H, br. s, CO<sub>2</sub>H), 10.39 (1H, s, NHCO), 8.01 (2H, d, *J* = 8.4 Hz, benzoate C-3), 7.94-7.86 (6H, m, benzoate C-2, ArH-3 and ArH-5, phenyl H-ortho), 7.75 (1H, app. t, *J* = 8.0 Hz, ArH-4), 7.59-7.51 (3H, m, phenyl H-meta and phenyl H-para); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) 166.7 (C-5), 165.2 (CO<sub>2</sub>H), 161.7 (CO<sub>2</sub>NH), 157.8 (C-3), 137.1 (benzoate C-4), 135.3 (PhC-quart), 133.8 (benzoate C-1), 133.6 (ArC-3), 132.2 (ArC-2), 130.9 (q, *J* = 30.9 Hz, ArC-6), 130.6 (benzoate C-4), 129.3 (benzoate C-2), 129.2 (PhC-ortho), 127.9 (PhC-meta), 126.4 (PhC-para), 125.9 (benzoate C-3), 125.7 (q, *J* = 5.0 Hz, ArC-5), 125.1 (ArC-1), 122.9 (q, *J* = 274.7 Hz, CF<sub>3</sub>), 114.5 (C-4). LC-MS (ESI): calc. for C<sub>24</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 487.06, observed 487.17 (*R*<sub>t</sub> = 6.73 min). HRMS (ESI): calc. for C<sub>24</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 487.0672 observed: 487.0677.

#### 4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazol-4-yl)amino) methyl) benzoic acid (**21**).

Methyl-4-formyl benzoate (93.0 mg, 0.570 mmol, 1.0 eq) and AcOH (0.5 mL) were added to a solution of amine **13** (0.201 g, 0.590 mmol, 1.05 eq) in MeOH (10.0 mL). The reaction mixture was heated at reflux for 18 h after which time it was cooled to room temperature and NaCNBH<sub>3</sub> (74.0 mg, 1.18 mmol, 2.0 eq) was added. The reaction mixture was heated at reflux for a further 18 h then concentrated *in vacuo*, diluted with EtOAc and washed with saturated aqueous NaHCO<sub>3</sub>, water and brine. The combined organic phase was dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography, eluting with 17% EtOAc in n-heptane, to furnish the ester (0.120 g, 43%). The intermediate product (0.106 g, 0.220 mmol) was subject to ester hydrolysis according to the General Procedure for ester hydrolysis and purified by trituration with Et<sub>2</sub>O to furnish **21** (0.094 g, 98%) as a white solid. *R*<sub>f</sub> = baseline (1:1 c-hexane/EtOAc); <sup>1</sup>H-NMR (400 MHz, DMSO-d<sub>6</sub>): δ (ppm) 12.82 (1H, br. s, CO<sub>2</sub>H), 7.95 (1H, d, *J* = 8.1 Hz, ArH-3), 7.88 (1H, d, *J* = 8.1 Hz, ArH-5), 7.84 (2H, d, *J* = 7.3 Hz, PhH-ortho), 7.80-7.78 (1H, m, ArH-4), 7.75 (2H, d, *J* = 8.1 Hz, benzoate H-2), 7.53-7.50 (2H, m, PhH-meta), 7.47-7.43 (1H, m, PhH-para), 7.15 (2H, d, *J* = 8.1 Hz, benzoate H-3), 5.11 (1H, app. t, *J* = 6.0 Hz, NHCH<sub>2</sub>), 4.00 (1H, dd, *J* = 15.5, 6.0 Hz, NHCH<sub>2</sub>a), 3.94 (1H, dd, *J* = 15.5, 6.0 Hz, NHCH<sub>2</sub>b); <sup>13</sup>C-NMR (100 MHz, DMSO-d<sub>6</sub>): δ (ppm) 167.2 (C-5), 155.3 (CO<sub>2</sub>H), 154.3 (C-3), 145.0 (benzoate C-4), 136.0 (ArC-2), 133.6 (ArC-3), 132.2 (PhC-quart.), 131.1 (q, *J* = 30.5 Hz, ArC-6), 129.22 (benzoate C-1), 129.17 (ArC-4), 129.0 (PhC-ortho), 128.9 (benzoate C-2), 127.6 (PhC-para), 127.1 (PhC-meta), 126.0 (ArC-5), 125.9 (C-4), 125.7 (benzoate C-3), 125.4 (q, *J* = 5.0 Hz, ArC-1), 122.9 (q, *J* = 274.6 Hz, CF<sub>3</sub>), 49.8 (CH<sub>2</sub>NH). LC-MS (ESI): calc. for C<sub>24</sub>H<sub>17</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 473.08, observed 473.17 (*R*<sub>t</sub> = 7.48 min). HRMS (ESI): calc. for C<sub>24</sub>H<sub>17</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 473.0880 observed: 473.0883.

**4-((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-phenyl-1,2-oxazol-4-yl)sulfamoyl) benzoic acid (22).** Methyl-4-(chlorosulfonyl)benzoate (0.464 g, 1.70 mmol, 3.0 eq) was added to a solution of amine **13** (0.192 g, 0.570 mmol, 1.0 eq) in pyridine (10.0 mL). The reaction mixture was stirred at 60 °C for 24 h then cooled to room temperature and concentrated *in vacuo*. The mixture was suspended in EtOAc and washed with 1 M aqueous HCl, saturated aqueous NaHCO<sub>3</sub>, water and brine. The organic phase was dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo* to furnish the ester (0.260 g, 71%) that was used without further purification. The intermediate product (0.188 g, 0.350 mmol) was subject to ester hydrolysis according to the General Procedure for ester hydrolysis and purified by trituration with Et<sub>2</sub>O to furnish **22** (0.079 g, 43%) as a yellow solid. *R*<sub>f</sub> = baseline (1:1 *n*-hexane/EtOAc); <sup>1</sup>H-NMR (400 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) 13.30 (1H, br. s, CO<sub>2</sub>H), 10.49 (1H, br. s, NHSO<sub>2</sub>), 7.87 (1H, d, *J* = 8.0 Hz, ArC-3), 7.79 (1H, d, *J* = 8.0 Hz, ArC-5), 7.75-7.68 (5H, m, ArH-4, PhH-ortho, benzoate H-3), 7.51 (2H, d, *J* = 8.2 Hz, benzoate H-2), 7.46-7.37 (3H, m, PhH-meta and PhH-para); <sup>13</sup>C-NMR (100 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) 165.9 (C-5), 164.8 (CO<sub>2</sub>H), 159.2 (C-3), 143.6 (benzoate C-4), 135.4 (ArC-2), 134.2 (ArC-3), 133.5 (benzoate C-1), 132.3 (PhC-quart.), 130.9 (q, *J* = 30.9 Hz, ArC-6), 130.8 (ArC-4), 129.7 (benzoate C-2), 128.8 (PhC-ortho), 126.2 (benzoate C-3), 126.1 (PhC-meta), 125.5 (q, *J* = 5.0 Hz, ArC-1), 125.2 (PhC-para), 124.4 (ArC-5), 122.8 (q, *J* = 274.5 Hz, CF<sub>3</sub>), 112.4 (C-4). LC-MS (ESI): calc. for C<sub>23</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>5</sub>S [M+H]<sup>+</sup>: 523.03, observed 523.00 (*R*<sub>t</sub> = 6.70 min). HRMS (ESI): calc. for C<sub>23</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>5</sub>S [M+H]<sup>+</sup>: 523.0342 observed: 523.0333.

**4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(furan-2-yl)-1,2-oxazol-4-yl)methyl)amino)benzoic acid (23).** According to the General Procedure for reductive amination, aldehyde **39** (0.180 g, 0.520 mmol) was reacted with *tert*-butyl-4-amino benzoate (0.102 g, 0.520 mmol) in MeOH. The crude product was purified by column chromatography, eluting with a gradient of 5-20% EtOAc in *n*-heptane, to furnish the intermediate imine (0.136 g) that was immediately subjected to the reduction step performed in EtOH at 85 °C. The crude product was purified by column chromatography, eluting with a gradient of 10-20% EtOAc in *n*-heptane, to furnish the intermediate amine (0.065 g, 24%). This product was subject to *tert*-butyl ester deprotection according to the General Procedure for *tert*-Butyl ester deprotection. The crude product was purified by column chromatography, eluting with 1% MeOH in CH<sub>2</sub>Cl<sub>2</sub>, to furnish the carboxylic acid **23** (0.040 g, 73%) as a white solid. *R*<sub>f</sub> = 0.13 (99:1 CH<sub>2</sub>Cl<sub>2</sub>/MeOH); <sup>1</sup>H-NMR (400 MHz, MeOD): δ (ppm) 7.84 (1H, d, *J* = 1.8 Hz, furanyl H-5), 7.75 (1H, d, *J* = 8.1 Hz, Ar H-3 or ArH-5), 7.74 (1H, d, *J* = 8.1 Hz, ArH-3 or ArH-5), 7.61 (1H, app. t, *J* = 8.1 Hz, ArH-4), 7.61 (2H, d, *J* = 8.8 Hz, benzoate H-2), 7.12 (1H, d, *J* = 3.5 Hz, furanyl H-3), 7.62 (1 H, dd, *J* = 3.5, 1.8 Hz, furanyl H-4), 6.34 (2 H, d, *J* = 8.8 Hz, benzoate H-3), 4.46 (1H, d, *J* = 15.3 Hz, CH<sub>2</sub>NH), 4.36 (1H, d, *J* = 15.3 Hz, CH<sub>2</sub>NH); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) 170.5 (C-5), 160.6 (CO<sub>2</sub>H), 159.9 (C-3), 153.4 (furanyl C-2), 146.5 (furanyl C-5), 144.1 (benzoate C-4), 137.6 (ArC-2), 134.6 (ArC-3), 132.9 (q, *J* = 32.5 Hz, ArC-6), 132.7 (ArC-4), 132.4 (benzoate C-2), 127.3 (ArC-1), 126.3 (q, *J* = 5.1 Hz, ArC-5), 124.4 (q, *J* = 273.8 Hz, CF<sub>3</sub>), 118.7 (benzoate C-1), 114.6 (C-4), 113.2 (furanyl C-4), 113.1 (furanyl C-3), 111.8 (benzoate C-3), 36.4 (CH<sub>2</sub>NH). LC-MS (ESI): calc. for C<sub>22</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 463.06, observed: 462.92 (*R*<sub>t</sub> = 6.67 min; HRMS (ESI): calc. for C<sub>22</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 463.0672, observed: 463.0661.

**4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(thiophen-2-yl)-1,2-oxazol-4-yl)methyl)amino)benzoic acid (24).** According to the General Procedure for reductive amination, aldehyde **40** (0.096 g, 0.270 mmol) was reacted with *tert*-butyl-4-amino benzoate (0.052 g, 0.27 mmol) in MeOH. The crude product was purified by column chromatography, eluting with a gradient of 5-10% EtOAc in *n*-heptane, to furnish the intermediate imine (0.125 g) that was immediately subjected to the reduction step performed in EtOH at 85 °C for 5 h. The crude product was purified by column chromatography, eluting with a gradient of 3-10% EtOAc in *n*-heptane, to furnish the intermediate amine (24.0 mg, 17%). This product was subject to *tert*-butyl ester deprotection according to the General Procedure for *tert*-Butyl ester deprotection. The crude product was purified by column chromatography, eluting with 3% MeOH in

CH<sub>2</sub>Cl<sub>2</sub>, to furnish the carboxylic acid **24** (12.0 mg, 56%) as a pale yellow solid. *R*<sub>f</sub> = 0.13 (98:2 CH<sub>2</sub>Cl<sub>2</sub>/MeOH); <sup>1</sup>H-NMR (400 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) 12.02 (1H, br. s, CO<sub>2</sub>H), 7.96 (1H, dd, *J* = 5.1 Hz, 1.1, thiophenyl H-5), 7.93 (1H, d, *J* = 8.0 Hz, ArH-3 or ArH-5), 7.87 (1H, d, *J* = 8.0 Hz, ArH-3 or ArH-5), 7.77-7.73 (2H, m, ArH-4, thiophenyl H-3), 7.55 (2H, d, *J* = 8.8 Hz, benzoate H-2), 7.33 (1H, dd, *J* = 5.1 Hz, 3.7, thiophenyl H-4), 6.48 (1H, app. t, *J* = 5.0 Hz, CH<sub>2</sub>NH), 6.42 (2H, d, *J* = 8.8 Hz, benzoate H-3), 4.20 (1H, dd, *J* = 14.8, 5.0 Hz, CH<sub>2</sub>NH), 4.13 (1H, dd, *J* = 14.8, 5.0 Hz, CH<sub>2</sub>NH); <sup>13</sup>C-NMR (100 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) 167.4 (C-5), 162.0 (CO<sub>2</sub>H), 159.5 (C-3), 151.6 (benzoate C-4), 135.5 (ArC-2), 133.8 (ArC-3), 132.3 (ArC-4), 130.8 (benzoate C-2), 130.6 (q, *J* = 30.7 Hz, ArC-6), 130.6 (thiophenyl C-5), 128.8 (thiophenyl C-3), 128.7 (thiophenyl C-4), 127.2 (thiophenyl C-2), 125.5 (q, *J* = 4.9 Hz, ArC-5), 125.3 (ArC-1), 122.8 (q, *J* = 274.5 Hz, CF<sub>3</sub>), 117.6 (benzoate C-1), 112.0 (C-4), 110.8 (benzoate C-3), 35.6 (CH<sub>2</sub>NH). LC-MS (ESI): calc. for C<sub>22</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>3</sub>S [M+H]<sup>+</sup>: 479.04, observed: 479.00 (*R*<sub>t</sub> = 7.23 min). HRMS (ESI): calc. for C<sub>22</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>3</sub>S [M+H]<sup>+</sup>: 479.0444, observed: 479.0429.

**4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-3-yl)-1,2-oxazol-4-yl)methyl)amino)benzoic acid (25).** According to the General Procedure for reductive amination, aldehyde **41** (0.060 g, 0.176 mmol) was reacted with methyl-4-amino benzoate (0.032 g, 0.211 mmol) in MeOH. The crude product was purified by column chromatography, eluting with 25% EtOAc in *n*-heptane, to furnish the intermediate imine (0.034 mg) that was immediately subjected to the reduction step performed in MeOH at reflux for 2 h. The crude product was purified by column chromatography, eluting with a gradient of 20% EtOAc in *n*-heptane, to furnish the intermediate amine (16.3 mg, 19%). This product was subject to ester hydrolysis according to the General Procedure for *tert*-Butyl ester deprotection. The crude product was purified by column chromatography, eluting with 1.5% MeOH in CH<sub>2</sub>Cl<sub>2</sub>, to furnish the carboxylic acid **25** (5.40 mg, 57%) as a white solid. *R*<sub>f</sub> = 0.10 (96:4 CH<sub>2</sub>Cl<sub>2</sub>/MeOH); <sup>1</sup>H-NMR (400 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) 11.98 (1H, br. s, CO<sub>2</sub>H), 11.52 (1H, s, pyrrole-NH), 7.90 (1H, d, *J* = 8.0 Hz, ArH-3), 7.84 (1H, d, *J* = 7.8 Hz, ArH-5), 7.72 (1H, app. t, *J* = 8.0 Hz, ArH-4), 7.53 (2H, d, *J* = 8.7 Hz, benzoate C-2), 7.39 (1H, m, pyrrole H-2), 6.98 (1H, m, pyrrole H-5), 6.55 (1H, m, pyrrole H-4), 6.41 (2H, d, *J* = 8.7 Hz, benzoate H-3), 6.34 (1H, app. t, *J* = 4.4 Hz, CH<sub>2</sub>NH), 4.07 (1H, dd, *J* = 14.4, 5.0 Hz, CH<sub>2</sub>NH), 4.02 (1H, dd, *J* = 14.5, 4.5 Hz, CH<sub>2</sub>NH); <sup>13</sup>C-NMR (100 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) 167.9 (CO<sub>2</sub>H), 165.7 (C-5), 159.2 (C-3), 152.3 (benzoate C-4), 136.0 (ArC-2), 134.1 (ArC-3), 132.4 (ArC-4), 131.3 (benzoate C-2), 130.9 (q, *J* = 30.4 Hz, ArC-6), 126.8 (ArC-1), 125.8 (q, *J* = 5.0 Hz, ArC-5), 122.0 (q, *J* = 274.3 Hz, CF<sub>3</sub>), 120.5 (pyrrole C-5), 119.2 (pyrrole C-2), 117.8 (benzoate C-1), 111.2 (benzoate C-3), 110.5 (pyrrole C-3), 109.2 (C-4), 106.7 (pyrrole C-4), 36.2 (CH<sub>2</sub>NH). LC-MS (ESI): calc. for C<sub>22</sub>H<sub>16</sub>ClF<sub>3</sub>N<sub>3</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 462.08, observed: 462.00 (*R*<sub>t</sub> = 6.20 min). HRMS (ESI): calc. for C<sub>22</sub>H<sub>16</sub>ClF<sub>3</sub>N<sub>3</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 462.0832, observed: 462.0834.

**4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(naphthalen-1-yl)-1,2-oxazol-4-yl)methyl)amino)benzoic acid (26).** According to the General Procedure for reductive amination, aldehyde **42** (0.034 g, 0.0850 mmol) was reacted with *tert*-butyl-4-amino benzoate (0.016 g, 0.085 mmol) in MeOH. The crude product was purified by column chromatography, eluting with a gradient of 2-10% EtOAc in *n*-heptane, to furnish the intermediate imine (0.010 g) that was immediately subjected to the reduction step performed in EtOH at 85 °C for 5 h. The crude product was purified by column chromatography, eluting with 20% EtOAc in *n*-heptane, to furnish the intermediate amine (8.00 mg, 16%). This product was subject to *tert*-butyl ester deprotection according to the General Procedure for *tert*-Butyl ester deprotection. The crude product was purified by column chromatography, eluting with 3% MeOH in CH<sub>2</sub>Cl<sub>2</sub>, to furnish the carboxylic acid **26** (3.00 mg, 48%) as a white solid. *R*<sub>f</sub> = 0.16 (96:4 CH<sub>2</sub>Cl<sub>2</sub>/MeOH); <sup>1</sup>H-NMR (400 MHz, DMSO-*d*<sub>6</sub>): δ (ppm) 11.93 (1H, br. s, CO<sub>2</sub>H), 8.23 (1H, d, *J* = 8.3 Hz, naphthyl-H), 8.12 (1H, d, *J* = 7.5 Hz, naphthyl-H), 7.94 (1H, d, *J* = 8.0 Hz, ArH-3 or ArH-5), 7.88 (1H, d, *J* = 8.0 Hz, ArH-3 or ArH-5), 7.84 (1H, d, *J* = 7.0 Hz, naphthyl-H), 7.76-7.66 (5H, m, ArH-4, naphthyl-H), 7.31 (2H, d, *J* = 8.3 Hz, benzoate C-2), 6.36 (1H, app. t, *J* = 5.8 Hz, CH<sub>2</sub>NH), 6.08 (2H, d, *J* = 8.3 Hz, benzoate C-3), 4.08 (1H, dd, *J* = 15.8,



5.8 Hz,  $CH_aNH$ ), 4.01 (1H, dd,  $J = 15.8, 5.8$  Hz,  $CH_bNH$ );  $^{13}C$ -NMR (100 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 167.3 (C-5), 166.9 ( $CO_2H$ ), 158.8 (C-3), 151.3 (benzoate C-4), 135.4 (ArC-2), 133.7 (ArC-3), 133.3 (naphthyl-C), 132.1 (ArC-4), 131.4 (naphthyl-C), 130.9 (naphthyl-C), 130.5 (q,  $J = 31.1$  Hz, ArC-6), 130.5 (benzoate C-2), 129.3 (naphthyl-C), 128.7 (naphthyl-C), 127.8 (naphthyl-C), 126.8 (naphthyl-C), 125.8 (ArC-1), 125.5 (q,  $J = 4.0$  Hz, ArC-5), 125.4 (naphthyl-C), 124.3 (naphthyl-C), 123.5 (naphthyl-C), 123.0 (q,  $J = 274.3$  Hz,  $CF_3$ ), 117.2 (benzoate C-1), 116.3 (C-4), 110.4 (benzoate C-3), 35.3 ( $CH_2NH$ ). LC-MS (ESI): calc. for  $C_{28}H_{19}ClF_3N_2O_3$   $[M+H]^+$ : 523.10, observed: 522.92 ( $R_t = 7.68$  min). HRMS (ESI): calc. for  $C_{28}H_{19}ClF_3N_2O_3$   $[M+H]^+$ : 523.1036, observed: 523.1046.

**4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(3-hydroxyphenyl)-1,2-oxazol-4-yl)methyl)amino)benzoic acid (27).** According to the General Procedure for reductive amination, aldehyde **43** (0.123 g, 0.255 mmol) was reacted with methyl-4-amino benzoate (0.038 g, 0.255 mmol) in MeOH. The crude product was purified by column chromatography, eluting with a gradient of 2–12% EtOAc in n-heptane, to furnish the intermediate imine (0.053 g) that was immediately subjected to the reduction step performed in MeOH at reflux for 3.5 h. This step occurred with concomitant loss of the silyl protecting group. The crude product was purified by column chromatography, eluting with a gradient of 15–35% EtOAc in n-heptane, to furnish the intermediate amine (21.0 mg, 16%). This product was subject to ester hydrolysis according to the General Procedure for *tert*-Butyl ester deprotection. The crude product was purified by column chromatography, eluting with 3% MeOH in  $CH_2Cl_2$ , to furnish the carboxylic acid **27** (17.4 mg, 99%) as a white solid.  $R_f = 0.10$  (96:4  $CH_2Cl_2$ /MeOH);  $^1H$ -NMR (400 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 7.90 (1H, d,  $J = 8.1$  Hz, ArH-3 or ArH-5), 7.85 (1H, d,  $J = 8.0$  Hz, ArH-3 or ArH-5), 7.72 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 7.52 (2H, d,  $J = 8.8$  Hz, benzoate H-2), 7.40 (1H, app. t,  $J = 8.0$  Hz, phenol H-5), 1.30–7.26 (2H, m, phenol H-2 and phenol H-4), 6.99 (1H, dd,  $J = 8.0, 1.5$  Hz, phenol H-6), 6.42 (1H, t,  $J = 5.0$  Hz,  $CH_2NH$ ), 6.36 (2H, d,  $J = 8.8$  Hz, benzoate H-3), 4.17 (1H, dd,  $J = 14.6, 5.0$  Hz,  $CH_aNH$ ), 4.09 (1H, dd,  $J = 14.6, 5.0$  Hz,  $CH_bNH$ );  $^{13}C$ -NMR (100 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 167.4 (C-5), 166.4 ( $CO_2H$ ), 159.6 (C-3), 157.9 (phenol C-3), 151.6 (benzoate C-4), 135.5 (ArC-2), 133.7 (ArC-3), 132.2 (ArC-4), 130.8 (benzoate C-2), 130.5 (q,  $J = 30.7$  Hz, ArC-6), 130.5 (phenol C-1), 127.7 (phenol C-5), 125.7 (ArC-1), 125.4 (q,  $J = 4.0$  Hz, ArC-5), 122.8 (q,  $J = 274.3$  Hz,  $CF_3$ ), 117.8 (phenol C-6), 117.6 (benzoate C-1), 113.6 (phenol C-2), 112.9 (C-4), 110.7 (benzoate C-3), 35.7 ( $CH_2NH$ ). LC-MS (ESI): calc. for  $C_{24}H_{17}ClF_3N_2O_4$   $[M+H]^+$ : 489.08, observed: 489.00 ( $R_t = 6.30$  min). HRMS (ESI): calc. for  $C_{24}H_{17}ClF_3N_2O_4$   $[M+H]^+$ : 489.0829, observed: 489.0823.

**Methyl 5-bromo-3-(2-chloro-6-(trifluoromethyl)phenyl)-1,2-oxazole-4-carboxylate (30a).** Methyl 3-bromopropionate (prepared according to ref. 44, 2.04 g, 12.5 mmol, 1.0 eq) was added to a solution of nitrile oxide **33** (2.75 g, 12.5 mmol, 1.0 eq) in THF (25 mL) and the reaction mixture was heated at reflux for 4 h. The reaction mixture was concentrated *in vacuo* to give a crude product as a 7:3 mixture of regioisomers. Purification by recrystallization from hot n-heptane furnished bromide **30a** (2.05 g, 43%) as a white solid (97:3 mixture of regioisomers – see Supporting Figure S2.1).  $R_f = 0.20$  (7:3 n-heptane/EtOAc);  $^1H$ -NMR (400 MHz,  $CDCl_3$ ):  $\delta$  (ppm) 7.73 (1H, d,  $J = 8.6$  Hz, ArH-3 or ArH-5), 7.71 (1H, d,  $J = 8.9$  Hz, ArH-3 or ArH-5), 7.58 (1H, app. t,  $J = 8.0$  Hz, ArH-3), 3.70 (3H, s,  $CO_2CH_3$ );  $^{13}C$ -NMR (100 MHz,  $CDCl_3$ ):  $\delta$  (ppm) 160.0 (C-5), 159.7 ( $CO_2Me$ ), 148.0 (C-3), 136.3 (ArC-2), 133.0 (ArC-3), 131.7 (q,  $J = 31.7$  Hz, ArC-6), 131.3 (ArC-4), 126.0 (ArC-1), 124.8 (q,  $J = 5.0$  Hz, ArC-5), 122.9 (q,  $J = 274.4$  Hz,  $CF_3$ ), 113.2 (C-4), 52.3 ( $CO_2CH_3$ ). LC-MS (ESI): calc. for  $C_{12}H_7BrClF_3NO_3$   $[M+H]^+$ : 383.92, observed: 386.00 ( $R_t = 7.12$  min).

**(2-chloro-6-(trifluoromethyl)phenyl)formonitrile oxide (33).**  $NEt_3$  (5.80 mL, 41.5 mmol, 1.2 eq) was added dropwise to a solution of imidoyl chloride **9a** (8.90 g, 34.6 mmol, 1.0 eq) in THF (110 mL). A white precipitate formed immediately. The resulting suspension was stirred vigorously at room temperature for 30 min and then filtered through a pad of  $SiO_2$  that was subsequently washed with THF

(250 mL). The solution was concentrated *in vacuo* to furnish **33** (8.25 g, 99%) as a white solid which was used immediately.  $R_f = 0.33$  (9:1 *n*-hexane/EtOAc);  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 7.70 (1H, d,  $J = 8.2$  Hz, H-3 or H-5), 7.67 (1H, d,  $J = 8.2$  Hz, H-3 or H-5), 7.54 (1H, app. t,  $J = 8.2$  Hz, H-4).

**Methyl 3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(furan-2-yl)-1,2-oxazole-4-carboxylate (34).**

According to the General Procedure for Suzuki coupling, bromide **30a** (0.150 g, 0.390 mmol) was coupled to furan-2-boronic acid pinacol ester (0.114 g, 0.585 mmol). The crude product was purified by flash column chromatography, eluting with a gradient of 10-50% EtOAc in *n*-heptane, to furnish **34** (0.088 g, 58%) as a white solid.  $R_f = 0.19$  (85:15 *n*-heptane/EtOAc);  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 7.82 (1H, dd,  $J = 3.6, 0.7$  Hz, furanyl H-5), 7.73 (1H, d,  $J = 8.0$  Hz, Ar H-3 or ArH-5), 7.72 (1H, dd,  $J = 1.8, 0.7$  Hz, furanyl H-3), 7.70 (1H, d,  $J = 8.0$  Hz, Ar H-3 or ArH-5), 7.56 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 6.66 (1H, dd,  $J = 3.6, 1.8$  Hz, furanyl H-4), 3.62 (3H, s,  $\text{CO}_2\text{CH}_3$ );  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 163.4 (C-5), 160.9 ( $\text{CO}_2\text{CH}_3$ ), 158.7 (C-3), 146.3 (furanyl C-3), 141.6 (furanyl C-2), 136.3 (ArC-2), 132.8 (ArC-3 and ArC-4), 131.6 (q,  $J = 31.3$  Hz, ArC-6), 130.8 (ArC-4), 127.1 (ArC-1), 124.7 (q,  $J = 5.0$  Hz, ArC-5), 123.0 (q,  $J = 274.5$  Hz,  $\text{CF}_3$ ), 118.3 (furanyl C-5), 112.6 (furanyl C-4), 107.3 (C-4), 51.9 ( $\text{CO}_2\text{CH}_3$ ). LC-MS (ESI): calc. for  $\text{C}_{16}\text{H}_{10}\text{ClF}_3\text{NO}_4$   $[\text{M}+\text{H}]^+$ : 371.02, observed: 372.08 ( $R_t = 7.77$  min).

**Methyl 3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(thiophen-2-yl)-1,2-oxazole-4-carboxylate (35).**

According to the General Procedure for Suzuki coupling, bromide **30a** (0.250 g, 0.650 mmol) was coupled to thiophene-2-boronic acid pinacol ester (0.273 g, 1.30 mmol). The crude product was purified by flash column chromatography, eluting with a gradient of 5-20% EtOAc in *n*-heptane, to furnish **35** (0.136 g, 54%) as a white solid.  $R_f = 0.32$  (4:1 *n*-heptane/EtOAc);  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 8.29 (1H, dd,  $J = 3.9, 1.2$  Hz, thiophenyl H-5), 7.75-7.69 (3H, m, ArH-3, ArH-5, thiophenyl H-3), 7.57 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 7.24 (1H, dd,  $J = 5.1, 3.9$  Hz, thiophenyl H-4), 3.64 (3H, s,  $\text{CO}_2\text{CH}_3$ );  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 167.9 (C-5), 161.4 ( $\text{CO}_2\text{CH}_3$ ), 159.2 (C-3), 136.3 (ArC-2), 132.8 (thiophenyl C-5), 132.7 (ArC-4), 132.2 (ArC-3 and thiophenyl C-3), 131.7 (q,  $J = 31.4$  Hz, ArC-6), 130.8 (ArH-4), 128.0 (thiophenyl C-4), 127.4 (thiophenyl C-2), 127.3 (ArC-1), 124.7 (q,  $J = 5.0$  Hz, ArC-5), 123.0 (q,  $J = 274.5$  Hz,  $\text{CF}_3$ ), 107.2 (C-4), 51.9 ( $\text{CO}_2\text{CH}_3$ ). LC-MS (ESI): calc. for  $\text{C}_{16}\text{H}_{10}\text{ClF}_3\text{NO}_3\text{S}$   $[\text{M}+\text{H}]^+$ : 387.99, observed: 388.25 ( $R_t = 7.62$  min).

**Methyl 5-(1-((*tert*-butoxy)carbonyl)-1H-pyrrol-3-yl)-3-(2-chloro-6-(trifluoromethyl)phenyl)-1,2-oxazole-4-carboxylate (36).** According to the General Procedure for Suzuki coupling, bromide **30a** (0.250 g, 0.650 mmol) was coupled to *tert*-butyl 3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-1H-pyrrole-1-carboxylate (0.381 g, 1.30 mmol). The crude product was purified by flash column chromatography, eluting with a gradient of 5-20% EtOAc in *n*-heptane, to furnish **36** (0.119 g, 39%) as a colourless oil.  $R_f = 0.30$  (4:1 *n*-heptane/EtOAc);  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 8.45 (1H, app. t,  $J = 2.0$  Hz, pyrrole H-2), 7.72-7.67 (2H, m, ArH-3 and ArH-5), 7.54 (1H, app. t,  $J = 7.9$  Hz, ArH-4), 7.35 (1H, dd,  $J = 3.4, 2.0$  Hz, pyrrole H-5), 6.98 (1H, dd,  $J = 3.4, 2.0$  Hz, pyrrole H-4), 3.61 (3H, s,  $\text{CO}_2\text{CH}_3$ ), 1.63 (9H, s,  $\text{C}(\text{CH}_3)_3$ );  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 168.9 (C-5), 161.6 ( $\text{CO}_2\text{CH}_3$ ), 159.0 (C-3), 148.2 ( $\text{NCO}_2$ ), 136.3 (ArC-2), 132.8 (ArC-3), 131.6 (q,  $J = 31.3$  Hz, ArC-6), 130.6 (ArC-4), 127.7 (ArC-1), 124.7 (q,  $J = 5.0$  Hz, ArC-5), 124.1 (pyrrole C-2), 123.0 (q,  $J = 274.4$  Hz,  $\text{CF}_3$ ), 121.0 (pyrrole C-5), 113.8 (pyrrole C-3), 111.5 (pyrrole C-4), 107.2 (C-4), 85.2 ( $\text{C}(\text{CH}_3)_3$ ), 51.7 ( $\text{CO}_2\text{CH}_3$ ), 28.0 ( $\text{C}(\text{CH}_3)_3$ ). LC-MS (ESI): calc. for  $\text{C}_{21}\text{H}_{19}\text{ClF}_3\text{N}_2\text{O}_5$   $[\text{M}+\text{H}]^+$ : 471.09, observed: 471.17 ( $R_t = 8.55$  min).

**Methyl 3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(naphthalen-1-yl)-1,2-oxazole-4-carboxylate (37).**

According to the General Procedure for Suzuki coupling, bromide **30a** (0.150 g, 0.390 mmol) was coupled to naphthalene-1-boronic acid pinacol ester (0.198 g, 0.780 mmol). The crude product was purified by flash column chromatography, eluting with a gradient of 10-20% EtOAc in *n*-heptane, to furnish **37** (0.059 g, 35%) as a colourless oil.  $R_f = 0.43$  (4:1 *n*-heptane/EtOAc);  $^1\text{H-NMR}$  (400 MHz,

CDCl<sub>3</sub>):  $\delta$  (ppm) 8.07 (1H, d,  $J$  = 8.3 Hz, naphthyl-H), 7.97-7.94 (1H, m, naphthyl-H), 7.81-7.75 (3H, m, naphthyl-H), 7.73-7.71 (1H, m, ArH-3 or ArH-5), 7.65-7.60 (2H, m, ArH-3 or ArH-5, naphthyl-H), 7.58-7.56 (2H, m, ArH-3 or ArH-5, naphthyl-H), 3.41 (3H, s, CO<sub>2</sub>CH<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 173.7 (C-5), 161.0 (CO<sub>2</sub>CH<sub>3</sub>), 158.9 (C-4), 136.4, 133.5, 132.9, 131.9, 131.8 (q,  $J$  = 31.4 Hz, ArC-6), 131.4, 130.9, 130.2, 129.4, 128.7, 127.6, 127.1 (ArC-1), 126.7, 124.9, 124.8, 124.2, 123.2 (q,  $J$  = 274.4 Hz, CF<sub>3</sub>), 112.2 (C-4), 51.8 (CO<sub>2</sub>CH<sub>3</sub>), (not all peaks could be precisely assigned with certainty). LC-MS (ESI): calc. for C<sub>22</sub>H<sub>14</sub>ClF<sub>3</sub>NO<sub>3</sub> [M+H]<sup>+</sup>: 432.05, observed: 432.25 ( $R_t$  = 8.14 min).

**Methyl 5-(3-((*tert*-butyldimethylsilyloxy)phenyl)-3-(2-chloro-6-(trifluoromethyl)phenyl)-1,2-oxazole-4-carboxylate (38).** According to the General Procedure for Suzuki coupling, bromide **30a** (0.400 g, 1.04 mmol) was coupled to 3-(4,4,5,5-Tetramethyl-1,3,2-dioxaborolan-2-yl)phenol (0.319 g, 1.45 mmol). The crude product was purified by flash column chromatography, eluting with a gradient of 10-20% EtOAc in *n*-heptane, to furnish the phenol (0.167 g, 40%) as a white solid. Imidazole (94.0 mg, 1.38 mmol, 3.0 eq) and *tert*-butyldimethylsilyl chloride (0.104 g, 0.690 mmol, 1.5 eq) were added to a solution of the phenol (0.183 g, 0.460 mmol, 1.0 eq) in DMF (3.0 mL). The reaction mixture was stirred at room temperature for 3 h, then diluted with saturated aqueous NH<sub>4</sub>Cl and extracted with EtOAc (2 x). The combined organic phase was washed with water (2 x) and brine, dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. The crude product was purified by column chromatography, eluting with a gradient of 0-20% EtOAc in *n*-heptane, to furnish silyl ether **38** (56.0 mg, 70%) as a colourless oil.  $R_f$  = 0.26 (9:1 *n*-heptane/EtOAc); <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 7.73 (2H, app. t,  $J$  = 8.6 Hz, ArH-3 and ArH-5), 7.70 (1H, ddd,  $J$  = 8.1, 1.7, 1.0 Hz, phenol H-4), 7.59-7.55 (2H, m, phenol H-2, ArH-4), 7.39 (1H, app. t,  $J$  = 8.1 Hz, phenol H-5), 7.04 (1H, ddd,  $J$  = 8.1 Hz, 2.4, 1.0, phenol H-6), 3.59 (3H, s, CO<sub>2</sub>CH<sub>3</sub>), 1.02 (9H, s, Si(CH<sub>3</sub>)<sub>2</sub>C(CH<sub>3</sub>)<sub>3</sub>), 0.26 (6H, s, Si(CH<sub>3</sub>)<sub>2</sub>C(CH<sub>3</sub>)<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 172.7 (C-5), 161.4 (CO<sub>2</sub>CH<sub>3</sub>), 159.4 (C-3), 155.9 (phenol C-3), 136.4 (ArC-2), 132.9 (ArC-3), 131.7 (q,  $J$  = 31.3 Hz, ArC-6), 130.8 (ArC-4), 129.7 (phenol C-5), 127.6 (ArC-1), 127.5 (phenol C-1), 124.8 (q,  $J$  = 5.0 Hz, ArC-5), 123.7 (phenol C-6), 123.0 (q,  $J$  = 274.4 Hz, CF<sub>3</sub>), 122.4 (phenol C-4), 120.8 (phenol C-2), 109.3 (C-4), 52.0 (CO<sub>2</sub>CH<sub>3</sub>), 25.8 (Si(CH<sub>3</sub>)<sub>2</sub>C(CH<sub>3</sub>)<sub>3</sub>), 18.3 (Si(CH<sub>3</sub>)<sub>2</sub>C(CH<sub>3</sub>)<sub>3</sub>), -4.3 (Si(CH<sub>3</sub>)<sub>2</sub>C(CH<sub>3</sub>)<sub>3</sub>). LC-MS (ESI): calc. for C<sub>24</sub>H<sub>26</sub>ClF<sub>3</sub>NO<sub>4</sub>Si [M+H]<sup>+</sup>: 512.12, observed: 512.25 ( $R_t$  = 9.62 min).

**3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(furan-2-yl)-1,2-oxazole-4-carbaldehyde (39).** Ester **34** (0.192 g, 0.516 mmol) was treated according to the General Procedure for conversion of esters to aldehydes, to furnish aldehyde **39** (0.169 g, 96%) as a colourless oil. The crude product was used without further purification.  $R_f$  = 0.51 (3:2 *n*-heptane/EtOAc); <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 10.20 (1H, s, CHO), 7.77 (1H, dd,  $J$  = 1.8, 0.8, furanyl H-5), 7.76 (1H, d,  $J$  = 8.0 Hz, Ar H-3 or ArH-5), 7.73 (1H, d,  $J$  = 7.8 Hz, ArH-3 or ArH-5), 7.60 (1H, app. t,  $J$  = 8.0 Hz, ArH-4), 7.55 (1H, dd,  $J$  = 3.6, 0.8 Hz, furanyl H-3), 6.72 (1H, dd,  $J$  = 3.6, 1.8 Hz, furanyl H-4); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 182.8 (CHO), 164.1 (C-5), 154.7 (C-3), 147.2 (furanyl C-3), 142.2 (furanyl C-2), 136.2 (ArC-2), 133.1 (ArC-3), 131.8 (q,  $J$  = 31.5 Hz, ArC-6), 131.3 (ArC-4), 125.6 (ArC-1), 125.0 (q,  $J$  = 5.0 Hz, ArC-5), 122.9 (q,  $J$  = 274.5 Hz, CF<sub>3</sub>), 116.9 (furanyl C-5), 115.2 (C-4), 112.9 (furanyl C-4). LC-MS (ESI): calc. for C<sub>15</sub>H<sub>8</sub>ClF<sub>3</sub>NO<sub>3</sub> [M+H]<sup>+</sup>: 342.01, observed: 342.08 ( $R_t$  = 6.93 min).

**3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(thiophen-2-yl)-1,2-oxazole-4-carbaldehyde (40).** Ester **35** (0.082 g, 0.210 mmol) was treated according to the General Procedure for conversion of esters to aldehydes. The crude product was purified by flash column chromatography, eluting with a gradient of 2-10% EtOAc in *n*-heptane, to furnish aldehyde **40** (0.048 g, 64%) as a white solid.  $R_f$  = 0.28 (4:1 *n*-heptane/EtOAc); <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 9.78 (1H, s, CHO), 8.32 (1H, dd,  $J$  = 3.9, 1.2 Hz, thiophenyl H-5), 7.80-7.75 (3H, m, ArH-3, ArH-5, thiophenyl H-3), 7.64 (1H, app. t,  $J$  = 8.1 Hz, ArH-4), 7.29 (1H, dd,  $J$  = 5.1, 3.9 Hz, thiophenyl H-4); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 182.2 (CHO), 168.0 (C-5), 159.0 (C-3), 136.6 (ArC-2), 133.4 (thiophenyl C-5), 132.9 (ArC-3), 132.8 (ArC-4), 132.3 (q,  $J$  = 31.5 Hz,

ArC-6), 131.6 (thiophenyl C-3), 128.7 (thiophenyl C-4), 127.0 (thiophenyl C-2), 125.3 (ArC-1), 125.2 (q,  $J = 5.0$  Hz, ArC-5), 122.9 (q,  $J = 274.5$  Hz,  $\text{CF}_3$ ), 114.4 (C-4). LC-MS (ESI): calc. for  $\text{C}_{15}\text{H}_8\text{ClF}_3\text{NO}_2\text{S}$   $[\text{M}+\text{H}]^+$ : 357.98, observed: 358.17 ( $R_t = 7.27$  min).

**3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-3-yl)-1,2-oxazole-4-carbaldehyde (41).** Ester **36** (0.228 g, 0.486 mmol) was treated according to the General Procedure for conversion of esters to aldehydes with a modification: the reduction step was performed with 3.0 equivalents of reducing agent and without cooling; this step occurred with concomitant loss of the Boc protecting group. The crude product was purified by flash column chromatography, eluting with 30% EtOAc in n-heptane, to furnish aldehyde **41** (0.084 g, 51%) as a brown solid.  $R_f = 0.20$  (7:3 n-heptane/EtOAc);  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 9.62 (1H, s, CHO), 8.77 (1H, br. s, NH), 8.26 (1H, app. dt,  $J = 3.3, 1.7$  Hz, pyrrole H-2), 7.78 (1H, d,  $J = 8.0$  Hz, ArH-3 or ArH-5), 7.75 (1H, d,  $J = 8.0$  Hz, ArH-3 or ArH-5), 7.61 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 7.02-6.95 (1H, m, pyrrole H-5), 6.95 (1H, dd,  $J = 2.7, 1.7$  Hz, pyrrole H-4);  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 182.8 (CHO), 170.7 (C-5), 160.0 (C-3), 133.6 (ArC-2), 133.3 (ArC-3), 132.3 (q,  $J = 31.5$  Hz, ArC-6), 131.4 (ArC-4), 125.7 (ArC-1), 125.1 (q,  $J = 5.0$  Hz, ArC-5), 123.8 (pyrrole C-2), 122.9 (q,  $J = 274.5$  Hz,  $\text{CF}_3$ ), 120.3 (pyrrole C-5), 113.2 (pyrrole C-3), 110.7 (C-4), 108.6 (pyrrole C-4). LC-MS (ESI): calc. for  $\text{C}_{15}\text{H}_9\text{ClF}_3\text{N}_2\text{O}_2$   $[\text{M}+\text{H}]^+$ : 341.02, observed: 341.08 ( $R_t = 6.33$  min).

**3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(naphthalen-1-yl)-1,2-oxazole-4-carbaldehyde (42).** Ester **37** (0.068 g, 0.160 mmol) was treated according to the General Procedure for conversion of esters to aldehydes. The crude product was purified by flash column chromatography, eluting with 17% EtOAc in n-heptane, to furnish aldehyde **42** (0.039 g, 61%) as a white solid.  $R_f = 0.28$  (4:1 n-heptane/EtOAc);  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 9.73 (1H, s, CHO), 8.14 (1H, d,  $J = 8.5$ , naphthyl-H), 8.02-7.98 (2H, m, naphthyl-H), 7.83 (1H, dd,  $J = 7.1, 1.2$ , naphthyl-H), 7.79 (1H, d,  $J = 8.0$ , ArH-3 or ArH-5), 7.76 (1H, d,  $J = 8.0$ , ArH-3 or ArH-5), 7.68-7.61 (4H, m, ArH-4 and naphthyl-H);  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 183.1 (CHO), 176.6 (C-5), 157.4 (C-4), 136.3, 133.8, 133.1, 132.9, 131.8 (q,  $J = 31.5$ , ArC-6), 131.4, 131.2, 130.1, 128.9, 128.5, 127.3, 126.1 (ArC-1), 125.1, 125.0 (q,  $J = 5.0$ , ArC-5), 124.7, 123.1 (q,  $J = 274.4$ ,  $\text{CF}_3$ ), 122.5, 118.7 (C-4), (not all peaks could be precisely assigned with certainty). LC-MS (ESI): calc. for  $\text{C}_{21}\text{H}_{12}\text{ClF}_3\text{NO}_2$   $[\text{M}+\text{H}]^+$ : 402.04, observed: 401.92 ( $R_t = 8.15$  min).

**5-(3-((*Tert*-butyldimethylsilyl)oxy)phenyl)-3-(2-chloro-6-(trifluoromethyl)phenyl)-1,2-oxazole-4-carbaldehyde (43).** Ester **38** (0.173 g, 0.340 mmol) was treated according to the General Procedure for conversion of esters to aldehydes. The crude product was purified by flash column chromatography, eluting with a gradient of 0-10% EtOAc in n-heptane, to furnish aldehyde **43** (0.123 g, 80%) as a white solid.  $R_f = 0.34$  (9:1 n-heptane/EtOAc);  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 9.94 (1H, s, CHO), 7.77 (1H, d,  $J = 8.0$  Hz, ArH-3 or ArH-5), 7.74 (1H, d,  $J = 8.0$  Hz, ArH-3 or ArH-5), 7.63-7.59 (2H, m, phenol H-4 and phenol H-5), 7.51 (1H, app. t,  $J = 2.0$  Hz, phenol H-2), 7.46 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 7.11 (1H, ddd,  $J = 8.2, 2.0, 1.0$  Hz, phenol H-6), 1.02 (9H, s,  $\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$ ), 0.27 (6H, s,  $\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$ );  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 182.8 (CHO), 174.4 (C-5), 158.5 (C-3), 156.6 (phenol C-3), 136.4 (ArC-2), 133.2 (ArC-3), 132.0 (q,  $J = 31.6$  Hz, ArC-6), 131.3 (ArC-4), 130.6 (phenol C-5), 126.8 (ArC-1), 126.0 (phenol C-1), 125.0 (q,  $J = 5.1$  Hz, ArC-5), 124.6 (phenol C-6), 123.0 (q,  $J = 274.5$  Hz,  $\text{CF}_3$ ), 122.0 (phenol C-4), 120.3 (phenol C-2), 116.3 (C-4), 25.6 ( $\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$ ), 18.4 ( $\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$ ), -4.2 ( $\text{Si}(\text{CH}_3)_2\text{C}(\text{CH}_3)_3$ ). LC-MS (ESI): calc. for  $\text{C}_{23}\text{H}_{24}\text{ClF}_3\text{NO}_3\text{Si}$   $[\text{M}+\text{H}]^+$ : 482.11, observed: 482.17 ( $R_t = 9.45$  min).

**Indazole 3 and thienopyrazole 5.** Indazole **3** and thienopyrazole **5** were prepared according to the synthetic routes described previously.<sup>31,48</sup>

## Biophysical assays

**ROR $\gamma$ t LBD expression and purification (used for TR-FRET assays).** A pET15b expression vector encoding the human ROR $\gamma$ t LBD (residues 265-518) with an N-terminal His<sub>6</sub>-tag was transformed by heat shock into BL21(DE3) *E. coli* cells. Single colonies were used to inoculate pre-cultures of 8 mL LB-media containing 100  $\mu$ g/mL ampicillin. After overnight incubation at 37 °C, each pre-culture was transferred to 1 L TB media supplemented with ampicillin (100  $\mu$ g/mL) and incubated at 37 °C until an OD<sub>600</sub> = 1.0 was reached. Protein expression was then induced with 0.5 mM isopropyl-b-d-thiogalactoside (IPTG) and cultures were incubated for 16 h at 18 °C. The cells were collected by centrifugation and suspended in lysis buffer (300 mM NaCl, 20 mM TrisHCl pH 8.0, 20 mM imidazole, 1 mM TCEP, 10% v/v glycerol, cComplete™, EDTA-free Protease Inhibitor Cocktail tablets (1 tablet/50 mL lysate) and benzonase (0.1  $\mu$ L/ 1 mL)). After lysis using a C3 Emulsiflex-C3 homogeniser (Avestin), the cell lysate was cleared by centrifugation at 4 °C and the protein was purified via Ni<sup>2+</sup> affinity column chromatography. Fractions containing the protein of interest were combined and dialysed against: 150 mM NaCl, 20 mM TrisHCl pH 8.0, 5 mM DTT and 10% v/v glycerol.

**TR-FRET coactivator recruitment assays.** Assays were conducted using 100 nM N-terminal biotinylated SRC-1 box-2 coactivator peptide (Biotin-N-PSSHSSLTARHKILHRLQEGSPSD-CONH<sub>2</sub>) and 20 nM His<sub>6</sub>-ROR $\gamma$ t-LBD in buffer containing 10 mM HEPES, 150 mM NaCl, 5 mM DTT, 0.1% BSA (w/v) and 0.1 mM CHAPS, pH 7.5. A terbium-labelled anti-His antibody (CisBio Bioassays, 61HISTLA) and d2-labelled streptavidin (CisBio Bioassays, 610SADLA) were used at the concentrations recommended by the supplier. Compounds (dissolved in DMSO) were titrated using a 2 x dilution series in Corning white low volume, low binding, 384-well plates at a final volume of 10  $\mu$ L. The final DMSO concentration was 2% v/v throughout. The plate was incubated at room temperature for 30 min and centrifuged before reading (excitation = 340 nm; emission = 665 nm and 620 nm) on a Tecan infinite F500 plate reader using the parameters recommended by CisBio Bioassays. The data were analyzed with GraphPad Prism 7.0 Software. The dose-response curves were fitted represented by:

$$y = A_1 + \frac{A_2 - A_1}{1 + \frac{x^p}{x_0^p}}$$

where y is the FRET ratio ((acceptor/donor)\*1000), A<sub>1</sub> is the bottom asymptote, A<sub>2</sub> is the top asymptote, p is the Hill slope, x is the ligand concentration in  $\mu$ M and x<sub>0</sub> is the IC<sub>50</sub> value in  $\mu$ M. Where dose-response curves did not reach a bottom asymptote this was fixed at the value of the negative control. Data were recorded in triplicate; error shown is standard deviation from the mean; curves are representative of > 3 independent experiments.

**Competition TR-FRET coactivator recruitment assays.** Competition assays were performed in an analogous fashion to that described above, only in the presence of fixed concentrations of cholesterol: 0.00  $\mu$ M (DMSO), 0.25  $\mu$ M, 1.00  $\mu$ M such that the final concentration of DMSO remained at 1.2% v/v.

**TR-FRET AlexaFluor-MRL-871 recruitment assays.** Assays were conducted using 100 nM AlexaFluor647-labelled MRL-871<sup>31</sup> (determined as the concentration at 50% of the maximum response when titrating the probe to the ROR $\gamma$ t LBD) and 20 nM His<sub>6</sub>-ROR $\gamma$ t-LBD in buffer as described above. A terbium-labelled anti-His antibody (CisBio Bioassays, 61HISTLA) was used at the concentration recommended by the supplier. The assay was carried out in the same manner as described above for the coactivator recruitment assay.

**Competition TR-FRET coactivator recruitment assays on PPAR $\gamma$ .** TR-FRET competition assays were performed in an analogous fashion to that described above, only using 100 nM His<sub>6</sub>-PPAR $\gamma$ -LBD instead of 20 nM His<sub>6</sub>-ROR $\gamma$ t-LBD. The assay was performed in the presence of 1  $\mu$ M rosiglitazone, in order to initially activate PPAR $\gamma$ .

### Protein X-ray crystallography

**RORyt LBD expression and purification (used for crystallography).** A pET15b expression vector was ordered from GenScript encoding for the RORyt LBD (residues 265–507) containing a C455H mutation (RORytC455H) and a C-terminal His-tag. The plasmid was transformed by heat shock into BL21(DE3) *E. coli* cells. A single colony was used to start three pre-cultures of 24 mL LB medium containing 100 µg/mL ampicillin. After overnight incubation at 37 °C, each pre-culture was transferred to 2 liters of 2 x YT medium supplied with 0.05% antifoam SE-15 (Sigma-Aldrich). These cultures were incubated until they reached an OD<sub>600</sub> = 0.6. Protein expression was induced by adding 0.25 mM IPTG. The temperature was decreased to 15 °C and protein expression proceeded overnight. The cells were collected by centrifugation at 10,000 RCF for 10 min at 4 °C. The resulting 30 grams of cell pellet was dissolved in lysis buffer (20 mM Tris, 500 mM NaCl, 2 mM TCEP, 0.1% Tween20, 10% glycerol, 10 cOmplete™ Protease Inhibitor Cocktail tablets (Roche) and 25 U/ml Bezonase® Nuclease (Millipore), adjusted to pH=8.0). After cell lysis using an Emulsiflex-C3 homogeniser (Avestin), the cell lysate was cleared by centrifugation at 40,000 RCF for 40 min at 4 °C and the supernatant was loaded on a 5 mL Ni-NTA Superflow cartridge (QIAGEN) pre-equilibrated with buffer A (20 mM Tris, 500 mM NaCl, 2 mM TCEP, 0.1% Tween20 and 10% glycerol). The column was washed with 10 CVs of buffer A supplied with 20 mM and sequentially with 10 CVs of Buffer A supplied with 50 mM imidazole. The protein was eluted from the resin using eight column volumes elution buffer (buffer A supplied with 200 mM imidazole). The purified protein was then dialysed overnight to buffer A containing 1.2 U of restriction-grade thrombin (Millipore) per milligram of purified protein to remove the His-tag. Next, the protein mixture was concentrated using an Amicon® Ultra centrifugal filter with a 10-kDa cutoff (Millipore) and loaded on a Superdex 75 pg 16/60 size-exclusion column (GE Life Sciences) using 20 mM Tris, 100 mM NaCl and 5 mM DTT (adjusted to pH=8.0) as running buffer. The flow-through was collected as 3 mL fractions which were analyzed using Q-TOF LC/MS. The fractions containing RORytC455H were combined and concentrated to a final concentration of 11.1 mg/mL. The concentrated protein sample was then aliquoted, flash-frozen and stored at -80 °C.

**X-ray crystallography.** The RORytC455H solution (11.1 mg/ml) was mixed with 2 equivalents of ligand and incubated on ice for 1 h. Next, the sample was centrifuged at 20,000 RCF for 20 min at 4 °C to remove precipitated proteins. MRC-2 well crystallization plates (Hampton Research, sitting drop) were prepared using a Mosquito pipetting robot (TTP Labtech). Well-diffracting crystals were obtained by mixing 0.9 µL of protein solution with 0.3 µL of 1.6–2.0 M ammonium sulfate and 0.1 M Tris (pH=8.5). The well was filled with 80 µL precipitant solution and plates were placed at 20 °C. Crystals could be observed after 1 h of incubation and grew to their final size overnight. The crystals were cryoprotected by transferring the crystals briefly to a solution containing 1.2 M AmSO<sub>4</sub>, 0.1 M Tris (pH=8.5) and 25% glycerol before being flash cooled in liquid N<sub>2</sub>. Diffraction data were collected at 100 K at the P11 beamline of the PETRA III facility at DESY (Hamburg, Germany) and processed using the CCP4 suite (version 7.0.075).<sup>49</sup> DIALS was used to integrate and scale the data.<sup>50</sup> The data was phased with PHASER using 5C4O as a molecular replacement model and ligand restraints of 25 were generated with AceDRG.<sup>51,52</sup> Sequential model building and refinement were performed with COOT and REFMAC, respectively.<sup>53,54</sup> PyMOL (version 2.2.3, Schrödinger) was used to make the figures.<sup>55</sup> The structure of RORytC455 in complex with 25 was deposited in the protein data bank (PDB) under code 6SAL.

**Quantitative IL-17a mRNA RT-PCR assay.** EL4 cells (Sigma-Aldrich) were grown in DMEM (Gibco) supplemented with 10% FBS and 1% penicillin-streptomycin. At 24 h after the cells were seeded onto a 12-well plate (300,000 cells/well), the cells were incubated with 10 µM compound (from a 10 mM stock in DMSO) or DMSO for 24 h, and activated with phorbol 12-myristate 13-acetate (PMA, 50 ng/mL; Sigma Aldrich) and ionomycin (1 µg/mL; Sigma Aldrich) for 5 h. The cells were then collected and RNA was isolated using a RNeasy Mini Kit (Qiagen) and reverse transcribed using the iScript Advanced cDNA Synthesis Kit (Bio-Rad). Quantitative RT-PCR was performed to analyze mRNA levels of mouse IL-17a

levels (in triplicate) using SYBR green technology (Bio-Rad) on a CFX Real-Time System (Bio-Rad). The following primer assays were purchased from Bio-Rad: IL-17a (qMmuCID0026592) and Gapdh (qMmuCED0027497). The level of IL-17a mRNA expression was normalized to that of Gapdh expression. The relative gene expression was calculated by the  $2^{-\Delta\Delta C_t}$  (Livak) method using the DMSO control as calibrator. Statistical analysis was performed using a one-way analysis of variance (ANOVA) comparing against the DMSO control following Dunnett's *post hoc* test (GraphPad Prism 7.0 software). A *P* value <0.05 was considered statistically significant. Data were recorded in triplicate; error shown is standard deviation from the mean; data are representative of three independent experiments.

### ADME experiments

**Chemical stability.** Chemical stability was determined by incubating test compounds at a final concentration of 2  $\mu$ M in aqueous buffer at pH 7.4 for 1, 7 and 24 h, respectively. The percentage of remaining compound (% remain) in relation to the zero time point was calculated following LC-MS-based measurement of sample aliquots of each time point.

**Kinetic solubility.** Aqueous solubility of compounds was determined by spectrophotometrical measurement of the kinetic solubility of a 500  $\mu$ M compound solution in HEPES buffer (pH 7.4) compared to a solution containing 50% of the organic solvent acetonitrile. To this end, saturated samples of the compounds in buffer were prepared starting from DMSO stocks, and samples were shaken for 90 min at room temperature. Precipitated material was removed by filtration and samples were further diluted with the same volume of acetonitrile. Absorbance spectra (250–500 nm) were recorded and relative solubility was calculated in comparison to the acetonitrile-buffer solution of the compounds using absorbance ratios.

**PAMPA.** Permeability through artificial membranes (PAMPA) was performed at an initial concentration of 500  $\mu$ M of the compound in the donor compartment. After an incubation period of 20 h, absorption of the receiver wells was measured by spectrophotometry and permeation was calculated by normalization of the compound flux across a blank filter.

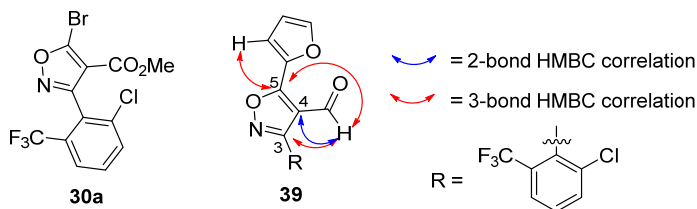
**Microsomal stability phase I.** Metabolic stability under oxidative conditions was measured using NADPH-supplemented human liver microsomes. Compound depletion was analysed by LC-MS/MS at a concentration of 1  $\mu$ M over time up to 50 min at 37 °C. Based on compound half-life  $t_{1/2}$ , *in vitro* intrinsic clearance  $Cl_{int}$  was calculated.  $Cl_{int} = \frac{V \times 0.693}{t_{1/2} \times mg}$ , with *V* = assay volume and *mg* = amount of microsomal protein.

**Microsomal stability phase II.** Metabolic stability under conjugative conditions was measured in the glucuronidation assay by LC-MS-based determination of % remaining of selected compounds. Prior to the assay, human liver microsomes were activated using alamethicin and further supplemented with 5 mM of the cofactor UDPGA. Compound (5  $\mu$ M) was added to the reaction mixture and it was incubated for 1 h at 37 °C.

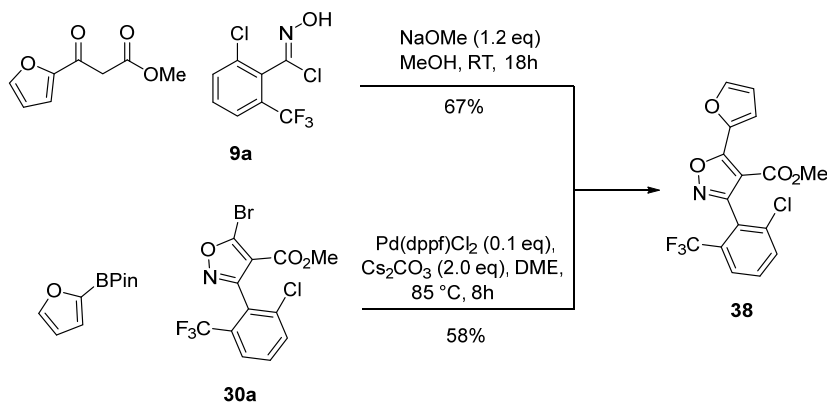
**Plasma stability.** Plasma stability was measured by LC-MS-based determination of % remaining of selected compounds at a concentration of 5  $\mu$ M after incubation in 50% plasma obtained from different species for 1 h at 37 °C.

**Plasma protein binding.** Plasma protein binding was determined by equilibrium dialysis. Plasma containing 5  $\mu$ M test compound was allowed to equilibrate with the buffer compartment for 6 h at 37 °C. Compound concentrations on both sides of the semi-permeable membrane were analysed by LC-MS/MS and % bound was calculated.

## Supporting Information

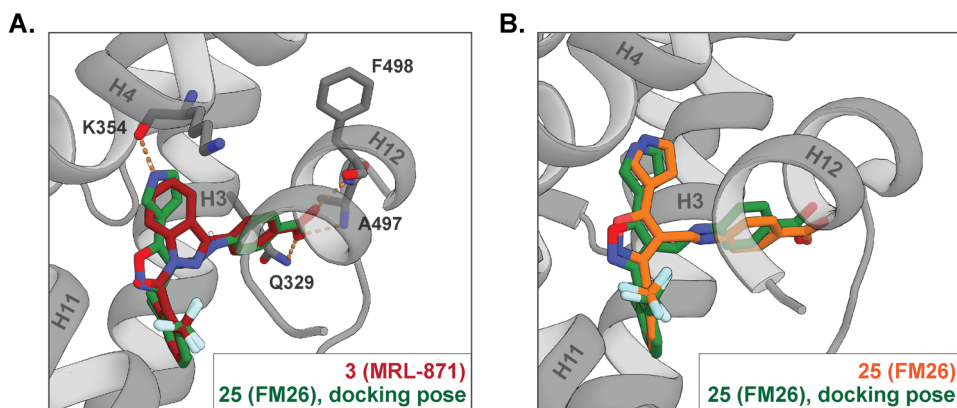


**Figure S2.1 | HMBC NMR correlations proving regiochemical assignment of bromo-isoxazole **30a** based on NMR analysis of aldehyde analogue **39**.** The regiochemistry of bromide **30a** was retrospectively determined by HMBC NMR analysis of aldehydes, exemplified by **39**. A 2-bond HMBC correlation between the aldehyde proton and distinctive C-4 carbon (112.9 ppm) was observed, as were 3-bond correlations with C-3 (154.7 ppm) and C-5 (164.1 ppm) carbon atoms. A HMBC 3-bond correlation was also observed between the 3-furyl proton and distinctive C-5 (164.1 ppm) carbon atom but not between any other isoxazole carbon atoms. Together, this is clear evidence for the aldehyde group being at the isoxazole C-4 position and is in agreement with literature precedent.<sup>45</sup> Such correlations were observed in all cases.

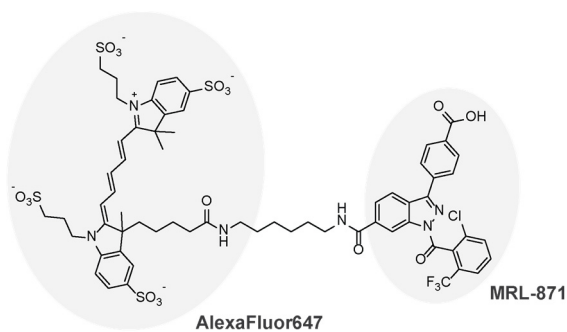


**Scheme S2.1 | Synthesis of isoxazole **38** via two independent routes.** To provide further evidence in support of the assignment in Figure S2.1, the 5-furyl isoxazole **38** was prepared via an independent synthetic route with good literature precedent for regioselectivity.<sup>56</sup> The NMR characterization data for the products of both synthetic routes matched exactly and thus provided further indication of the correct regiochemical assignment.

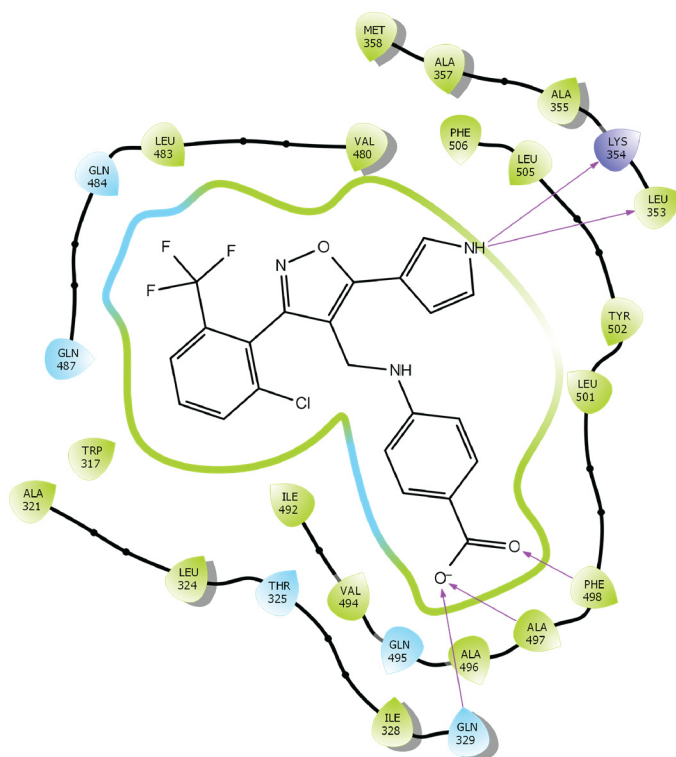




**Figure S2.2** | *In silico* modelled docking pose of 25 in complex with the ROR $\gamma$ t LBD (PDB: 4YPQ). **A)** Overlay of the docking pose of 25 (green sticks) with the co-crystal structure of the ROR $\gamma$ t in complex with 3 (red sticks) (PDB: 4YPQ). **B)** Overlay of the docking pose of 25 (green sticks) with the co-crystal structure of ROR $\gamma$ t in complex with 25 (orange sticks) (PDB: 6SAL).



**Figure S2.3** | Chemical structure of AlexaFluor647-labelled MRL-871 probe, used for TR-FRET AlexaFluor-MRL recruitment assays.



**Figure S2.4 | 2D Protein-ligand interaction plot showing the interactions between 25 and the surrounding amino acids in the allosteric binding site of RORyt.** Hydrophobic amino acid residues are indicated in green, polar residues in blue and positively charged residues in purple. Purple arrows indicate hydrogen bond interactions. The interaction diagram was generated with Maestro (Schrödinger LLC, version 12.3 (2020-1)), using the co-crystal structure of RORyt in complex with compound 25 (PDB: 6SAL).

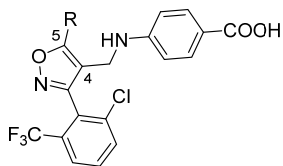
Table S2.1 & Table S2.2 | Top 10 ranked ligands based on Glide Scores from a docking screen of the Chem-T generated C-5 SAR virtual library (2.1) and the *de novo* C-5 SAR virtual library (2.2).

Table S2.1














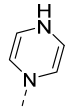
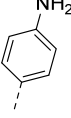
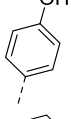




Rank	R (C-5 substituent)	Glide Score
1	 25 (FM26)	-15.74
2		-14.29
3	 23	-14.27
4		-14.26
5		-14.20
6	 24	-14.17
7		-14.13
8		-14.07
9		-14.00
10	 19	-13.99

Table S2.2

Rank	R (C-5 substituent)	Glide Score
1		-15.36
2		-14.61
3	 27	-14.60
4	3 (MRL-871)	-14.58
5		-14.52
6		-14.39
7		-14.09
8		-13.84
9		-13.80
10		-13.79
24	5 (Glenmark's compound 13)	-13.11
40	 26	-10.84

**Table S2.3 | Data collection and refinement statistics for the crystal structure of RORyt in complex with ligand 25 (FM26) (PDB: 6SAL).**

RORyt in complex with ligand 25	
<i>Data collection</i>	
Space group	P 6 <sub>1</sub> 2 2
Cell dimensions	
a, b, c (Å)	107.6, 107.6, 100.1
α, β, γ (°)	90, 90, 120
Resolution (Å)	47.38 – 1.61 (1.67 – 1.61)
<i>I</i> / σ( <i>I</i> )	25.0 (1.2)
Completeness (%)	97.8 (81.6)
Redundancy	28.3 (7.4)
CC <sub>1/2</sub>	0.992 (0.439)
<i>Refinement</i>	
No. reflections	43698
R <sub>work</sub> /R <sub>free</sub>	0.181/0.203
No. atoms	
Protein	1992
Ligand/ion	32
Water	257
<i>B</i> -factors	
Protein	37.26
Ligand/ion	29.11
Water	49.88
R.m.s. deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.13

## References

1. Wang, J. *et al.* RORy drives androgen receptor expression and represents a therapeutic target in castration-resistant prostate cancer. *Nat. Med.* **22**, 488–496 (2016).
2. Zhang, Y. *et al.* Discovery and Characterization of XY101, a Potent, Selective, and Orally Bioavailable RORy Inverse Agonist for Treatment of Castration-Resistant Prostate Cancer. *J. Med. Chem.* **62**, 4716–4730 (2019).
3. Ivanov, I. I. *et al.* The Orphan Nuclear Receptor RORyt Directs the Differentiation Program of Proinflammatory IL-17+ T Helper Cells. *Cell* **126**, 1121–1133 (2006).
4. Manel, N., Unutmaz, D. & Littman, D. R. The differentiation of human TH-17 cells requires TGF- $\beta$  and induction of the nuclear receptor RORyt. *Nat. Immunol.* **9**, 641–649 (2008).
5. Yang, X. O. *et al.* T Helper 17 Lineage Differentiation Is Programmed by Orphan Nuclear Receptors ROR $\alpha$  and RORy. *Immunity* **28**, 29–39 (2008).
6. Miossec, P. & Kolls, J. K. Targeting IL-17 and TH17 cells in chronic inflammation. *Nat. Rev. Drug Discov.* **11**, 763–776 (2012).
7. Burkett, P. R. & Kuchroo, V. K. IL-17 Blockade in Psoriasis. *Cell* **167**, 1669 (2016).
8. Lock, C. *et al.* Gene-microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis. *Nat. Med.* **8**, 500–508 (2002).
9. Duerr, R. H. *et al.* A Genome-Wide Association Study Identifies IL23R as an Inflammatory Bowel Disease Gene. *Science* **314**, 1461–1463 (2006).
10. Hueber, W. *et al.* Effects of AIN457, a fully human antibody to interleukin-17A, on psoriasis, rheumatoid arthritis, and uveitis. *Sci. Transl. Med.* **2**, 52ra72 (2010).
11. Attia, A. *et al.* Safety and Efficacy of Brodalumab for Moderate-to-Severe Plaque Psoriasis: A Systematic Review and Meta-Analysis. *Clin. Drug Investig.* **37**, 439–451 (2017).
12. Mease, P. J. *et al.* Ixekizumab, an IL-17A specific monoclonal antibody, for the treatment of biologic-naïve patients with active psoriatic arthritis. *Ann. Rheum. Dis.* **76**, 79–87 (2017).
13. Fauber, B. P. & Magnuson, S. Modulators of the nuclear receptor retinoic acid receptor-related orphan receptor- $\gamma$  (RORy or RORc). *J. Med. Chem.* **57**, 5871–5792 (2014).
14. Bronner, S. M., Zbieg, J. R. & Crawford, J. J. RORgamma antagonists and inverse agonists: a patent review. *Expert Opin. Ther. Pat.* **27**, 101–112 (2017).
15. Pandya, V. B., Kumar, S., Sachchidanand, Sharma, R. & Desai, R. C. Combating Autoimmune Diseases With RORy Inhibitors: Hits and Misses. *J. Med. Chem.* **61**, 10976–10995 (2018).
16. Tanis, V. M. *et al.* 3-Substituted Quinolines as RORyt Inverse Agonists. *Bioorg. Med. Chem. Lett.* **29**, 1463–1470 (2019).
17. Imura, C. *et al.* A novel RORyt inhibitor is a potential therapeutic agent for the topical treatment of psoriasis with low risk of thymic aberrations. *J. Dermatol. Sci.* **93**, 176–185 (2019).
18. Tian, J. *et al.* Discovery of N-indanyl benzamides as potent RORyt inverse agonists. *Eur. J. Med. Chem.* **167**, 37–48 (2019).
19. Kotoku, M. *et al.* Discovery of Second Generation RORy Inhibitors Composed of an Azole Scaffold. *J. Med. Chem.* **62**, 2837–2842 (2019).
20. Duan, J. J.-W. *et al.* Structure-based Discovery of Phenyl (3-Phenylpyrrolidin-3-yl)sulfones as Selective, Orally Active RORyt Inverse Agonists. *ACS Med. Chem. Lett.* **10**, 367–373 (2019).
21. Gronemeyer, H., Gustafsson, J.-Å. & Laudet, V. Principles for modulation of the nuclear receptor superfamily. *Nat. Rev. Drug Discov.* **3**, 950–964 (2004).
22. Li, X. *et al.* Structural studies unravel the active conformation of apo RORyt NR and a common inverse agonism of two diverse classes of RORyt inhibitors. *J. Biol. Chem.* **292**, 11618–11630 (2017).
23. Harris, J. M., Lau, P., Chen, S. L. & Muscat, G. E. O. Characterization of the retinoid orphan-related receptor- $\alpha$  coactivator binding interface: structural basis for ligand-independent transcription. *Mol. Endocrinol.* **16**, 998–1012 (2002).
24. Hu, X. *et al.* Sterol metabolism controls TH17 differentiation by generating endogenous RORy agonists. *Nat. Chem. Biol.* **11**, 141–147 (2015).
25. Huh, J. R. *et al.* Digoxin and its derivatives suppress TH17 cell differentiation by antagonizing RORyt activity. *Nature* **472**, 486–490 (2011).
26. Kumar, N. *et al.* The benzenesulfoamide To901317 is a novel retinoic acid receptor-related orphan receptor- $\alpha$ /gamma inverse agonist. *Mol. Pharmacol.* **77**, 228–236 (2010).
27. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nat. Rev. Drug Discov.* **5**, 993–996 (2006).
28. Meijer, F. A., Leijten-van de Gevel, I. A., de Vries, R. M. J. M. & Brunsveld, L. Allosteric small molecule modulators of nuclear receptors. *Mol. Cell. Endocrinol.* **485**, 20–34 (2019).

29. Tice, C. M. & Zheng, Y.-J. Non-canonical modulators of nuclear receptors. *Bioorg. Med. Chem. Lett.* **26**, 4157–4164 (2016).
30. Moore, T. W., Mayne, C. G. & Katzenellenbogen, J. a. Minireview: Not picking pockets: nuclear receptor alternate-site modulators (NRAMs). *Mol. Endocrinol.* **24**, 683–695 (2010).
31. Scheepstra, M. *et al.* Identification of an allosteric binding site for ROR $\gamma$ t inhibition. *Nat. Commun.* **6**, e8833 (2015).
32. Ouvry, G. *et al.* Discovery of phenoxyindazoles and phenylthioindazoles as ROR $\gamma$  inverse agonists. *Bioorganic Med. Chem. Lett.* **26**, 5802–5808 (2016).
33. Guo, Y. *et al.* Inhibition of ROR $\gamma$ T Skews TCR $\alpha$  Gene Rearrangement and Limits T Cell Repertoire Diversity. *Cell Rep.* **17**, 3206–3218 (2016).
34. Zhong, C. & Zhu, J. Small-Molecule ROR $\gamma$ t Antagonists: One Stone Kills Two Birds. *Trends Immunol.* **38**, 229–231 (2017).
35. Chaudari, S. S. *et al.* Bicyclic Heterocyclic Compounds as ROR gamma Modulators. *PCT Int. Appl.* WO2015/008234 (2015).
36. de Vries, R. M. J. M., Meijer, F. A., Doveston, R. G. & Brunsvel, L. Elucidation of an Allosteric Mode of Action for a Thienopyrazole ROR $\gamma$ t Inverse Agonist. *ChemMedChem* **15**, 561–565 (2020).
37. Dixon, S. L. *et al.* PHASE: A new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J. Comput. Aided. Mol. Des.* **20**, 647–671 (2006).
38. Dixon, S. L., Smondyrev, A. M. & Rao, S. N. PHASE: A novel approach to pharmacophore modeling and 3D database searching. *Chem. Biol. Drug Des.* **67**, 370–372 (2006).
39. Asinex Corp. Gold and platinum collections, release: 1994-2004, update: 2016-10. [www.asinex.com/libraries-html/libraries\\_gold\\_platinum-html](http://www.asinex.com/libraries-html/libraries_gold_platinum-html).
40. Kalin, J. H., Zhang, H., Gaudrel-Grosay, S., Vistoli, G. & Kozikowski, A. P. Chiral mercaptoacetamides display enantioselective inhibition of histone deacetylase 6 and exhibit neuroprotection in cortical neuron models of oxidative stress. *ChemMedChem* **7**, 425–39 (2012).
41. Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).
42. Halgren, T. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **47**, 1750–1759 (2004).
43. Abreu, R. M. V., Froufe, H. J. C., Daniel, P. O. M., Queiroz, M. J. R. P. & Ferreira, I. C. F. R. ChemT, an open-source software for building template-based chemical libraries. *SAR QSAR Environ. Res.* **22**, 603–610 (2011).
44. Leroy, J. Preparation of 3-bromopropiolic esters: methyl and tert-butyl 3-bromopropiolates. *Org. Synth.* **74**, 212 (1997).
45. Oakdale, J. S., Sit, R. K. & Fokin, V. V. Ruthenium-Catalyzed Cycloadditions of 1-Haloalkynes with Nitrile Oxides and Organic Azides: Synthesis of 4-Haloisoxazoles and 5-Halotriazoles. *Chem. - A Eur. J.* **20**, 11101–11110 (2014).
46. Buckman, D. O. *et al.* Lysophosphatidic acid receptor antagonists. *PCT Int. Appl.* WO2014/1113485 (2014).
47. Cuzzocrea, S. *et al.* Rosiglitazone, a ligand of the peroxisome proliferator-activated receptor-gamma, reduces acute inflammation. *Eur. J. Pharmacol.* **483**, 79–93 (2004).
48. Gege, C. Retinoid-related orphan receptor  $\gamma$  t modulators: comparison of Glenmark's me-too patent application (WO2015008234) with the originator application from Merck Sharp and Dohme (WO2012106995). *Expert Opin. Ther. Pat.* **25**, 1215–1221 (2015).
49. Potterton, L. *et al.* CCP412: The new graphical user interface to the CCP 4 program suite. *Acta Crystallogr. D Struct. Biol.* **74**, 68–84 (2018).
50. Clabbers, M. T. B. *et al.* Electron diffraction data processing with DIALS. *Acta Crystallogr. D Struct. Biol.* **74**, 506–518 (2018).
51. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D. Struct. Biol.* **63**, 32–41 (2007).
52. Long, F. *et al.* AceDRG : a stereochemical description generator for ligands. *Acta Crystallogr. D Struct. Biol.* **73**, 112–122 (2017).
53. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Struct. Biol.* **66**, 486–501 (2010).
54. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Struct. Biol.* **67**, 355–67 (2011).
55. The PyMOL Molecular Graphics System, Version 2.2.3. (Schrödinger LLC, 2015).
56. Huang, H. *et al.* Design, Synthesis, and Biological Evaluation of Novel Nonsteroidal FXR Antagonists: Molecular Basis of FXR Antagonism. *ChemMedChem* **10**, 1184–1199 (2015).

# Chapter 3

## Structure Activity Relationship Studies of Trisubstituted Isoxazoles as Selective Allosteric Ligands for ROR $\gamma$ t

### Abstract

Inhibition of ROR $\gamma$ t is a promising strategy in the treatment of autoimmune diseases. Next to the classical binding pocket, ROR $\gamma$ t also features a unique allosteric binding site in its ligand binding domain. The majority of allosteric ROR $\gamma$ t modulators identified thus far are all based on closely related chemotypes. Chapter 2 described the identification of trisubstituted isoxazoles as a novel class of allosteric ROR $\gamma$ t inverse agonists, with **FM26** as the most potent hit compound. Utilizing a combination of X-ray crystallography studies and biochemical assays, we explore the structure activity relationship (SAR) profile of these isoxazoles, focusing on the underexplored C-4 and C-5 positions. Analysis showed changing the C-4 amine linker to a more lipophilic linker leads to a ~10 fold increase in potency, resulting in low nanomolar lead compounds. Additionally, the study highlights the importance of featuring a hydrogen bond donor at the C-5 position. The most potent compounds showed a significant cellular activity, an increased selectivity profile, and promising pharmacokinetic properties. The combination of these results and the co-crystal structures of ROR $\gamma$ t in complex with several of the novel analogues provide essential insights into the SAR of this trisubstituted isoxazole class of allosteric ligands for ROR $\gamma$ t. Additionally, these results provide entry points for optimization of allosteric ROR $\gamma$ t ligands with other scaffolds and highlight the potential for further research on allosteric targeting of ROR $\gamma$ t.

**This chapter has been published as:** F.A. Meijer, A.O.W.M. Saris, R.G. Doveston, G.J.M. Oerlemans, R.M.J.M. de Vries, B.A. Somsen, A. Unger, B. Klebl, C. Ottmann, P.J. Cossar & L. Brunsveld. Structure Activity Relationship Studies of Trisubstituted Isoxazoles as Selective Allosteric Ligands for ROR $\gamma$ t. *J. Med. Chem.* In press (2021).

**A patent application has been granted for a selection of ligands shown in this chapter:** L. Brunsveld, R.G. Doveston, S. Leysen, L.-G. Milroy, F.A. Meijer, M. Scheepstra & C. Ottmann. Substituted Heterocyclic Compounds and their Use as ROR $\gamma$ t Inhibitors. *PCT Int. Appl.* WO2020/149740 (2020).

## Introduction

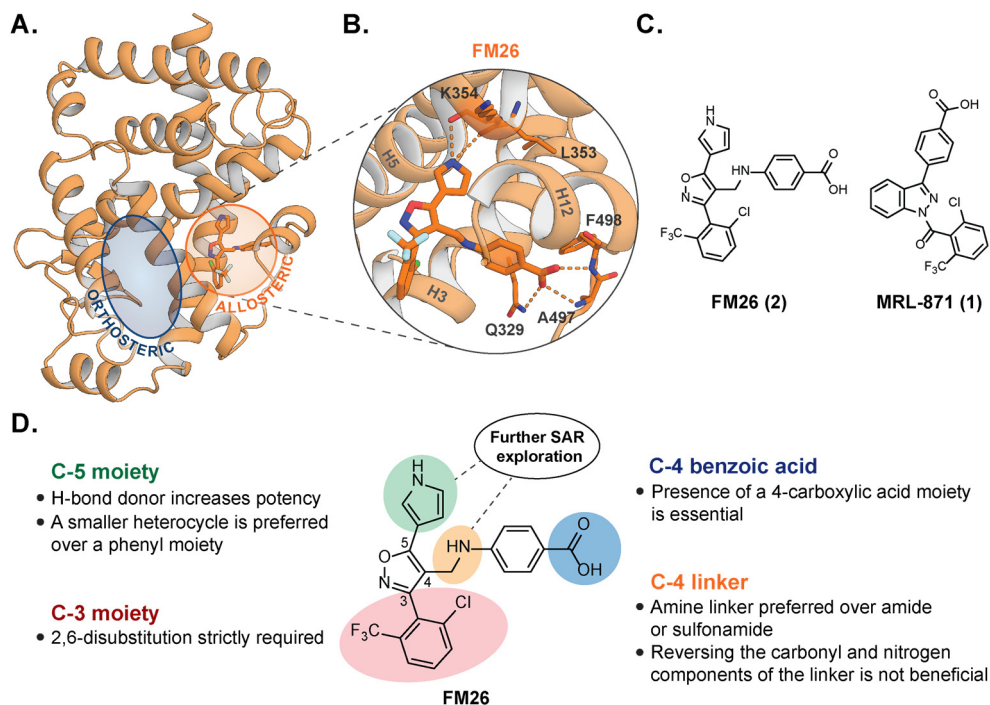
Nuclear receptors (NRs) are a family of ligand-dependent transcription factors that are highly attractive drug targets because of their central role in several regulatory processes in the human body.<sup>1–4</sup> Within the NR family, ROR $\gamma$ t has received increased attention because of its essential role in the immune system.<sup>5</sup> ROR $\gamma$ t is a key regulator in the differentiation of naïve CD4<sup>+</sup> T cells into T helper 17 (Th17) cells, and the production of the pro-inflammatory cytokine IL-17a.<sup>5–7</sup> Elevated IL-17a levels are associated with the development of autoimmune diseases, including psoriasis, multiple sclerosis and rheumatoid arthritis.<sup>8–11</sup> Disrupting the Th17/IL-17a pathway shows high potential for the treatment of autoimmune disorders, which has already been validated by the clinical success of anti-IL-17a monoclonal antibodies.<sup>8,12,13</sup> Inhibition of ROR $\gamma$ t could pose an attractive alternative strategy to decrease IL-17a production in the treatment of autoimmune diseases.<sup>12</sup> Many research groups have shown significant interest in the identification of small molecule ROR $\gamma$ t inhibitors (or, more specifically, inverse agonists), with several synthetic compounds progressed into clinical trials.<sup>14–18</sup>

With the exception of **RTA-1701**<sup>18,19</sup> (see Chapter 1), all of the ROR $\gamma$ t inverse agonists that have entered clinical trials likely target the highly conserved ligand binding pocket, termed orthosteric binding site, within the ligand binding domain (LBD) of ROR $\gamma$ t (Figure 3.1A).<sup>5,14</sup> While orthosteric targeting has been highly successful, novel molecular modalities with alternative modes of action, such as allosteric modulators, offer an interesting alternative for targeting NRs and could circumvent issues related to promiscuity between NRs.<sup>20–23</sup> Recently, an allosteric binding site was identified in the LBD of ROR $\gamma$ t, at a location that is topographically distinct to the orthosteric site (Figure 3.1A).<sup>24</sup> The indazole **MRL-871** (**1**) (Figure 3.1C) was identified as a prototypical allosteric ligand for ROR $\gamma$ t, acting as a potent inverse agonist and decreasing coactivator binding with similar efficacy as orthosteric inverse agonists.<sup>24,25</sup> The allosteric pocket of ROR $\gamma$ t is unique within the NR family and thought not to be the target of endogenous ligands.<sup>24,26</sup> Therapeutic compounds targeting the allosteric site potentially have a higher selectivity profile and do not act in competition, but in synergy, with endogenous agonists.<sup>20,21,27</sup> Therefore, these compounds would be greatly beneficial for drug discovery and chemical biology applications.

Despite the high potential of allosteric inverse agonists for ROR $\gamma$ t in drug discovery, the number of examples have remained limited to compounds based on the indazole core of **1** or similar chemotypes, and the previously mentioned allosteric ligand **RTA-1701**.<sup>19,24,25,28–33</sup> Follow-up studies improved the pharmacokinetic (PK) and selectivity profile of **1** and its derivatives. However, off-target effects towards other NRs, most notably towards PPAR $\gamma$ , and



challenges regarding cytotoxicity and metabolic stability have been observed.<sup>24,28–31</sup> The scarcity and lack of scaffold diversity of allosteric RORyt ligands, and the limitations regarding specificity and drug-like properties, necessitates further exploration of the allosteric pocket. Specifically, establishing structure activity relationship (SAR) profiles for allosteric chemotypes distinct from the indazole-analogous ligand classes is of great importance.



**Figure 3.1 | Trisubstituted isoxazoles, exemplified by hit compound FM26, have been discovered as novel types of allosteric ligands for RORyt.** **A)** The co-crystal structure of the RORyt LBD in complex with allosteric ligand **FM26** (shown as orange sticks) (PDB: 6SAL),<sup>34</sup> where the allosteric site (orange circle) and the orthosteric site (blue circle) are indicated. **B)** Enlarged view of **FM26** (shown as orange sticks) in the allosteric pocket, where hydrogen bond interactions are indicated with orange dashes. **C)** Chemical structures of allosteric RORyt ligands **FM26 (2)** and **MRL-871 (1)**. **D)** Exploration of hit-to-lead structure activity relationship (SAR) profile of **FM26**.

In Chapter 2, we used an *in silico* pharmacophore search approach to identify a novel class of RORyt allosteric inverse agonists, featuring a trisubstituted isoxazole core.<sup>34–35</sup> SAR studies around this novel chemotype resulted in the discovery of **FM26 (2)** (Figure 3.1C), which shows sub-micromolar potency as a RORyt inverse agonist and significant inhibition of cellular

IL-17a expression levels. Initial SAR studies of **2** identified two key pharmacophore features, where the 2,6-disubstituted phenyl ring at the C-3 position and a C-4 benzoic acid moiety were shown to be optimal (Figure 3.1D).<sup>24,25,29–32,34</sup> Less knowledge has been garnered on the linker at the C-4 position and the C-5 substituent. Regarding the C-4 linker, an amine linker (as is used for **2**) was found to be preferred for potency over some more rigid linkers.<sup>34</sup> Additionally, the co-crystal structure (Figure 3.1B) showed that the presence of a hydrogen bond donating *N*-heterocycle at the C-5 isoxazole position (as is the case for the pyrrole in **2**) significantly increased the potency towards RORyt, by the formation of an additional polar interaction with the backbone carbonyls of residues Leu353 and Lys354.<sup>34</sup> In order to identify compounds suitable for focused lead optimization, crucial explorations around the C-4 and C-5 positions are needed to further explore the SAR, improve the potency of the isoxazole series, increase RORyt specificity and direct PK optimization.

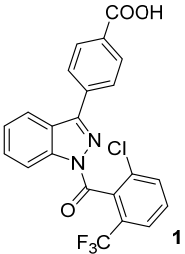
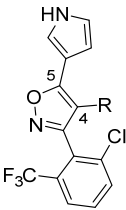

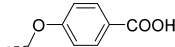



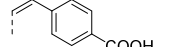
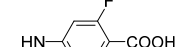
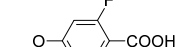
In this chapter, we report the findings of a structure-based SAR study on the isoxazole class of RORyt allosteric ligands with a specific focus on improved RORyt activity, mitigating off-target PPAR $\gamma$  activity, and a first exploration of PK profiles. Utilizing X-ray crystallography to guide focused library development, two libraries were synthesized based on the C-4 and C-5 isoxazole position (Figure 3.1D). Compounds that showed promising RORyt inhibition were also assessed for off-target binding to PPAR $\gamma$  and their PK properties were evaluated using an ADME panel. The most promising lead compounds showed significantly improved activity *in vitro*, decreased cross-reactivity with PPAR $\gamma$  and a promising PK profile.

### ***In silico* docking studies guide a SAR study at the isoxazole C-4 position**

The initial investigation started with expanding the SAR data around the C-4 position of the isoxazole series, guided by *in silico* docking studies. A virtual library of derivatives around compound **2**, containing different C-4 linkers and benzoic acid substituents, was designed with a focus on diversity and synthetic feasibility (see Supporting Table S3.1). This library was docked into the allosteric site of RORyt (PDB: 4YPQ<sup>24</sup>), using Glide, the molecular docking tool in Schrödinger.<sup>36,37</sup> For each ligand, the docking pose was evaluated and given a 'Glide Score'.<sup>36,37</sup> Table 3.1 shows the compounds that resulted in the most promising Glide Scores, indicating that linkers responsible for an increased cLogP of the compounds (i.e. ether (**3**), thioether (**4**), methylated amine (**5**) and alkene (**6** & **7**)) would be beneficial for affinity. Additionally, the data suggests that a fluoro substitution at the *ortho* position of the benzoic acid moiety would be well tolerated (**8** & **9**).

**Table 3.1 | Structure activity relationship studies around the C-4 isoxazole position.** cLogP values<sup>a</sup>, Glide Scores, TR-FRET IC<sub>50</sub> values ( $\mu$ M) from coactivator recruitment and AlexaFluor-MRL-871 recruitment assays and  $\Delta$ T<sub>m</sub> values ( $^{\circ}$ C) from thermal shift assays. TR-FRET and TSA data are representative of three independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD.

<sup>a</sup>The cLogP values were predicted using MarvinSketch (20.10).

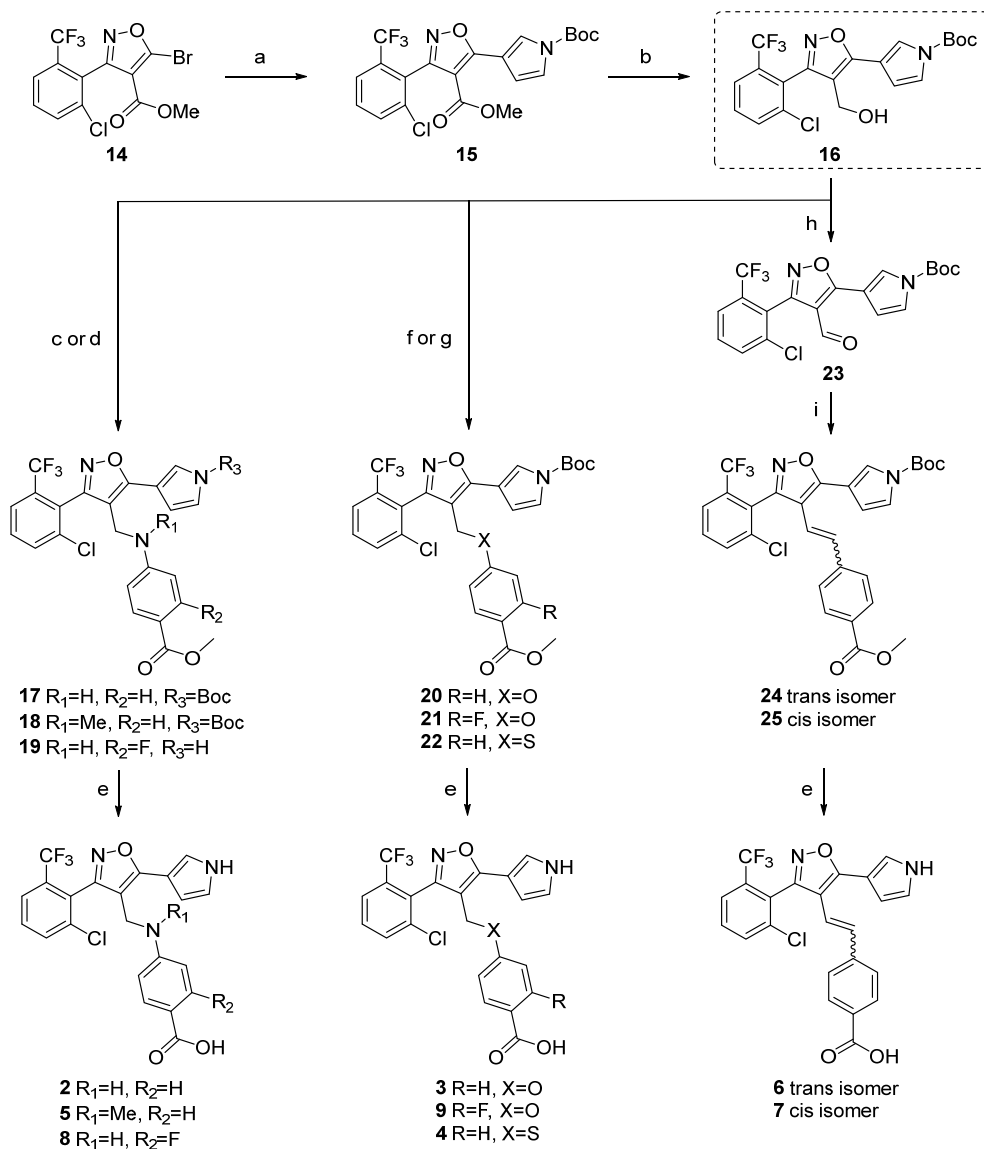
<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  <p><b>1</b></p> </div> <div style="text-align: center;">  <p><b>2-9</b></p> </div> </div>						
Cmpd	R (C-4)	cLogP <sup>a</sup>	Glide Score	IC <sub>50</sub> ( $\mu$ M) TR-FRET Coactivator	IC <sub>50</sub> ( $\mu$ M) TR-FRET AlexaFluor-MRL	$\Delta$ T <sub>m</sub> ( $^{\circ}$ C)
<b>1</b>	—	5.64	− 14.99	0.013 $\pm$ 0.001	0.0070 $\pm$ 0.0007	7.7 $\pm$ 0.0
<b>2</b>		5.35	− 14.26	0.27 $\pm$ 0.02	0.10 $\pm$ 0.01	2.4 $\pm$ 0.2
<b>3</b>		5.72	− 15.17	0.031 $\pm$ 0.003	0.012 $\pm$ 0.001	4.9 $\pm$ 0.0
<b>4</b>		6.41	− 14.95	6.6 $\pm$ 0.8	2.5 $\pm$ 0.3	1.3 $\pm$ 0.0
<b>5</b>		5.99	− 15.16	>100	6.2 $\pm$ 0.6	0.0 $\pm$ 0.2
<b>6</b>		6.50	− 15.32	0.020 $\pm$ 0.002	0.0074 $\pm$ 0.0009	6.4 $\pm$ 0.0
<b>7</b>		6.50	− 13.86	0.49 $\pm$ 0.03	0.29 $\pm$ 0.03	0.7 $\pm$ 0.0
<b>8</b>		5.50	− 14.26	3.6 $\pm$ 0.3	1.5 $\pm$ 0.3	0.7 $\pm$ 0.0
<b>9</b>		5.87	− 15.50	0.22 $\pm$ 0.01	0.074 $\pm$ 0.009	3.0 $\pm$ 0.2

## An optimized synthesis route allows the efficient synthesis of C-4 isoxazole derivatives

Isoxazoles **2-9** (Table 3.1) were synthesized (Scheme 3.1) in order to biochemically evaluate the predictions from the *in silico* docking experiments. Common intermediate **14** was synthesized via a 1,3-dipolar cycloaddition between a nitrile oxide and alkynyl bromide, as described previously (Supporting Scheme S3.1).<sup>34</sup> In order to prepare derivatives **2-9**, it was important to obtain core intermediate **16** with an *N*-Boc protected pyrrole moiety, instead of the free pyrrole as was the case in the original synthesis route for **2** (Supporting Scheme S3.2).<sup>34</sup> This was achieved via a Suzuki cross-coupling reaction, employing a dppf instead of tetrakis palladium catalyst (used in previous research) which proved to be essential for maintaining the Boc-protected pyrrole **15** in an acceptable yield (Scheme 3.1). DIBAL was then used to selectively reduce the ethyl ester, yielding **16** without concomitant loss of the Boc group (Scheme 3.1), which was previously observed with LiAlH<sub>4</sub> (Supporting Scheme S3.2).<sup>34</sup>

The primary alcohol of **16** was then used as a functional handle for derivatization of the isoxazole C-4 position. Critical to the synthesis of this series of isoxazoles was optimization of the reductive amination step, which provided sub-optimal yields in the synthesis of **2** (Supporting Scheme S3.2).<sup>34</sup> In order to improve the total yield of the synthesis, the reductive amination was substituted for a nucleophilic substitution reaction. Final compounds **2** and **5** were accessed in an efficient manner by mesylation of alcohol **16** with methane sulfonic anhydride (monitored by NMR), *in situ* addition of the substituted aniline, followed by hydrolysis of the benzoate ester. In contrast, **8** (fluoro substituent) was synthesized via the original reductive amination route (Scheme 3.1).

Compounds with an ether and thioether linkage (**3**, **9** and **4**) were synthesized from core intermediate **16** via a Mitsunobu reaction with hydroxy- or mercaptobenzoate, respectively, affording the products in acceptable yields (Scheme 3.1). For the *cis* and *trans* alkene linker (**6** and **7**), the alcohol of compound **16** was oxidized to the aldehyde, which was subsequently used in a Wittig reaction, obtaining a 3:1 mixture of the *cis/trans* isomers (Scheme 3.1). These were separated via preparative HPLC-UV to afford **24** and **25**, and followed by ester hydrolysis, afforded final compounds **6** (*trans*) and **7** (*cis*) (for which the stereochemistry was determined based on the relative <sup>1</sup>H-NMR *J*-coupling values between the alkene protons).

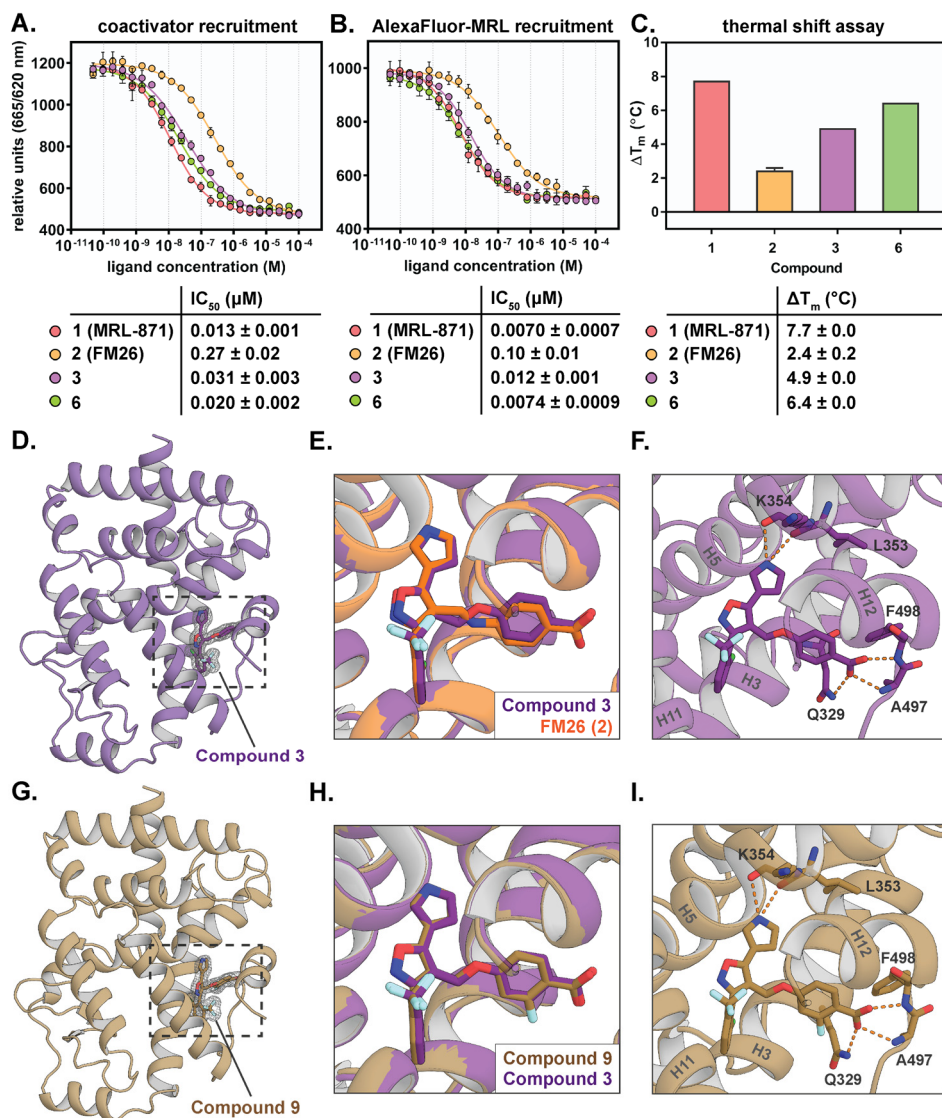


**Scheme 3.1 | Synthesis route for the different trisubstituted isoxazoles (C-4 library).** Reagents and conditions: (a) *N*-Boc-pyrrole-B(pin), Pd(dppf)Cl<sub>2</sub>, Cs<sub>2</sub>CO<sub>3</sub>, DME, 85 °C, 8 h, 50%; (b) DIBAL, CH<sub>2</sub>Cl<sub>2</sub>, -78 °C, 3 h, 80%; (c) i) (MeSO<sub>2</sub>)<sub>2</sub>O, Et<sub>3</sub>N, CH<sub>2</sub>Cl<sub>2</sub>, 0 °C → rt, 3 h; ii) add aniline or *N*-Methylaniline, rt, 3 h, 36% (**17**), 52% (**18**); (d) i) DMP, CH<sub>2</sub>Cl<sub>2</sub>, rt, 3 h, 44%; ii) Methyl 4-amino benzoate, MeOH, AcOH, reflux, 24 h, 42%; iii) NaBH<sub>4</sub>, EtOH, reflux, 3 h, 17% (**19**); (e) LiOH, EtOH, H<sub>2</sub>O, 95 °C, 3 h, 29-81%; (f) Methyl 4-hydroxybenzoate or Methyl 2-fluoro-4-hydroxybenzoate, DIAD, PPh<sub>3</sub>, Et<sub>3</sub>N, THF, reflux, 3 h, 55% (**20**), 25% (**21**); (g) Methyl 4-mercaptobenzoate, DIAD, PPh<sub>3</sub>, Et<sub>3</sub>N, THF, reflux, 3 h, 27% (**22**); (h) DMP, CH<sub>2</sub>Cl<sub>2</sub>, rt, 3 h, quant.; (i) (4-(methoxycarbonyl)benzyl)triphenylphosphonium (see experimental for the synthetic procedure of the triphenylphosphonium), LiHMDS, THF, -78 °C → rt, 24 h, 7% (**24**), 27% (**25**).

## Biochemical assays and X-ray crystallography reveal a positive correlation between the C-4 linker lipophilicity and potency

The potency of the C-4 modified compounds for ROR $\gamma$ t was investigated in a time-resolved FRET (TR-FRET) coactivator recruitment assay (Table 3.1, Figure 3.2A, a schematic representation of the assay setup is shown in Chapter 2, Figure 2.3C).<sup>38</sup> ROR $\gamma$ t is constitutively active, which means that it shows a basic level of transcriptional activity, thus allowing the partial recruitment of coactivators in the absence of an agonist.<sup>39</sup> Reference compounds **1** and **2** showed potent inhibition of coactivator binding, with IC<sub>50</sub> values comparable to previous studies.<sup>24,34</sup> Interestingly, compound **3** (**FM156**) containing an ether linkage demonstrated a 9-fold increase in potency compared to **2** with an amine linker (IC<sub>50</sub> of  $31 \pm 3$  nM vs.  $270 \pm 20$  nM, respectively), resulting in an IC<sub>50</sub> value in the same range as indazole **1**. The *trans*-alkene linker (**6**, **FM260**) resulted in a low nanomolar IC<sub>50</sub> value as well (IC<sub>50</sub> of  $20 \pm 2$  nM), in contrast to the *cis*-alkene linker (**7**) which was 25 times less potent (IC<sub>50</sub> of  $490 \pm 30$  nM), as predicted by the docking scores. In contrast, a thioether linkage (**4**) resulted in a significantly decreased potency compared to **2** (IC<sub>50</sub> of  $6.6 \pm 0.8$   $\mu$ M), whereas a methylated amine linker (**5**) resulted in a loss of activity (IC<sub>50</sub> >100  $\mu$ M). The introduction of a fluorine substituent at the *ortho* position of the benzoic acid moiety proved detrimental to coactivator inhibition with a 7-13 fold decrease in potency compared to the compound without a substituent (**8** vs. **2** and **9** vs. **3**), which was also the case for an isoxazole analogue in a previous SAR study (described in Chapter 2).<sup>34</sup>

In order to prove an allosteric mode of action, the compounds were also tested in two other TR-FRET assay formats. First, a previously described AlexaFluor647-labelled MRL-871 probe<sup>24</sup> was used, which upon binding to ROR $\gamma$ t shows FRET pairing with an anti-His terbium cryptate donor on the protein (a schematic representation of the assay setup is shown in Chapter 2, Figure 2.3C).<sup>24</sup> All compounds showed displacement of the AlexaFluor-MRL allosteric probe (Table 3.1, Figure 3.2B), proving an allosteric binding mode, with the IC<sub>50</sub> values correlating with the IC<sub>50</sub> values from the TR-FRET coactivator recruitment assay (Table 3.1). In particular, ether **3** and *trans*-alkene **6** demonstrated efficient displacement of the probe, with IC<sub>50</sub> values of  $12 \pm 1$  nM and  $7.4 \pm 0.9$  nM, respectively, significantly increased as compared to **2** (IC<sub>50</sub> value of  $100 \pm 10$  nM).



**Figure 3.2 | Biochemical analysis and X-ray crystallography data for isoxazole compounds with C-4 modifications.** A/B) Dose-response curves from the TR-FRET coactivator recruitment assays (A) and AlexaFluor-MRL recruitment assays (B) on ROR $\gamma$ t, including an overview of the IC<sub>50</sub> values. C) Shift in melting temperature ( $\Delta T_m$  in °C) as measured in thermal shift assays. Data are representative of three independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD. D/G) The co-crystal structure of the ROR $\gamma$ t LBD in complex with **3** (PDB: 7NPC) (D) and with **9** (PDB: 7NP5) (G). The final 2Fo – Fc electron density map of the compounds is shown as an isomesh contoured at 1 $\sigma$ . E/H) Overlay of the co-crystal structures of ROR $\gamma$ t in complex with FM26 (**2**) (PDB: 6SAL) and with **3** (E), and an overlay of ROR $\gamma$ t in complex with **3** and with **9** (H). F/I) Enlarged view of the allosteric pocket of ROR $\gamma$ t showing the interactions between **3** (F)/**9** (I) and the protein. Hydrogen bond interactions are shown with orange dashes.

The most potent ligands **3** and **6** were also measured in a competitive TR-FRET coactivator recruitment assay, where the compounds were titrated in the presence of the orthosteric agonist cholesterol<sup>40</sup> (Supporting Figure S3.1, a schematic representation of the assay setup is shown in Chapter 2, Figure 2.4E). The binding curves for compounds **3** and **6** show that the IC<sub>50</sub> values did not decrease upon increasing concentrations of cholesterol, which demonstrates that their binding mode is independent to that of cholesterol, supporting their allosteric binding mode. In fact, the presence of cholesterol slightly enhanced the potency of both **3** and **6**, suggesting cooperativity between the orthosteric and allosteric site, as was observed previously.<sup>27,34</sup>

A thermal shift assay (TSA) was performed as an orthogonal assay to investigate the effect of the compounds on the thermal denaturation of ROR $\gamma$ t. Ligand binding typically improves the thermal stability of a protein by stabilizing the protein fold, as indicated by the change in melting temperature,  $\Delta T_m$ .<sup>41-43</sup> The C-4 isoxazole derivatives **3-9** showed a thermal stabilization effect according to their potency, as observed in the TR-FRET assays (Table 3.1, Figure 3.2C). The most potent compounds **3** and **6** showed particularly high  $\Delta T_m$  values (4.9 and 6.4 °C, respectively). These values are improved compared to that for **2**, which provides further indication that **3** and **6** have a high binding affinity for ROR $\gamma$ t.

Crystallization studies provided the co-crystal structure of the ROR $\gamma$ t LBD in complex with compounds **3** (ether linker) and **9** (fluoro substituent) with resolutions of 1.46 Å and 1.55 Å, respectively (Supporting Table S3.4). The co-crystal structures of ROR $\gamma$ t with **3** and **9** showed clear electron density for the compounds in the allosteric binding site between helices 3, 4, 11 and 12 (Figure 3.2D & 3.2G). In Figure 3.2E, an overlay of ROR $\gamma$ t in complex with **2** and with **3** is shown, demonstrating a highly similar binding pose, with the C-3, C-4 and C-5 isoxazole substituents anchored at the same position. The pyrrole moiety of **3** forms a hydrogen bond interaction with the main chain carbonyls of residues Leu353 and Lys354 (Figure 3.2F). Additionally, the carboxylic acid moiety forms hydrogen bond interactions with the side chain of Gln329 and the backbone amide hydrogen atoms of Phe498 and Ala497, as is also the case for **1** and **2**. Interestingly, the electron density for **3** and **2** would allow two different conformations of the C-4 linker,<sup>34</sup> but the preferred conformation for the ether linker in **3** is opposite to the preferred conformation for the amine linker in **2** (Figure 3.2E). For compound **9**, the electron density and binding mode are highly similar to **3**, with the same conformation for the ether linker present (Figure 3.2G-I). The *ortho*-fluoro substituent on the benzoic acid moiety does not significantly influence the conformation of the compound, except for the



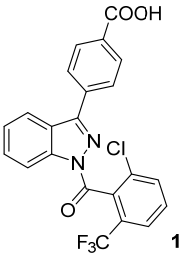
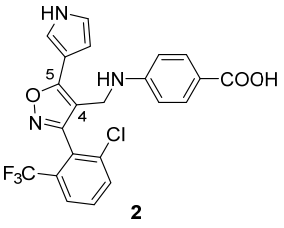
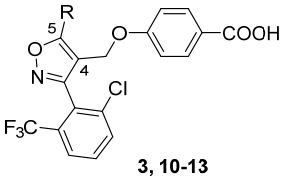
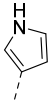

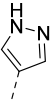
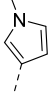
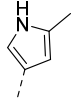
disubstituted phenyl ring at the isoxazole C-3 position, which is normally fixed at the same position but is now slightly shifted (Figure 3.2H).

The increased potency of ether **3** and alkene **6** could be the result of increased hydrophobic effects of these more lipophilic compounds towards the hydrophobic allosteric binding pocket (see protein-ligand interaction plot in Chapter 2, Figure S2.4). The different preferred conformation of the ether linker for **3**, as seen in the co-crystal structure, might also be contributing to the potency. In contrast, the compound with a thioether linkage (**4**) shows a significantly lower potency, which might be caused by a slight change in bond angle and length of the linker. The docking pose predicts that the isoxazole core and pyrrole moiety are clearly shifted compared to the compound with an ether linker (**3**) (Supporting Figure S3.2A). The drop in potency for **5**, featuring a methylated amine linker, likely results from restricted rotation of the linker. The electron density for the isoxazoles shows that two conformations of the linker are present, but when the linker is more rotationally restricted, as is the case for **5**, only one of these linker conformations will most likely occur, potentially resulting in a higher entropic penalty and in turn a decrease in binding affinity. Additionally, the methylated linker could change the conformation of the benzoic acid moiety, preventing optimal hydrogen bond formation with the protein. The introduction of an *ortho*-fluoro substituent at the benzoic acid leads to a slight decrease in potency (**8** vs. **2** and **9** vs. **3**). The co-crystal structure in complex with **9** shows that the fluoro substituent is placed in the same plane as the carboxylic acid moiety (Figure 3.2I). This unfavorable conformation is believed to be caused in order to both fit the fluoro substituent in the pocket and also allow hydrogen bond interactions between the carboxylic acid and the protein, which could lead to charge repulsion between the two moieties and thus a decreased potency.

### **SAR studies show the necessity of a hydrogen bond donor moiety at the isoxazole C-5 position**

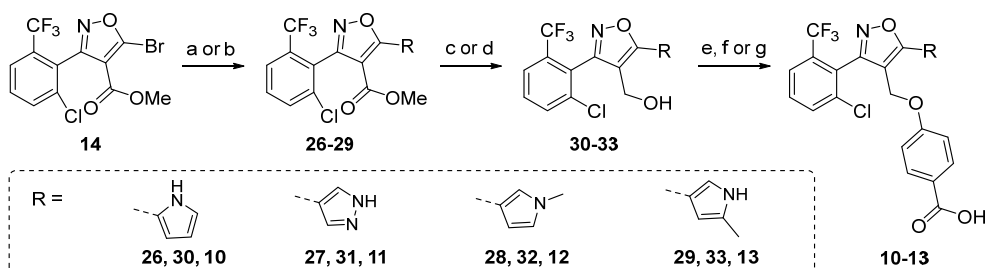
The SAR was further explored at the isoxazole C-5 position to investigate the hydrogen bonding character of the pyrrole moiety of **2**. Compound **3** with the ether linkage was chosen as core scaffold, based on the combination of its high potency, availability of the co-crystal structure, and synthetic feasibility. A focused library of C-5 derivatives (compounds **10-13**, Table 3.2) was designed and synthesized (the full library used for docking is shown in Supporting Table S3.2). Specifically, the effect of the nitrogen position (**10**), pKa (**11**) (pKa = 10.17 for **11** vs. 14.99 for **3**) and substitution (**12**) was investigated, as well as hydrophobic space around the pyrrole ring (**13**).

**Table 3.2 | Structure activity relationship studies around the C-5 isoxazole position.** Glide Scores, TR-FRET  $IC_{50}$  values ( $\mu M$ ) from coactivator recruitment and AlexaFluor-MRL-871 recruitment assays and  $\Delta T_m$  values ( $^{\circ}C$ ) from thermal shift assays. TR-FRET and TSA data are representative of three independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD.

<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  <p><b>1</b></p> </div> <div style="text-align: center;">  <p><b>2</b></p> </div> <div style="text-align: center;">  <p><b>3, 10-13</b></p> </div> </div>					
Compound	R (C-5)	Glide Score	$IC_{50}$ ( $\mu M$ ) TR-FRET Coactivator	$IC_{50}$ ( $\mu M$ ) TR-FRET AlexaFluor-MRL	$\Delta T_m$ ( $^{\circ}C$ )
<b>1</b>	—	−14.99	$0.013 \pm 0.001$	$0.0070 \pm 0.0007$	$7.7 \pm 0.0$
<b>2</b>	—	−14.26	$0.27 \pm 0.02$	$0.10 \pm 0.01$	$2.4 \pm 0.2$
<b>3</b>		−15.17	$0.031 \pm 0.003$	$0.012 \pm 0.001$	$4.9 \pm 0.0$
<b>10</b>		−14.76	$0.14 \pm 0.01$	$0.046 \pm 0.004$	$2.8 \pm 0.0$
<b>11</b>		−14.85	$0.11 \pm 0.01$	$0.033 \pm 0.003$	$2.7 \pm 0.2$
<b>12</b>		−15.12	$3.3 \pm 0.3$	$1.9 \pm 0.3$	$0.0 \pm 0.2$
<b>13</b>		−14.67	$2.9 \pm 0.2$	$1.8 \pm 0.3$	$0.1 \pm 0.0$

Compounds **10-13** were synthesized using an analogous synthesis route to that used for the C-4 isoxazole compounds (Scheme 3.2). Since the C-5 moiety is incorporated early in the synthesis, it was not possible to diversify from a late-stage intermediate, as for the C-4 modifications. Bromide **14** was subjected to a Suzuki reaction with the associated pinacol ester, containing a Boc-, THP-, or methyl-protected heterocycle, followed by ester hydrolysis with DIBAL to yield **30-33** (Scheme 3.2). For the *N*-methyl-pyrrole, the original conditions

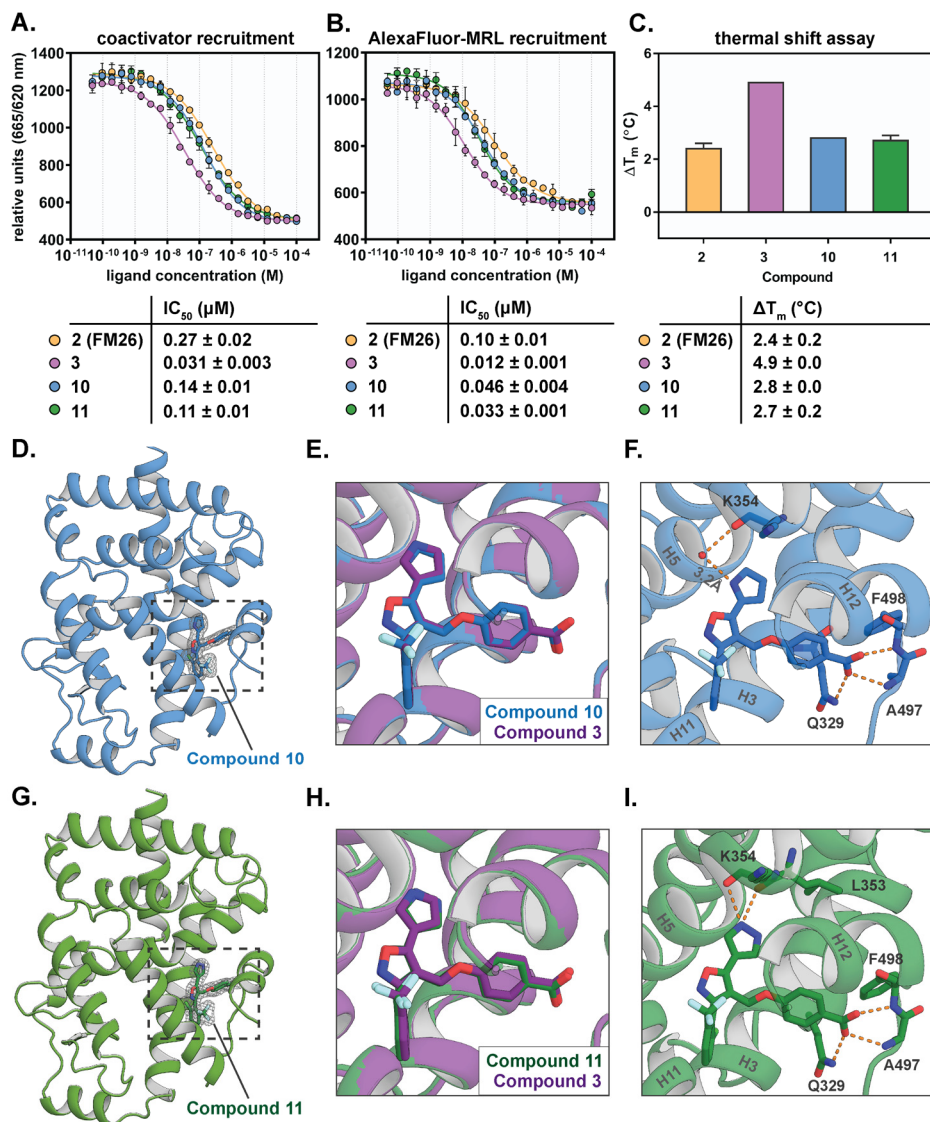
were used (Supporting Scheme S3.2) (the tetrakis palladium catalyst in the Suzuki reaction and  $\text{LiAlH}_4$  for the reduction) because no labile protecting group was used in this synthesis route. Next, a Mitsunobu reaction with methyl 4-hydroxybenzoate, and ester hydrolysis combined with heterocycle deprotection were conducted to afford the final compounds **10-13**.



**Scheme 3.2 | Synthesis route for the different trisubstituted isoxazoles (C-5 library).** Reagents and conditions: (a) *N*-Boc-pyrrole-B(pin), *N*-THP-pyrazole-B(pin) or 5-methyl *N*-Boc-pyrrole-B(pin) (see experimental for the synthetic procedure of this pinacol ester),  $\text{Pd}(\text{dppf})\text{Cl}_2$ ,  $\text{Cs}_2\text{CO}_3$ , DME,  $85^\circ\text{C}$ , 8 h, 49% (**26**), 34% (**27**), 24% (**29**); (b) *N*-Methyl-pyrrole-B(pin),  $\text{Pd}(\text{PPh}_3)_4$ ,  $\text{Na}_2\text{CO}_3$ , DME,  $85^\circ\text{C}$ , 8 h, 55% (**28**); (c) DIBAL,  $\text{CH}_2\text{Cl}_2$ ,  $-78^\circ\text{C}$ , 3 h, 69% (**30**), 71% (**31**), 58% (**33**); (d)  $\text{LiAlH}_4$ , THF,  $0^\circ\text{C} \rightarrow \text{rt}$ , 2 h, 76% (**32**); (e) i) Methyl 4-hydroxybenzoate, DIAD,  $\text{PPh}_3$ , THF, reflux, 3 h, 21%; ii)  $\text{LiOH}$ , EtOH,  $\text{H}_2\text{O}$ ,  $95^\circ\text{C}$ , 3 h, 41% (**10**); (f) i) Methyl 4-hydroxybenzoate, DIAD,  $\text{PPh}_3$ ,  $\text{Et}_3\text{N}$ , THF, reflux, 3 h, 44-45%; ii)  $\text{LiOH}$ , EtOH,  $\text{H}_2\text{O}$ ,  $95^\circ\text{C}$ , 3 h, 86% (**12**), 83% (**13**); (g) i) Methyl 4-hydroxybenzoate, DIAD,  $\text{PPh}_3$ ,  $\text{Et}_3\text{N}$ , THF, reflux, 3 h, 49%; ii) TFA,  $\text{CH}_2\text{Cl}_2$ ,  $40^\circ\text{C}$ , 2 h, 87%; iii)  $\text{LiOH}$ , EtOH,  $\text{H}_2\text{O}$ ,  $95^\circ\text{C}$ , 3 h, 89% (**11**).

To assess the SAR around the C-5 position, analogues **10-13** were evaluated in the TR-FRET coactivator recruitment assay (Figure 3.3A, Table 3.2). A nitrogen shift (2- instead of 3-position) in the pyrrole ring (**10**) resulted in a 4.5-fold decrease in potency ( $\text{IC}_{50} = 140 \pm 10$  nM) compared to **3**. Pyrazole **11** (**FM257**) showed a slight drop in potency, with an  $\text{IC}_{50}$  value of  $110 \pm 10$  nM. For compound **12** (containing a methylated pyrrole) and compound **13** (methylation of the pyrrole at the 5-position), a lower activity was observed with  $\text{IC}_{50}$  values of  $3.3 \pm 0.3$   $\mu\text{M}$  and  $2.9 \pm 0.2$   $\mu\text{M}$ , respectively.

Additionally, the TR-FRET AlexaFluor-MRL assay was performed, validating an allosteric binding mode for all compounds (Figure 3.3B, Table 3.2). Analysis of the  $\text{IC}_{50}$  values for this series showed that these were in line with the  $\text{IC}_{50}$  values from the TR-FRET coactivator assay (Table 3.2). In the TSAs (Figure 3.3C, Table 3.2), compounds **10** and **11** induced thermal stabilization with  $\Delta T_m$  values of 2.8 and 2.7  $^\circ\text{C}$ , respectively, which is lower than for **3** ( $\Delta T_m = 4.9^\circ\text{C}$ ) but slightly higher than for **2** ( $\Delta T_m = 2.4^\circ\text{C}$ ). Compounds **12** and **13** did not show any response.



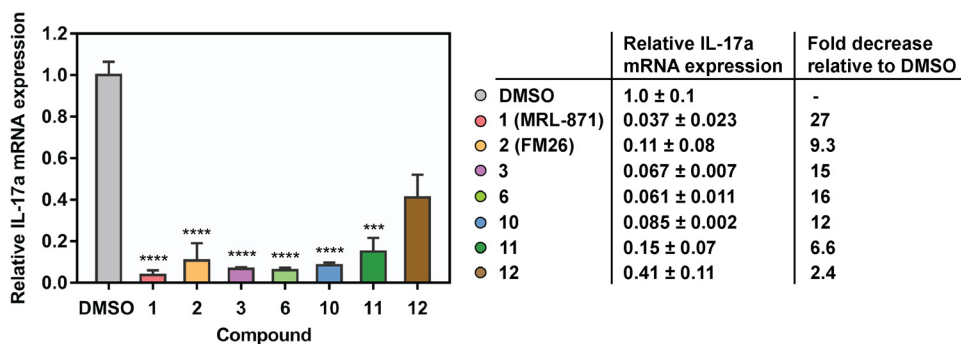
**Figure 3.3 | Biochemical analysis and X-ray crystallography data for isoxazole compounds with C-5 modifications.** A/B) Dose-response curves from the TR-FRET coactivator recruitment assays (A) and AlexaFluor-MRL recruitment assays (B) on RORγt, including an overview of the IC<sub>50</sub> values. C) Shift in melting temperature (ΔT<sub>m</sub> in °C) as measured in thermal shift assays. Data are representative of three independent experiments (recorded in triplicate) and are presented as mean ± SD. D/G) The co-crystal structure of the RORγt LBD in complex with 10 (PDB: 7NEC) (D) and with 11 (PDB: 7NP6) (G). The final 2Fo – Fc electron density map of the compounds is shown as an isomesh contoured at 1σ. E/H) Overlay of the co-crystal structures of RORγt in complex with 3 (PDB: 7NPC) and with 10 (E), or with 3 and with 11 (H). F/I) Enlarged view of the allosteric pocket of RORγt showing the interactions between 10 (F)/11 (I) and the protein. Hydrogen bond interactions are shown with orange dashes.

Crystallization studies provided the co-crystal structures of the ROR $\gamma$ t LBD in complex with C-5 modified compounds **10** (2-pyrrole) and **11** (pyrazole) with resolutions of 1.95Å and 1.84Å, respectively (Supporting Table S3.4). The co-crystal structure of ROR $\gamma$ t with **10** revealed a clear ligand electron density in the allosteric site with the ether linker in the same preferred conformation as for **3** (Figure 3.3D,E). The 2-pyrrole at the C-5 position, with the nitrogen at the different position, does not establish a direct hydrogen bond interaction with the protein, as expected, potentially explaining the lower potency of the compound for ROR $\gamma$ t. Interestingly, the NH of the pyrrole substituent of **10** forms an alternative hydrogen bond network via a water molecule in the binding pocket at a distance of 3.2Å from the NH of the pyrrole (Figure 3.3F) (this water molecule appears to be present in all crystal structures with an allosteric ligand, and thus has a structural role). For pyrazole **11**, the binding mode is highly similar to **3**, again establishing a hydrogen bond interaction between the NH of the pyrazole and the backbone carbonyls of ROR $\gamma$ t (Figure 3.3G-I).

These results show that small changes at the pyrrole C-5 moiety lead to relevant changes in the potency of the compounds. The 4.5-fold decrease in potency for the compound with a changed position of the nitrogen in the pyrrole ring (**10**) (compared to **3**) is most probably due to the loss of the characteristic hydrogen bond interaction between the pyrrole and the backbone of the protein. The additional hydrogen bond interaction with a water molecule explains why **10** is still active on ROR $\gamma$ t with an IC<sub>50</sub> value <150 nM, despite the lack of the characteristic C-5 hydrogen bond interaction. For the pyrazole (**11**), the 3.5-fold lower potency indicates that the lower pKa of the pyrazole is not beneficial for hydrogen bond formation, although other factors might also be involved, e.g. the existence of the pyrazole in two tautomeric forms for which only one can establish the hydrogen bond interaction with the protein. The much lower potency for compound **12** (*N*-methyl-pyrrole) suggests that the space around the pyrrole is limited, which is also supported by the low potency of the methyl-substitution of the pyrrole at the 5-position (**13**). The docking score predicts compound **12** to be highly active, based on the docking pose where the pyrrole ring is rotated and the methyl substituent could point towards a small cavity (Supporting Figure S3.2B). However, based on the biochemical data, the methylated pyrrole does not appear to have enough space for this rotation in the allosteric pocket. Together, these data show the need for a hydrogen bond donor at the right position of the heterocycle at the C-5 position of the isoxazole scaffold, while additional substituents at the ring appear to be too bulky to be tolerated.

### The new isoxazole series show inhibition of IL-17a expression in EL4 cells

A selection of the novel isoxazole compounds was tested in a quantitative reverse transcriptase PCR (RT-PCR) assay to investigate their functional effect. EL4 cells were treated with 10  $\mu$ M of compound or DMSO, after which the IL-17a mRNA levels were measured (Figure 3.4). Hit compound **2** showed a 9.3-fold reduction, in accordance with our previous reports.<sup>34</sup> The optimized compounds **3** and **6** (IC<sub>50</sub> values of 31 nM and 20 nM, respectively) showed a 15- and 16-fold reduction of IL-17a levels compared to DMSO, which is in line with their high biochemical potency and indicates good cellular uptake and activity. Compounds **10** and **11** induced a 12- and 6.6-fold decrease in IL-17a mRNA levels, correlating with their slightly lower biochemical potency compared to **3** and **6**. Compound **12** only resulted in a 2.4-fold decrease compared to DMSO, as expected, given its significantly lower potency in the TR-FRET coactivator recruitment assay.



**Figure 3.4 | IL-17a mRNA expression levels in EL4 cells treated with a selection of the isoxazole ligands (10  $\mu$ M, 24 h) or DMSO and fold decrease of IL-17a expression relative to DMSO.** The level of IL-17a expression was normalized to that of GAPDH expression. Data are representative of two independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD. The relative gene expression was calculated by the  $2^{-\Delta\Delta C_t}$  (Livak) method using the DMSO control as a calibrator. Statistical analysis was performed using a one-way analysis of variance (ANOVA) compared against the DMSO control following Dunnett's *post hoc* test; \*\*\*  $P < 0.001$  and \*\*\*\*  $P < 0.0001$ .

### The isoxazole series show an improved selectivity profile for ROR $\gamma$ t

Previous studies have shown that **1** is selective for ROR $\gamma$ t over other NRs, except for PPAR $\gamma$  on which it shows a significant cross-reactivity to the orthosteric binding site, as was also supported by the co-crystal structure.<sup>24,44</sup> Isoxazole **2** already showed a higher selectivity for ROR $\gamma$ t over PPAR $\gamma$  than **1**<sup>34</sup>, so we were interested to establish if the novel isoxazole compounds further improved this promising selectivity profile.

In a TR-FRET coactivator recruitment assay, for which the results are shown in Table 3.3, compound **1** showed PPAR $\gamma$  agonism with an EC<sub>50</sub> value of  $0.34 \pm 0.02$   $\mu$ M, while the EC<sub>50</sub> value for **2** was >20-fold higher (EC<sub>50</sub> =  $8.2 \pm 0.3$   $\mu$ M).<sup>24,34</sup> All novel isoxazole ligands developed in this study were at least five times less active on PPAR $\gamma$  than **1** (EC<sub>50</sub> values >1.7  $\mu$ M), except for compounds **12** and **13** (both bearing a methyl substituent at the pyrrole moiety) which showed a higher cross-reactivity on PPAR $\gamma$  (EC<sub>50</sub> values of  $1.2 \pm 0.1$   $\mu$ M and  $0.79 \pm 0.10$   $\mu$ M, respectively). This might be due to the more bulky substitution pattern around the pyrrole C-5 substituent.

**Table 3.3 | EC<sub>50</sub> values observed in the TR-FRET coactivator recruitment assays on PPAR $\gamma$ , IC<sub>50</sub> values from the competition TR-FRET assays on PPAR $\gamma$ , and  $\Delta T_m$  values (°C) from thermal shift assays. TR-FRET and TSA data are representative of three independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD.**

Compound	EC <sub>50</sub> ( $\mu$ M) PPAR $\gamma$	IC <sub>50</sub> ( $\mu$ M) PPAR $\gamma$ (competition rosiglitazone)	$\Delta T_m$ (°C) PPAR $\gamma$	Fold-selective for ROR $\gamma$ t over PPAR $\gamma$	
				TR-FRET	TSA
<b>1</b>	$0.34 \pm 0.02$	$6.3 \pm 0.7$	$2.5 \pm 0.2$	26	3.1
<b>2</b>	$8.2 \pm 0.3$	$81 \pm 7$	$0.3 \pm 0.0$	30	8.0
<b>3</b>	$3.1 \pm 0.1$	$63 \pm 8$	$0.6 \pm 0.0$	100	8.1
<b>4</b>	$18 \pm 1$	> 100	$0.7 \pm 0.0$	2.7	0.5
<b>5</b>	> 100	> 100	$0.0 \pm 0.0$	–	–
<b>6</b>	$2.5 \pm 0.1$	$64 \pm 3$	$1.1 \pm 0.2$	125	5.8
<b>7</b>	$7.5 \pm 0.3$	>100	$0.1 \pm 0.0$	15	7.0
<b>8</b>	$15 \pm 1$	>100	$0.1 \pm 0.0$	1.6	7.0
<b>9</b>	$5.8 \pm 0.2$	$67 \pm 6$	$0.4 \pm 0.0$	67	7.5
<b>10</b>	$1.7 \pm 0.1$	$29 \pm 4$	$1.2 \pm 0.0$	7.1	2.3
<b>11</b>	$3.6 \pm 0.1$	$79 \pm 11$	$0.7 \pm 0.0$	33	3.9
<b>12</b>	$1.2 \pm 0.1$	$20 \pm 2$	$1.5 \pm 0.0$	0.36	<0.1
<b>13</b>	$0.79 \pm 0.10$	$19 \pm 3$	$1.5 \pm 0.0$	0.27	<0.1

In order to further explore the binding site of the isoxazole compounds on PPAR $\gamma$ , a TR-FRET assay was performed whereby the competition between the isoxazole ligands and a known orthosteric PPAR $\gamma$  agonist, rosiglitazone, was measured (Table 3.3).<sup>34-45</sup> All compounds showed competition with rosiglitazone at high concentrations, indicating binding to the orthosteric site of PPAR $\gamma$ . Interestingly, at least a 3-fold lower competition with rosiglitazone was observed for the isoxazole compounds compared to **1** (i.e. at least a 3-fold increase of the IC<sub>50</sub> values compared to **1**), with most analogues demonstrating IC<sub>50</sub> values >50  $\mu$ M. Overall, a similar trend was observed for the IC<sub>50</sub> values in this competition assay as for the EC<sub>50</sub> values in the previous TR-FRET coactivator recruitment assay (Table 3.2). Again, compounds **12** and **13** appeared to be less selective for ROR $\gamma$ t than the more potent ROR $\gamma$ t-binding isoxazoles (IC<sub>50</sub> values of  $20 \pm 2$  and  $19 \pm 3$   $\mu$ M, respectively). Compound **10** showed some cross-reactivity with PPAR $\gamma$  as well, although with an IC<sub>50</sub> value >20  $\mu$ M.

Next, the compounds were evaluated in a PPAR $\gamma$  TSA (Table 3.3). Compounds **10**, **12** and **13** showed the highest thermal stabilization effect ( $\Delta T_m$  of 1.2 and 1.5  $^{\circ}$ C), in line with the TR-FRET data, and also **6** appeared to show a significant stabilization effect ( $\Delta T_m$  of 1.1  $^{\circ}$ C). Compound **1** again showed the most significant activity on PPAR $\gamma$ , with a  $\Delta T_m$  of 2.5  $^{\circ}$ C.

The selectivity of the compounds for ROR $\gamma$ t over PPAR $\gamma$  was then calculated (Table 3.3) by comparing the data for PPAR $\gamma$  and ROR $\gamma$ t for both the TR-FRET coactivator recruitment assay and TSA (see Table 3.1, 3.2 and 3.3). Interestingly, compounds **3** and **6**, which are highly potent ROR $\gamma$ t inverse agonists (IC<sub>50</sub> values <35 nM), show a 100- and 125-fold selectivity for ROR $\gamma$ t over PPAR $\gamma$  respectively, based on the TR-FRET results, which is a significant improvement compared to **2** (30-fold selectivity). Contrastingly, **12** and **13**, containing a methyl substituent at the pyrrole moiety, show higher cross-reactivity towards PPAR $\gamma$ , in the same range as for **1**. Together, the cross-reactivity of the isoxazoles on PPAR $\gamma$  can be modulated by only small changes in the pyrrole substitution pattern. Gratifyingly, the most potent compounds for ROR $\gamma$ t (**3** and **6**) also feature the highest selectivity for ROR $\gamma$ t over PPAR $\gamma$ .

## ADME profile

Previous studies on the ADME parameters of the isoxazoles showed that these were promising but not optimal yet for compound **2**, most probably governed by the pyrrole moiety and the amine linker, which are both likely to be prone to oxidation.<sup>34-46</sup> The ADME profiling of the novel isoxazoles (Table 3.4) demonstrates that compounds **3**, **6**, **9** and **10** in general show comparable properties to hit compound **2**, with promising values for the solubility and



(passive) membrane permeability, especially because the compounds contain a carboxylic acid moiety which generally leads to lower permeability values. The metabolic stability in human liver microsomes showed that the clearance values in phase I metabolism are rather high, similar to **2**. Phase II stability showed promising improved values (% remain) compared to **1**, especially for compound **10**, though only glucuronidation (as the most important phase II pathway) was investigated. The plasma stability (except for **6**) was slightly lower than for **1** and relatively high levels of blood plasma protein binding were measured. Contrastingly, compound **11**, containing a pyrazole instead of a pyrrole substituent at the C-5 position, showed a rather different ADME profile. It demonstrated a promising stability in phase I metabolism and was highly stable in plasma, although the phase II stability and permeability were lower than for the other isoxazoles.

The open-source program OSIRIS Property Explorer<sup>47</sup> was used to predict the drug-likeness and toxicity of the compounds (Supporting Table S3.3) to have a first indication on these aspects. All isoxazole compounds (except **5**) were predicted to have no side effects (i.e. mutagenic, tumorigenic, irritant and reproductive effects), whereas **1** was predicted to have a potentially high risk of reproductive toxicity. Additionally, all isoxazole ligands showed higher drug-scores than **1**, with pyrazole **11** showing the best properties, again highlighting its potential for further development.

Combined, the change of an amine to an ether or alkene linker (**2** vs. **3** and **6**) does not significantly affect the ADME(T) properties (except for the phase II metabolism of **3**), while the change of a pyrrole to a pyrazole moiety (**11**) results in a differentiated, promising profile. The C-4/C-5 isoxazole modifications thus provide valuable entry points for optimization of the ADME properties.

**Table 3.4 | ADME properties for compounds **1**, **2**, **3**, **6**, **9**, **10** and **11**.**

Compound	Solubility ( $\mu$ M)	PAMPA (% Flux)	Microsomal Stability		Plasma stability (% remain)	Plasma protein binding (% bound)
			Phase I (Cl <sub>int</sub> , $\mu$ L/min/mg)	Phase II (% remain)		
<b>1</b>	423	29.7	0.4	37.7	95.5	99.7
<b>2</b>	460	36.1	31.2	73.0	79.1	99.9
<b>3</b>	436	40.9	57.8	30.7	74.2	99.7
<b>6</b>	392	40.3	69.0	58.3	59.5	99.9
<b>9</b>	445	37.3	77.0	54.6	97.1	99.8
<b>10</b>	421	50.5	68.0	68.9	89.5	99.3
<b>11</b>	476	23.0	8.8	39.1	100.0	98.8

## Conclusions

The nuclear receptor ROR $\gamma$ t has an important regulatory role in the immune system and inhibition of ROR $\gamma$ t via allosteric inverse agonists could be a promising strategy in the treatment of autoimmune disorders. Apart from indazoles, with **1** (**MRL-871**) as a prominent example, there has been little chemical diversity of allosteric ROR $\gamma$ t modulators. Trisubstituted isoxazoles constitute a distinct novel chemotype in this matter, with **2** (**FM26**) as an exemplary potent hit compound.

Here, we further optimized this isoxazole series, specifically focusing on the linker at the C-4 position and the heterocycle at the C-5 position to deliver lead compounds for further optimization. Optimization of the linker at the C-4 position revealed a clear correlation between the potency and the lipophilicity and flexibility of the linker, with an ether or alkene linker resulting in the highest potency. The co-crystal structure of ROR $\gamma$ t in complex with ether compound **3** revealed the role of hydrophobic interactions and conformation of the linker at the C-4 position. The most potent compounds **3** (**FM156**) and **6** (**FM260**) were highly selective for ROR $\gamma$ t over PPAR $\gamma$ , which is a valuable improvement compared to the exemplary indazole **1**.

The focused library of C-5 isoxazoles revealed several routes for chemical exploration. Changing the position of the hydrogen bond donor via repositioning of the pyrrole ring (**10**) led only to a 3-fold lower potency, with the formation of another hydrogen bond interaction network via a water molecule in the ROR $\gamma$ t pocket. Changing the pyrrole to a pyrazole (with a lower pKa) in compound **11** (**FM257**) was also tolerated and it resulted in a promising PK profile, opening up routes for affinity and ADME optimization.

In conclusion, by probing the effect of the linker at the C-4 position and the importance of the polar character of the C-5 moiety, this study has led to valuable leads as allosteric inverse agonists for ROR $\gamma$ t with improved potency, selectivity against PPAR $\gamma$  and initial ADME properties. The results not only provide new insights into the SAR for this specific isoxazole class of allosteric ROR $\gamma$ t ligands, but can also most probably similarly be translated to other allosteric ROR $\gamma$ t chemotypes. Overall, the trisubstituted isoxazole class of allosteric ROR $\gamma$ t ligands shows high potential for chemical biology approaches as well as for future development in drug discovery programs against autoimmune diseases.

## Acknowledgements

Annet Saris is kindly acknowledged for her efforts on the optimization of the synthesis route. Rens de Vries, Bente Somsen and Guido Oerlemans are gratefully acknowledged for performing X-ray crystallography studies. Anke Unger and Bert Klebl from the Lead Discovery Center in Dortmund are acknowledged for performing ADME measurements and discussing the data. Richard Doveston and Peter Cossar are kindly thanked for the helpful discussions and suggestions for the chemical synthesis.

This chapter was reprinted (adapted) with permission from *J. Med. Chem.* Copyright 2021 American Chemical Society. Further permissions related to this material should be directed to the ACS.

## Experimental Section

**Molecular docking studies.** *In silico* docking studies were performed as described in Chapter 2, using Maestro (version 12.3 (2020-1), Schrödinger LLC).

***In silico* drug-likeness and toxicity predictions.** OSIRIS Property Explorer utilizes the database of traded drugs and commercially available compounds (Fluka), assumable as non-drug-like dataset, to assess the occurrence frequency of each fragment in the individual structure.<sup>47</sup> The program was used to estimate the risks of side effects, such as mutagenic, tumorigenic, irritant and reproductive effects, as well as druglikeness and overall drug-score (by combining the outcome of cLogP, LogS (solubility), MW, toxicity risks and druglikeness). The drug-score is a measure of the compound's potential to meet the criteria of a possible drug candidate.

**General chemistry.** All non-aqueous reactions were performed under an argon atmosphere unless otherwise stated. Water-sensitive reactions were performed in oven-dried glassware, cooled under argon before use. All solvents were supplied by Biosolve or Sigma-Aldrich and used without further purification. Dry solvents were obtained from a MBRAUN Solvent Purification System (MB-SPS-800). Water was purified by a Millipore purification train. Deuterated solvents were obtained from Cambridge Isotope Laboratories. Solvents were removed *in vacuo* using a Büchi rotary evaporator and a diaphragm pump. Commercially available starting materials were obtained from Sigma-Aldrich, TCI Chemicals and Fluorochem. Proton ( $^1\text{H}$ ) NMR (400 MHz), carbon ( $^{13}\text{C}$ ) NMR (100 MHz) and 2D NMR (400 MHz) were recorded on a Bruker Avance 400 MHz NMR spectrometer. Proton spectra are referenced to tetramethyl silane (TMS). Carbon spectra are referenced to TMS or the solvent peak of the deuterated spectrum. NMR spectra are reported as follows: chemical shift ( $\delta$ ) in parts per million (ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublet, td = triplet of doublets, app. = apparent), coupling constant ( $J$ ) in Hertz (Hz) (if applicable) and integration (proton spectra only). Peak assignments are based on additional 2D NMR techniques (COSY, HMBC, HSQC). Analytical Liquid Chromatography coupled with Mass Spectrometry (LC-MS) was performed on a C<sub>4</sub> Jupiter SuC4300A 150 x 2.0 mm column using ultrapure water with 0.1% formic acid (FA) and acetonitrile with 0.1% FA, in general with a gradient of 5% to 100% acetonitrile over 10 min, connected to a Thermo Fisher LCQ Fleet Ion Trap Mass Spectrometer. The purity of the samples was assessed using a UV detector at 254 nm. Unless otherwise stated all final compounds were >95% pure as judged by HPLC. High resolution mass spectra (HRMS) were recorded using a Waters ACQUITY UPLC I-Class LC system coupled to a Xevo G2 Quadrupole Time of Flight (Q-TOF) mass spectrometer. Preparative HP-LC was performed on a Gemini S4 110A 150 x 21.20 mm column using ultrapure water with 0.1% FA and acetonitrile with 0.1% F.A. with various gradients (mentioned for each compound specifically).

Column chromatography was either performed manually using silica gel (60-200  $\mu\text{m}$  particle size, 60 Å) or using an automated Grace Reveleris X2 chromatograph with pre-packed silica columns supplied by Buchi/Grace (40  $\mu\text{m}$  particle size). Reaction progress was monitored by thin-layer liquid chromatography (TLC) using Merck TLC silica gel 60 F254 plates. Visualization of the plates was achieved using an ultraviolet lamp ( $\lambda_{\text{max}} = 254 \text{ nm}$ ).

### Synthetic procedures

**General Procedure for Suzuki Coupling.** Under an inert atmosphere the pinacol boronate (2.0 eq),  $\text{Cs}_2\text{CO}_3$  (2.0 eq) and  $\text{Pd}(\text{dppf})\text{Cl}_2$  (0.1 eq) were added to a solution of bromide **14** (1.0 eq) in de-gassed DME (0.1 M). The reaction mixture was heated at 85 °C for 8 h, cooled to room temperature, diluted with  $\text{H}_2\text{O}$  and extracted with  $\text{EtOAc}$  (3 x). The combined organic phase was washed with brine, dried over  $\text{MgSO}_4$ , filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography using the specified eluent to afford the desired product.

**General Procedure for Reduction of Esters to Alcohols.** Under an inert atmosphere, DIBAL (1.23 g/mL in cyclohexane, 15.0 eq) was added dropwise to a solution of ester (1.0 eq) in anhydrous  $\text{CH}_2\text{Cl}_2$  (0.1 M) at -78 °C. The reaction mixture was followed by TLC analysis. Upon complete consumption of the starting material, the reaction mixture was warmed to room temperature, quenched by addition of a saturated Rochelle Salt ( $\text{KNaC}_4\text{H}_4\text{O}_6 \cdot 4\text{H}_2\text{O}$ ) solution and stirred vigorously for 60 min. Subsequently, the mixture was extracted with  $\text{CH}_2\text{Cl}_2$  (3 x), and separated. The combined organic phase was washed with brine, dried over  $\text{MgSO}_4$ , filtered and concentrated *in vacuo* to afford the title compound which was purified as described.

**General Procedure for Mesylation and Substitution.** Under an inert atmosphere, triethylamine (3.0 eq) was added to a solution of alcohol compound **16** (1.0 eq) in anhydrous  $\text{CH}_2\text{Cl}_2$  (0.1 M). Methanesulfonic anhydride (1.5 eq) was added to the flask and the reaction mixture was heated at 40 °C. The reaction was monitored using NMR analysis. Subsequently, aniline (5.0 eq) was added and the reaction mixture was stirred for 1 h, cooled to room temperature, and  $\text{CH}_2\text{Cl}_2$  was removed *in vacuo*. The crude product was purified by flash column chromatography using the specified eluent.

**General Procedure for Ester Hydrolysis and N-Boc Deprotection.**  $\text{LiOH} \cdot \text{H}_2\text{O}$  (10.0 eq) was added to a suspension of ester (1.0 eq) in a 4:1 mixture of  $\text{EtOH}/\text{H}_2\text{O}$  (0.05 M). The reaction mixture was heated to 95 °C and followed by TLC analysis. Upon complete consumption of the starting material,  $\text{EtOH}$  was removed *in vacuo* and the resulting aqueous mixture was acidified to pH 3 using 10% v/v aqueous  $\text{HCl}$  and extracted with a 9:1 mixture of  $\text{CH}_2\text{Cl}_2/\text{MeOH}$  (5 x). The combined organic phase was dried ( $\text{MgSO}_4$ ), filtered and concentrated *in vacuo* to furnish the final compound which was purified as described.

**General Procedure for Mitsunobu Coupling.** Under an inert atmosphere, triphenylphosphine (2.0 eq) and DIAD (2.0 eq) were dissolved in anhydrous THF (0.05 M) and stirred for 15 min at 0 °C. Subsequently, the alcohol compound (1.0 eq), the benzoate (1.1 eq) and triethylamine (1.0 eq) were added and the reaction mixture was heated at 80 °C for 3 h, cooled to room temperature and THF was removed *in vacuo*. The crude product was purified by flash column chromatography using the specified eluent.

**4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-3-yl)isoxazol-4-yl)methyl)amino) benzoic acid (**2**).** According to the General Procedure for Ester Hydrolysis and N-Boc Deprotection, ester **17** (0.035 g, 0.067 mmol) was hydrolyzed with concomitant loss of the Boc protecting group. The crude product was purified by column chromatography, eluting with 1.5%  $\text{MeOH}$  in  $\text{CH}_2\text{Cl}_2$ , to furnish the carboxylic acid **2** (10.3 mg, 33%) as a white solid.  $^1\text{H-NMR}$  (400 MHz,  $\text{DMSO-d}_6$ ):  $\delta$  (ppm) 11.98 (1H, br. s,  $\text{CO}_2\text{H}$ ), 11.52 (1H, s, pyrrole-NH), 7.90 (1H, d,  $J = 8.0 \text{ Hz}$ , ArH-3), 7.84 (1H, d,  $J = 7.8 \text{ Hz}$ , ArH-5), 7.72 (1H, app.

t,  $J = 8.0$  Hz, ArH-4), 7.53 (2H, d,  $J = 8.7$  Hz, benzoate C-2), 7.39 (1H, m, pyrrole H-2), 6.98 (1H, m, pyrrole H-5), 6.55 (1H, m, pyrrole H-4), 6.41 (2H, d,  $J = 8.7$  Hz, benzoate H-3), 6.34 (1H, app. t,  $J = 4.4$  Hz, CH<sub>2</sub>NH), 4.07 (1H, dd,  $J = 14.4, 5.0$  Hz, CH<sub>A</sub>NH), 4.02 (1H, dd,  $J = 14.5, 4.5$  Hz, CH<sub>B</sub>NH); <sup>13</sup>C-NMR (100 MHz, DMSO-d<sub>6</sub>):  $\delta$  (ppm) 167.9 (CO<sub>2</sub>H), 165.7 (C-5), 159.2 (C-3), 152.3 (benzoate C-4), 136.0 (ArC-2), 134.1 (ArC-3), 132.4 (ArC-4), 131.3 (benzoate C-2), 130.9 (q,  $J = 30.4$  Hz, ArC-6), 126.8 (ArC-1), 125.8 (q,  $J = 5.0$  Hz, ArC-5), 122.0 (q,  $J = 274.3$  Hz, CF<sub>3</sub>), 120.5 (pyrrole C-5), 119.2 (pyrrole C-2), 117.8 (benzoate C-1), 111.2 (benzoate C-3), 110.5 (pyrrole C-3), 109.2 (C-4), 106.7 (pyrrole C-4), 36.2 (CH<sub>2</sub>NH). LC-MS (ESI): calc. for C<sub>22</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>3</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 462.08, observed: 462.17 ( $R_t = 6.13$  min). HRMS (ESI): calc. for C<sub>22</sub>H<sub>15</sub>ClF<sub>3</sub>N<sub>3</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 462.0832, observed: 462.0835.

**4-((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-3-yl)isoxazol-4-yl)methoxy) benzoic acid (3).** According to the General Procedure for Ester Hydrolysis and *N*-Boc Deprotection, ester **20** (0.020 g, 0.035 mmol) was hydrolyzed with concomitant loss of the Boc protecting group. The crude product was purified by column chromatography, eluting with 1.5% MeOH in CH<sub>2</sub>Cl<sub>2</sub>, to furnish the carboxylic acid **3** (12.0 mg, 75%) as a white solid. <sup>1</sup>H-NMR (400 MHz, acetone-d<sub>6</sub>):  $\delta$  (ppm) 10.78 (1H, br. s, CO<sub>2</sub>H), 7.92 (4H, m, ArH-3, ArH-5 and benzoate C-2), 7.77 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 7.47 (1H, m, pyrrole H-2), 7.03 (1H, m, pyrrole H-5), 6.94 (2H, d,  $J = 8.8$  Hz, benzoate H-3), 6.64 (1H, m, pyrrole H-4), 5.04 (2H, s, CH<sub>2</sub>O); <sup>13</sup>C-NMR (100 MHz, acetone-d<sub>6</sub>):  $\delta$  (ppm) 167.6 (C-5), 167.3 (benzoate C-4), 163.1 (CO<sub>2</sub>H), 159.8 (C-3), 137.2 (ArC-2), 134.4 (ArC-3), 132.8 (q,  $J = 30.4$  Hz, ArC-6), 132.6 (benzoate C-2), 132.5 (ArC-4), 127.9 (ArC-1), 126.1 (q,  $J = 5.0$  Hz, ArC-5), 124.2 (benzoate C-1), 122.7 (q,  $J = 274.3$  Hz, CF<sub>3</sub>), 121.0 (pyrrole C-5), 120.0 (pyrrole C-2), 115.1 (benzoate C-3), 111.6 (pyrrole C-3), 108.3 (C-4), 107.4 (pyrrole C-4), 60.8 (CH<sub>2</sub>O). LC-MS (ESI): calc. for C<sub>22</sub>H<sub>14</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 463.06, observed: 463.25 ( $R_t = 4.62$  min). HRMS (ESI): calc. for C<sub>22</sub>H<sub>14</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 463.0672, observed: 463.0650.

**4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-3-yl)isoxazol-4-yl)methyl)thio) benzoic acid (4).** According to the General Procedure for Ester Hydrolysis and *N*-Boc Deprotection, ester **22** (0.020 g, 0.034 mmol) was hydrolyzed with concomitant loss of the Boc protecting group. The crude product was purified by column chromatography, eluting with 2% MeOH in CH<sub>2</sub>Cl<sub>2</sub>, to furnish the carboxylic acid **4** (8.00 mg, 49%) as a white solid. <sup>1</sup>H-NMR (400 MHz, acetone-d<sub>6</sub>):  $\delta$  (ppm) 10.77 (1H, br. s, CO<sub>2</sub>H), 7.89 (4H, m, ArH-3, ArH-5 and benzoate C-2), 7.79 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 7.52 (1H, m, pyrrole H-2), 7.31 (2H, d,  $J = 8.5$  Hz, benzoate H-3), 7.03 (1H, m, pyrrole H-5), 6.68 (1H, m, pyrrole H-4), 4.22 (2H, s, CH<sub>2</sub>S); <sup>13</sup>C-NMR (100 MHz, acetone-d<sub>6</sub>):  $\delta$  (ppm) 167.2 (CO<sub>2</sub>H), 166.8 (C-5), 159.8 (C-3), 144.6 (benzoate C-4), 137.2 (ArC-2), 134.5 (ArC-3), 132.7 (ArC-4), 132.5 (q,  $J = 30.4$  Hz, ArC-6), 131.4 (benzoate C-1), 130.9 (benzoate C-2), 128.3 (ArC-1), 127.3 (benzoate C-3), 126.3 (q,  $J = 5.0$  Hz, ArC-5), 122.7 (q,  $J = 274.3$  Hz, CF<sub>3</sub>), 120.9 (pyrrole C-5), 119.7 (pyrrole C-2), 111.6 (pyrrole C-3), 107.3 (C-4), 107.3 (pyrrole C-4), 26.2 (CH<sub>2</sub>S). LC-MS (ESI): calc. for C<sub>22</sub>H<sub>14</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>3</sub>S [M+H]<sup>+</sup>: 479.04, observed: 479.25 ( $R_t = 4.86$  min). HRMS (ESI): calc. for C<sub>22</sub>H<sub>14</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>3</sub>S [M+H]<sup>+</sup>: 479.0444, observed: 479.0438.

**4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-3-yl)isoxazol-4-yl)methyl) (methyl)amino) benzoic acid (5).** According to the General Procedure for Ester Hydrolysis and *N*-Boc Deprotection, ester **18** (0.040 g, 0.068 mmol) was hydrolyzed with concomitant loss of the Boc protecting group. The crude product was purified by column chromatography, eluting with 2% MeOH in CH<sub>2</sub>Cl<sub>2</sub>, to furnish the carboxylic acid **5** (19.3 mg, 60%) as a white solid. <sup>1</sup>H-NMR (400 MHz, DMSO-d<sub>6</sub>):  $\delta$  (ppm) 12.08 (1H, br. s, CO<sub>2</sub>H), 11.60 (1H, s, pyrrole-NH), 7.80 (1H, d,  $J = 8.1$  Hz, ArH-3), 7.76 (1H, d,  $J = 7.8$  Hz, ArH-5), 7.64 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 7.53 (1H, m, pyrrole H-2), 7.49 (2H, d,  $J = 8.9$  Hz, benzoate C-2), 7.02 (1H, m, pyrrole H-5), 6.63 (1H, m, pyrrole H-4), 6.33 (2H, d,  $J = 8.9$  Hz, benzoate H-3), 4.48 (2H, s, CH<sub>2</sub>NCH<sub>3</sub>), 2.54 (3H, s, CH<sub>3</sub>NCH<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, DMSO-d<sub>6</sub>):  $\delta$  (ppm) 167.3 (CO<sub>2</sub>H), 165.0 (C-5), 158.5 (C-3), 151.7 (benzoate C-4), 135.5 (ArC-2), 133.6 (ArC-3), 131.8 (ArC-4), 130.5 (benzoate C-2), 130.1 (q,  $J = 30.4$  Hz, ArC-6), 126.5 (ArC-1), 125.4 (q,  $J = 5.0$  Hz, ArC-5), 124.2 (q,  $J = 274.3$  Hz, CF<sub>3</sub>), 120.2

(pyrrole C-5), 118.8 (pyrrole C-2), 117.3 (benzoate C-1), 110.5 (benzoate C-3), 109.7 (pyrrole C-3), 109.0 (C-4), 106.3 (pyrrole C-4), 43.9 ( $\text{CH}_2\text{NCH}_3$ ), 37.0 ( $\text{CH}_2\text{NCH}_3$ ). LC-MS (ESI): calc. for  $\text{C}_{23}\text{H}_{17}\text{ClF}_3\text{N}_3\text{O}_3$   $[\text{M}+\text{H}]^+$ : 476.09, observed: 476.17 ( $R_t = 6.33$  min). HRMS (ESI): calc. for  $\text{C}_{23}\text{H}_{17}\text{ClF}_3\text{N}_3\text{O}_3$   $[\text{M}+\text{H}]^+$ : 476.0989, observed: 476.0971.

**(E)-4-(2-(3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-3-yl)isoxazol-4-yl)vinyl)benzoic acid (6).**

According to the General Procedure for Ester Hydrolysis and *N*-Boc Deprotection, ester **24** (0.012 g, 0.021 mmol) was hydrolyzed with concomitant loss of the Boc protecting group. The crude product was purified via preparative HPLC (gradient of 70–73% acetonitrile in  $\text{H}_2\text{O}$ ) to furnish the carboxylic acid **6** (6.1 mg, 63%) as a white solid.  $^1\text{H}$ -NMR (400 MHz, acetone- $d_6$ ):  $\delta$  (ppm) 10.85 (1H, br. s,  $\text{CO}_2\text{H}$ ), 8.00 (2H, ArH-3 and ArH-5), 7.92 (3H, m, benzoate C-2 and ArH-4), 7.62 (1H, m, pyrrole H-2), 7.39 (3H, m, benzoate H-3 and  $\text{C}_4\text{-HC=CH}$ ), 7.08 (1H, m, pyrrole H-5), 6.74 (1H, m, pyrrole H-4), 6.15 (1H, s, benzoate  $\text{C}_4\text{-HC=CH}$ );  $^{13}\text{C}$ -NMR (100 MHz, acetone- $d_6$ ):  $\delta$  (ppm) 166.3 ( $\text{CO}_2\text{H}$ ), 165.2 (C-5), 156.8 (C-3), 141.5 (benzoate C-4), 136.1 (ArC-2), 133.8 (ArC-3), 131.9 (ArC-4), 131.4 (q,  $J = 30.4$  Hz, ArC-6), 130.0 (benzoate C-2), 129.4 (benzoate C-1), 128.1 (benzoate  $\text{C}_4\text{-HC=CH}$ ), 128.0 (ArC-1), 126.0 (benzoate C-3), 125.6 (q,  $J = 5.0$  Hz, ArC-5), 121.9 (q,  $J = 274.3$  Hz,  $\text{CF}_3$ ), 120.2 (pyrrole C-5), 119.3 (pyrrole C-2), 118.8 ( $\text{C}_4\text{-HC=CH}$ ), 110.9 (pyrrole C-3), 110.3 (C-4), 106.7 (pyrrole C-4). LC-MS (ESI): calc. for  $\text{C}_{23}\text{H}_{14}\text{ClF}_3\text{N}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 459.06, observed: 459.25 ( $R_t = 4.90$  min). HRMS (ESI): calc. for  $\text{C}_{23}\text{H}_{14}\text{ClF}_3\text{N}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 459.0723, observed: 459.0726.

**(Z)-4-(2-(3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-3-yl)isoxazol-4-yl)vinyl)benzoic acid (7).**

According to the General Procedure for Ester Hydrolysis and *N*-Boc Deprotection, ester **25** (0.040 g, 0.070 mmol) was hydrolyzed with concomitant loss of the Boc protecting group. The crude product was purified via preparative HPLC (gradient of 70–73% acetonitrile in  $\text{H}_2\text{O}$ ) to furnish the carboxylic acid **7** (26.0 mg, 81%) as a white solid.  $^1\text{H}$ -NMR (400 MHz, acetone- $d_6$ ):  $\delta$  (ppm) 10.50 (1H, br. s,  $\text{CO}_2\text{H}$ ), 7.80 (5H, m, ArH-3, ArH-5, ArH-4 and benzoate C-2), 7.36 (2H, d,  $J = 8.2$  Hz, benzoate H-3), 7.21 (1H, m, pyrrole H-2), 6.79 (1H, m, pyrrole H-5), 6.74 (1H, d,  $J = 12.2$  Hz, benzoate  $\text{C}_4\text{-HC=CH}$ ), 6.49 (1H, m, pyrrole H-4), 6.29 (1H, d,  $J = 12.2$  Hz,  $\text{C}_4\text{-HC=CH}$ );  $^{13}\text{C}$ -NMR (100 MHz, acetone- $d_6$ ):  $\delta$  (ppm) 166.3 ( $\text{CO}_2\text{H}$ ), 163.7 (C-5), 158.4 (C-3), 141.1 (benzoate C-4), 136.1 (ArC-2), 133.5 (benzoate  $\text{C}_4\text{-HC=CH}$ ), 133.5 (ArC-3), 131.6 (ArC-4), 131.3 (q,  $J = 30.4$  Hz, ArC-6), 129.4 (benzoate C-2), 128.4 (benzoate C-3), 127.2 (benzoate C-1), 125.3 (q,  $J = 5.0$  Hz, ArC-5), 124.5 (ArC-1), 121.8 (q,  $J = 274.3$  Hz,  $\text{CF}_3$ ), 119.3 (pyrrole C-5), 118.8 (pyrrole C-2), 118.7 ( $\text{C}_4\text{-HC=CH}$ ), 111.2 (pyrrole C-3), 108.6 (C-4), 106.3 (pyrrole C-4). LC-MS (ESI): calc. for  $\text{C}_{23}\text{H}_{14}\text{ClF}_3\text{N}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 459.06, observed: 459.25 ( $R_t = 4.87$  min). HRMS (ESI): calc. for  $\text{C}_{23}\text{H}_{14}\text{ClF}_3\text{N}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 459.0723, observed: 459.0715.

**4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-3-yl)isoxazol-4-yl)methyl) amino)-2-fluorobenzoic acid (8).** According to the General Procedure for Ester Hydrolysis and *N*-Boc Deprotection, ester **19** (0.008 g, 0.016 mmol) was hydrolyzed. The crude product was purified by column chromatography, eluting with 1.5% MeOH in  $\text{CH}_2\text{Cl}_2$ , to furnish the carboxylic acid **8** (2.2 mg, 29%) as a white solid.  $^1\text{H}$ -NMR (400 MHz, DMSO- $d_6$ ):  $\delta$  12.09 (1H, br. s,  $\text{CO}_2\text{H}$ ), 11.5 (1H, s, pyrrole-NH), 7.87 (1H, d,  $J = 8.0$  Hz, ArH-3), 7.82 (1H, d,  $J = 7.9$  Hz, ArH-5), 7.72 (1H, app. t,  $J = 7.9$  Hz, ArH-4), 7.47 (1H, app. t,  $J = 8.8$  Hz, benzoate H-2), 7.38 (1H, m, pyrrole H-2), 6.97 (1H, m, pyrrole H-5), 6.64 (1H, app. t,  $J = 4.8$  Hz,  $\text{CH}_2\text{NH}$ ), 6.55 (1H, m, pyrrole H-4), 6.24 (1H, dd,  $J = 8.8, 2.1$  Hz, benzoate H-3a), 6.10 (1H, dd,  $J = 14.4, 1.8$  Hz, benzoate H-3b), 4.06 (2H, m,  $\text{CH}_2\text{NH}$ ). LC-MS (ESI): calc. for  $\text{C}_{22}\text{H}_{14}\text{ClF}_4\text{N}_3\text{O}_3$   $[\text{M}+\text{H}]^+$ : 480.07, observed: 480.25 ( $R_t = 6.21$  min). HRMS (ESI): calc. for  $\text{C}_{22}\text{H}_{14}\text{ClF}_4\text{N}_3\text{O}_3$   $[\text{M}+\text{H}]^+$ : 480.0738, observed: 480.0759.

**4-((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-3-yl)isoxazol-4-yl)methoxy)-2-fluorobenzoic acid (9).** According to the General Procedure for Ester Hydrolysis and *N*-Boc Deprotection, ester **21** (0.017 mg, 0.029 mmol) was hydrolyzed with concomitant loss of the Boc protecting group. The crude product was purified by column chromatography, eluting with 1.5% MeOH in CH<sub>2</sub>Cl<sub>2</sub>, to furnish the carboxylic acid **9** (8.0 mg, 57%) as a white solid. <sup>1</sup>H-NMR (400 MHz, acetone-d<sub>6</sub>):  $\delta$  (ppm) 10.81 (1H, br. s, CO<sub>2</sub>H), 7.98 (3H, m, ArH-3, ArH-5 and benzoate C-2), 7.78 (1H, app. t,  $J$  = 8.0 Hz, ArH-4), 7.48 (1H, m, pyrrole H-2), 7.02 (1H, m, pyrrole H-5), 6.73 (2H, m, benzoate H-3), 6.64 (1H, m, pyrrole H-4), 5.06 (2H, s, CH<sub>2</sub>O); <sup>13</sup>C-NMR (100 MHz, acetone-d<sub>6</sub>):  $\delta$  (ppm) 167.4 (C-5), 165.5 (CO<sub>2</sub>H), 164.0 (benzoate C-4), 162.9 (benzoate C-F), 159.7 (C-3), 137.2 (ArC-2), 134.4 (ArC-3), 132.7 (q,  $J$  = 30.4 Hz, ArC-6), 132.6 (ArC-4), 132.4 (benzoate C-2), 127.8 (ArC-1), 126.1 (q,  $J$  = 5.0 Hz, ArC-5), 125.4 (benzoate C-1), 122.7 (q,  $J$  = 274.3 Hz, CF<sub>3</sub>), 121.0 (pyrrole C-5), 120.0 (pyrrole C-2), 111.7 (benzoate C-3a), 111.5 (pyrrole C-3), 108.0 (C-4), 107.4 (pyrrole C-4), 103.7 (benzoate C-3b), 61.3 (CH<sub>2</sub>O). LC-MS (ESI): calc. for C<sub>22</sub>H<sub>13</sub>ClF<sub>4</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 481.05, observed: 481.25 ( $R_t$  = 4.92 min). HRMS (ESI): calc. for C<sub>22</sub>H<sub>13</sub>ClF<sub>4</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 481.0578, observed: 481.0572.

**4-((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-2-yl)isoxazol-4-yl)methoxy) benzoic acid (10).** According to the General Procedure for Mitsunobu Coupling (without the addition of triethylamine), alcohol **30** (0.080 g, 0.180 mmol) was reacted with methyl 4-hydroxybenzoate (0.030 g, 0.200 mmol). The crude product was purified by column chromatography, eluting with a gradient of 5-25% EtOAc in *n*-heptane, to furnish the ether compound (22.0 mg, 21%). The resulting product (0.012 g, 0.021 mmol) was subject to ester hydrolysis with concomitant loss of the Boc protecting group according to the General Procedure for Ester Hydrolysis and *N*-Boc Deprotection. The crude product was purified by column chromatography, eluting with 1.5% MeOH in CH<sub>2</sub>Cl<sub>2</sub>, to furnish the carboxylic acid **10** (4.0 mg, 41%) as a white solid. <sup>1</sup>H-NMR (400 MHz, acetone-d<sub>6</sub>):  $\delta$  (ppm) 11.13 (1H, br. s, CO<sub>2</sub>H), 7.91 (4H, m, ArH-3, ArH-5 and benzoate C-2), 7.80 (1H, app. t,  $J$  = 8.0 Hz, ArH-4), 7.19 (1H, m, pyrrole H-5), 6.93 (2H, d,  $J$  = 8.4 Hz, benzoate H-3), 6.79 (1H, m, pyrrole H-3), 6.34 (1H, m, pyrrole H-4), 5.09 (2H, s, CH<sub>2</sub>O); <sup>13</sup>C-NMR (100 MHz, acetone-d<sub>6</sub>):  $\delta$  (ppm) 162.4 (C-5), 162.0 (CO<sub>2</sub>H), 161.2 (benzoate C-4), 159.1 (C-3), 136.2 (ArC-2), 133.5 (ArC-3), 131.9 (benzoate C-2), 131.6 (ArC-4), 131.5 (q,  $J$  = 30.4 Hz, ArC-6), 126.5 (ArC-1), 125.3 (q,  $J$  = 5.0 Hz, ArC-5), 123.2 (benzoate C-1), 122.7 (pyrrole C-5), 121.8 (q,  $J$  = 274.3 Hz, CF<sub>3</sub>), 118.7 (pyrrole C-2), 114.2 (benzoate C-3), 111.6 (pyrrole C-3), 111.3 (pyrrole C-4), 107.6 (C-4), 59.8 (CH<sub>2</sub>O). LC-MS (ESI): calc. for C<sub>22</sub>H<sub>14</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 463.06, observed: 463.17 ( $R_t$  = 4.87 min). HRMS (ESI): calc. for C<sub>22</sub>H<sub>14</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 463.0672, observed: 463.0665.

**4-((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrazol-4-yl)isoxazol-4-yl)methoxy) benzoic acid (11).** According to the General Procedure for Mitsunobu Coupling, alcohol **31** (0.100 g, 0.234 mmol) was reacted with methyl 4-hydroxybenzoate (0.039 g, 0.257 mmol). The crude product was purified by column chromatography, eluting with a gradient of 10-25% EtOAc in *n*-heptane, to furnish the ether compound (65.0 mg, 49%). The ether compound (0.045 g, 0.080 mmol), containing a THP-protected pyrazole, was diluted in a mixture of TFA/CH<sub>2</sub>Cl<sub>2</sub> (50:50) (0.05 M) and the reaction mixture was stirred at 40 °C for 2 h. Afterwards, it was cooled to room temperature and the solvent was removed *in vacuo*. The crude product was purified by column chromatography, eluting with 25-40% EtOAc in *n*-heptane, to furnish the pyrazole product (28.0 mg, 88%). The resulting product was subject to ester hydrolysis according to the General Procedure for Ester Hydrolysis and *N*-Boc Deprotection. The crude product was purified via preparative HPLC (gradient of 62-67% acetonitrile in H<sub>2</sub>O) to furnish the carboxylic acid **11** (26.0 mg, 89%) as a white solid. <sup>1</sup>H-NMR (400 MHz, acetone-d<sub>6</sub>):  $\delta$  (ppm) 8.24 (2H, s, pyrazole H-3 and H-5), 7.92 (4H, m, ArH-3, ArH-5 and benzoate C-2), 7.80 (1H, app. t,  $J$  = 7.9 Hz, ArH-4), 6.94 (2H, d,  $J$  = 8.1 Hz, benzoate H-3), 5.12 (2H, s, CH<sub>2</sub>O); <sup>13</sup>C-NMR (100 MHz, acetone-d<sub>6</sub>):  $\delta$  (ppm) 167.2 (CO<sub>2</sub>H), 164.6 (C-5), 162.9 (benzoate C-4), 159.9 (C-3), 137.1 (ArC-2), 134.4 (ArC-3), 133.6 (pyrazole C-3 and C-5), 133.0 (q,  $J$  = 30.4 Hz, ArC-6), 132.8 (ArC-4), 132.4 (benzoate C-2), 127.4 (ArC-1), 126.2 (q,  $J$  = 5.0 Hz, ArC-

5), 124.3 (benzoate C-1), 122.7 (q,  $J = 274.3$  Hz,  $\text{CF}_3$ ), 115.1 (benzoate C-3), 110.2 (C-4), 109.6 (pyrrole C-4). LC-MS (ESI): calc. for  $\text{C}_{21}\text{H}_{13}\text{ClF}_3\text{N}_3\text{O}_4$   $[\text{M}+\text{H}]^+$ : 464.05, observed: 464.08 ( $R_t = 4.45$  min). HRMS (ESI): calc. for  $\text{C}_{21}\text{H}_{13}\text{ClF}_3\text{N}_3\text{O}_4$   $[\text{M}+\text{H}]^+$ : 464.0625, observed: 464.0610.

**4-((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1-methyl-1H-pyrrol-3-yl)isoxazol-4-yl)methoxy)benzoic acid (12).** According to the General Procedure for Mitsunobu Coupling, alcohol **32** (0.175 g, 0.490 mmol) was reacted with methyl 4-hydroxybenzoate (0.082 g, 0.540 mmol). The crude product was purified by column chromatography, eluting with a gradient of 5–20% EtOAc in *n*-heptane, to furnish the ether compound (109 mg, 45%). The resulting product (0.060 g, 0.122 mmol) was subject to ester hydrolysis according to the General Procedure for Ester Hydrolysis and *N*-Boc Deprotection. The crude product was purified by column chromatography, eluting with 1.5% MeOH in  $\text{CH}_2\text{Cl}_2$ , to furnish the carboxylic acid **12** (50.0 mg, 86%) as a white solid.  $^1\text{H}$ -NMR (400 MHz, acetone- $d_6$ ):  $\delta$  (ppm) 7.92 (4H, m, ArH-3, ArH-5 and benzoate C-2), 7.76 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 7.38 (1H, m, pyrrole H-2), 6.92 (2H, d,  $J = 8.8$  Hz, benzoate H-3), 6.88 (1H, m, pyrrole H-5), 6.56 (1H, m, pyrrole H-4), 5.04 (2H, s,  $\text{CH}_2\text{O}$ ), 3.78 (3H, s,  $\text{NCH}_3$ );  $^{13}\text{C}$ -NMR (100 MHz, acetone- $d_6$ ):  $\delta$  (ppm) 167.3 (C-5), 167.1 ( $\text{CO}_2\text{H}$ ), 163.1 (benzoate C-4), 159.7 (C-3), 137.1 (ArC-2), 134.3 (ArC-3), 132.7 (q,  $J = 30.4$  Hz, ArC-6), 132.5 (ArC-4), 132.4 (benzoate C-2), 127.9 (ArC-1), 126.1 (q,  $J = 5.1$  Hz, ArC-5), 124.8 (pyrrole C-5), 124.1 (benzoate C-1), 123.3 (pyrrole C-2), 122.7 (q,  $J = 274.0$  Hz,  $\text{CF}_3$ ), 115.1 (benzoate C-3), 111.6 (pyrrole C-3), 108.2 (C-4), 107.8 (pyrrole C-4), 60.8 ( $\text{CH}_2\text{O}$ ), 36.7 ( $\text{NCH}_3$ ). LC-MS (ESI): calc. for  $\text{C}_{23}\text{H}_{16}\text{ClF}_3\text{N}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 477.08, observed: 477.33 ( $R_t = 5.06$  min). HRMS (ESI): calc. for  $\text{C}_{23}\text{H}_{16}\text{ClF}_3\text{N}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 477.0829, observed: 477.0813.

**4-((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(5-methyl-1H-pyrrol-3-yl)isoxazol-4-yl)methoxy)benzoic acid (13).** According to the General Procedure for Mitsunobu Coupling, alcohol **33** (0.175 g, 0.383 mmol) was reacted with methyl 4-hydroxybenzoate (0.064 g, 0.421 mmol). The crude product was purified by column chromatography, eluting with a gradient of 10–40% EtOAc in *n*-heptane, to furnish the ether compound (100 mg, 44%). The resulting product (0.030 g, 0.051 mmol) was subject to ester hydrolysis with concomitant loss of the Boc protecting group according to the General Procedure for Ester Hydrolysis and *N*-Boc Deprotection. The crude product was purified via preparative HPLC (gradient of 65–70% acetonitrile in  $\text{H}_2\text{O}$ ) to furnish the carboxylic acid **13** (20.0 mg, 83%) as a white solid.  $^1\text{H}$ -NMR (400 MHz, acetone- $d_6$ ):  $\delta$  (ppm) 10.47 (1H, br. s,  $\text{CO}_2\text{H}$ ), 7.90 (4H, m, ArH-3, ArH-5 and benzoate C-2), 7.77 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 7.29 (1H, m, pyrrole H-2), 6.93 (2H, d,  $J = 8.5$  Hz, benzoate H-3), 6.44 (1H, m, pyrrole H-4), 5.02 (2H, s,  $\text{CH}_2\text{O}$ ), 2.29 (3H, s,  $\text{NCCH}_3$ );  $^{13}\text{C}$ -NMR (100 MHz, acetone- $d_6$ ):  $\delta$  (ppm) 167.7 (C-5), 167.3 (benzoate C-4), 163.1 ( $\text{CO}_2\text{H}$ ), 159.7 (C-3), 137.2 (ArC-2), 134.4 (ArC-3), 132.7 (q,  $J = 30.4$  Hz, ArC-6), 132.5 (benzoate C-2), 132.4 (ArC-4), 130.7 (pyrrole C-5), 128.0 (ArC-1), 126.1 (q,  $J = 5.0$  Hz, ArC-5), 124.2 (benzoate C-1), 122.7 (q,  $J = 274.3$  Hz,  $\text{CF}_3$ ), 119.7 (pyrrole C-2), 115.1 (benzoate C-3), 111.7 (pyrrole C-3), 108.0 (C-4), 105.2 (pyrrole C-4), 60.8 ( $\text{CH}_2\text{O}$ ), 12.6 ( $\text{NCCH}_3$ ). LC-MS (ESI): calc. for  $\text{C}_{23}\text{H}_{16}\text{ClF}_3\text{N}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 477.08, observed: 477.25 ( $R_t = 4.88$  min). HRMS (ESI): calc. for  $\text{C}_{23}\text{H}_{16}\text{ClF}_3\text{N}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 477.0829, observed: 477.0830.

**Methyl 5-bromo-3-(2-chloro-6-(trifluoromethyl)phenyl)isoxazole-4-carboxylate (15).** According to the General Procedure for Suzuki coupling, bromide **14** (0.900 g, 2.34 mmol) was reacted with *tert*-butyl 3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-1H-pyrrole-1-carboxylate (1.37 g, 4.68 mmol). The crude product was purified by flash column chromatography, eluting with a gradient of 5–10% EtOAc in *n*-heptane, to furnish **15** (0.553 g, 50%) as a colourless oil.  $^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 8.44 (1H, app. t,  $J = 2.0$  Hz, pyrrole H-2), 7.76–7.65 (2H, m, ArH-3 and ArH-5), 7.55 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 7.36 (1H, dd,  $J = 3.4, 2.2$  Hz, pyrrole H-5), 6.99 (1H, dd,  $J = 3.4, 1.6$  Hz, pyrrole H-4), 3.61 (3H, s,  $\text{CO}_2\text{CH}_3$ ), 1.64 (9H, s,  $\text{C}(\text{CH}_3)_3$ );  $^{13}\text{C}$ -NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 168.9 (C-5), 161.5 ( $\text{CO}_2\text{CH}_3$ ), 158.9 (C-3), 148.1 ( $\text{NCO}_2$ ), 136.2 (ArC-2), 132.7 (ArC-3), 131.4 (q,  $J = 31.4$  Hz, ArC-6), 130.5 (ArC-4), 127.7 (ArC-1), 124.6 (q,  $J = 5.0$  Hz, ArC-5), 124.1 (pyrrole C-2), 121.6 (q,  $J = 274.4$  Hz,  $\text{CF}_3$ ), 120.9 (pyrrole C-5), 113.7



(pyrrole C-3), 111.5 (pyrrole C-4), 107.1 (C-4), 85.1 (C(CH<sub>3</sub>)<sub>3</sub>), 51.6 (CO<sub>2</sub>CH<sub>3</sub>), 27.9 (C(CH<sub>3</sub>)<sub>3</sub>). LC-MS (ESI): calc. for C<sub>21</sub>H<sub>18</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 471.09, observed: 470.92 (R<sub>t</sub> = 6.04 min).

**Tert-butyl 3-(3-(2-chloro-6-(trifluoromethyl)phenyl)-4-(hydroxymethyl)isoxazol-5-yl)-1H-pyrrole-1-carboxylate (16).** Ester **15** (0.100 g, 0.212 mmol) was treated according to the General Procedure for Reduction of esters to alcohols. The crude product was purified by flash column chromatography, eluting with a gradient of 10-20% EtOAc in n-heptane, to furnish alcohol **16** (0.075 g, 80%) as a colourless oil. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 7.88 (1H, s, pyrrole H-2), 7.78-7.68 (2H, m, ArH-3 and ArH-5), 7.57 (1H, app. t,  $J$  = 8.0 Hz, ArH-4), 7.37-7.33 (1H, m, pyrrole H-5), 6.79-6.73 (1H, m, pyrrole H-4), 4.42 (2H, br. s, CH<sub>2</sub>OH), 1.63 (9H, s, C(CH<sub>3</sub>)<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 164.6 (C-5), 158.7 (C-3), 148.2 (NCO<sub>2</sub>), 136.5 (ArC-2), 133.1 (ArC-3), 132.2 (q,  $J$  = 31.2 Hz, ArC-6), 131.0 (ArC-4), 126.7 (ArC-1), 124.9 (q,  $J$  = 5.1 Hz, ArC-5), 121.5 (q,  $J$  = 274.5 Hz, CF<sub>3</sub>), 121.4 (pyrrole C-5), 120.0 (pyrrole C-2), 114.5 (pyrrole C-3), 113.1 (C-4), 110.3 (pyrrole C-4), 84.9 (C(CH<sub>3</sub>)<sub>3</sub>), 53.8 (CH<sub>2</sub>OH), 27.9 (C(CH<sub>3</sub>)<sub>3</sub>). LC-MS (ESI): calc. for C<sub>20</sub>H<sub>18</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 443.09, observed: 443.08 (R<sub>t</sub> = 7.23 min).

**Tert-butyl 3-(3-(2-chloro-6-(trifluoromethyl)phenyl)-4-(((4-(methoxycarbonyl) phenyl)amino)methyl)isoxazol-5-yl)-1H-pyrrole-1-carboxylate (17).** According to the General Procedure for Mesylation and Substitution, alcohol **16** (0.072 g, 0.163 mmol) was mesylated and reacted with methyl 4-aminobenzoate (0.122 g, 0.805 mmol) *in situ*. The crude product was purified by column chromatography, eluting with a gradient of 5-20% EtOAc in n-heptane, to furnish **17** (32.2 mg, 36%) as a colourless oil. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  7.83 (2H, d,  $J$  = 8.2 Hz, benzoate H-2), 7.69 (3H, m, ArH-3, ArH-5 and pyrrole H-2), 7.53 (1H, app. t,  $J$  = 8.0 Hz, ArH-4), 7.34 (1H, m, pyrrole H-5), 6.65 (1H, m, pyrrole H-4), 6.51 (2H, d,  $J$  = 8.6 Hz, benzoate H-3), 4.09 (2H, m, CH<sub>2</sub>NH), 4.06 (1H, m, CH<sub>2</sub>NH), 3.84 (3H, s, OCH<sub>3</sub>), 1.52 (9H, s, C(CH<sub>3</sub>)<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 167.1 (COOCH<sub>3</sub>), 164.3 (C-5), 159.1 (C-3), 150.9 (benzoate C-4), 148.0 (NCO<sub>2</sub>), 136.5 (ArC-2), 133.2 (ArC-3), 132.1 (q,  $J$  = 31.2 Hz, ArC-6), 131.5 (benzoate C-2), 131.2 (ArC-4), 126.4 (ArC-1), 125.0 (q,  $J$  = 5.0 Hz, ArC-5), 121.6 (pyrrole C-5), 121.5 (q,  $J$  = 274.5 Hz, CF<sub>3</sub>), 119.7 (pyrrole C-2), 119.2 (benzoate C-1), 114.2 (pyrrole C-3), 110.3 (C-4), 111.5 (benzoate C-3), 109.9 (pyrrole C-4), 85.0 (C(CH<sub>3</sub>)<sub>3</sub>), 51.6 (OCH<sub>3</sub>), 36.5 (CH<sub>2</sub>NH), 27.8 (C(CH<sub>3</sub>)<sub>3</sub>). LC-MS (ESI): calc. for C<sub>28</sub>H<sub>25</sub>ClF<sub>3</sub>N<sub>3</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 576.14, observed: 576.08 (R<sub>t</sub> = 8.23 min).

**Tert-butyl 3-(3-(2-chloro-6-(trifluoromethyl)phenyl)-4-(((4-(methoxycarbonyl) phenyl)(methyl)amino)methyl)isoxazol-5-yl)-1H-pyrrole-1-carboxylate (18).** According to the General Procedure for Mesylation and Substitution, alcohol **16** (0.120 g, 0.271 mmol) was mesylated and reacted with methyl 4-(methylamino)benzoate (0.190 g, 1.15 mmol) *in situ*. The crude product was purified by column chromatography, eluting with a gradient of 5-20% EtOAc in n-heptane, to furnish **18** (70.3 mg, 52%) as a colourless oil. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  7.70 (3H, m, benzoate H-2 and pyrrole H-2), 7.60 (1H, d,  $J$  = 7.8 Hz, ArH-5), 7.50 (1H, d,  $J$  = 8.0 Hz, ArH-3), 7.41 (1H, app. t,  $J$  = 8.0 Hz, ArH-4), 7.38 (1H, m, pyrrole H-4), 6.68 (1H, m, pyrrole H-5), 6.39 (2H, d,  $J$  = 8.8 Hz, benzoate H-3), 4.42 (2H, m, CH<sub>2</sub>NCH<sub>3</sub>), 3.84 (3H, s, OCH<sub>3</sub>), 2.66 (3H, s, CH<sub>2</sub>NCH<sub>3</sub>), 1.52 (9H, s, C(CH<sub>3</sub>)<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 167.2 (COOCH<sub>3</sub>), 163.3 (C-5), 159.0 (C-3), 152.2 (benzoate C-4), 148.1 (NCO<sub>2</sub>), 136.7 (ArC-2), 133.0 (ArC-3), 131.7 (q,  $J$  = 31.2 Hz, ArC-6), 131.0 (benzoate C-2), 130.8 (ArC-4), 126.9 (ArC-1), 124.9 (q,  $J$  = 5.0 Hz, ArC-5), 121.6 (q,  $J$  = 274.5 Hz, CF<sub>3</sub>), 121.5 (pyrrole C-5), 119.5 (pyrrole C-2), 117.8 (benzoate C-1), 114.3 (pyrrole C-3), 111.5 (C-4), 110.9 (benzoate C-3), 110.2 (pyrrole C-4), 85.1 (C(CH<sub>3</sub>)<sub>3</sub>), 51.5 (OCH<sub>3</sub>), 44.2 (CH<sub>2</sub>NCH<sub>3</sub>), 36.8 (CH<sub>2</sub>NCH<sub>3</sub>), 27.9 (C(CH<sub>3</sub>)<sub>3</sub>). LC-MS (ESI): calc. for C<sub>29</sub>H<sub>27</sub>ClF<sub>3</sub>N<sub>3</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 590.16, observed: 590.00 (R<sub>t</sub> = 8.64 min).

**Methyl 4-(((3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-3-yl)isoxazol-4-yl) methyl)amino)-2-fluorobenzoate (19).** Alcohol compound **16** (0.256 g, 0.580 mmol) was dissolved in anhydrous CH<sub>2</sub>Cl<sub>2</sub> (8 mL). To this was added Dess-Martin Periodinane (0.369 mg, 0.870 mmol) and the reaction mixture

was stirred at room temperature for 3 h. The reaction mixture was quenched by the addition of 10% aqueous  $\text{Na}_2\text{S}_2\text{O}_3$  solution and extracted with  $\text{CH}_2\text{Cl}_2$  (3 x). The combined organic phase was washed with saturated aqueous  $\text{NaHCO}_3$  and  $\text{H}_2\text{O}$ , dried over  $\text{MgSO}_4$ , filtered and concentrated *in vacuo*. The crude product was purified by column chromatography, eluting with a gradient of 20–30% EtOAc in n-heptane, to furnish the aldehyde (112.0 mg, 44%). Methyl 4-amino-2-fluorobenzoate (0.052 g, 0.310 mmol) was added to a solution of the aldehyde (0.112 g, 0.250 mmol) and AcOH (1.43  $\mu\text{L}$ , 0.025 mmol) in MeOH (2 mL). The reaction mixture was heated at reflux for 24 h and then concentrated *in vacuo*. The intermediate imine was isolated by flash column chromatography, eluting with a gradient of 10–30% EtOAc in n-heptane (62.3 mg, 42%), and then dissolved in EtOH (1 mL), cooled to 0 °C (ice) and treated with  $\text{NaBH}_4$  (0.020 mg, 0.526 mmol). The reaction mixture was stirred at 86 °C for 4 h. The solvent was removed *in vacuo* and the mixture was dissolved in  $\text{CH}_2\text{Cl}_2$  and washed with water. The aqueous phase was further extracted with  $\text{CH}_2\text{Cl}_2$  (3 x), dried ( $\text{MgSO}_4$ ), filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography, eluting with a gradient of 30–40% EtOAc in n-heptane, to furnish compound **19** (10.5 mg, 17%).  $^1\text{H-NMR}$  (400 MHz,  $\text{DMSO-d}_6$ ):  $\delta$  11.4 (1H, s, pyrrole-NH), 7.89 (1H, d,  $J$  = 8.0 Hz, ArH-5), 7.84 (1H, d,  $J$  = 7.5 Hz, ArH-3), 7.72 (1H, app. t,  $J$  = 8.0 Hz, ArH-4), 7.47 (1H, app. t,  $J$  = 8.7 Hz, benzoate H-2), 7.40 (1H, m, pyrrole H-2), 6.99 (1H, m, pyrrole H-5), 6.76 (1H, app. t,  $J$  = 4.6 Hz,  $\text{CH}_2\text{NH}$ ), 6.55 (1H, m, pyrrole H-4), 6.25 (1H, dd,  $J$  = 8.9, 1.9 Hz, benzoate H-3a), 6.10 (1H, d,  $J$  = 14.6 Hz, benzoate H-3b), 4.17 (2H, m,  $\text{CH}_2\text{NH}$ ). LC-MS (ESI): calc. for  $\text{C}_{23}\text{H}_{16}\text{ClF}_4\text{N}_3\text{O}_3$  [ $\text{M}+\text{H}$ ] $^{+}$ : 494.08, observed: 494.25 ( $R_t$  = 7.14 min).

**tert-butyl 3-(3-(2-chloro-6-(trifluoromethyl)phenyl)-4-((4-(methoxycarbonyl) phenoxy)methyl)isoxazol-5-yl)-1H-pyrrole-1-carboxylate (20).** According to the General Procedure for Mitsunobu Coupling, alcohol **16** (0.023 g, 0.053 mmol) was reacted with methyl 4-hydroxybenzoate (0.010 g, 0.058 mmol). The crude product was purified by column chromatography, eluting with a gradient of 5–25% EtOAc in n-heptane, to furnish **20** (16.6 mg, 55%) as a white solid.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.92 (2H, d,  $J$  = 8.9 Hz, benzoate H-2), 7.71 (3H, m, ArH-3, ArH-5 and pyrrole H-2), 7.56 (1 H, app. t,  $J$  = 8.0 Hz, ArH-4), 7.35 (1H, m, pyrrole H-5), 6.82 (2H, d,  $J$  = 8.9 Hz, benzoate H-3), 6.66 (1H, m, pyrrole H-4), 4.83 (2H, s,  $\text{CH}_2\text{O}$ ), 3.87 (3H, s,  $\text{OCH}_3$ ), 1.55 (9H, s, ( $\text{C}(\text{CH}_3)_3$ );  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 166.7 ( $\text{COOCH}_3$ ), 165.1 (C-5), 161.8 (benzoate C-4), 158.9 (C-3), 148.1 ( $\text{NCO}_2$ ), 136.7 (ArC-2), 133.1 (ArC-3), 132.3 (q,  $J$  = 31.2 Hz, ArC-6), 131.6 (benzoate C-2), 131.1 (ArC-4), 126.4 (ArC-1), 124.9 (q,  $J$  = 5.0 Hz, ArC-5), 123.2 (benzoate C-1), 121.6 (pyrrole C-5), 121.4 (q,  $J$  = 274.5 Hz,  $\text{CF}_3$ ), 120.0 (pyrrole C-2), 114.2 (pyrrole C-3), 114.1 (benzoate C-3), 110.2 (pyrrole C-4), 109.3 (C-4), 85.0 ( $\text{C}(\text{CH}_3)_3$ ), 59.4 ( $\text{CH}_2\text{O}$ ), 51.9 ( $\text{OCH}_3$ ), 27.8 ( $\text{C}(\text{CH}_3)_3$ ). LC-MS (ESI): calc. for  $\text{C}_{28}\text{H}_{24}\text{ClF}_3\text{N}_2\text{O}_6$  [ $\text{M}+\text{H}$ ] $^{+}$ : 577.13, observed: 577.17 ( $R_t$  = 8.57 min).

**tert-butyl 3-(3-(2-chloro-6-(trifluoromethyl)phenyl)-4-(3-fluoro-4-(methoxy carbonyl)phenoxy)methyl)isoxazol-5-yl)-1H-pyrrole-1-carboxylate (21).** According to the General Procedure for Mitsunobu Coupling, alcohol **16** (0.050 g, 0.113 mmol) was reacted with methyl 2-fluoro-4-hydroxybenzoate (0.021 g, 0.124 mmol). The crude product was purified by column chromatography, eluting with a gradient of 5–25% EtOAc in n-heptane, to furnish **21** (17.0 mg, 25%) as a white solid.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.85 (1H, app. t,  $J$  = 8.6 Hz, benzoate H-2), 7.73 (3H, m, ArH-3, ArH-5 and pyrrole H-2), 7.57 (1 H, app. t,  $J$  = 8.0 Hz, ArH-4), 7.36 (1H, m, pyrrole H-5), 6.65 (2H, m, pyrrole H-4 and benzoate H-3a), 6.52 (1H, dd,  $J$  = 12.4, 2.3 Hz, benzoate H-3b), 4.82 (2H, s,  $\text{CH}_2\text{O}$ ), 3.89 (3H, s,  $\text{OCH}_3$ ), 1.57 (9H, s, ( $\text{C}(\text{CH}_3)_3$ );  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 165.2 (C-5), 164.5 ( $\text{COOCH}_3$ ), 162.9 (benzoate C-F), 161.9 (benzoate C-4), 158.9 (C-3), 148.0 ( $\text{NCO}_2$ ), 136.7 (ArC-2), 133.6 (benzoate C-2), 133.2 (ArC-3), 132.3 (q,  $J$  = 31.2 Hz, ArC-6), 131.2 (ArC-4), 126.2 (ArC-1), 125.0 (q,  $J$  = 5.0 Hz, ArC-5), 121.7 (pyrrole C-5), 121.4 (q,  $J$  = 274.5 Hz,  $\text{CF}_3$ ), 120.0 (pyrrole C-2), 114.1 (pyrrole C-3), 111.6 (benzoate C-3a), 110.4 (benzoate C-1), 110.2 (pyrrole C-4), 108.8 (C-4), 103.1 (benzoate C-3b), 85.1 ( $\text{C}(\text{CH}_3)_3$ ), 59.9 ( $\text{CH}_2\text{O}$ ), 52.0 ( $\text{OCH}_3$ ), 27.8 ( $\text{C}(\text{CH}_3)_3$ ). LC-MS (ESI): calc. for  $\text{C}_{28}\text{H}_{23}\text{ClF}_4\text{N}_2\text{O}_6$  [ $\text{M}+\text{H}$ ] $^{+}$ : 595.12, observed: 595.00 ( $R_t$  = 6.14 min).

**Tert-butyl 3-(3-(2-chloro-6-(trifluoromethyl)phenyl)-4-(((4-(methoxycarbonyl) phenyl)thio)methyl)isoxazol-5-yl)-1H-pyrrole-1-carboxylate (22).** According to the General Procedure for Mitsunobu Coupling, alcohol **16** (0.055 g, 0.124 mmol) was reacted with methyl 4-mercaptobenzoate (0.024 g, 0.137 mmol). The crude product was purified by column chromatography, eluting with a gradient of 0–20% EtOAc in n-heptane, to furnish **21** (20.0 mg, 27%) as a white solid.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.84 (2H, d,  $J$  = 8.5 Hz, benzoate H-2), 7.72 (3H, m, ArH-3, ArH-5 and pyrrole H-2), 7.57 (1 H, app. t,  $J$  = 8.5 Hz, ArH-4), 7.34 (1H, m, pyrrole H-5), 7.18 (2H, d,  $J$  = 8.9 Hz, benzoate H-3), 6.68 (1H, m, pyrrole H-4), 4.01 (2H, s,  $\text{CH}_2\text{O}$ ), 3.89 (3H, s,  $\text{OCH}_3$ ), 1.58 (9H, s,  $(\text{C}(\text{CH}_3)_3)$ ). LC-MS (ESI): calc. for  $\text{C}_{28}\text{H}_{24}\text{ClF}_3\text{N}_2\text{O}_5\text{S}$   $[\text{M}+\text{H}]^+$ : 593.10, observed: 592.92 ( $R_t$  = 6.28 min).

**Tert-butyl 3-(3-(2-chloro-6-(trifluoromethyl)phenyl)-4-formylisoxazol-5-yl)-1H-pyrrole-1-carboxylate (23).** Alcohol compound **16** (0.150 g, 0.407 mmol) was dissolved in anhydrous  $\text{CH}_2\text{Cl}_2$  (5 mL). To this was added Dess-Martin Periodinane (0.216 mg, 0.610 mmol) and the reaction mixture was stirred at room temperature for 3 h. The reaction mixture was quenched by the addition of 10% aqueous  $\text{Na}_2\text{S}_2\text{O}_3$  solution and extracted with  $\text{CH}_2\text{Cl}_2$  (3 x). The combined organic phase was washed with saturated aqueous  $\text{NaHCO}_3$  and  $\text{H}_2\text{O}$ , dried over  $\text{MgSO}_4$ , filtered and concentrated *in vacuo*. The crude product was purified by column chromatography, eluting with a gradient of 15–30% EtOAc in n-heptane, to furnish **21** (150 mg, 100%) as a colourless oil.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 9.67 (1H, s, CHO), 8.54 (1H, s, pyrrole H-2), 7.81–7.73 (2H, m, ArH-3 and ArH-5), 7.63 (1H, app. t,  $J$  = 8.0 Hz, ArH-4), 7.43–7.38 (1H, m, pyrrole H-5), 6.97–6.92 (1H, m, pyrrole H-4), 1.66 (9H, s,  $\text{C}(\text{CH}_3)_3$ ). LC-MS (ESI): calc. for  $\text{C}_{20}\text{H}_{16}\text{ClF}_3\text{N}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 441.08, observed: 440.92 ( $R_t$  = 5.87 min).

**Tert-butyl (E) or (Z)-3-(3-(2-chloro-6-(trifluoromethyl)phenyl)-4-(4-(methoxy carbonyl)styryl)isoxazol-5-yl)-1H-pyrrole-1-carboxylate (24 & 25).** Under an inert atmosphere, triphenylphosphine (0.890 g, 3.39 mmol) and methyl 4-(bromomethyl)benzoate (0.519 g, 2.27 mmol) were dissolved in acetonitrile (20 mL). The reaction mixture was stirred at 82 °C for 3 h, when TLC indicated full conversion of the starting material. The reaction mixture was cooled to room temperature and precipitated in toluene (150 mL). The solid was filtered, collected and dried under reduced pressure to yield (4-(methoxycarbonyl)benzyl) triphenylphosphonium as a crystalline white solid (930 mg, 100%). The resulting product (0.152 g, 0.369 mmol) was dissolved in anhydrous THF (2 mL) under an inert atmosphere. The solution was cooled to –78 °C, LiHMDS (0.4 mL, 1M in hexane) was added and the solution was stirred at –78 °C for 2 h. Aldehyde **23** (0.125 g, 0.284 mmol), dissolved in anhydrous THF (2 mL), was added dropwise. The reaction mixture was slowly warmed to room temperature and stirred for 24 h. The reaction was quenched with  $\text{H}_2\text{O}$  (10 mL) and extracted with EtOAc (3 x). The combined organic phase was washed with  $\text{H}_2\text{O}$  and brine, dried over  $\text{MgSO}_4$ , filtrated and concentrated *in vacuo*. The crude product was purified by column chromatography, eluting with a gradient of 5–20% EtOAc in n-heptane, to furnish a mixture of *cis/trans* isomers (3:1) (70 mg). The mixture of *cis/trans* isomers was separated via preparative HPLC (gradient of 80–100% acetonitrile in  $\text{H}_2\text{O}$ ) to furnish **24** (*trans* isomer, 12.0 mg, 7%) and **25** (*cis* isomer, 44.0 mg, 27%), both as a white solid.

**Trans isomer (24):**  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.92 (2H, d,  $J$  = 7.9 Hz, benzoate H-2), 7.80 (3H, m, ArH-3, ArH-5 and pyrrole H-2), 7.63 (1 H, app. t,  $J$  = 7.8 Hz, ArH-4), 7.39 (1H, m, pyrrole H-5), 7.24 (2H, d,  $J$  = 7.9 Hz, benzoate H-3), 7.03 (1H, d,  $J$  = 16.5 Hz,  $\text{C}_4\text{-HC=CH}$ ), 6.73 (1H, m, pyrrole H-4), 6.17 (1H, d,  $J$  = 16.4 Hz, benzoate  $\text{C}_4\text{-HC=CH}$ ), 3.90 (3H, s,  $\text{OCH}_3$ ), 1.64 (9H, s,  $(\text{C}(\text{CH}_3)_3)$ );  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 166.7 ( $\text{COOCH}_3$ ), 162.7 (C-5), 157.1 (C-3), 148.1 ( $\text{NCO}_2$ ), 141.3 (benzoate C-4), 136.7 (ArC-2), 133.2 (ArC-3), 132.2 (q,  $J$  = 31.5 Hz, ArC-6), 131.1 (ArC-4), 129.9 (benzoate C-2), 129.8 (benzoate  $\text{C}_4\text{-HC=CH}$ ), 129.2 (benzoate C-1), 127.6 (ArC-1), 126.0 (benzoate C-3), 125.1 (q,  $J$  = 5.0 Hz, ArC-5), 121.5 (pyrrole C-5), 121.4 (q,  $J$  = 274.5 Hz,  $\text{CF}_3$ ), 119.7 (pyrrole C-2), 118.0 ( $\text{C}_4\text{-HC=CH}$ ), 114.6 (pyrrole C-3), 112.6 (C-4), 110.4 (pyrrole C-4), 85.1 ( $\text{C}(\text{CH}_3)_3$ ), 52.1 ( $\text{OCH}_3$ ), 27.9 ( $\text{C}(\text{CH}_3)_3$ ). LC-MS (ESI): calc. for  $\text{C}_{29}\text{H}_{24}\text{ClF}_3\text{IN}_2\text{O}_5$   $[\text{M}+\text{H}]^+$ : 573.13, observed: 573.00 ( $R_t$  = 6.44 min).

**Cis isomer (25):** <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 7.77 (2H, d, *J* = 8.0 Hz, benzoate H-2), 7.68 (2H, m, ArH-3 and ArH-5), 7.53 (1 H, app. t, *J* = 8.0 Hz, ArH-4), 7.38 (1H, br. s, pyrrole H-2), 7.28 (2H, d, *J* = 8.0 Hz, benzoate H-3), 7.08 (1H, br. s, pyrrole H-5), 6.62 (1H, d, *J* = 12.1 Hz, C<sub>4</sub>-HC=CH), 6.50 (1H, br. s, pyrrole H-4), 6.18 (1H, d, *J* = 12.1 Hz, benzoate C<sub>4</sub>-HC=CH), 3.87 (3H, s, OCH<sub>3</sub>), 1.57 (9H, s, (C(CH<sub>3</sub>)<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) 166.7 (COOCH<sub>3</sub>), 161.3 (C-5), 159.0 (C-3), 148.1 (NCO<sub>2</sub>), 140.8 (benzoate C-4), 136.6 (ArC-2), 133.9 (C<sub>4</sub>-HC=CH), 133.1 (ArC-3), 132.0 (q, *J* = 31.5 Hz, ArC-6), 130.8 (ArC-4), 129.5 (benzoate C-2), 129.1 (benzoate C-1), 128.2 (benzoate C-3), 126.8 (ArC-1), 124.9 (q, *J* = 5.0 Hz, ArC-5), 121.5 (q, *J* = 274.5 Hz, CF<sub>3</sub>), 120.6 (pyrrole C-5), 119.7 (pyrrole C-2), 118.2 (benzoate C<sub>4</sub>-HC=CH), 114.6 (pyrrole C-3), 110.8 (C-4), 110.0 (pyrrole C-4), 84.7 (C(CH<sub>3</sub>)<sub>3</sub>), 52.1 (OCH<sub>3</sub>), 27.9 (C(CH<sub>3</sub>)<sub>3</sub>). LC-MS (ESI): calc. for C<sub>29</sub>H<sub>24</sub>ClF<sub>3</sub>IN<sub>2</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 573.13, observed: 573.00 (*R*<sub>t</sub> = 6.26 min).

**Methyl 3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-2-yl)isoxazole-4-carboxylate (26).** According to the General Procedure for Suzuki coupling, bromide **14** (0.500 g, 1.30 mmol) was coupled to *tert*-butyl 2-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-1H-pyrrole-1-carboxylate (0.762 g, 2.60 mmol). The crude product was purified by flash column chromatography, eluting with a gradient of 5–10% EtOAc in *n*-heptane, to furnish **26** (0.298 g, 49%) as a colourless oil. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) 7.77–7.67 (2H, m, ArH-3 and ArH-5), 7.60–7.55 (2H, m, ArH-4 and pyrrole H-5), 6.88 (1H, dd, *J* = 3.6, 1.7 Hz, pyrrole H-3), 6.36 (1H, app. t, *J* = 3.4 Hz, pyrrole H-4), 3.57 (3H, s, CO<sub>2</sub>CH<sub>3</sub>), 1.52 (9H, s, C(CH<sub>3</sub>)<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) 166.7 (C-5), 161.1 (CO<sub>2</sub>CH<sub>3</sub>), 158.4 (C-3), 148.3 (NCO<sub>2</sub>), 136.3 (ArC-2), 132.8 (ArC-3), 131.5 (q, *J* = 31.4 Hz, ArC-6), 130.7 (ArC-4), 127.1 (ArC-1), 126.2 (pyrrole C-5), 124.6 (q, *J* = 5.0 Hz, ArC-5), 121.6 (q, *J* = 274.4 Hz, CF<sub>3</sub>), 121.2 (pyrrole C-3), 118.5 (pyrrole C-2), 111.1 (pyrrole C-4), 110.2 (C-4), 85.3 (C(CH<sub>3</sub>)<sub>3</sub>), 51.7 (CO<sub>2</sub>CH<sub>3</sub>), 27.7 (C(CH<sub>3</sub>)<sub>3</sub>). LC-MS (ESI): calc. for C<sub>21</sub>H<sub>18</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 471.09, observed: 470.92 (*R*<sub>t</sub> = 5.73 min).

**Methyl 3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrazol-4-yl)isoxazole-4-carboxylate (27).** According to the General Procedure for Suzuki coupling, bromide **14** (0.500 g, 1.30 mmol) was reacted with *tert*-butyl 1-(3,4-dihydro-2H-pyran-6-yl)-4-(4,4,5-trimethyl-1,3,2-dioxaborolan-2-yl)-1H-pyrazole (0.723 g, 2.60 mmol). The crude product was purified by flash column chromatography, eluting with a gradient of 5–10% EtOAc in *n*-heptane, to furnish **27** (0.203 g, 34%) as a colourless oil. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) 8.78 (1H, s, pyrazole H-5), 8.34 (1H, s, pyrazole H-3), 7.75–7.67 (2H, app. t, *J* = 9.2 Hz, ArH-3 and ArH-5), 7.56 (1H, app. t, *J* = 8.0 Hz, ArH-4), 5.51–5.46 (1H, m, NCHCO), 4.14–4.04 (1H, m, THP-H), 3.78–3.72 (1H, m, THP-H), 3.60 (3H, s, CO<sub>2</sub>CH<sub>3</sub>), 2.20–2.12 (2H, m, THP-H), 1.75–1.62 (4H, m, THP-H); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) 167.8 (C-5), 161.6 (CO<sub>2</sub>CH<sub>3</sub>), 158.6 (C-3), 139.7 (pyrazole C-3), 136.2 (ArC-2), 132.7 (ArC-3), 131.4 (q, *J* = 31.3 Hz, ArC-6), 130.6 (ArC-4), 130.6 (pyrazole C-5), 127.5 (ArC-1), 124.6 (q, *J* = 4.9 Hz, ArC-5), 121.6 (q, *J* = 274.4 Hz, CF<sub>3</sub>), 109.6 (pyrazole C-4), 106.8 (C-4), 88.0 (NCHCO), 67.8 (THP-C), 51.7 (CO<sub>2</sub>CH<sub>3</sub>), 30.6 (THP-C), 24.8 (THP-C), 20.1 (THP-C). LC-MS (ESI): calc. for C<sub>20</sub>H<sub>17</sub>ClF<sub>3</sub>N<sub>3</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 456.09, observed: 456.08 (*R*<sub>t</sub> = 5.37 min).

**Methyl 3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1-methyl-1H-pyrrol-3-yl)isoxazole-4-carboxylate (28).** Under an inert atmosphere 1-methyl-3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-1H-pyrrole (0.404 g, 1.95 mmol), Na<sub>2</sub>CO<sub>3</sub> (0.276 g, 2.6 mmol) and Pd(PPh<sub>3</sub>)<sub>4</sub> (0.150 g, 0.13 mmol) were added to a solution of bromide **14** (0.500 g, 1.3 mmol) in de-gassed DME/H<sub>2</sub>O (4:1) (15 mL). The reaction mixture was heated at 85 °C for 8 h, cooled to room temperature, diluted with H<sub>2</sub>O and extracted with EtOAc (3 x). The combined organic phase was dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo*. The crude product was purified by column chromatography, eluting with 5–15% EtOH in *n*-Heptane, to furnish **28** (0.275 g, 55%) as a yellow solid. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) 7.97 (1H, s, pyrrole H-2), 7.75–7.64 (2H, m, ArH-3 and ArH-5), 7.52 (1H, app. t, *J* = 8.0 Hz, ArH-4), 6.92 (1H, br. s, pyrrole H-5), 6.69 (1H, br. s, pyrrole H-4), 3.74 (3H, s, NCH<sub>3</sub>), 3.56 (3H, s, CO<sub>2</sub>CH<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) 170.4 (C-5), 162.1 (CO<sub>2</sub>CH<sub>3</sub>), 158.6 (C-3), 136.2 (ArC-2), 132.6 (ArC-3), 131.3 (q, *J* = 31.3 Hz, ArC-6), 130.3 (ArC-

4), 128.2 (ArC-1), 126.8 (pyrrole C-2), 124.5 (q,  $J = 5.0$  Hz, ArC-5), 123.1 (pyrrole C-5), 121.6 (q,  $J = 274.4$  Hz, CF<sub>3</sub>), 110.6 (pyrrole C-3), 109.4 (pyrrole C-4), 104.9 (C-4), 51.4 (CO<sub>2</sub>CH<sub>3</sub>), 36.7 (NCH<sub>3</sub>). LC-MS (ESI): calc. for C<sub>17</sub>H<sub>12</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>3</sub>[M+H]<sup>+</sup>: 385.05, observed: 385.17 ( $R_t = 5.27$  min).

**Methyl 3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(5-methyl-1H-pyrrol-3-yl)isoxazole-4-carboxylate (29).** Under an inert atmosphere, a Schlenk tube was charged with *tert*-butyl 2,4-dimethyl-1H-pyrrole-1-carboxylate (1.10 g, 6.05 mmol). Two separate flasks under Argon were charged with [Ir(OMe)(COD)]<sub>2</sub> (0.060 g, 0.091 mmol) and dtbpy (0.049 g, 0.182 mmol). HBPIn (1.16 g, 9.08 mmol) was added to the [Ir(OMe)(COD)]<sub>2</sub>-containing flask. Anhydrous THF (18 mL) was added to the dtbpy-containing flask and the solution was mixed with the [Ir(OMe)(COD)]<sub>2</sub> and HBPIn mixture. The resulting solution was transferred to the Schlenk flask and the reaction mixture was stirred at 60 °C for 6 h. Afterwards, the reaction mixture was cooled to room temperature and THF was removed *in vacuo*. The crude product was purified by flash column chromatography, eluting with a gradient of 0-5% EtOAc in *n*-heptane, to furnish *tert*-butyl 2-methyl-4-(4,4,5-trimethyl-1,3,2-dioxaborolan-2-yl)-1H-pyrrole-1-carboxylate (1.30 g, 70%).<sup>48</sup> Subsequently, according to the General Procedure for Suzuki coupling, bromide **14** (1.09 g, 2.82 mmol) was reacted with the synthesized pinacol boronate (1.30 g, 4.23 mmol). The crude product was purified by flash column chromatography, eluting with a gradient of 30-60% CH<sub>2</sub>Cl<sub>2</sub> in *n*-heptane, to furnish **29** (0.325 g, 24%) as a colourless oil. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 8.39 (1H, s, pyrrole H-2), 7.74-7.65 (2H, m, ArH-3 and ArH-5), 7.54 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 6.70 (1H, s, pyrrole H-4), 3.61 (3H, s, CO<sub>2</sub>CH<sub>3</sub>), 2.51 (3H, s, NCCH<sub>3</sub>), 1.64 (9H, s, C(CH<sub>3</sub>)<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 169.0 (C-5), 161.6 (CO<sub>2</sub>CH<sub>3</sub>), 158.9 (C-3), 148.9 (NCO<sub>2</sub>), 136.2 (ArC-2), 132.8 (pyrrole C-5), 132.6 (ArC-3), 131.3 (q,  $J = 31.4$  Hz, ArC-6), 130.5 (ArC-4), 127.7 (ArC-1), 124.9 (pyrrole C-2), 124.5 (q,  $J = 4.9$  Hz, ArC-5), 121.6 (q,  $J = 274.4$  Hz, CF<sub>3</sub>), 111.6 (pyrrole C-3), 110.9 (pyrrole C-4), 106.9 (C-4), 84.7 (C(CH<sub>3</sub>)<sub>3</sub>), 51.6 (CO<sub>2</sub>CH<sub>3</sub>), 28.0 (C(CH<sub>3</sub>)<sub>3</sub>), 15.3 (NCCH<sub>3</sub>). LC-MS (ESI): calc. for C<sub>22</sub>H<sub>20</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>5</sub>[M+H]<sup>+</sup>: 485.10, observed: 485.00 ( $R_t = 6.22$  min).

**(3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrrol-2-yl)isoxazol-4-yl)methanol (30).** Ester **26** (0.260 g, 0.552 mmol) was treated according to the General Procedure for Reduction of esters to alcohols. The crude product was purified by flash column chromatography, eluting with a gradient of 10-20% EtOAc in *n*-heptane, to furnish alcohol **30** (169 mg, 69%) as a colourless oil. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 7.77-7.69 (2H, m, ArH-3 and ArH-5), 7.57 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 7.51 (1H, dd,  $J = 3.2, 1.7$  Hz, pyrrole H-5), 6.70 (1H, dd,  $J = 3.4, 1.7$  Hz, pyrrole H-3), 6.34 (1H, app. t,  $J = 3.4$  Hz, pyrrole H-4), 4.34 (2H, br. s, CH<sub>2</sub>OH), 1.53 (9H, s, C(CH<sub>3</sub>)<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 161.9 (C-5), 158.6 (C-3), 148.7 (NCO<sub>2</sub>), 136.5 (ArC-2), 133.2 (ArC-3), 132.1 (q,  $J = 31.2$  Hz, ArC-6), 131.0 (ArC-4), 126.8 (ArC-1), 125.5 (pyrrole C-5), 125.0 (q,  $J = 5.1$  Hz, ArC-5), 121.5 (q,  $J = 274.5$  Hz, CF<sub>3</sub>), 119.9 (pyrrole C-3), 118.7 (pyrrole C-2), 116.9 (C-4), 111.3 (pyrrole C-4), 85.5 (C(CH<sub>3</sub>)<sub>3</sub>), 54.1 (CH<sub>2</sub>OH), 27.8 (C(CH<sub>3</sub>)<sub>3</sub>). LC-MS (ESI): calc. for C<sub>20</sub>H<sub>18</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub>[M+H]<sup>+</sup>: 443.09, observed: 443.00 ( $R_t = 5.27$  min).

**(3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1H-pyrazol-4-yl)isoxazol-4-yl)methanol (31).** Ester **27** (0.200 g, 0.434 mmol) was treated according to the General Procedure for Reduction of esters to alcohols. The crude product was purified by flash column chromatography, eluting with a gradient of 30-50% EtOAc in *n*-heptane, to furnish alcohol **31** (0.100 g, 71%) as a colourless oil. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 8.25 (1H, s, pyrazole H-5), 8.09 (1H, s, pyrazole H-3), 7.80-7.69 (2H, m, ArH-3 and ArH-5), 7.59 (1H, app. t,  $J = 8.0$  Hz, ArH-4), 5.47 (1H, t,  $J = 6.1$  Hz, NCHCO), 4.41 (1H, br. s, CH<sub>2</sub>OH), 4.14-3.98 (1H, m, THP-H), 3.76-3.71 (1H, m, THP-H), 2.17-1.98 (3H, m, THP-H), 1.78-1.58 (3H, m, THP-H); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 163.3 (C-5), 158.6 (C-3), 137.9 (pyrazole C-3), 136.5 (ArC-2), 133.2 (ArC-3), 132.2 (q,  $J = 31.2$  Hz, ArC-6), 131.1 (ArC-4), 127.5 (pyrazole C-5), 126.6 (ArC-1), 125.0 (q,  $J = 5.1$  Hz, ArC-5), 121.5 (q,  $J = 274.4$  Hz, CF<sub>3</sub>), 112.8 (C-4), 110.2 (pyrazole C-4), 87.9 (NCHCO), 67.8 (THP-C),

53.9 (CH<sub>2</sub>OH), 30.7 (THP-C), 24.9 (THP-C), 22.1 (THP-C). LC-MS (ESI): calc. for C<sub>19</sub>H<sub>17</sub>ClF<sub>3</sub>N<sub>3</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 428.09, observed: 428.08 (R<sub>t</sub> = 4.64 min).

**(3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(1-methyl-1H-pyrrol-3-yl)isoxazol-4-yl) methanol (32).** Under an inert atmosphere, LiAlH<sub>4</sub> (1 M in THF, 1.07 mL, 1.01 mmol) was added dropwise to a solution of ester **28** (0.275 g, 0.710 mmol) in anhydrous THF (6 mL) at 0 °C. The reaction mixture was warmed to room temperature and stirred for 2 h. The reaction mixture was quenched by the addition of saturated aqueous NH<sub>4</sub>Cl solution and stirred vigorously for 60 min. Subsequently, the mixture was extracted with EtOAc (3 x). The combined organic phase was washed with brine, dried over MgSO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography, eluting with a gradient of 10-20% EtOAc in n-heptane, to furnish alcohol **32** (0.192 g, 76%) as a colourless oil. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) 7.76-7.68 (2H, m, ArH-3 and ArH-5), 7.56 (1H, app. t, J = 8.3 Hz, ArH-4), 7.30 (1H, app. t, J = 1.9 Hz, pyrrole H-2), 6.70 (1H, app. t, J = 2.5 Hz, pyrrole H-5), 6.68 (1H, app. t, J = 2.5 Hz, pyrrole H-4), 4.38 (2H, br. s, CH<sub>2</sub>OH), 3.73 (3H, s, NCH<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) 166.3 (C-5), 158.6 (C-3), 136.6 (ArC-2), 133.1 (ArC-3), 132.2 (q, J = 31.3 Hz, ArC-6), 130.9 (ArC-4), 127.1 (ArC-1), 124.9 (q, J = 5.0 Hz, ArC-5), 123.4 (pyrrole C-5), 122.5 (pyrrole C-2), 121.5 (q, J = 274.4 Hz, CF<sub>3</sub>), 111.3 (pyrrole C-3), 110.9 (C-4), 107.5 (pyrrole C-4), 54.2 (CH<sub>2</sub>OH), 36.6 (NCH<sub>3</sub>). LC-MS (ESI): calc. for C<sub>16</sub>H<sub>12</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>2</sub> [M+H]<sup>+</sup>: 357.05, observed: 357.17 (R<sub>t</sub> = 4.63 min).

**(3-(2-chloro-6-(trifluoromethyl)phenyl)-5-(5-methyl-1H-pyrrol-3-yl)isoxazol-4-yl) methanol (33).** Ester **29** (0.400 g, 0.825 mmol) was treated according to the General Procedure for Reduction of esters to alcohols. The crude product was purified by flash column chromatography, eluting with a gradient of 10-25% EtOAc in n-heptane, to furnish alcohol **33** (0.220 g, 58%) as a colourless oil. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) 7.83 (1H, s, pyrrole H-2), 7.77-7.58 (2H, m, ArH-3 and ArH-5), 7.58 (1H, app. t, J = 8.0 Hz, ArH-4), 6.48 (1H, s, pyrrole H-4), 4.42 (2H, br. s, CH<sub>2</sub>OH), 2.50 (3H, s, NCCH<sub>3</sub>), 1.63 (9H, s, C(CH<sub>3</sub>)<sub>3</sub>); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) 164.8 (C-5), 158.7 (C-3), 149.1 (NCO<sub>2</sub>), 136.5 (ArC-2), 133.4 (pyrrole C-5), 133.1 (ArC-3), 132.2 (q, J = 31.4 Hz, ArC-6), 131.0 (ArC-4), 126.8 (ArC-1), 124.9 (q, J = 4.9 Hz, ArC-5), 121.5 (q, J = 274.4 Hz, CF<sub>3</sub>), 120.7 (pyrrole C-2), 112.9 (C-4), 112.3 (pyrrole C-3), 109.9 (pyrrole C-4), 84.6 (C(CH<sub>3</sub>)<sub>3</sub>), 53.9 (CH<sub>2</sub>OH), 28.0 (C(CH<sub>3</sub>)<sub>3</sub>), 15.4 (NCCH<sub>3</sub>). LC-MS (ESI): calc. for C<sub>21</sub>H<sub>20</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 457.11, observed: 456.92 (R<sub>t</sub> = 5.53 min).

## Biophysical assays

**RORyt LBD expression and purification (used for TR-FRET assays).** His<sub>6</sub>-RORyt-LBD was expressed and purified as described in Chapter 2.

**TR-FRET coactivator recruitment assays, competition assays and AlexaFluor-MRL-871 recruitment assays on RORyt.** TR-FRET assays and data analysis were performed as described in Chapter 2. Data were recorded in triplicate; error shown is standard deviation from the mean; curves are representative of > 3 independent experiments.

**Thermal shift assays on RORyt.** Thermal shift assays were performed using 40 μL samples containing 5 μM RORyt LBD, 10 μM compound and 2.5 x SYPRO® Orange (Sigma) in buffer containing 150 mM NaCl, 10 mM HEPES (pH 7.5) and 1% DMSO. Hard-Shell® 96-Well PCR Plates (low profile, thin wall, skirted, green/white #hsp9645) were used. The samples were heated from 25 °C to 80 °C at a rate of 0.5 °C per 5 s in a CFX96 Touch Real-Time PCR Detection System (Bio-Rad). After each increment, Sypro Orange intensity was measured using the plate read option in Bio-Rad CFX Manager 3.1. Melting temperatures were determined by the Bio-Rad CFX Manager 3.1 software in negative mode. Data were recorded in triplicate; error shown is standard deviation from the mean; data are representative of three independent experiments.

**Protein X-ray crystallography**

**ROR $\gamma$ t LBD expression and purification (used for crystallography).** His<sub>6</sub>-ROR $\gamma$ t-LBD containing a C<sub>455</sub>H mutation was expressed and purified as described in Chapter 2.

**X-ray crystallography.** X-ray crystallography experiments and data analysis were performed as described in Chapter 2. Diffraction data were collected at the P11 beamline of the PETRA III facility at DESY (Hamburg, Germany), the ID30B beamline of ESRF (Grenoble, France) and the I04 beamline of Diamond Light Source (Oxford, United Kingdom). The structures of ROR $\gamma$ tC<sub>455</sub>H in complex with **3**, **9**, **10** and **11** were deposited in the protein data bank (PDB) under codes 7NPC, 7NP5, 7NEC and 7NP6.

**Quantitative IL-17a mRNA RT-PCR assay.** Cell culture, RT-PCR experiments and data analysis were performed as described in Chapter 2. Statistical analysis was performed using a one-way analysis of variance (ANOVA) comparing against the DMSO control following Dunnett's *post hoc* test (GraphPad Prism 7.0 software). A *P* value < 0.05 was considered statistically significant. Data were recorded in triplicate; error shown is standard deviation from the mean; data are representative of two independent experiments.

**Selectivity TR-FRET coactivator recruitment assays**

**TR-FRET coactivator recruitment assays on PPAR $\gamma$ .** TR-FRET competition assays were performed in an analogous fashion to that described above, only using 10 nM His<sub>6</sub>-PPAR $\gamma$ -LBD instead of 20 nM His<sub>6</sub>-ROR $\gamma$ t-LBD and 200 nM *N*-terminal biotinylated PGC1 $\alpha$  coactivator peptide (Biotin-N-GTTPPPQEAEPSLLKLLAPANTQ-CONH<sub>2</sub>) instead of 100 nM SRC-1 box-2 coactivator peptide.

**Competition TR-FRET coactivator recruitment assays on PPAR $\gamma$ .** TR-FRET competition assays were performed in an analogous fashion to that described above, only using 100 nM His<sub>6</sub>-PPAR $\gamma$ -LBD instead of 20 nM His<sub>6</sub>-ROR $\gamma$ t-LBD. The assay was performed in the presence of 1  $\mu$ M rosiglitazone, in order to initially activate PPAR $\gamma$ .

**Thermal shift assays on PPAR $\gamma$ .** Thermal shift assays were performed in an analogous fashion to that described above, only using PPAR $\gamma$  LBD instead of ROR $\gamma$ t LBD.

**Absorption, distribution, metabolism and excretion experiments.** ADME measurements were performed as described in Chapter 2.

## Supporting Information

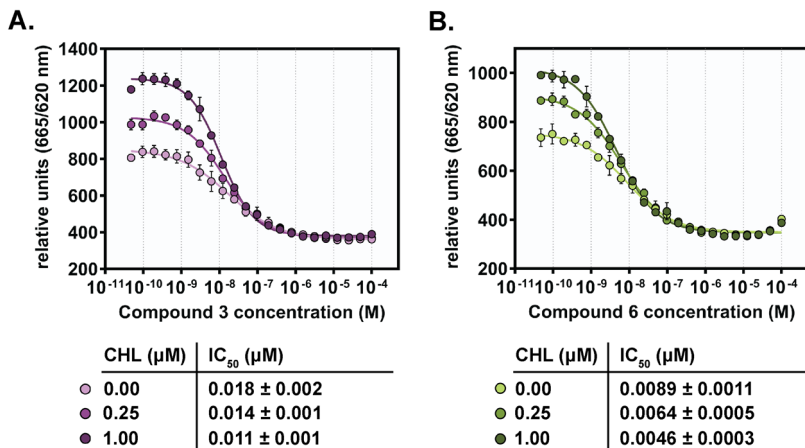


Figure S3.1 | Dose-response curves from the competitive TR-FRET coactivator recruitment assays on ROR $\gamma$ t. Titration of compound 3 (A) and compound 6 (B) to ROR $\gamma$ t in the presence of fixed concentrations of cholesterol (CHL) (0.00  $\mu$ M, 0.25  $\mu$ M and 1.00  $\mu$ M).

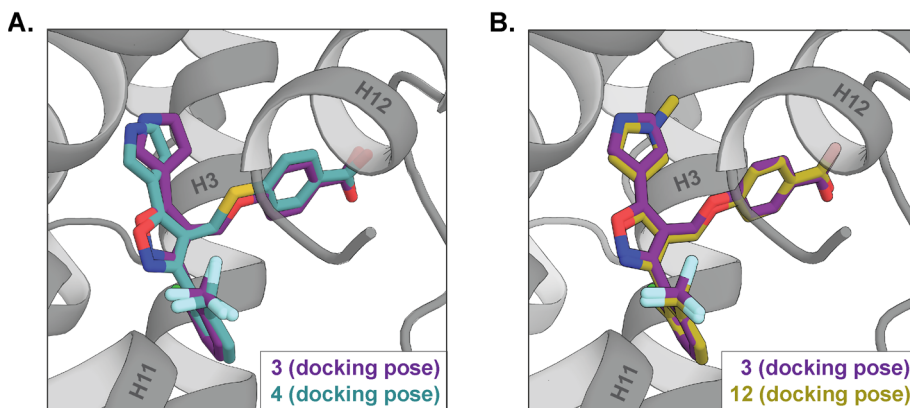
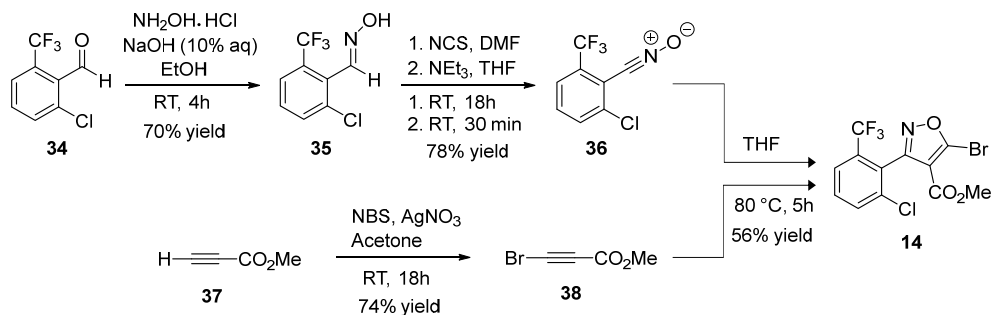
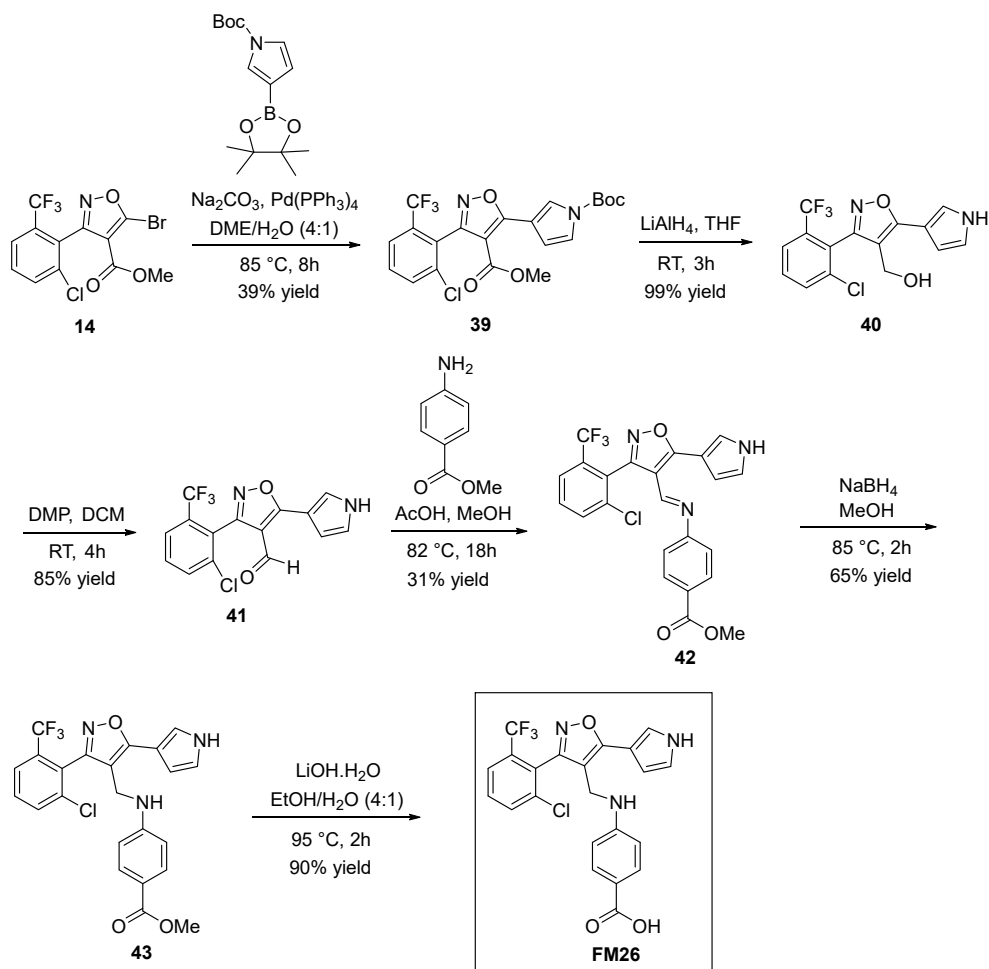


Figure S3.2 | Comparison of the *in silico* modelled docking poses of different isoxazole compounds in complex with the ROR $\gamma$ t LBD. A) Overlay of the docking pose of 3 (purple sticks) with the docking pose of 4 (cyan sticks) in ROR $\gamma$ t (PDB: 4YPQ). B) Overlay of the docking pose of 3 (purple sticks) with the docking pose of 12 (yellow sticks) in ROR $\gamma$ t (PDB: 4YPQ).





**Scheme S3.1 | Synthesis route for isoxazole scaffold 14.** Details on the synthetic methods have been described in Chapter 2.<sup>34</sup>



**Scheme S3.2 | Original synthesis route for isoxazole compound FM26.** Details on the synthetic methods have been described in Chapter 2.<sup>34</sup>

Table S3.1 | Library of compounds with modifications at the C-4 position of isoxazole FM26, ranked based on the Glide Scores from an *in silico* docking screen.

Rank	R (C-4)	Glide Score	Rank	R (C-4)	Glide Score
1		-15.50	12		-14.27
2		-15.26	13		-14.26
3		-15.32	14		-14.26
4		-15.17	15		-13.86
5		-15.16	16		-13.80
6	1 (MRL-871)	-14.99	17		-13.65
7		-14.95	18		-13.48
8		-14.56	19		-13.42
9		-14.51	20		-13.13
10		-14.49	21		-12.94
11		-14.38			

Table S3.2| Library of compounds with modifications at the C-5 position of isoxazole FM156, ranked based on the Glide Scores from an *in silico* docking screen.

Rank	R (C-5)	Glide Score	Rank	R (C-5)	Glide Score
1		-15.18	12		-14.67
2		-15.17	13		-14.66
3		-15.11	14		-14.65
4		-15.12	15		-14.65
5		-15.06	16		-14.47
6		-15.01	17		-14.35
7	1 (MRL-871)	-14.99	18		-14.33
8		-14.85	19	2 (FM26)	-14.26
9		-14.81	20		-13.88
10		-14.78	21		-9.96
11		-14.76			

Table S3.3 | Toxicity risks, druglikeness and drug-score for all compounds, predicted by OSIRIS.<sup>47</sup>

Compound	Toxicity risks				Drug-likeness	Drug-score
	Mutagenic	Tumorigenic	Irritant	Reproductive effect		
1	+	+	+	-	- 7.84	0.17
2	+	+	+	+	- 4.61	0.34
3	+	+	+	+	- 5.19	0.32
4	+	+	+	+	- 5.45	0.28
5	+	-	+	+	- 3.22	0.20
6	+	+	+	+	- 9.78	0.28
7	+	+	+	+	- 9.78	0.28
8	+	+	+	+	- 7.70	0.32
9	+	+	+	+	- 7.29	0.31
10	+	+	+	+	- 4.53	0.32
11	+	+	+	+	- 3.65	0.37
12	+	+	+	+	- 4.35	0.31
13	+	+	+	+	- 4.56	0.30

Table S3.4 | Data collection and refinement statistics for the crystal structure of ROR $\gamma$ t in complex with compounds 3, 9, 10 and 11.

ROR $\gamma$ t with:	Compound 3	Compound 9	Compound 10	Compound 11
<i>Data collection</i>				
Space group	P 6 <sub>1</sub> 2 2	P 6 <sub>1</sub> 2 2	P 6 <sub>1</sub> 2 2	P 6 <sub>1</sub> 2 2
Cell dimensions				
a, b, c (Å)	107.6, 107.6, 100.2	108.2, 108.2, 98.8	108.3, 108.3, 99.4	108.7, 108.7, 98.6
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	68.24 – 1.47 (1.50 – 1.47)	68.00 – 1.55 (1.58 – 1.55)	93.83 – 1.95 (2.00 – 1.95)	98.65 – 1.84 (1.88 – 1.84)
<i>I</i> / $\sigma$ ( <i>I</i> )	25.7 (0.9)	18.3 (0.4)	12.4 (1.0)	5.4 (0.7)
Completeness(%)	100.0 (99.9)	99.9 (99.5)	100.0 (99.9)	100.0 (100.0)
Redundancy	113.1 (105.6)	37.3 (38.3)	39.1 (40.9)	36.3 (36.5)
CC <sub>1/2</sub>	1.000 (0.382)	1.000 (0.386)	0.990 (0.449)	0.990 (0.390)
<i>Refinement</i>				
No. reflections	58535	49561	25654	30277
R <sub>work</sub> /R <sub>free</sub>	0.168/0.197	0.234/0.269	0.195/0.235	0.215/0.245
No. atoms				
Protein	1982	2018	1947	1958
Ligand/ion	32	33	32	32
Water	141	123	120	145
B-factors				
Protein	34.12	39.73	39.44	32.90
Ligand/ion	29.53	30.47	33.48	25.14
Water	44.80	47.20	46.09	40.35
R.m.s. deviations				
Bond lengths (Å)	0.013	0.007	0.007	0.008
Bond angles (°)	1.56	0.93	0.82	1.00
PDB ID	7NPC	7NP5	7NEC	7NP6

## References

1. Robinson-Rechavi, M., Escriva Garcia, H. & Laudet, V. The nuclear receptor superfamily. *J. Cell Sci.* **116**, 585–586 (2003).
2. Huang, P., Chandra, V. & Rastinejad, F. Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. *Annu. Rev. Physiol.* **72**, 247–272 (2010).
3. Gronemeyer, H., Gustafsson, J.-Å. & Laudet, V. Principles for modulation of the nuclear receptor superfamily. *Nat. Rev. Drug Discov.* **3**, 950–964 (2004).
4. Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 (2017).
5. Solt, L. A. & Burris, T. P. Action of RORs and their ligands in (patho)physiology. *Trends Endocrinol. Metab.* **23**, 619–627 (2012).
6. Ivanov, I. I. *et al.* The Orphan Nuclear Receptor ROR $\gamma$ t Directs the Differentiation Program of Proinflammatory IL-17+ T Helper Cells. *Cell* **126**, 1121–1133 (2006).
7. Manel, N., Unutmaz, D. & Littman, D. R. The differentiation of human TH-17 cells requires TGF- $\beta$  and induction of the nuclear receptor ROR $\gamma$ t. *Nat. Immunol.* **9**, 641–649 (2008).
8. Miossec, P. & Kolls, J. K. Targeting IL-17 and TH17 cells in chronic inflammation. *Nat. Rev. Drug Discov.* **11**, 763–776 (2012).
9. Yang, J., Sundrud, M. S., Skepner, J. & Yamagata, T. Targeting Th17 cells in autoimmune diseases. *Trends Pharmacol. Sci.* **35**, 493–500 (2014).
10. Burkett, P. R. & Kuchroo, V. K. IL-17 Blockade in Psoriasis. *Cell* **167**, 1669 (2016).
11. Lock, C. *et al.* Gene-microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis. *Nat. Med.* **8**, 500–508 (2002).
12. Isono, F., Fujita-Sato, S. & Ito, S. Inhibiting ROR $\gamma$ t/Th17 axis for autoimmune disorders. *Drug Discov. Today* **19**, 1205–1211 (2014).
13. Hueber, W. *et al.* Effects of AIN457, a fully human antibody to interleukin-17A, on psoriasis, rheumatoid arthritis, and uveitis. *Sci. Transl. Med.* **2**, 52ra72 (2010).
14. Fauber, B. P. & Magnuson, S. Modulators of the nuclear receptor retinoic acid receptor-related orphan receptor- $\gamma$  (ROR $\gamma$  or RORc). *J. Med. Chem.* **57**, 5871–5792 (2014).
15. Cyr, P., Bronner, S. M. & Crawford, J. J. Recent progress on nuclear receptor ROR $\gamma$  modulators. *Bioorganic Med. Chem. Lett.* **26**, 4387–4393 (2016).
16. Pandya, V. B., Kumar, S., Sachchidanand, Sharma, R. & Desai, R. C. Combating Autoimmune Diseases With ROR $\gamma$  Inhibitors: Hits and Misses. *J. Med. Chem.* **61**, 10976–10995 (2018).
17. Bronner, S. M., Zbieg, J. R. & Crawford, J. J. ROR $\gamma$  antagonists and inverse agonists: a patent review. *Expert Opin. Ther. Pat.* **27**, 101–112 (2017).
18. Sun, N., Guo, H. & Wang, Y. Retinoic acid receptor-related orphan receptor gamma-t (ROR $\gamma$ t) inhibitors in clinical development for the treatment of autoimmune diseases: a patent review (2016-present). *Expert Opin. Ther. Pat.* **29**, 663–674 (2019).
19. Jiang, X. *et al.* A novel series of cysteine-dependent, allosteric inverse agonists of the nuclear receptor ROR $\gamma$ t. *Bioorg. Med. Chem. Lett.* **30**, e126967 (2020).
20. Meijer, F. A., Leijten-van de Gevel, I. A., de Vries, R. M. J. M. & Brunsveld, L. Allosteric small molecule modulators of nuclear receptors. *Mol. Cell. Endocrinol.* **485**, 20–34 (2019).
21. Moore, T. W., Mayne, C. G. & Katzenellenbogen, J. a. Minireview: Not picking pockets: nuclear receptor alternate-site modulators (NRAMs). *Mol. Endocrinol.* **24**, 683–695 (2010).
22. Caboni, L. & Lloyd, D. G. Beyond the ligand-binding pocket: targeting alternate sites in nuclear receptors. *Med. Res. Rev.* **33**, 1081–1118 (2013).
23. Tice, C. M. & Zheng, Y.-J. Non-canonical modulators of nuclear receptors. *Bioorg. Med. Chem. Lett.* **26**, 4157–4164 (2016).
24. Scheepstra, M. *et al.* Identification of an allosteric binding site for ROR $\gamma$ t inhibition. *Nat. Commun.* **6**, e8833 (2015).
25. Karstens, W. F. J. *et al.* ROR $\gamma$ t Inhibitors. *PCT Int. Appl.* WO2012/106995 (2012).
26. Leijten-van de Gevel, I. A. & Brunsveld, L. Delineation of the molecular determinants of the unique allosteric binding site of the orphan nuclear receptor ROR $\gamma$ t. *J. Biol. Chem.* **295**, 9183–9191 (2020).
27. de Vries, R. M. J. M.; Meijer, F. A.; Doveston, R. G.; Leijten-van de Gevel, I. A.; Brunsveld, L. Cooperativity between the Orthosteric and Allosteric Ligand Binding Sites of ROR $\gamma$ t. *Proc. Natl. Acad. Sci.* **118**, e2021287118 (2020).
28. Fauber, B. P. *et al.* Discovery of imidazo[1,5-a]pyridines and -pyrimidines as potent and selective RORc inverse agonists. *Bioorganic Med. Chem. Lett.* **25**, 2907–2912 (2015).
29. Ouvry, G. *et al.* Discovery of phenoxyindazoles and phenylthioindazoles as ROR $\gamma$  inverse agonists. *Bioorganic Med. Chem. Lett.* **26**, 5802–5808 (2016).

30. Shaikh, N. S. *et al.* Discovery and pharmacological evaluation of indole derivatives as potent and selective ROR $\gamma$ t inverse agonist for multiple autoimmune conditions. *Bioorg. Med. Chem. Lett.* **29**, 2208–2217 (2019).
31. Zhang, H. *et al.* Discovery of N-(Indazol-3-yl)piperidine-4-carboxylic Acids as ROR $\gamma$ t Allosteric Inhibitors for Autoimmune Diseases. *ACS Med. Chem. Lett.* **11**, 114–119 (2020).
32. Chaudari, S. S. *et al.* Bicyclic Heterocyclic Compounds as ROR gamma Modulators. *PCT Int. Appl.* WO2015/008234 (2015).
33. de Vries, R. M. J. M., Meijer, F. A., Doveston, R. G. & Brunsveld, L. Elucidation of an Allosteric Mode of Action for a Thienopyrazole ROR $\gamma$ t Inverse Agonist. *ChemMedChem* **15**, 561–565 (2020).
34. Meijer, F. A. *et al.* Ligand-Based Design of Allosteric Retinoic Acid Receptor-Related Orphan Receptor  $\gamma$ t (ROR $\gamma$ t) Inverse Agonists. *J. Med. Chem.* **63**, 241–259 (2020).
35. Brunsveld, L. *et al.* Substituted heterocyclic compounds and their use as retinoic acid receptor-related orphan receptor (ROR) gamma-t inhibitors. *PCT Int. Appl.* WO2020/149740 (2020).
36. Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).
37. Halgren, T. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **47**, 1750–1759 (2004).
38. Degorce, F. HTRF: A Technology Tailored for Drug Discovery - A Review of Theoretical Aspects and Recent Applications. *Curr. Chem. Genomics* **3**, 22–32 (2009).
39. Li, X. *et al.* Structural studies unravel the active conformation of apo ROR $\gamma$ t NR and a common inverse agonism of two diverse classes of ROR $\gamma$ t inhibitors. *J. Biol. Chem.* **292**, 11618–11630 (2017).
40. Soroosh, P. *et al.* Oxysterols are agonist ligands of ROR $\gamma$ t and drive Th17 cell differentiation. *Proc. Natl. Acad. Sci.* **111**, 12163–12168 (2014).
41. DeSantis, K. A. & Reinking, J. L. Use of Differential Scanning Fluorimetry to Identify Nuclear Receptor Ligands. *Methods Mol. Biol.* **1443**, 21–30 (2016).
42. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212–2221 (2007).
43. Gao, K., Oerlemans, R. & Groves, M. R. Theory and applications of differential scanning fluorimetry in early-stage drug discovery. *Biophys. Rev.* **12**, 85–104 (2020).
44. Leijten-van de Gevel, I. A., van Herk, K. H. N., de Vries, R. M. J. M. & Brunsveld, L. PDB ID: 6TDC. PPAR gamma ligand binding domain in complex with MRL-871. (2019). doi:10.2210/pdb6tdc/pdb
45. Cuzzocrea, S. *et al.* Rosiglitazone, a ligand of the peroxisome proliferator-activated receptor-gamma, reduces acute inflammation. *Eur. J. Pharmacol.* **483**, 79–93 (2004).
46. Dalvie, D. K., Kalgutkar, A. S., Khojasteh-Bakht, S. C., Obach, R. S. & O'Donnell, J. P. Biotransformation Reactions of Five-Membered Aromatic Heterocyclic Rings. *Chem. Res. Toxicol.* **15**, 269–299 (2002).
47. Molecular property explorer OSIRIS. 2021 (accessed 2021 26th of February). Available from <http://www.organic-chemistry.org/prog/peo/drugScore>.
48. Kallepalli, V. A. *et al.* Boc Groups as Protectors and Directors for Ir-Catalyzed C–H Borylation of Heterocycles. *J. Org. Chem.* **74**, 9199–9201 (2009).

# Chapter 4

## Cooperativity between the Orthosteric and Allosteric Ligand Binding Sites of ROR $\gamma$ t

### Abstract

Cooperative ligand binding is an important phenomenon in biological systems where ligand binding influences the binding of another ligand at a different site of the protein via an intramolecular network of interactions. The underlying mechanisms behind cooperative binding remain poorly understood, primarily due to the lack of structural data of these ternary complexes. Using TR-FRET studies, we showed that cooperative ligand binding occurs for ROR $\gamma$ t, a nuclear receptor associated with the pathogenesis of autoimmune diseases. To provide the crucial structural insights, we solved twelve crystal structures of ROR $\gamma$ t simultaneously bound to various orthosteric and allosteric ligands. The presence of the orthosteric ligand induces a clamping motion of the allosteric pocket via helix 4-5. Molecular dynamics simulations revealed the unusual mechanism behind this clamping motion, with Ala355 switching between helix 4 and 5. The orthosteric ROR $\gamma$ t agonists regulate the conformation of Ala355, thereby stabilizing the conformation of the allosteric pocket and cooperatively enhancing the affinity of the allosteric inverse agonists.

**This chapter has been published as:** F.A. Meijer<sup>#</sup>, R.M.J.M. de Vries<sup>#</sup>, R.G. Doveston, I.A. Leijten-van de Gevel & L. Brunsveld. Cooperativity between the Orthosteric and Allosteric Ligand Binding Sites of ROR $\gamma$ t. *Proc. Natl. Acad. Sci.* 118, e2021287118 (2021).

<sup>#</sup> These authors contributed equally

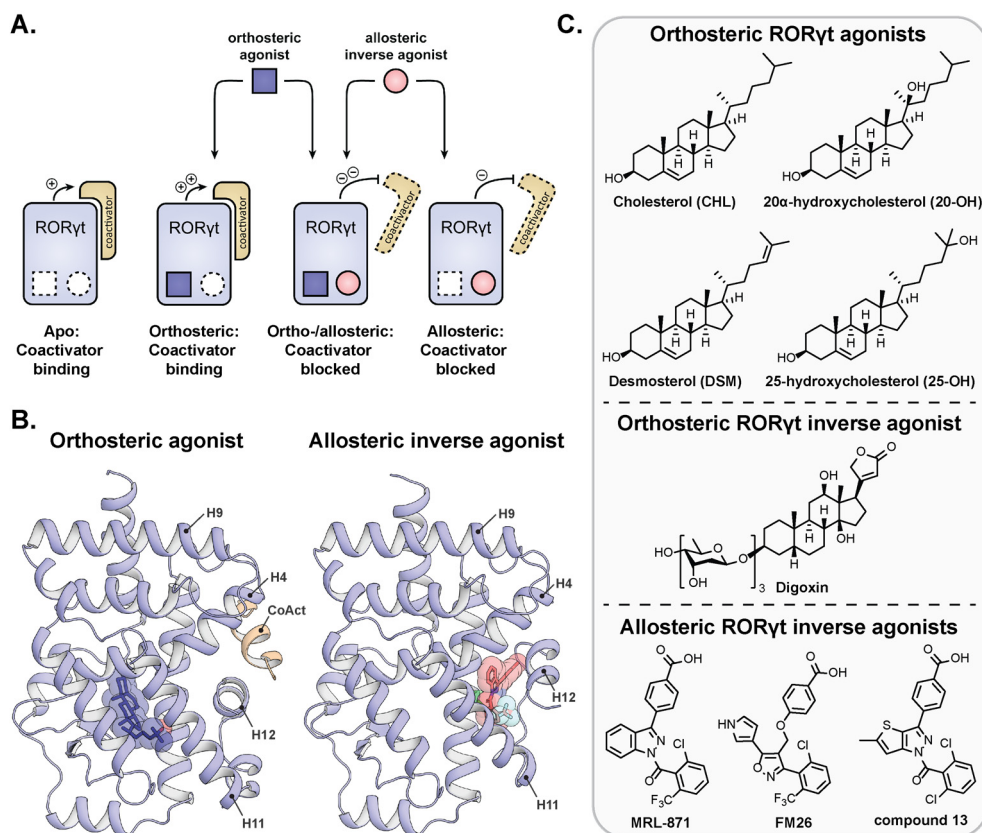
## Introduction

Allosteric ligands bind to pockets on proteins that typically do not overlap with the canonical, orthosteric binding pockets that are usually targeted by endogenous ligands.<sup>1,2</sup> Therefore, allosteric ligands exert their effects via different structural modes of action.<sup>1-3</sup> This can convey advantages over orthosteric ligands in terms of potency, because competition with endogenous ligands is removed, and selectivity, because allosteric sites are less conserved across protein families.<sup>1</sup> Molecules that target allosteric pockets are therefore of high interest for drug development. Over the last decades, allosteric ligands have been identified for several important protein classes, like G-protein coupled receptors (GPCRs) and kinases,<sup>4-6</sup> with some of those compounds developed into marketed drugs.<sup>7,8</sup>

Simultaneous binding of an endogenous, orthosteric ligand and an allosteric drug at different binding sites (dual ligand binding) is a fascinating pharmacological concept, since this can modulate the overall physiological effect of the drug. Of particular significance are cooperative dual ligand binding events where binding of one ligand enhances that of the other,<sup>9-11</sup> as observed for GPCR ligands in particular.<sup>12,13</sup> However, detailed structural insights into the mechanisms of cooperative ligand binding remain scarce.<sup>14</sup> This, in part, results from the absence of high-resolution structural data, required to visualize the effects of dual ligand binding. A better structural understanding of cooperativity in dual ligand binding is therefore required to accelerate the development of new allosteric drugs.

Dual ligand binding has occasionally been observed for nuclear receptors (NRs), but there is no clear mechanistic understanding of connected cooperative effects.<sup>15,16</sup> NRs are a class of transcription factors that can be modulated by endogenous and synthetic small molecules<sup>17</sup> and constitute attractive drug targets, with 16% of all marketed drugs targeting this protein class.<sup>18</sup> Mechanistic understanding and exploitation of cooperative dual ligand binding in NRs harbors great potential for drug development. An interesting NR in this context is ROR $\gamma$ t, that plays an essential role in the differentiation of T helper 17 (Th17) cells, associated with the pathogenesis of autoimmune diseases.<sup>19-21</sup> Inhibition of ROR $\gamma$ t with small molecules, in order to disrupt the Th17/IL-17 pathway, is a promising strategy toward reducing the inflammatory response.<sup>20,22-24</sup> The ROR $\gamma$ t ligand binding domain (LBD) contains both a clearly defined canonical, orthosteric binding site, accessible for endogenous and synthetic compounds<sup>24-27</sup>, and a second binding site, termed allosteric pocket, which was shown to bind allosteric inverse agonists (Figure 4.1A,B).<sup>28-35</sup> Comparison of the crystal structures of the ROR $\gamma$ t LBD in complex with either an orthosteric or allosteric ligand, indicates the possibility for dual ligand binding (Figure 4.1B).<sup>32,34</sup>





**Figure 4.1 | Dual ligand binding in ROR $\gamma$ t.** **A)** Schematic representation of the ROR $\gamma$ t LBD which is intrinsically active in the apo state, inducing coactivator binding. In the presence of an orthosteric agonist (blue), coactivator binding is further increased, while in the presence of an allosteric inverse agonist (pink), coactivator binding is blocked. The presence of both an orthosteric and allosteric ligand, leads to more efficient inhibition of coactivator recruitment because of cooperative dual ligand binding. **B)** Crystal structure of the ROR $\gamma$ t LBD in complex with the agonist 25-hydroxycholesterol (blue sticks) bound to the orthosteric site (PDB: 3LoL), and crystal structure of the ROR $\gamma$ t LBD in complex with inverse agonist **MRL-871** (red sticks) bound to the allosteric site (PDB: 5C4O). The structures show that the orthosteric and allosteric binding sites do not overlap and they highlight the prominently different orientation of H12. **C)** Chemical structures of ROR $\gamma$ t orthosteric agonists (cholesterol (CHL), 20 $\alpha$ -hydroxycholesterol (20-OH), desmosterol (DSM) and 25-hydroxycholesterol (25-OH)), orthosteric inverse agonist (digoxin) and allosteric inverse agonists (**MRL-871**, **FM26** and **compound 13**) used in this study.

In this chapter, we reveal the biochemical and structural proof of ROR $\gamma$ t cooperative dual ligand binding, using a variety of orthosteric and allosteric ligand combinations. Our study also provides a detailed mechanistic explanation of cooperativity between the two binding sites of ROR $\gamma$ t. Extensive dual ligand binding studies, combined with dual ligand protein

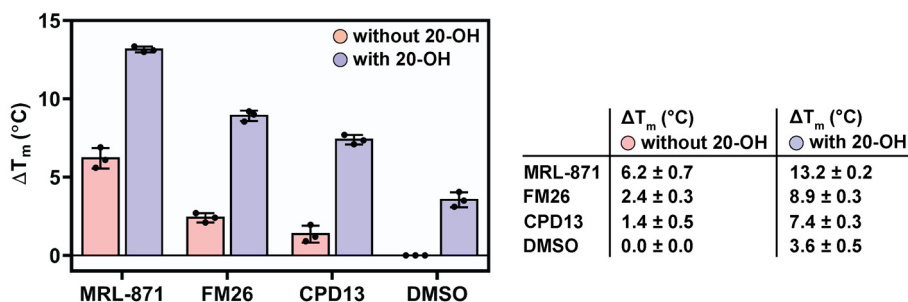
co-crystallography and molecular dynamics (MD) simulations highlight the cooperative binding events and shed light on the underlying molecular mechanism controlling protein conformation and enhanced dual ligand affinity (Figure 4.1A).

### Selection of ROR $\gamma$ t orthosteric and allosteric ligands

The transcriptional activity of ROR $\gamma$ t is correlated with binding of coactivator proteins to its LBD (Figure 4.1A). This interaction is controlled by the conformation of ROR $\gamma$ t helix 12 (H12), which can adopt a stabilized agonistic conformation that promotes coactivator binding, or a destabilized inverse agonistic state that inhibits intrinsic coactivator recruitment. A diverse set of ROR $\gamma$ t ligands (Figure 4.1C) was selected to evaluate the apparent dual ligand binding behavior. Although there is no singularly defined endogenous ligand, cholesterol (CHL) and its derivatives including 20 $\alpha$ -hydroxycholesterol (20-OH), 25-hydroxycholesterol (25-OH), and desmosterol (DSM) are known to be orthosteric agonists that promote coactivator binding by stabilizing H12 in an agonistic conformation (Figure 4.1A,B).<sup>25,36</sup> Digoxin is an exemplary orthosteric inverse agonist, that destabilizes the folding of H12, thereby inhibiting intrinsic coactivator recruitment.<sup>27</sup> **MRL-871**<sup>29,28</sup>, **FM26**<sup>32</sup>, and Glenmark's **compound 13**<sup>33,34,37</sup> are all allosteric inverse agonists, with varying potencies, that reposition H12 into a distinct conformation that prevents coactivator binding (Figure 4.1A,B).<sup>29</sup>

### Dual ligand binding enhances the stability of the ROR $\gamma$ t LBD

Ligand binding typically improves the thermal stability of NRs via structural and dynamic changes to the protein fold.<sup>38,39</sup> Thermal shift assays (TSAs) were performed on ROR $\gamma$ t to investigate the effect of single and dual ligand binding, by measuring the shift in melting temperature ( $\Delta T_m$ ) for the protein.<sup>39,40</sup> In the presence of the orthosteric agonist 20 $\alpha$ -hydroxycholesterol, the  $T_m$  of ROR $\gamma$ t increased by 3.6 °C, relative to the DMSO control, indicating enhanced thermal stability of the protein upon orthosteric agonist binding (Figure 4.2, DMSO, blue bar). Similarly, the binding of the allosteric ligands resulted in  $\Delta T_m$  values between 1 and 7 °C (Figure 4.2, pink bars). When the allosteric ligands were used in combination with 20 $\alpha$ -hydroxycholesterol,  $\Delta T_m$  values between 7 and 14 °C were measured, strongly exceeding the individual and additive effects of the two types of ligands. The significant synergistic enhancement of the thermal stability of ROR $\gamma$ t upon dual ligand binding is a strong indication for a cooperative behavior between the two binding sites.

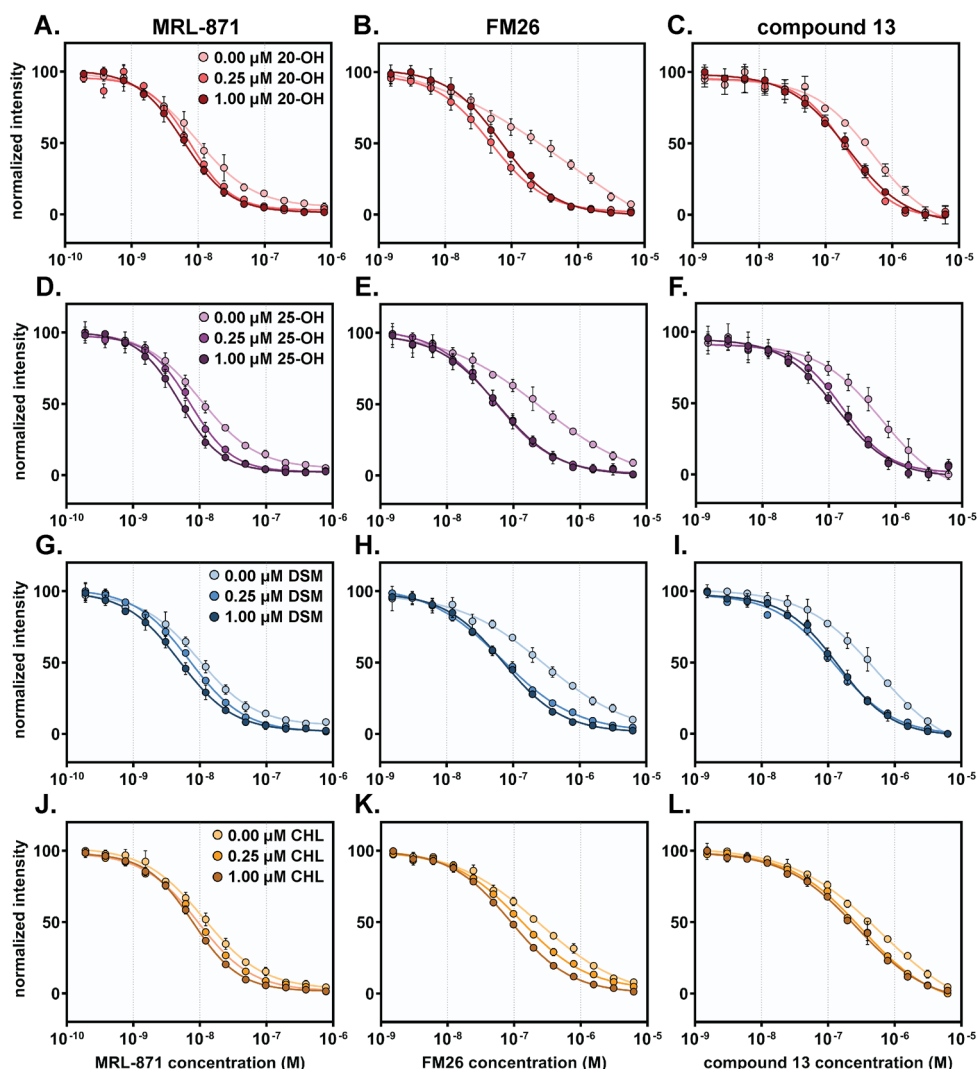


**Figure 4.2 | Thermal stability of ROR $\gamma$ t using a thermal shift assay (TSA) format, in the presence of allosteric ligands MRL-871, FM26 and compound 13, as well as DMSO.** The  $\Delta T_m$  values (shift in melting temperature in °C, relative to DMSO) are shown in absence (pink bars) and presence (blue bars) of orthosteric ligand 20 $\alpha$ -hydroxycholesterol (20-OH). Data recorded in triplicate from three independent experiments. Data are presented as mean  $\Delta T_m \pm$  SD.

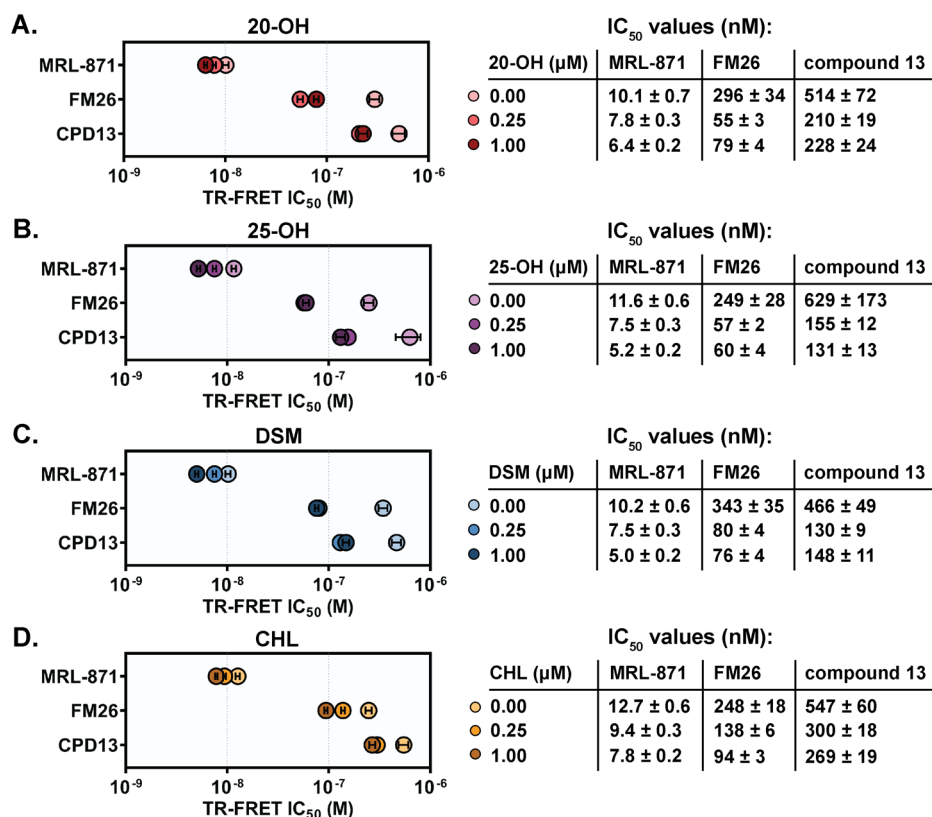
### Orthosteric ROR $\gamma$ t ligands enhance the potency of allosteric ROR $\gamma$ t ligands

Time-resolved FRET (TR-FRET) assays were used to determine the effect of dual ligand binding on coactivator displacement.<sup>41</sup> First, the individual ligands were tested in a TR-FRET coactivator recruitment assay (see Supporting Figures S4.1 and S4.2). The orthosteric ligands cholesterol, 20 $\alpha$ -hydroxycholesterol and desmosterol all showed dose-dependent agonistic behavior as expected.<sup>21,36</sup> Surprisingly, 25-hydroxycholesterol was found to be a partial inverse agonist, contrary to previous reports that showed it would be a ROR $\gamma$ t agonist.<sup>21,36</sup> The orthosteric ligand digoxin exhibited dose-dependent inverse agonistic behavior,<sup>27</sup> which was also the case for the allosteric ligands, as was reported in previous studies.<sup>29,32,34</sup>

Next, the effect of dual ligand binding was investigated by obtaining the dose-response curves for the allosteric ligands in presence and absence of orthosteric ligands, as shown in Figure 4.3 (a schematic representation of the assay setup is shown in Chapter 2, Figure 2.4E). Interestingly, when the allosteric ligands were titrated to assay mixtures containing different fixed concentrations of orthosteric ligands, their IC<sub>50</sub> values decreased (i.e. they became more potent) as the concentration of orthosteric ligand increased (Figure 4.3, Figure 4.4 and Supporting Table S4.1). This synergistic effect is demonstrated by the dose-response curves shifting to the left and increased Hill slopes (Figure 4.3 and Supporting Table S4.1). Although 25-hydroxycholesterol was found to be an inverse agonist, the enhancing effect on the potency of the allosteric ligands was still observed. These data provide compelling evidence for cooperative dual ligand binding to ROR $\gamma$ t, enhancing the potency of the allosteric inverse agonists.



**Figure 4.3 | Dose-response curves from the competitive TR-FRET coactivator recruitment assays.** A-L) Titration of allosteric ligands MRL-871 (A, D, G and J), FM26 (B, E, H and K) and compound 13 (C, F, I and L) to ROR $\gamma$ t in the presence of fixed concentrations (0.00  $\mu$ M, 0.25  $\mu$ M and 1.00  $\mu$ M) of 20 $\alpha$ -hydroxycholesterol (20-OH) (A-C), 25-hydroxycholesterol (25-OH) (D-F), desmosterol (DSM) (G-I) and cholesterol (CHL) (J-L). The data were normalized with regards to plateau levels (non-normalized data are provided in Supporting Figure S4.3). Data are representative of three independent experiments (recorded in triplicate) and error bars represent the SD of the mean.



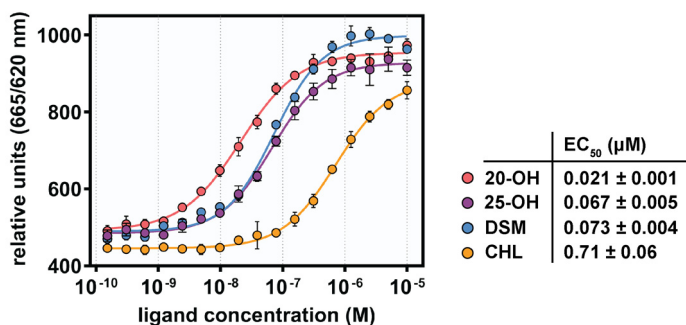
**Figure 4.4 | Overview of  $IC_{50}$  values for MRL-871, FM26 and compound 13 in the presence of different fixed concentrations of orthosteric ligands 20 $\alpha$ -hydroxycholesterol (20-OH) (A), 25-hydroxycholesterol (25-OH) (B), desmosterol (DSM) (C) and cholesterol (CHL) (D). The  $IC_{50}$  values for the allosteric ligands decreased as the concentration of orthosteric ligand increased. Data are representative of three independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD.**

Differences between the cooperative responses were observed for the orthosteric ligands. 25-hydroxycholesterol and desmosterol had the most profound effect on decreasing the allosteric inverse agonist  $IC_{50}$  values (i.e. increasing their potency) compared to 20 $\alpha$ -hydroxycholesterol and cholesterol (Figure 4.3 and Figure 4.4). Furthermore, for 25-hydroxycholesterol and desmosterol, the maximum cooperative response was generally established at a concentration of 0.25  $\mu$ M, whereas cholesterol and 20 $\alpha$ -hydroxycholesterol showed this maximum effect at a higher concentration (1.00  $\mu$ M) (Figure 4.4). These differences are likely to be a result of a combination of slightly differing agonist binding modes and affinities to the orthosteric pocket.

## Orthosteric ligands increase the binding affinity of an allosteric MRL-871 probe

An orthogonal TR-FRET assay format was used to evaluate the effect of orthosteric ligands on the binding affinity of an AlexaFluor647-labelled MRL-871 probe to the allosteric pocket of the ROR $\gamma$ t LBD (see Supporting Figure S4.5 for a schematic representation of the assay setup). Titration of all four orthosteric ligands to a fixed concentration of ROR $\gamma$ t and the MRL-871 probe resulted in an increased FRET signal in a dose-dependent manner (Figure 4.5). This shows that orthosteric ligand binding increases the affinity of the allosteric probe and thus provides further compelling and consistent evidence for cooperative dual ligand binding to ROR $\gamma$ t.

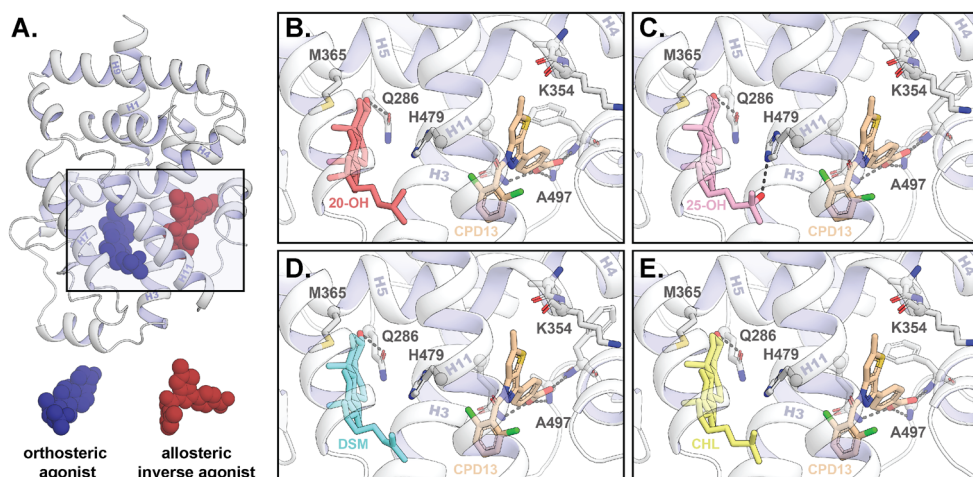
As seen in the coactivator recruitment assay, the orthosteric ligands increased the binding affinity of the allosteric probe by different extents. Here, 20 $\alpha$ -hydroxycholesterol had the most profound effect followed by 25-hydroxycholesterol and desmosterol. Cholesterol was least effective at inducing allosteric probe binding. The pattern of activity was not entirely consistent with that observed in the coactivator recruitment assay where 20 $\alpha$ -hydroxycholesterol showed a lower cooperative effect compared to 25-hydroxycholesterol and desmosterol. Coactivator binding and allosteric ligand binding are discrete events, and thus it is possible that subtle structural differences have a greater impact on one binding event compared to the other.



**Figure 4.5 | Dose-response curves from the TR-FRET assay with an AlexaFluor647-labelled MRL-871 allosteric probe, including an overview of the EC<sub>50</sub> values.** Titration of the orthosteric ligands 20 $\alpha$ -hydroxycholesterol (20-OH), 25-hydroxycholesterol (25-OH), desmosterol (DSM) and cholesterol (CHL) to a fixed concentration of ROR $\gamma$ t (20 nM) and allosteric MRL-871 probe (100 nM). Data are representative of three independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD.

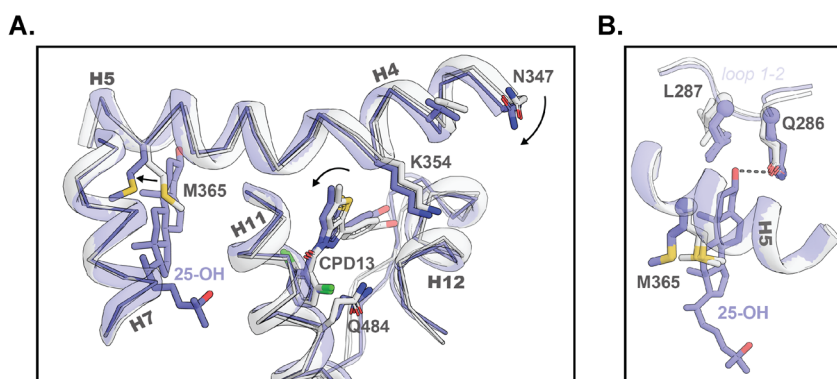
## Co-crystal structures provide molecular insights into the simultaneous binding of orthosteric and allosteric ROR $\gamma$ t ligands

X-ray protein crystallography was used to examine the impact of cooperative dual ligand binding on protein flexibility, protein folding, and ligand binding modes. In order to allow for proper comparison of the crystal structures and minimizing the possibility of crystallization artifacts, conditions were screened that established identical crystal packing. Using this approach, the first ternary complexes of ROR $\gamma$ t bound to both an orthosteric and allosteric ligand were crystallized. In total, twelve novel high-resolution crystal structures were solved, including all combinations of the four orthosteric ligands (20 $\alpha$ -hydroxycholesterol, 25-hydroxy-cholesterol, desmosterol and cholesterol) and three allosteric ligands (MRL-871, FM26 and **compound 13**) (Figure 4.6B-E, Supporting Figure S4.4 and Supporting Tables S4.2-S4.4).



**Figure 4.6 | Crystal structures of the ROR $\gamma$ t LBD in complex with orthosteric and allosteric ligands.** A) Cartoon representation of ROR $\gamma$ t in complex with an orthosteric ligand (blue spheres) and an allosteric ligand (red spheres). The rectangle indicates the location of the enlarged orthosteric and allosteric ligand binding pocket. B-E) The orthosteric and allosteric ligand binding pocket of ROR $\gamma$ t in the presence of various orthosteric ligands (20 $\alpha$ -hydroxycholesterol (20-OH) in red, 25-hydroxycholesterol (25-OH) in pink, desmosterol (DSM) in blue and cholesterol (CHL) in yellow) and the allosteric ligand **compound 13** (brown).

All twelve ternary ROR $\gamma$ t crystal structures revealed the protein folded into a conformation where H12 is positioned over the allosteric ligand and thus physically preventing potential coactivator binding. This is consistent with the binary crystal structures of ROR $\gamma$ t in complex with an allosteric modulator only.<sup>29,32,34</sup> Previously, we reported that allosteric ligands **FM26** and **compound 13** introduce more bulk towards helix 4 of ROR $\gamma$ t compared to **MRL-871** (Supporting Figure S4.6), resulting in a modest shift of helix 4 (towards helix 9).<sup>32,34</sup> Interestingly, the additional binding of an orthosteric ligand was seen to reverse this process in all the ternary crystal structures containing an orthosteric ligand and **FM26** or **compound 13** (Figure 4.7A and Supporting Figure S4.6). For the structures containing **MRL-871**, a similar but less pronounced effect was observed.



**Figure 4.7 | Comparison of the crystal structures of ROR $\gamma$ t in the presence (blue, PDB: 6T5o) or absence (white, PDB: 6TLM) of an orthosteric modulator. A)** The presence of the orthosteric ligand shifts helix 4 towards the allosteric pocket, thereby clamping the allosteric ligand. **B)** Focused view of the orthosteric binding pocket. Side chains of Gln286, Leu287 and Met365 are shown for all crystal structures containing ligands in both pockets (twelve structures in blue) as well as in absence of an orthosteric modulator (three structures in white). The presence of the orthosteric ligand locks Met365 into a defined state, which is conserved for all twelve crystal structures containing orthosteric ligands.

Superposition of the crystal structures showed that the binding modes of the cholesterol derivatives are comparable to those seen in earlier binary structures of ROR $\gamma$ t in complex with the orthosteric ligands 20 $\alpha$ - and 25-hydroxycholesterol (PDB: 3KYT and 3LoL, respectively).<sup>36</sup> Comparison of the orthosteric pockets showed that the conformation of surrounding residues is predominantly unaffected by binding of the cholesterol derivatives, which is the same for the allosteric pocket. However, an altered conformation was observed for Met365 which, in the absence of an orthosteric ligand, was oriented towards the center of the orthosteric pocket



(Figure 4.7B). The cholesterol derivatives occupy this part of the pocket, thereby ‘locking’ Met365 in a distinctly repositioned conformation (Figure 4.7B). Interestingly, this Met365 repositioning is correlated with a restriction of the movement of helix 5 and leads to a conformational change of helix 4 towards the allosteric ligand, a clamping effect that is seen in all the crystal structures (Figure 4.7A). The distance between the  $\alpha$ -carbons of Asn347 on helix 4 and Gln484 on helix 11 is used as a measure for the clamping motion (Supporting Figure S4.7). Although the movement of helix 4 is restricted by the allosteric ligand, the presence of the orthosteric ligands can reduce this distance by 0.1 nm. To accommodate the altered conformation of helix 4, the allosteric ligands slightly twist, following the motion of helix 4 (Figure 4.7A and Supporting Figure S4.6). A larger movement of helix 4 towards the allosteric ligand is correlated with a larger twist of the allosteric ligand (Supporting Figure S4.6). Consequently, the conformation of the loop between helix 11 and helix 12 changes due to the polar interactions of the conserved carboxylic acid of the allosteric ligands with the protein backbone. Altogether, the structural data reveal a molecular mechanism of how the presence of an orthosteric ligand influences the binding behavior of the allosteric ligand.

### **Orthosteric ligands restrict the conformational flexibility of ROR $\gamma$ t Met 365 and alter the conformation of helix 7 and 11**

MD simulations were performed to investigate the interplay between orthosteric and allosteric ligand binding. For this, co-crystal structures of ROR $\gamma$ t in complex with both ligands were compared to the respective structure with only the allosteric modulator bound. To improve the reliability of the simulations, five independent simulations were performed per complex, each starting from a random initial velocity distribution.

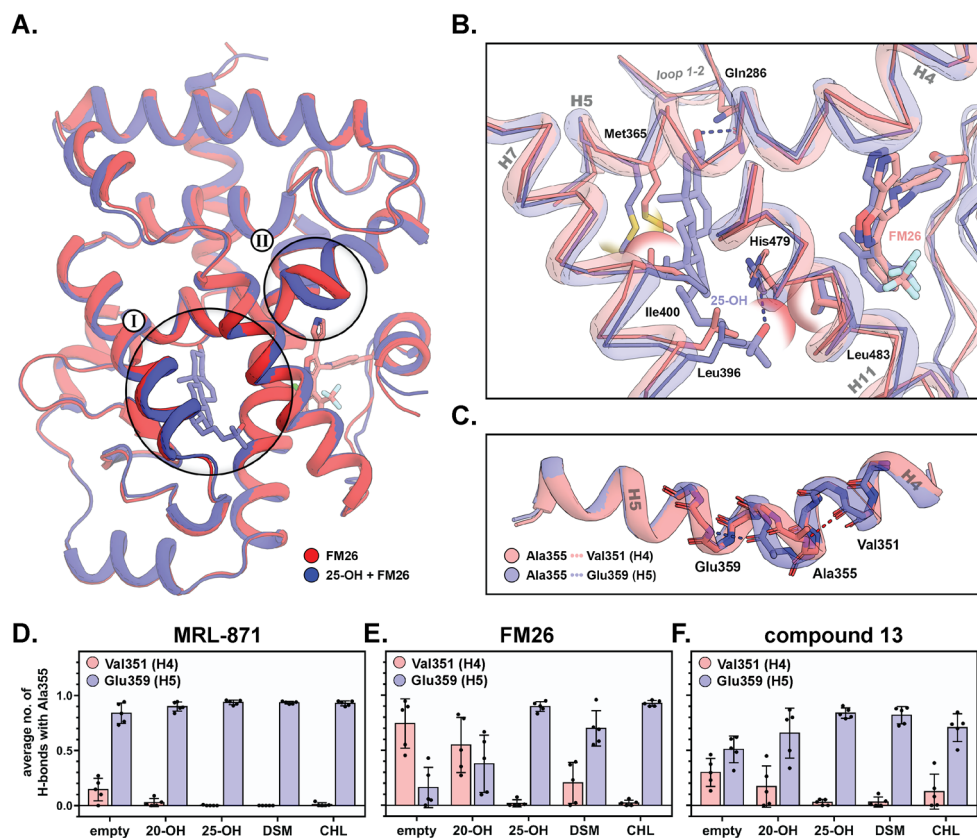
For all simulations, no large conformational changes in the tertiary structure of the protein or ligand conformation were observed. The presence of any orthosteric ligand significantly reduced the overall flexibility of the complete protein backbone (Supporting Figure S4.8). Consistent with the crystal structures, Met365 of ROR $\gamma$ t showed limited conformational freedom in the orthosteric pocket due to steric hindrance with the C-ring of the cholesterol derivatives. The specific conformation of Met365 leads to a repositioning of Ile400 on helix 7, thereby shifting this helix away from helix 5 (Figure 4.8B). The aliphatic tail of the cholesterol derivatives is oriented towards Leu483 on helix 11, causing this helix to move towards the allosteric ligand, restricting the overall mobility of both the allosteric ligand and helix 11 (Figure 4.8B and Supporting Figure S4.8).

## Orthosteric ligands influence the helix participation of Ala355 resulting in clamping of the allosteric binding pocket

We investigated in detail the helix 4 shift which was observed in our crystallographic data upon binding of the orthosteric ligand. The MD simulations showed an unusual mechanism by which this helix movement takes place. In the crystal structures, Ala355 is located at the end of helix 4, but during our simulations, Ala355 showed the ability to exchange its participation between helix 4 and helix 5 (Figure 4.8C). The RORyt structures containing both ligands significantly bias Ala355 towards the helix 5 conformation compared to the complexes with only the allosteric ligand present. This altered equilibrium of conformations is more distinct for complexes with the bulkier allosteric ligands **FM26** and **compound 13** since these ligands promote the helix 4 conformation for Ala355 in absence of an orthosteric ligand. The participation in helix 5 by Ala355 induces a shift of helix 4 towards the allosteric ligand, moving the ligand deeper into the allosteric binding pocket (Figure 4.8B,C). In agreement with the crystal structures, an apparent clamping motion of helix 4 can be observed for all structures containing an orthosteric ligand, illustrated by the distance between the  $\alpha$ -carbons of Asn347 and Gln484 (Supporting Figure S4.9).

The effect of different orthosteric ligands on the conformation of Ala355 was determined by measuring the average number of hydrogen bonds of Ala355 with the backbone of Val351 (helix 4) and Glu359 (helix 5), respectively (Figure 4.8D-F). The structures containing **MRL-871** showed that, together with binding of any of the orthosteric ligands, Ala355 is almost exclusively in the helix 5 conformation (Figure 4.8D). For the complex of RORyt with **FM26** alone, Ala355 is primarily in the helix 4 conformation but in the presence of an orthosteric ligand, the equilibrium completely shifts towards helix 5 (Figure 4.8E). A similar behavior was observed for **compound 13** (Figure 4.8F). For all structures containing 25-hydroxy-cholesterol, Ala355 showed to be almost exclusively in the helix 5 conformation, independent of which allosteric ligand is bound. In contrast, 20 $\alpha$ -hydroxycholesterol only had a minor effect on the conformational equilibrium. It appears that the flexibility of helix 5 plays a key role in defining the conformation of Ala355. Without an orthosteric ligand present, helix 5 acts as a spring, allowing movement of helix 4-5 and allowing Ala355 to sample both conformations. Upon binding of an orthosteric ligand, the conformation of Met365 is locked making the spring more rigid, promoting the helix 5 conformation. This provides an explanation of why orthosteric ligands with less conformational freedom in the ligand binding pocket, due to additional polar interactions or more rigid alkene bonds for 25-hydroxy-cholesterol and desmosterol respectively, more effectively induce the helix 5 conformation.

The absolute conformation of Ala355 is directly correlated to the binding mode of the allosteric ligand. Therefore, the extent by which the orthosteric ligand adjusts the conformational equilibrium of this residue will define the cooperative binding behavior.



**Figure 4.8 | Comparison of the ROR $\gamma$ t complexes bound to an allosteric ligand in the presence (blue) or absence (red) of an orthosteric ligand using molecular dynamics. A)** Superposition of the average structure of FM26 with and without 25-hydroxycholesterol (25-OH). The orthosteric ligand binding pocket (I) and the transition between helix 4 and 5 of ROR $\gamma$ t (II) are highlighted. **B)** Focused view of the orthosteric pocket of ROR $\gamma$ t. Polar interactions are shown as dashed lines and steric clashes as semi-transparent spheres. **C)** Isolated helix 4-5 showing the conformational switch of Ala355 from helix 4 to helix 5 upon orthosteric ligand binding. **D-F)** The average number of hydrogen bonds of Ala355 with Val351 (helix 4; red) or Glu359 (helix 5; blue) in the presence of different orthosteric and allosteric modulators over the course of the simulation. Bars represent the average value over five independent simulations with the individual values represented as black spheres and the error bar showing the standard deviation.

## Conclusions

Cooperative dual ligand binding is a relevant but poorly understood concept in drug discovery. Instead of competing with an endogenous ligand, the endogenous ligand and an allosteric ligand can collaborate to produce a modulated pharmacological response. It is difficult to rationally design ligands that show a predictable cooperative binding behavior, which is primarily caused by a lack of structural understanding of the underlying cooperativity. In this work, we used a combination of biochemical data, protein crystallography, and MD simulations to produce a mechanistic explanation of how cooperative dual ligand binding occurs for the NR ROR $\gamma$ t. The thermal shift data indicated cooperative stabilization of ROR $\gamma$ t folding by dual ligand binding. The TR-FRET coactivator recruitment assays demonstrated the functional effect of the cooperative binding by an enhanced inhibitory potency of the allosteric ligands on coactivator binding, in the presence of an orthosteric ligand. Although all orthosteric ligands showed cooperative behavior in these TR-FRET data, they all did this to different extents. Desmosterol and 25-hydroxycholesterol gave the highest cooperative response, showing a significant decrease in IC<sub>50</sub> values of the allosteric ligands, whereas cholesterol and 20 $\alpha$ -hydroxycholesterol showed only minor effects. In an orthogonal TR-FRET assay format we showed that orthosteric ligand binding directly increased the binding affinity of an allosteric probe, which provided additional evidence for cooperative dual ligand binding.

The determination of the ternary crystal structures of ROR $\gamma$ t with all combinations of orthosteric and allosteric ligands allowed for the elucidation of a mechanistic explanation for the cooperative dual ligand binding behavior. Orthosteric ligands lock Met365 in a distinct conformation that leads to a conformational change of helix 4-5, which results in a clamping effect of the allosteric binding pocket. This results in a modest conformational change of the allosteric ligand. Frequently, these small differences between crystal structures go unnoticed because of a lack of proper reference structures, but they can be critical to explain protein functioning.<sup>42</sup> The generation of the twelve closely related ternary structures allowed for the clear delineation of these differences.

MD simulations further confirmed the clamping behavior of the allosteric pocket upon orthosteric ligand binding. The clamping motion was achieved by a yet, to our knowledge, unknown characteristic of Ala355 that showed the ability to transition between the end of helix 4 and the beginning of helix 5, which plays a significant role in the clamping effect. Restricting the conformational flexibility of Met365 with an orthosteric ligand limits the movement of helix 4-5, thereby promoting the helix 5 conformation of Ala355, which results in a

conformational change of helix 4 towards the allosteric ligand. Desmosterol and 25-hydroxycholesterol more effectively directed the absolute conformation of Ala355 to helix 5 compared to cholesterol and 20 $\alpha$ -hydroxycholesterol. A similar trend was observed for these compounds in the TR-FRET data where desmosterol and 25-hydroxycholesterol also showed the largest effect on the increase in potency for the allosteric ligands. This is likely to be caused by the reduced flexibility of these ligands in the orthosteric binding pocket, and thus a more effective locking of Met365. Ultimately, the conformation of Ala355 is directly correlated to the binding mode and, as a result, the binding affinity of the allosteric ligand. Considering the TR-FRET data, the helix 5 conformation of Ala355, and the associated clamping motion of helix 4, shows to have a positive effect on the binding affinity of the allosteric ligand and provides an explanation for the cooperative binding behavior.

Kojetin and coworkers already demonstrated the essential role of helix 4-5 in the allosteric regulation of dimerization and the AF-2 site of RXR.<sup>42</sup> The bent conformation of helix 4-5 is a common characteristic within the NR family. Like ROR $\gamma$ t, most of the NR family members contain a conformationally flexible residue at the transition between these two helices.<sup>43</sup> Therefore, it is likely that orthosteric ligand binding also has a significant effect on the dynamics and conformation of helix 4-5 of other NRs.<sup>42</sup> In addition to helix 12, also helix 4 is essential for the recruitment of coactivators.<sup>17</sup> An altered behavior of helix 4 as a result of orthosteric ligand binding could therefore potentially influence the cofactor binding behavior across all NRs.

In summary, our data provide the first mechanistic explanation for cooperative dual ligand binding in NRs, via a mechanism in ROR $\gamma$ t that operates via an internal conformational change of the LBD. The specific ROR $\gamma$ t cooperativity data in this study lets speculate that similar mechanistic concepts can also be found to govern other NRs and protein classes containing two binding sites. These mechanistic insights bring the pharmacological concept of cooperative dual ligand binding for NRs a step closer to implementation in NR drug discovery. The potential to further enhance the pharmacological effects of allosteric ligands by an interplay with the endogenous orthosteric NR ligands, provides a highly attractive entry for a novel NR pharmacology.

## Acknowledgements

Rens de Vries performed X-ray crystallography studies and molecular dynamics (MD) simulations and designed the corresponding figures, for which he is greatly acknowledged. Furthermore, Iris Leijten-van de Gevel is thanked for performing thermal shift assays and Richard Doveston is kindly acknowledged for the valuable discussions.

This chapter was reprinted (adapted) with permission from *Proc. Natl. Acad. Sci.* 2021, 118, e2021287118. Copyright 2021 National Academy of Sciences.

## Experimental Section

### Biophysical assays

**RORyt LBD expression and purification (used for TR-FRET assays).** His<sub>6</sub>-RORyt-LBD was expressed and purified as described in Chapter 2.

**TR-FRET coactivator recruitment assays, competition assays and AlexaFluor-MRL-871 recruitment assays.** TR-FRET assays and data analysis were performed as described in Chapter 2. The TR-FRET competition data was normalized with regards to plateau levels (positive and negative control values). Data were recorded in triplicate; error shown is standard deviation from the mean; curves are representative of > 3 independent experiments.

**Thermal shift assays.** Thermal shift assays and data analysis were performed as described in Chapter 3, using 5  $\mu$ M RORyt LBD and single allosteric ligands (15  $\mu$ M **MRL-871**, 60  $\mu$ M **FM26** and 20  $\mu$ M **compound 13** (lowest concentrations giving a maximal  $\Delta T_m$ ) or compound combinations (combining the previous allosteric ligand concentrations with 60  $\mu$ M 20 $\alpha$ -hydroxycholesterol (20-OH)).  $\Delta T_m$  values were determined as mean  $\pm$  standard deviation from three independent experiments performed in triplicate and normalized to DMSO.

### Protein X-ray crystallography

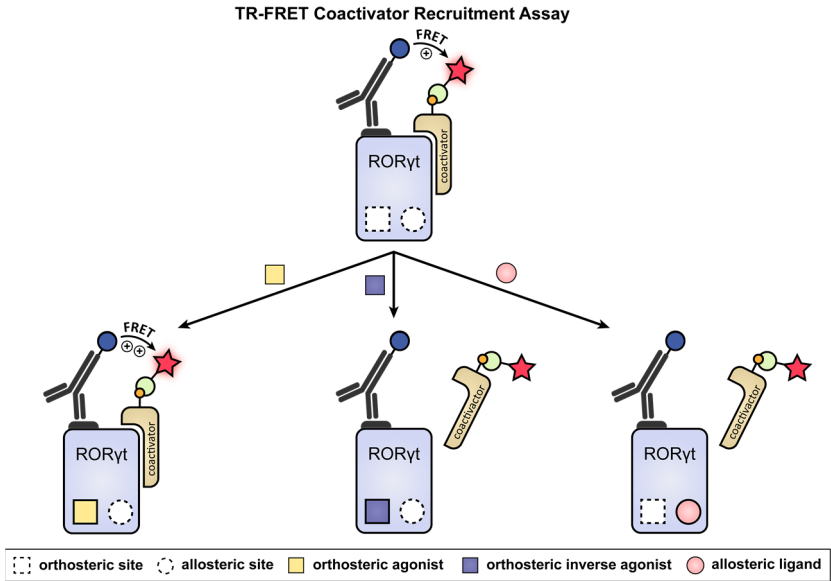
**RORyt LBD expression and purification (used for crystallography).** His<sub>6</sub>-RORyt-LBD containing a Cys455His mutation was expressed and purified as described in Chapter 2.

**X-ray crystallography.** Allosteric ligands **MRL-871**, **FM26** and **compound 13** were dissolved in DMSO to a final concentration of 40 mM, 30 mM and 20 mM, respectively. The cholesterol derivatives were poorly soluble in DMSO and were therefore dissolved in EtOH to a final concentration of 40 mM. All ligands were aliquoted to prevent freeze-thaw cycles and to prevent evaporation of the ligand solution. For both ligands, 1.7-2.5 equivalents were added to the RORytC455H solution (11.1 mg/mL) and the mixture was incubated on ice. After 1 h, the sample was centrifuged at 20,000 RCF for 20 min at 4 °C to eliminate ligand and protein precipitate. All crystals were produced using a sitting drop crystallization method. MRC-2 well (Hampton Research) plates were prepared using a Mosquito pipetting robot (TTP Labtech) and stored at room temperature. Dependent on the ligand combinations, different crystallization and cryo-protection conditions were used. In general, crystals grew to their final size overnight and nucleated at the bottom of the well, thereby attaching to the plastic surface. An Ultra Micro-Needle (HR4-849, Hampton Research) was used to dent the plastic right next to the crystal to release the crystal. Diffraction data of the crystals containing 20 $\alpha$ -hydroxycholesterol were collected at the P11 beamline of the Positron Electron Tanden Ring Anlage III (PETRA III) facility at DESY (Hamburg, Germany) while the other crystals were measured at the i03 beamline of the Diamond Light Source (Oxford, United Kingdom). All crystals were measured at 100K using a wavelength of 1 Å. Initial data processing was performed using the CCP4i2 suite (version 7.0.077).<sup>44</sup> Diffraction Integration for Advanced Light Sources (DIALS) was used to integrate the data and Aimless was used for scaling.<sup>45,46</sup> Using the RORyt crystal structure in complex with allosteric ligand **FM26** (PDB: 6SAL) as a search

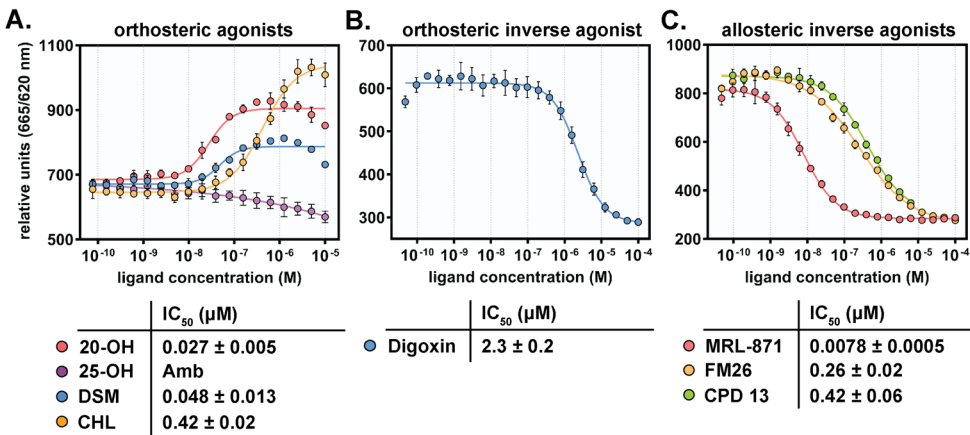
model for molecular replacement, Phaser was used to phase the data and ligand restraints were generated using AceDRG.<sup>47,48</sup> REFMAC and Crystallographic Object-Oriented Toolkit (COOT) were used for sequential refinement and model building.<sup>49,50</sup> Final refinement was performed using phenix.refine from the Phenix software suite (version 1.16\_3459).<sup>51</sup> For all structures, no Ramachandran outlier were observed, except for one in 6TLT. The Ramachandran statistics showed that 98-99% of the residues are in the preferred conformation and 1-2% are in the allowed conformation. Figures were made with PyMOL (version 2.2.3, Schrödinger).<sup>52</sup>

**Molecular dynamics studies.** The GROMACS 2019.3 molecular dynamics package was used to perform the simulations.<sup>53</sup> X-ray structures of ROR $\gamma$ t in complex with an allosteric ligand (PDB entries: 5C4O, 6SAL and 6TLM) and both the orthosteric and allosteric ligand (PDB entries: 6T4G, 6T4I, 6T4J, 6T4K, 6T4T, 6T4U, 6T4W, 6T4X, 6T4Y, 6T5o, 6TLQ and 6TLT) were used. Whenever necessary, the protein was N-terminally truncated to Thr268 in order to use the same protein sequence for all simulations. The FF14SB force field was used to parameterize the protein.<sup>54</sup> Ligands were parameterized separately using the General Amber Force Field (GAFF).<sup>55</sup> The complex was immersed in a cubic box with approximately 22500 TIP3P waters extending 20 Å away from the protein surface.<sup>56</sup> The system charge was neutralized using one Cl<sup>-</sup> ion. The system was first energy minimized using the steepest decent minimization algorithm using a maximum number of 50000 steps. Next, the system was progressively equilibrated by performing three heavy-atom restrained in the isothermal-isovolumetric (NVT) simulations for 100 ps at 100 K, 200 K and 300 K consecutively (Velocity-rescale thermostat) with a time (coupling) constant of 0.1 ns.<sup>57</sup> The final step of equilibration was performed for 100 ps in the isothermal-isobaric (NPT) ensemble at 300 K (Parrinello-Rahman barostat) with a time (coupling) constant of 2.0 ns).<sup>58</sup> During all stages, the maximum force on the protein and ligand atoms was set to 1000 kJ mol<sup>-1</sup> nm<sup>-2</sup> and the bonds were restrained using the Linear Constraint Solver (LINCS) algorithm.<sup>59</sup> The long-range electrostatics were calculated using the Particle Mesh Ewald (PME) method with a short range cutoff of 1.0 nm and a grid spacing of 0.16 nm.<sup>60</sup> Five independent simulation runs of 100 ns were performed for each system, with every run starting from a random initial velocity distribution.

Supporting Information

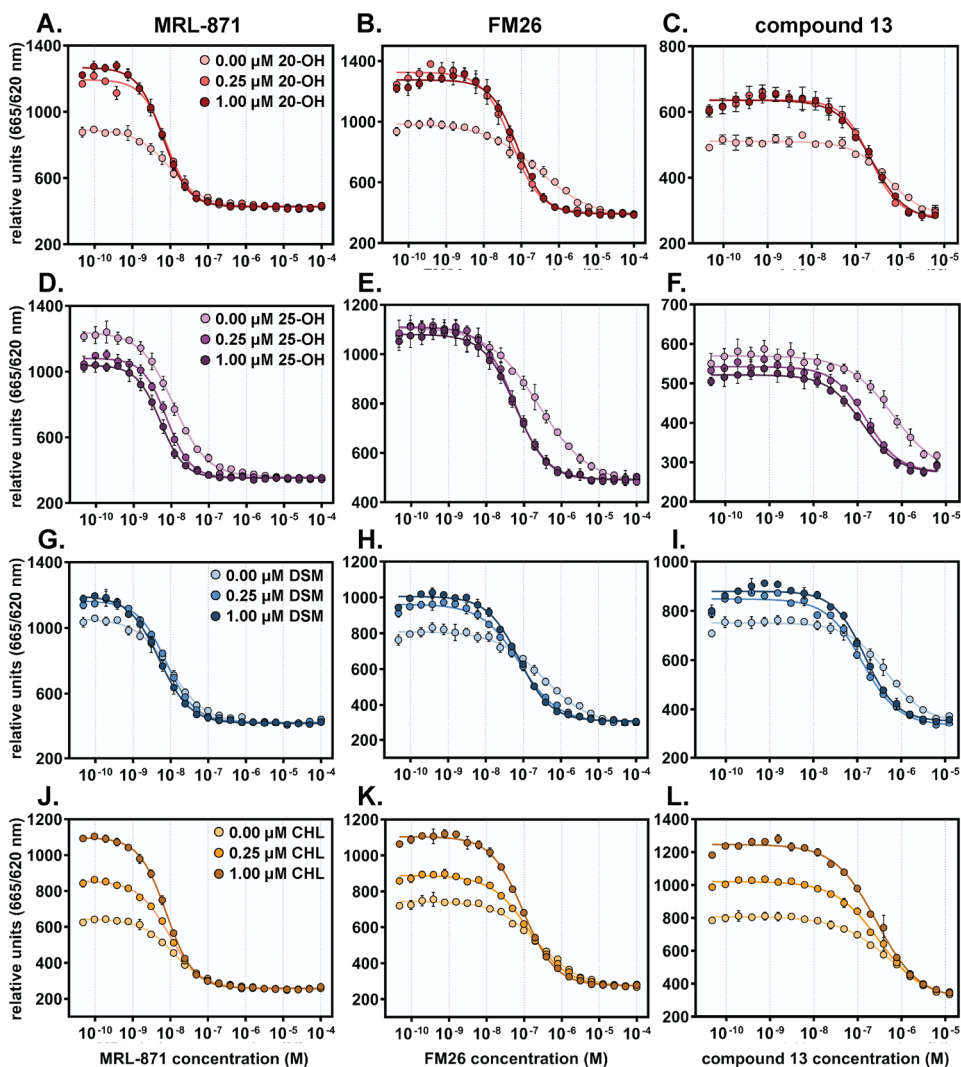


**Figure S4.1 | Schematic representation of the TR-FRET coactivator recruitment assay.** RORγt is intrinsically active in the apo state, inducing coactivator binding, which results in FRET pairing of an anti-His terbium cryptate donor with the d2-labelled streptavidin acceptor. Binding of an agonist further increases coactivator binding, resulting in a higher FRET pairing, whereas the binding of an orthosteric or allosteric inverse agonist results in coactivator displacement thus a lower FRET pairing.

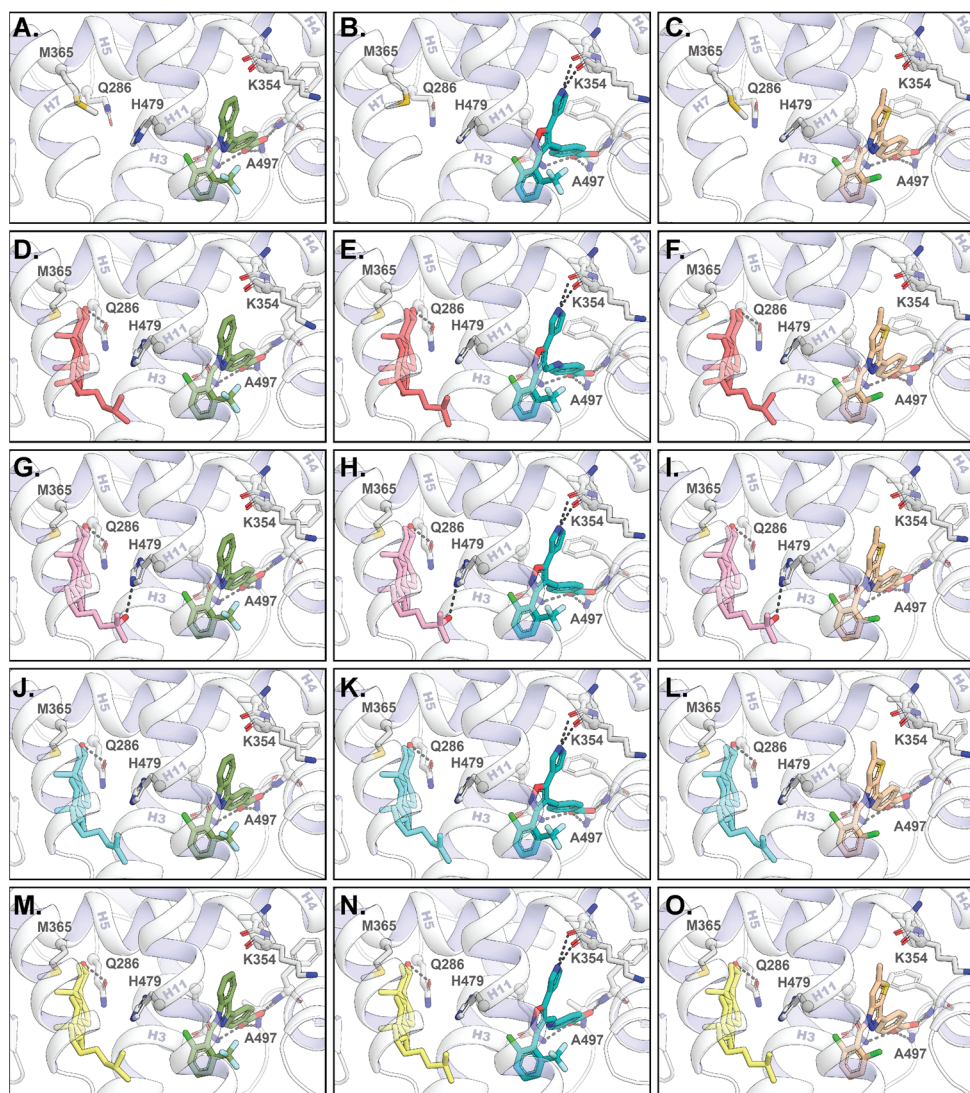


**Figure S4.2 | Dose-response curves from the TR-FRET coactivator recruitment assays, including an overview of the IC<sub>50</sub> values.** A-C) Titration of all ligands to RORγt: orthosteric agonists 20α-hydroxy-cholesterol (20-OH), 25-hydroxycholesterol (25-OH), desmosterol (DSM) and cholesterol (CHL) (A), orthosteric inverse agonist digoxin (B) and allosteric inverse agonists MRL-871, FM26 and compound 13 (CPD 13) (C).



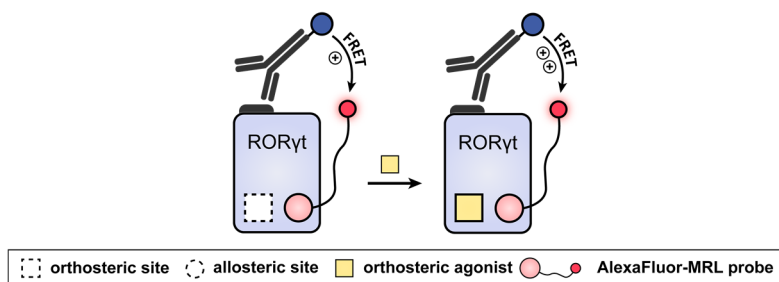


**Figure S4.3 |** Non-normalized data of the data shown in Figure 4.3. Dose-response curves from the competitive TR-FRET coactivator recruitment assays. A-L) Titration of allosteric ligands MRL-871 (A, D, G and J), FM26 (B, E, H and K) and compound 13 (C, F, I and L) to ROR $\gamma$ t in the presence of fixed concentrations (0.00  $\mu$ M, 0.25  $\mu$ M and 1.00  $\mu$ M) of 20 $\alpha$ -hydroxycholesterol (20-OH) (A-C), 25-hydroxycholesterol (25-OH) (D-F), desmosterol (DSM) (G-I) and cholesterol (CHL) (J-L).

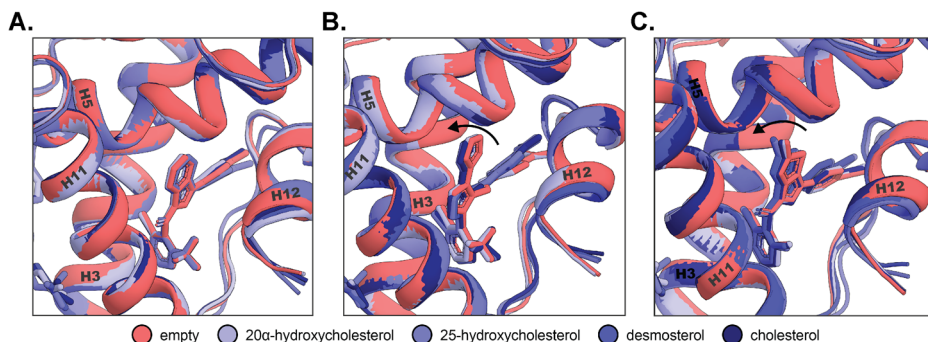


**Figure S4.4 | Crystal structures of the ROR $\gamma$ t LBD in complex with orthosteric and allosteric ligands.** A-C) Focused view of the orthosteric and allosteric ligand binding pockets from the previously published crystal structures containing only an allosteric ligand (**MRL-871** (PDB: 5C4O) in green, **FM26** (PDB: 6SAL) in cyan or **compound 13** (PDB: 6TLM) in brown). D-O) The orthosteric and allosteric ligand binding pocket of ROR $\gamma$ t in the presence of twelve combinations of orthosteric and allosteric ligands (20 $\alpha$ -hydroxycholesterol in red, 25-hydroxycholesterol in pink, desmosterol in blue and cholesterol in yellow).

## TR-FRET AlexaFluor-MRL Recruitment Assay



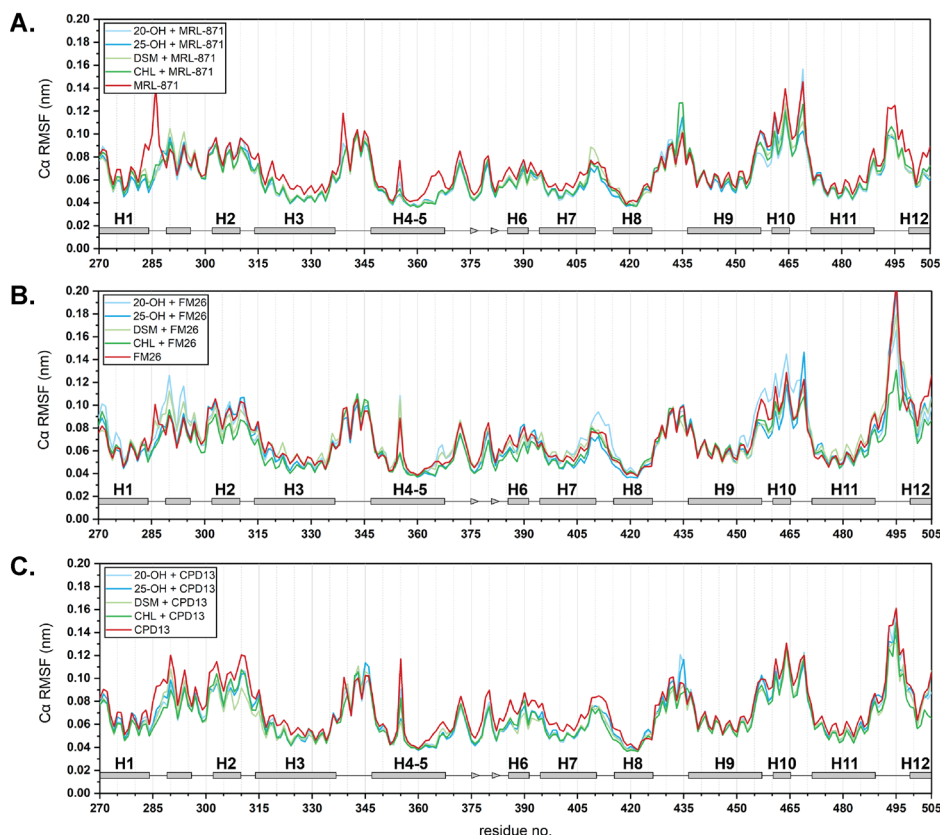
**Figure S4.5 | Schematic representation of the TR-FRET AlexaFluor-MRL recruitment assay, using the AlexaFluor647-labelled MRL-871 probe.** In the absence of ligand, the AlexaFluor-MRL probe binds to ROR $\gamma$ t, resulting in FRET pairing between an anti-His terbium cryptate done and the probe. The binding of an orthosteric agonist increases the binding affinity of the probe (cooperative effect), resulting in a higher FRET pairing.



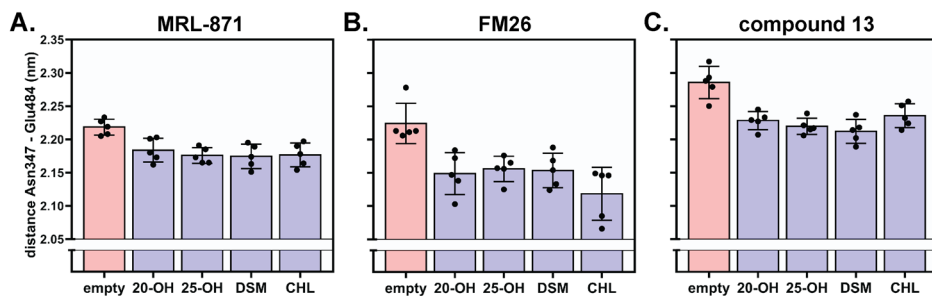
**Figure S4.6 | Comparison of the allosteric ligand binding mode in the crystal structures in absence (red) or presence of orthosteric ligands (blue-tones).** A-C) Structural overlay of crystal structures containing MRL-871 (A), FM26 (B) and compound 13 (C).

		distance C $\alpha$ Asn347-Gln484 (nm)				
allosteric ligands	CPD13	2.257 ML = 0.037	2.232 ML = 0.019	2.145 ML = 0.026	2.166 ML = 0.029	2.236 ML = 0.021
	FM26	2.245 ML = 0.021	2.211 ML = 0.015	2.194 ML = 0.017	2.214 ML = 0.023	2.159 ML = 0.023
	MRL-871	2.173 ML = 0.017	2.138 ML = 0.021	2.134 ML = 0.024	2.116 ML = 0.022	2.123 ML = 0.028
		2.173 ML = 0.017	2.138 ML = 0.021	2.134 ML = 0.024	2.116 ML = 0.022	2.123 ML = 0.028
		empty	20-OH	25-OH	DSM	CHL
		orthosteric ligands				

**Figure S4.7 | Distance (in nm) between the  $\alpha$ -carbons of Asn347 (helix 4) and Gln484 (helix 11) in the crystal structures.** The maximum-likelihood coordinate error (ML; in nm) is provided for every structure. The cartoon of ROR $\gamma$ t shows the positions of the  $\alpha$ -carbons of Asn347 and Gln484 as blue spheres.



**Figure S4.8 | A-C) Average root mean square fluctuation (RMSF) of the  $\alpha$ -carbons of ROR $\gamma$ t in complex with different orthosteric and allosteric ligands (MRL-871 (A), FM26 (B) and compound 13 (C)) derived from five simulations per complex. The red lines show the RMSF in absence of an orthosteric modulator. The secondary structure of the protein is represented as a rectangle, triangle and a line for  $\alpha$ -helices,  $\beta$ -sheets and loops, respectively.**



**Figure S4.9 | Distance between the  $\alpha$ -carbons of Asn347 and Gln484 in MD simulations. A-C) Bars represent the average distance between the  $\alpha$ -carbons of Asn347 and Gln484 (Figure S4.7) over five independent simulations with the individual values represented as black spheres and the error bars showing the standard deviation.**

**Table S4.1 | IC<sub>50</sub> and Hill slope values from the competitive TR-FRET coactivator recruitment assays with fixed concentrations of orthosteric ligands.** Abbreviations: 2 $\alpha$ -hydroxycholesterol (2 $\alpha$ -OH), 25-hydroxycholesterol (25-OH), desmosterol (DSM) and cholesterol (CHL).

orthosteric ligand conc. (μM)	allosteric ligand						
	MRL-871		FM26		compound 13		
	IC <sub>50</sub> (nM)	Hill slope	IC <sub>50</sub> (nM)	Hill slope	IC <sub>50</sub> (nM)	Hill slope	
0.00	10.1 ± 0.7	-0.95 ± 0.06	296 ± 34	-0.60 ± 0.04	514 ± 72	-1.10 ± 0.13	2α-OH
0.25	7.8 ± 0.3	-1.34 ± 0.07	55 ± 3	-1.13 ± 0.07	210 ± 19	-1.34 ± 0.14	
1.00	6.4 ± 0.2	-1.30 ± 0.05	79 ± 4	-1.23 ± 0.07	228 ± 24	-1.13 ± 0.11	
0.00	11.6 ± 0.6	-0.98 ± 0.04	249 ± 28	-0.65 ± 0.04	629 ± 173	-0.84 ± 0.12	25-OH
0.25	7.5 ± 0.3	-1.37 ± 0.06	57 ± 2	-1.03 ± 0.04	155 ± 12	-1.16 ± 0.09	
1.00	5.2 ± 0.2	-1.36 ± 0.06	60 ± 4	-1.02 ± 0.06	131 ± 13	-1.09 ± 0.11	
0.00	10.2 ± 0.6	-0.93 ± 0.04	343 ± 35	-0.70 ± 0.05	466 ± 49	-0.98 ± 0.08	DSM
0.25	7.5 ± 0.3	-1.13 ± 0.04	80 ± 4	-0.86 ± 0.04	130 ± 9	-1.06 ± 0.07	
1.00	5.0 ± 0.2	-1.05 ± 0.04	76 ± 4	-1.02 ± 0.04	148 ± 11	-1.21 ± 0.10	
0.00	12.7 ± 0.6	-0.97 ± 0.04	248 ± 18	-0.77 ± 0.04	547 ± 60	-0.74 ± 0.06	CHL
0.25	9.4 ± 0.3	-1.04 ± 0.03	138 ± 6	-0.86 ± 0.03	300 ± 18	-0.87 ± 0.04	
1.00	7.8 ± 0.2	-1.20 ± 0.03	94 ± 3	-1.01 ± 0.03	269 ± 19	-0.90 ± 0.05	

**Table S4.2 | Data collection and refinement statistics for the crystal structure of RORyt in complex with allosteric ligand MRL-871 and different orthosteric ligands.** Abbreviations: 20 $\alpha$ -hydroxy-cholesterol (20-OH), 25-hydroxycholesterol (25-OH), desmosterol (DSM) and cholesterol (CHL).

RORyt with:	20-OH+MRL-871	25-OH+MRL-871	DSM+MRL-871	CHL+MRL-871
<i>Data collection</i>				
Space group	P 6 <sub>1</sub> 2 2	P 6 <sub>1</sub> 2 2	P 6 <sub>1</sub> 2 2	P 6 <sub>1</sub> 2 2
Cell dimensions				
a, b, c (Å)	108.5, 108.5, 105.9	108.6, 108.6, 107.5	108.3, 108.3, 108.5	108.0, 108.0, 107.4
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	54.26 – 2.00 (2.07 – 2.00)	93.97 – 1.95 (2.02 – 1.95)	48.46 – 1.89 (1.95 – 1.89)	46.75 – 1.84 (1.91 – 1.84)
<i>I</i> / $\sigma$ ( <i>I</i> )	10.1 (0.3)	28.4 (1.6)	35.7 (1.7)	29.4 (1.2)
Completeness(%)	97.9 (82.7)	100.0 (100.0)	100.0 (100.0)	98.8 (92.6)
Redundancy	27.6 (8.2)	38.9 (38.0)	39.1 (40.4)	32.8 (13.7)
CC <sub>1/2</sub>	0.985 (0.551)	1.000 (0.741)	1.000 (0.777)	1.000 (0.527)
<i>Refinement</i>				
No. reflections	24921	27179	30693	32217
R <sub>work</sub> /R <sub>free</sub>	0.192/0.228	0.186/0.215	0.185/0.206	0.191/0.214
No. atoms				
Protein	1978	2040	2175	2053
Ligand/ion	66	66	71	65
Water	64	47	71	113
B-factors				
Protein	56.31	61.26	59.92	52.32
Ligand/ion	46.63	50.67	52.19	45.29
Water	54.58	57.05	56.28	54.27
R.m.s. deviations				
Bond lengths (Å)	0.007	0.014	0.014	0.015
Bond angles (°)	0.82	1.71	1.76	1.94
PDB ID	6T4U	6T4Y	6T4K	6T4I

**Table S4.3 | Data collection and refinement statistics for the crystal structure of ROR $\gamma$ t in complex with allosteric ligand FM26 and different orthosteric ligands.** Abbreviations: 20 $\alpha$ -hydroxy-cholesterol (20-OH), 25-hydroxycholesterol (25-OH), desmosterol (DSM) and cholesterol (CHL).

ROR $\gamma$ t with:	20-OH+FM26	25-OH+FM26	DSM+FM26	CHL+FM26
<i>Data collection</i>				
Space group	P 6 <sub>1</sub> 2 2	P 6 <sub>1</sub> 2 2	P 6 <sub>1</sub> 2 2	P 6 <sub>1</sub> 2 2
Cell dimensions				
a, b, c (Å)	108.5, 108.5, 99.3	108.9, 108.9, 98.5	108.9, 108.9, 98.5	108.5, 108.5, 105.0
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	54.25 – 1.62 (1.68 – 1.62)	94.29 – 1.48 (1.53 – 1.48)	47.66 – 1.79 (1.85 – 1.79)	48.50 – 1.93 (2.00 – 1.93)
<i>I</i> / $\sigma$ ( <i>I</i> )	23.0 (2.7)	25.5 (1.9)	22.7 (1.8)	28.0 (1.8)
Completeness(%)	99.6 (97.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Redundancy	137.1 (73.2)	39.1 (39.5)	39.0 (40.1)	39.0 (40.0)
CC <sub>1/2</sub>	1.000 (0.977)	1.000 (0.757)	1.000 (0.689)	1.000 (0.848)
<i>Refinement</i>				
No. reflections	44254	30639	32984	28750
R <sub>work</sub> /R <sub>free</sub>	0.148/0.189	0.175/0.185	0.175/0.197	0.178/0.213
No. atoms				
Protein	2053	2055	2046	2030
Ligand/ion	73	73	66	66
Water	198	255	186	105
B-factors				
Protein	29.44	28.79	36.46	52.30
Ligand/ion	26.11	24.03	32.77	45.89
Water	42.05	42.38	48.35	53.92
R.m.s. deviations				
Bond lengths (Å)	0.024	0.017	0.009	0.016
Bond angles (°)	1.850	1.790	1.190	1.890
PDB ID	6T4T	6T4X	6T4J	6T4G

**Table S4.4 | Data collection and refinement statistics for the crystal structure of RORyt in complex with allosteric ligand MRL-871 and different orthosteric ligands.** Abbreviations: 20 $\alpha$ -hydroxy-cholesterol (20-OH), 25-hydroxycholesterol (25-OH), desmosterol (DSM), cholesterol (CHL) and compound 13 (CPD13).

RORyt with:	20-OH+CPD13	25-OH+CPD13	DSM+CPD13	CHL+CPD13
<i>Data collection</i>				
Space group	P 6 <sub>1</sub> 2 2	P 6 <sub>1</sub> 2 2	P 6 <sub>1</sub> 2 2	P 6 <sub>1</sub> 2 2
Cell dimensions				
a, b, c (Å)	108.3, 108.3, 99.3	108.3, 108.3, 108.5	108.7, 108.7, 104.7	108.9, 108.9, 98.6
$\alpha, \beta, \gamma$ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	54.16 – 1.71 (1.77 – 1.71)	48.46 – 1.87 (1.94 – 1.87)	94.16 – 2.11 (2.16 – 2.10)	98.64 – 1.75 (1.78 – 1.75)
<i>I</i> / $\sigma$ ( <i>I</i> )	31.0 (0.9)	35.8 (1.6)	10.2 (0.5)	8.7 (0.5)
Completeness(%)	99.9 (99.6)	100.0 (100.0)	99.4 (92.0)	100.0 (99.8)
Redundancy	37.4 (37.3)	39.1 (39.4)	37.6 (38.8)	37.3 (35.7)
CC <sub>1/2</sub>	0.991 (0.219)	1.000 (0.799)	0.999 (0.414)	1.000 (0.297)
<i>Refinement</i>				
No. reflections	37646	31584	21797	35249
R <sub>work</sub> /R <sub>free</sub>	0.160/0.200	0.186/0.211	0.196/0.235	0.178/0.212
No. atoms				
Protein	2016	2040	2005	2042
Ligand/ion	69	63	62	74
Water	176	241	17	181
B-factors				
Protein	34.44	55.11	66.01	35.61
Ligand/ion	31.76	42.73	58.22	39.38
Water	4.48	53.75	56.45	46.39
R.m.s. deviations				
Bond lengths (Å)	0.013	0.015	0.016	0.017
Bond angles (°)	1.82	1.78	2.10	1.93
PDB ID	6T4W	6T50	6TLT	6TLQ



## References

- Meijer, F. A., Leijten-van de Gevel, I. A., de Vries, R. M. J. M. & Brunsveld, L. Allosteric small molecule modulators of nuclear receptors. *Mol. Cell. Endocrinol.* **485**, 20–34 (2019).
- van Westen, G. J. P., Gaulton, A. & Overington, J. P. Chemical, Target, and Bioactive Properties of Allosteric Modulation. *PLoS Comput. Biol.* **10**, e1003559 (2014).
- Changeux, J. P. & Christopoulos, A. Allosteric Modulation as a Unifying Mechanism for Receptor Function and Regulation. *Cell* **166**, 1084–1102 (2016).
- Thal, D. M., Glukhova, A., Sexton, P. M. & Christopoulos, A. Structural insights into G-protein-coupled receptor allostery. *Nature* **559**, 45–53 (2018).
- Wu, P., Clausen, M. H. & Nielsen, T. E. Allosteric small-molecule kinase inhibitors. *Pharmacol. Ther.* **156**, 59–68 (2015).
- Lewis, J. A., Lebois, E. P. & Lindsley, C. W. Allosteric modulation of kinases and GPCRs: design principles and structural diversity. *Curr. Opin. Chem. Biol.* **12**, 269–280 (2008).
- Jacobsen, S. E., Gether, U. & Bräuner-Osborne, H. Investigating the molecular mechanism of positive and negative allosteric modulators in the calcium-sensing receptor dimer. *Sci. Rep.* **7**, 46355 (2017).
- Capdeville, R., Buchdunger, E., Zimmermann, J. & Matter, A. Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug. *Nat. Rev. Drug Discov.* **1**, 493–502 (2002).
- Whitty, A. Cooperativity and biological complexity. *Nat. Chem. Biol.* **4**, 435–439 (2008).
- Stefan, M. I. & Le Novère, N. Cooperative Binding. *PLoS Comput. Biol.* **9**, e1003106 (2013).
- Cattoni, D. I., Chara, O., Kaufman, S. B. & González Flecha, F. L. Cooperativity in Binding Processes: New Insights from Phenomenological Modeling. *PLoS One* **10**, e0146043–e0146043 (2015).
- Wootten, D., Christopoulos, A. & Sexton, P. M. Emerging paradigms in GPCR allostery: Implications for drug discovery. *Nat. Rev. Drug Discov.* **12**, 630–644 (2013).
- Lu, J. *et al.* Structural basis for the cooperative allosteric activation of the free fatty acid receptor GPR40. *Nat. Struct. Mol. Biol.* **24**, 570–577 (2017).
- Cui, Q. & Karplus, M. Allostery and cooperativity revisited. *Protein Sci.* **17**, 1295–1307 (2008).
- Shang, J. *et al.* Cooperative cobinding of synthetic and natural ligands to the nuclear receptor PPAR $\gamma$ . *Elife* **7**, (2018).
- Putcha, B.-D. K., Wright, E., Brunzelle, J. S. & Fernandez, E. J. Structural basis for negative cooperativity within agonist-bound TR:RXR heterodimers. *Proc. Natl. Acad. Sci.* **109**, 6084–6087 (2012).
- Darimont, B. D. *et al.* Structure and specificity of nuclear receptor-coactivator interactions. *Genes Dev.* **12**, 3343–3356 (1998).
- Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 (2017).
- Ivanov, I. I. *et al.* The Orphan Nuclear Receptor ROR $\gamma$ t Directs the Differentiation Program of Proinflammatory IL-17+ T Helper Cells. *Cell* **126**, 1121–1133 (2006).
- Fauber, B. P. & Magnuson, S. Modulators of the nuclear receptor retinoic acid receptor-related orphan receptor- $\gamma$  (ROR $\gamma$  or RORc). *J. Med. Chem.* **57**, 5871–5792 (2014).
- Solt, L. A. & Burris, T. P. Action of RORs and their ligands in (patho)physiology. *Trends Endocrinol. Metab.* **23**, 619–627 (2012).
- Bronner, S. M., Zbieg, J. R. & Crawford, J. J. ROR $\gamma$  antagonists and inverse agonists: a patent review. *Expert Opin. Ther. Pat.* **27**, 101–112 (2017).
- Pandya, V. B., Kumar, S., Sachchidanand, Sharma, R. & Desai, R. C. Combating Autoimmune Diseases With ROR $\gamma$  Inhibitors: Hits and Misses. *J. Med. Chem.* **61**, 10976–10995 (2018).
- Cyr, P., Bronner, S. M. & Crawford, J. J. Recent progress on nuclear receptor ROR $\gamma$  modulators. *Bioorganic Med. Chem. Lett.* **26**, 4387–4393 (2016).
- Hu, X. *et al.* Sterol metabolism controls TH17 differentiation by generating endogenous ROR $\gamma$  agonists. *Nat. Chem. Biol.* **11**, 141–147 (2015).
- Kumar, N. *et al.* The benzenesulfoamide To901317 is a novel retinoic acid receptor-related orphan receptor- $\alpha$ /gamma inverse agonist. *Mol. Pharmacol.* **77**, 228–236 (2010).
- Huh, J. R. *et al.* Digoxin and its derivatives suppress TH17 cell differentiation by antagonizing ROR $\gamma$ t activity. *Nature* **472**, 486–490 (2011).
- Karstens, W. F. J. *et al.* RORgammaT Inhibitors. *PCT Int. Appl.* WO2012/106995 (2012).
- Scheepstra, M. *et al.* Identification of an allosteric binding site for ROR $\gamma$ t inhibition. *Nat. Commun.* **6**, e8833 (2015).
- Ouvry, G. *et al.* Discovery of phenoxyindazoles and phenylthioindazoles as ROR $\gamma$  inverse agonists. *Bioorganic Med. Chem. Lett.* **26**, 5802–5808 (2016).

31. Fauber, B. P. *et al.* Discovery of imidazo[1,5-a]pyridines and -pyrimidines as potent and selective ROR $\gamma$  inverse agonists. *Bioorganic Med. Chem. Lett.* **25**, 2907–2912 (2015).
32. Meijer, F. A. *et al.* Ligand-Based Design of Allosteric Retinoic Acid Receptor-Related Orphan Receptor  $\gamma$ t (ROR $\gamma$ t) Inverse Agonists. *J. Med. Chem.* **63**, 241–259 (2020).
33. Chaudari, S. S. *et al.* Bicyclic Heterocyclic Compounds as ROR gamma Modulators. *PCT Int. Appl. WO2015/008234* (2015).
34. de Vries, R. M. J. M., Meijer, F. A., Doveston, R. G. & Brunsvel, L. Elucidation of an Allosteric Mode of Action for a Thienopyrazole ROR $\gamma$ t Inverse Agonist. *ChemMedChem* **15**, 561–565 (2020).
35. Jiang, X. *et al.* A novel series of cysteine-dependent, allosteric inverse agonists of the nuclear receptor ROR $\gamma$ t. *Bioorg. Med. Chem. Lett.* **30**, e126967 (2020).
36. Jin, L. *et al.* Structural Basis for Hydroxycholesterols as Natural Ligands of Orphan Nuclear Receptor ROR $\gamma$ . *Mol. Endocrinol.* **24**, 923–929 (2010).
37. Gege, C. Retinoid-related orphan receptor  $\gamma$  t modulators: comparison of Glenmark's me-too patent application (WO2015008234) with the originator application from Merck Sharp and Dohme (WO2012106995). *Expert Opin. Ther. Pat.* **25**, 1215–1221 (2015).
38. DeSantis, K. A. & Reinking, J. L. Use of Differential Scanning Fluorimetry to Identify Nuclear Receptor Ligands. *Methods Mol. Biol.* **1443**, 21–30 (2016).
39. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212–2221 (2007).
40. Gao, K., Oerlemans, R. & Groves, M. R. Theory and applications of differential scanning fluorimetry in early-stage drug discovery. *Biophys. Rev.* **12**, 85–104 (2020).
41. Degorce, F. HTRF: A Technology Tailored for Drug Discovery - A Review of Theoretical Aspects and Recent Applications. *Curr. Chem. Genomics* **3**, 22–32 (2009).
42. Kojetin, D. J. *et al.* Structural mechanism for signal transduction in RXR nuclear receptor heterodimers. *Nat. Commun.* **6**, 8013 (2015).
43. Leijten-van de Gevel, I. A. & Brunsvel, L. Delineation of the molecular determinants of the unique allosteric binding site of the orphan nuclear receptor ROR $\gamma$ t. *J. Biol. Chem.* **295**, 9183–9191 (2020).
44. Potterton, L. *et al.* CCP412: The new graphical user interface to the CCP 4 program suite. *Acta Crystallogr. D Struct. Biol.* **74**, 68–84 (2018).
45. Clabbers, M. T. B. *et al.* Electron diffraction data processing with DIALS. *Acta Crystallogr. D Struct. Biol.* **74**, 506–518 (2018).
46. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
47. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D. Struct. Biol.* **63**, 32–41 (2007).
48. Long, F. *et al.* AceDRG : a stereochemical description generator for ligands. *Acta Crystallogr. D Struct. Biol.* **73**, 112–122 (2017).
49. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Struct. Biol.* **67**, 355–67 (2011).
50. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Struct. Biol.* **66**, 486–501 (2010).
51. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 352–367 (2012).
52. The PyMOL Molecular Graphics System, Version 2.2.3. (Schrödinger LLC, 2015).
53. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
54. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
55. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general Amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
56. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
57. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 14101 (2007).
58. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
59. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
60. Holden, Z. C., Richard, R. M. & Herbert, J. M. Periodic boundary conditions for QM/MM calculations: Ewald summation for extended Gaussian basis sets. *J. Chem. Phys.* **139**, 244108 (2013).

# Chapter 5

## Covalent Occlusion of the RORyt Ligand Binding Pocket Allows Unambiguous Targeting of an Allosteric Site

### Abstract

Inhibition of RORyt with small molecules holds great potential as a therapeutic strategy for autoimmune diseases. RORyt has a unique allosteric ligand binding site in the ligand binding domain which is distinct from the canonical, orthosteric binding site. Allosteric modulation of RORyt shows high potential, but the targeted discovery of novel allosteric ligands is highly challenging via currently available methods. Here, we introduce covalent orthosteric probes for RORyt with the aim of occluding the binding of canonical, orthosteric ligands and allow for the selective screening of allosteric ligands. The reference compound **GW9662** was found to ligate to the native Cys320 residue in the orthosteric site of RORyt, however it completely inhibited coactivator binding. A library of **GW9662** derivatives was designed and screened, which resulted in the discovery of a set of chemical probes that act as partial inverse agonists, not significantly affecting coactivator recruitment. The probes demonstrated complete occlusion of orthosteric ligand binding, while allosteric ligands could still induce activity on RORyt. Ultimately, these covalent probes could be used to underpin screening approaches for the unambiguous and rapid identification of novel allosteric RORyt ligands.

**This chapter has been published as:** F.A. Meijer<sup>#</sup>, M.C.M. van den Oetelaar<sup>#</sup>, R.G. Doveston, E.N.R. Sampers & L. Brunsveld. Covalent Occlusion of the RORyt Ligand Binding Pocket Allows Unambiguous Targeting of an Allosteric Site. *ACS Med. Chem. Lett.* 12, 631-639 (2021).

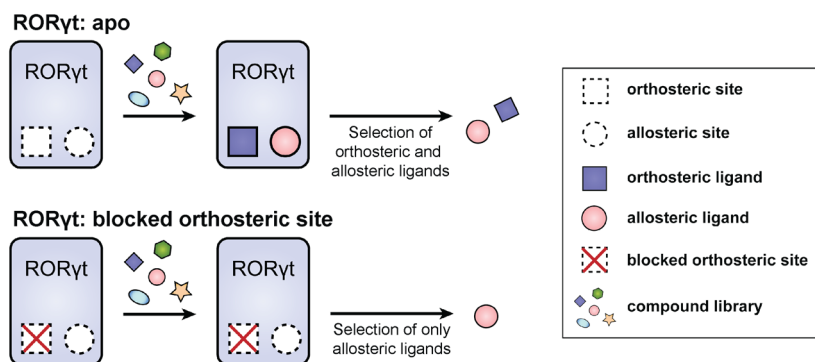
<sup>#</sup> These authors contributed equally

## Introduction

ROR $\gamma$ t is a nuclear receptor (NR) that plays an important regulatory role in the immune system via the Th17/IL-17a pathway.<sup>1-3</sup> Inhibition of ROR $\gamma$ t by inverse agonists has been shown to be a promising strategy for the treatment of autoimmune diseases and thus has been the focus of several drug discovery programs.<sup>4,5</sup> The majority of ROR $\gamma$ t ligands bind to a highly conserved canonical binding pocket, termed the orthosteric binding site, within the ligand binding domain (LBD) of ROR $\gamma$ t (Figure 5.2B).<sup>6-10</sup> ROR $\gamma$ t appears to be transcriptionally active, even in the absence of an agonist ligand.<sup>11</sup> However, agonists (e.g. cholesterol and its derivatives) that bind to the orthosteric site enhance ROR $\gamma$ t transcriptional activity further by stabilizing the active conformation of helix 12 (H12) in a way that promotes the recruitment of transcriptional coactivators.<sup>12</sup> ROR $\gamma$ t inverse agonists binding to this site (e.g. digoxin) destabilize the active conformation of H12, inhibiting coactivator recruitment.<sup>13</sup>

The conserved nature of the ROR $\gamma$ t orthosteric binding pocket presents some challenges for drug discovery, for example because of competition with endogenous ligands.<sup>14</sup> Most interestingly, and unique among the NR family, ROR $\gamma$ t has a second binding site, termed allosteric binding site (located at a topographically distinct place in the LBD) (Figure 5.2B).<sup>15</sup> This allosteric site offers ample opportunities for innovative NR drug discovery. The indazole **MRL-871**, thiazole **compound 13** (Glenmark), and isoxazole **FM26** are examples of highly potent ROR $\gamma$ t inverse agonists that bind to this allosteric site.<sup>15-19</sup> These ligands exert their effect via a reorientation of H12 into a conformation that prevents coactivator binding.<sup>15,18,19</sup>

Despite the potential of NR allosteric inverse agonists, the number of examples and their chemical diversity have remained rather limited.<sup>15-17,19-22</sup> Furthermore, understanding of the structure activity relationships (SARs) and scaffold diversity of allosteric ROR $\gamma$ t inverse agonists is of importance to tune the potency, selectivity and pharmacokinetic profiles of potential pharmaceutical lead molecules. However, unambiguous screening for allosteric ligands is challenging, since both orthosteric and allosteric ligands can show an inhibitory response on ROR $\gamma$ t (Figure 5.1, top). Currently, discrimination between orthosteric and allosteric ligands is not trivial and the orthosteric pocket can host a plethora of chemically diverse compounds. Illustrative of the above, the mode of action for **MRL-871** was only described in retrospect.<sup>15</sup> Even in a targeted program, the methods currently available require a series of biophysical and structural experiments to discriminate between orthosteric and allosteric inverse agonism.<sup>15</sup> Therefore, the development of a molecular approach for the specific identification of allosteric ligands is highly desirable.

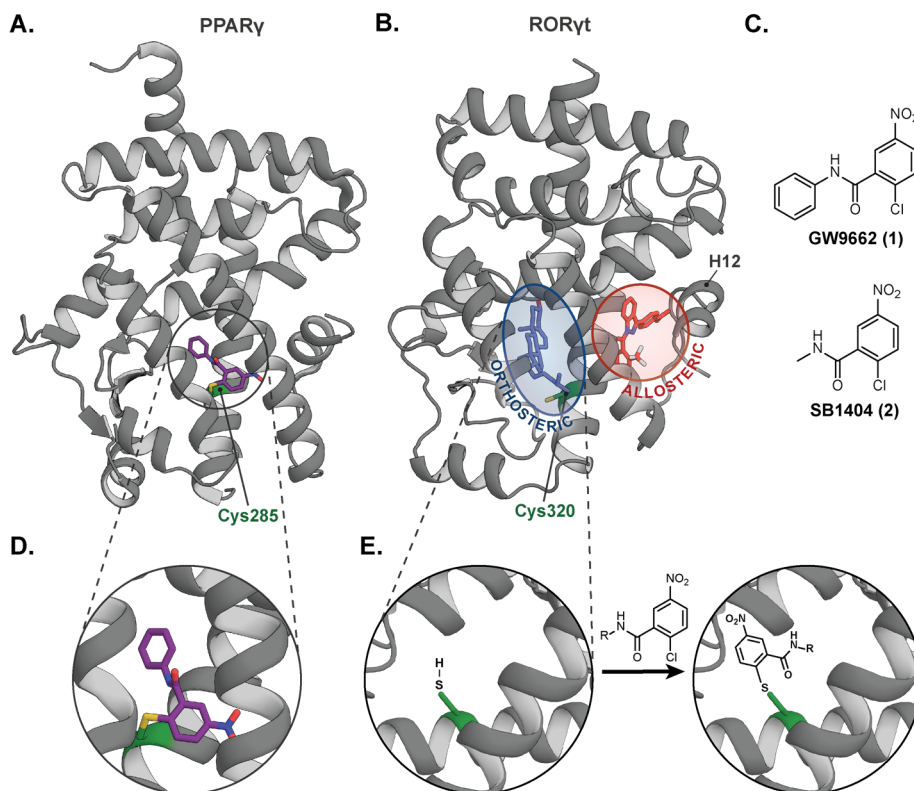


**Figure 5.1 | Schematic representation of the effect of ROR $\gamma$  orthosteric site occlusion.** When both binding sites are available, both orthosteric and allosteric ligands will be identified in screening assays. When the orthosteric binding site is occluded, allosteric ligands can be unambiguously identified.

A potential strategy to achieve the goal defined above would be via a competitive binding assay, whereby the displacement of a well-characterized allosteric probe ligand can be monitored, such as the previously reported AlexaFluor647-labelled MRL-871 ligand.<sup>15,19</sup> However, competition assays are notoriously problematic for identification of the weaker binding initial chemical entities. Additionally, orthosteric inverse agonists could potentially lower the allosteric probe affinity, which would lead to false positive results that diminish the reliability of the assay. We therefore postulated that using a molecular probe to occlude the ROR $\gamma$  orthosteric binding site, in combination with an established coactivator recruitment assay, would provide a more robust approach (Figure 5.1, bottom). An important feature of the prospective orthosteric probe is that it should have minimal effect on the characteristics of the allosteric pocket and on ROR $\gamma$  coactivator binding.

Our approach took inspiration from studies of the structurally related NR peroxisome proliferator-activated receptor  $\gamma$  (PPAR $\gamma$ ). PPAR $\gamma$  contains a cysteine residue within the orthosteric ligand binding pocket (Cys285) (Figure 5.2A,D), which has previously been targeted by covalent ligands.<sup>23–25</sup> In particular, covalent modification of Cys285 by the electron deficient aryl chlorides **GW9662** (**1**) and **SB1404** (**2**) (via a nucleophilic aromatic substitution reaction) (Figure 5.2C) occlude the binding, and thus also the activity, of certain PPAR $\gamma$  agonists.<sup>24–28</sup> ROR $\gamma$  contains an analogous cysteine residue within the orthosteric ligand binding pocket (Cys320 (ROR $\gamma$  numbering), see Figure 5.2B and Supporting Figure S5.1). Therefore, we were interested to determine if this residue could also be targeted with covalent molecular probes (Figure 5.2E) that would block ligand binding to the orthosteric site without

significantly affecting coactivator binding. Examples of ROR $\gamma$ t covalent inverse agonists have been reported, but they, thus far, target other cysteine residues that are not in proximity to the orthosteric binding pocket.<sup>22</sup>



**Figure 5.2 | Covalent attachment of a chemical probe to a cysteine residue in the orthosteric binding site of PPAR $\gamma$  and ROR $\gamma$ t.** **A)** Crystal structure of PPAR $\gamma$  ligated to GW9662 (purple) at the Cys285 residue (green) in the orthosteric binding site (PDB: 3BoR). **B)** Crystal structure of ROR $\gamma$ t (PDB: 6T4I) showing the orthosteric (blue) and allosteric (red) ligand binding site. The Cys320 residue in the orthosteric site (analogous to the Cys285 residue in PPAR $\gamma$ ) is shown in green. **C)** Chemical structures of PPAR $\gamma$  covalent ligands GW9662 (1) and SB1404 (2). **D)** Enlarged view of GW9662 ligated to PPAR $\gamma$ . **E)** The thiol of Cys320 in ROR $\gamma$ t is postulated to ligate to the electron-deficient aryl ring of GW9662 or derivatives.

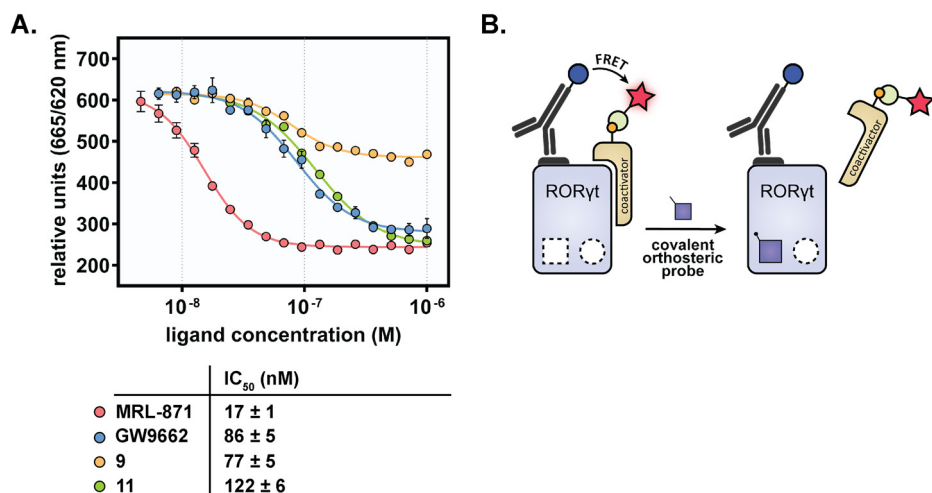
In this chapter, we describe the design, synthesis and biochemical effect of a series of covalent ligands for ROR $\gamma$ t that occlude orthosteric ligand binding. The evaluation of a library of GW9662 derivatives resulted in the discovery of a set of covalent probes that showed full ligation to the Cys320 residue in the orthosteric site of ROR $\gamma$ t, and acted as partial inverse

agonists. Time-resolved FRET (TR-FRET) coactivator recruitment assays and thermal shift assays (TSAs) demonstrated complete occlusion of the orthosteric binding site, while the allosteric site was still accessible for the binding of allosteric ligands. Ultimately, these orthosteric covalent probes could be used as tool compounds in screening approaches, that will allow for the rapid identification of novel allosteric RORyt inverse agonists.

### GW9662 is an orthosteric covalent full inverse agonist for RORyt

The PPAR $\gamma$  covalent ligands **GW9662** and **SB1404** (Figure 5.2C) were used as starting points in the search for covalent probes for RORyt. The compounds were synthesized via an amide coupling reaction between the commercially available 2-chloro-5-nitrobenzoyl chloride and the appropriate amine (Supporting Scheme S5.1).<sup>24,27</sup> A single **GW9662** molecule fully ligated to RORyt (Supporting Figure S5.2A), following optimization of the ligation conditions<sup>24</sup> (see Supporting Table S5.1), as shown by quadrupole time-of-flight mass spectrometry (Q-TOF MS). In order to prove binding to the Cys320 residue, a RORyt Cys320Ala mutant was generated. No ligation of **GW9662** to this RORyt mutant was observed which verified the expected Cys320 ligation site (Supporting Figure S5.2B). Surprisingly, **SB1404**, which has a smaller methyl group instead of the phenyl substituent, did not show any ligation to RORyt, while it did fully ligate to PPAR $\gamma$ . This might be because **SB1404** has a lower affinity for the RORyt orthosteric binding pocket compared to **GW9662**. This observation highlights the difference between the binding pockets of the two NRs.

A TR-FRET coactivator recruitment assay<sup>29</sup> was used to investigate the effect of **GW9662** on the activity of RORyt, i.e. the ability of RORyt to recruit coactivators (see Figure 5.3B). The dose-response curve in Figure 5.3A shows inhibition of coactivator recruitment upon titration of **GW9662** to the protein (similar behavior to the full allosteric inverse agonist **MRL-871**), demonstrating that the probe acts as a full inverse agonist for RORyt with an IC<sub>50</sub> value of 86  $\pm$  5 nM. A possible reason for this inverse agonistic character could be derived from the docking pose, which shows a conformation of **GW9662** where the warhead (nitro-phenyl moiety) points toward H11 (Supporting Figure S5.3A), which could result in destabilization of the active conformation of H12, inhibiting coactivator recruitment.<sup>30</sup> Alignment of the docking pose of **GW9662** to the crystal structure of RORyt in complex with the known RORyt orthosteric inverse agonist **To901317**<sup>31,32</sup> shows that the nitro-moiety of **GW9662** and the CF<sub>3</sub>-groups of **To901317** have the same orientation and distance towards H11, which could explain the inverse agonistic behavior of **GW9662** (Supporting Figure S5.5).

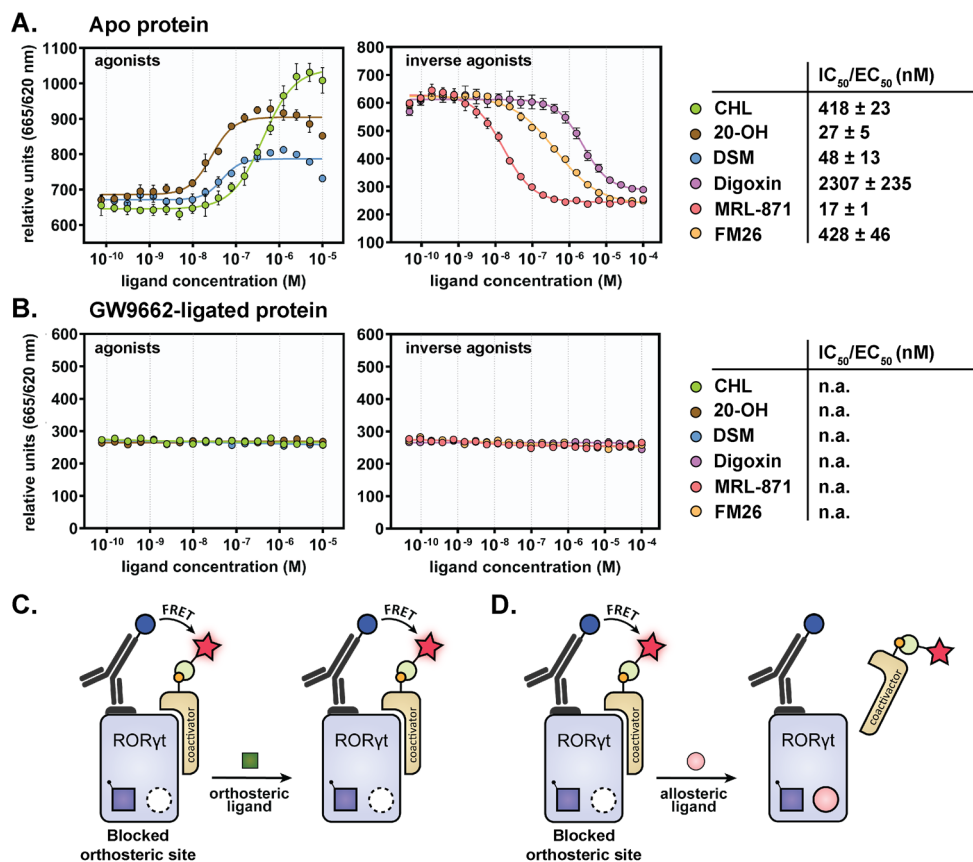


**Figure 5.3 | Dose-response curves from the TR-FRET coactivator recruitment assays on RORyt, including an overview of the IC<sub>50</sub> values. A)** Titration of different orthosteric covalent probes (GW9662, **9** and **11**) and MRL-871. Data are representative of two independent experiments (recorded in triplicate) and are presented as mean ± SD. **B)** Schematic representation of the TR-FRET coactivator recruitment assay. When the covalent orthosteric probe ligates to the protein, coactivator recruitment is blocked by its (partial) inverse agonistic character, resulting in a lower FRET pairing.

A ligand binding TR-FRET coactivator recruitment assay was used to determine if GW9662-ligation effectively occluded orthosteric ligand binding to RORyt, without affecting allosteric ligand binding (see Figure 5.4C,D). To probe this, three orthosteric agonists (cholesterol (CHL), 20 $\alpha$ -hydroxycholesterol (20-OH) and desmosterol (DSM)), an orthosteric inverse agonist (digoxin) and two allosteric inverse agonists (MRL-871 and FM26) were tested (see chemical structures in Chapter 4, Figure 4.1C).<sup>12,13,15,19</sup> For the apo RORyt protein, the cholesterol derivatives showed agonistic character, increasing coactivator recruitment to the LBD (Figure 5.4A), consistent with literature reports.<sup>2,12</sup> The orthosteric inverse agonist digoxin and the allosteric inverse agonists reduced coactivator recruitment in a dose-dependent manner as expected (Figure 5.4A).<sup>13,19</sup> These titration experiments were then repeated with the GW9662-ligated RORyt (Figure 5.4B). In this experiment, the orthosteric agonists did not cause any increase in coactivator recruitment to RORyt, indicating occlusion of the orthosteric binding site. However, the allosteric ligands MRL-871 and FM26 also did not demonstrate a change in coactivator recruitment, due to the full inverse agonistic behavior of GW9662. Because of its full inverse agonistic behavior, GW9662 is thus not suitable for the screening of allosteric inverse agonists. We therefore set out to identify a GW9662



analogue that would act either as true covalent antagonist (/agonist), or as a partial covalent inverse agonist, that would maintain some ROR $\gamma$ t sensitivity to allosteric ligand binding.



**Figure 5.4 | TR-FRET coactivator recruitment ligand binding assays with ROR $\gamma$ t (unligated and ligated) by titration of various orthosteric and allosteric ligands. A/B) Ligand binding assays for apo (unligated) ROR $\gamma$ t (A) and GW9662-ligated ROR $\gamma$ t (B). Abbreviations: n.a., not active. Data are representative of two independent experiments (recorded in triplicate) and are presented as mean ± SD. C/D) Schematic representation of the TR-FRET coactivator recruitment assay, using the protein ligated to an orthosteric chemical probe. Orthosteric ligand binding will be occluded, showing no effect on the initial coactivator recruitment capacity (C). Allosteric ligand binding will result in reduced coactivator binding and therefore a lower FRET pairing (D).**

## GW9662 SAR exploration leads to covalent partial inverse agonists

An initial library of twelve **GW9662** derivatives was synthesized (compounds **3-14**, Table 5.1), containing the same warhead, but with modifications in place of the phenyl moiety, to investigate its influence on the ligation efficiency and activity of ROR $\gamma$ t. These modifications were made with the aim of identifying a probe that ligates to the ROR $\gamma$ t Cys320, but does not behave as a full inverse agonist. The modifications were varied in terms of bulkiness,  $\pi$ - $\pi$  stacking capacity, aromaticity and substitution pattern. The compounds were synthesized in a similar manner to **GW9662** (Supporting Scheme S5.1B).<sup>24,33-35</sup> First, the ligation efficiency was explored for all derivatives via Q-TOF analysis of the reacted protein. Although **GW9662** fully ligated to the protein, its derivatives ligated with varying efficiency (Table 5.1) (ligation conditions are shown in Supporting Table S5.1). From the ligation data, it can be concluded that a ring system (preferably aromatic) is necessary at the phenyl position of **GW9662** in order to obtain full ligation to the protein. A methyl substitution on the phenyl ring is only fully tolerated at the ortho position (compound **9**), indicating that bulk at *meta/para* positions lowers the ligation efficiency (compound **7** and **8**). Furthermore, an electron-withdrawing substituent on the phenyl ring improves the ligation efficiency compared to an electron-donating group (compound **13** vs. **10**), most probably caused by the higher electrophilicity of the warhead. Interestingly, all probes (except for compounds **12** and **14**) also showed full covalent attachment to PPAR $\gamma$ , indicating that PPAR $\gamma$  shows less differentiation in the ligation of the compounds than ROR $\gamma$ t, which is probably due to the larger size of the PPAR $\gamma$  binding pocket.

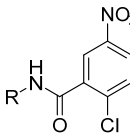
The compounds that fully ligated to ROR $\gamma$ t (**9** and **11**) were taken forward for evaluation of their binding behavior in a TR-FRET coactivator recruitment assay. As shown in Figure 5.3A, compound **11** showed the same full inverse agonistic behavior as **GW9662**, but was slightly less potent ( $IC_{50}$  value of  $122 \pm 6$  nM). More interestingly, in contrast to the other two probes, compound **9** showed partial inverse agonistic behavior, thus not completely blocking coactivator recruitment to the LBD, with 55% remaining activity (Figure 5.3A). The difference in activity between **GW9662** and compound **9** could be explained by an inversed binding conformation of **9** (with the warhead pointing towards the orthosteric pocket instead of towards H11) which was supported by an *in silico* docking experiment (Supporting Figure S5.3B).

Table 5.1 | Chemical structures of GW9662 derivatives and their % ligation to RORyt.





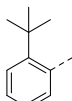

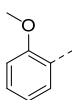





Compound	Chemical structure	RORyt ligation	Compound	Chemical structure	RORyt ligation
<b>1</b> (GW9662)		100%	<b>8</b>		75-90%
<b>2</b> (SB1404)		0%	<b>9</b>		100%
<b>3</b>		25-50%	<b>10</b>		0-25%
<b>4</b>		25-50%	<b>11</b>		100%
<b>5</b>		75-90%	<b>12</b>		0%
<b>6</b>		0%	<b>13</b>		50-75%
<b>7</b>		25-50%	<b>14</b>		0%

A focused SAR study was performed around compound **9** to obtain more insight into the desirable partial behavior of this probe. A library of nine compound **9** derivatives (compounds **15-23**, Table 5.2) was designed, by varying the size, polarity, and electron density of the ortho-substituent. The probes were synthesized as described above (Supporting Scheme S5.1B). All the probes from this set of derivatives fully ligated to RORyt (Table 5.2), except for compound **16** that has a *tert*-butyl substituent. Although an ethyl substituent and a naphthalene moiety appeared to be tolerated (**15** and **21**), the *tert*-butyl moiety of **16** is likely too bulky for optimal binding. Furthermore, the polarity of the substituent (**22** and **23**) did not affect the ligation behavior and even a double-ortho methyl substitution was tolerated (**19**).

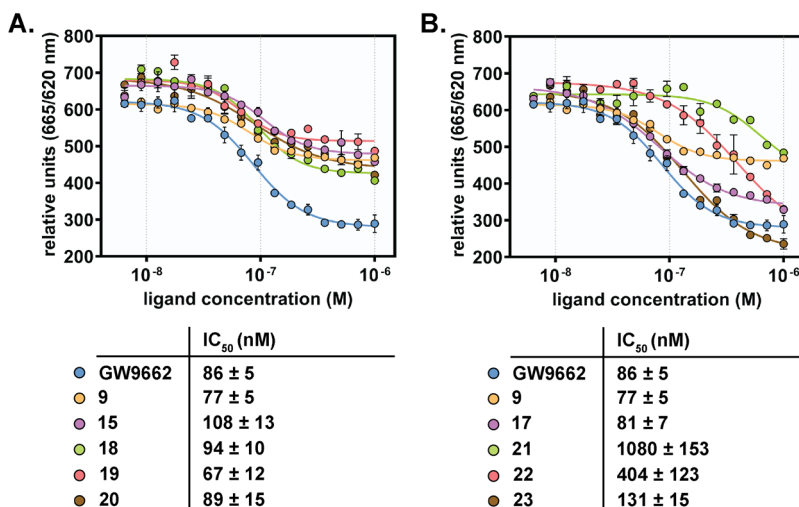
**Table 5.2 | Chemical structures of GW9662 derivatives, their % ligation tot ROR $\gamma$ t, TR-FRET IC<sub>50</sub> values (nM) and full/partial inverse agonistic character.** Abbreviations: n.d., not determined. TR-FRET data are representative of two independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD.



**1, 9, 11, 15-23**

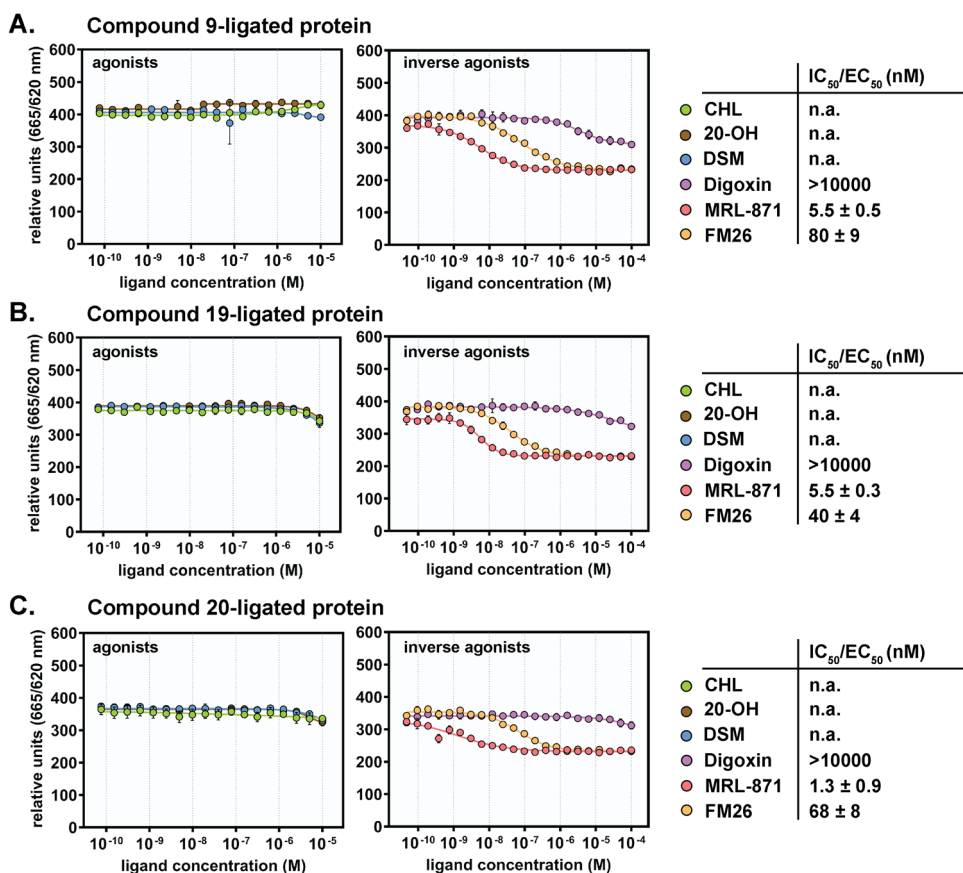
Compound	R	ROR $\gamma$ t ligation	IC <sub>50</sub> (nM)	Profile
<b>1</b> (GW9662)		100%	86 $\pm$ 5	Full inverse agonist
<b>9</b>		100%	77 $\pm$ 5	Partial inverse agonist 55% remaining activity
<b>11</b>		100%	122 $\pm$ 6	Full inverse agonist
<b>15</b>		100%	108 $\pm$ 13	Partial inverse agonist 51% remaining activity
<b>16</b>		50-75%	n.d.	n.d.
<b>17</b>		100%	81 $\pm$ 7	Partial inverse agonist 12% remaining activity
<b>18</b>		100%	94 $\pm$ 10	Partial inverse agonist 36% remaining activity
<b>19</b>		100%	67 $\pm$ 12	Partial inverse agonist 61% remaining activity
<b>20</b>		100%	89 $\pm$ 15	Partial inverse agonist 41% remaining activity
<b>21</b>		100%	1080 $\pm$ 153	n.d.
<b>22</b>		100%	404 $\pm$ 123	Full inverse agonist
<b>23</b>		100%	131 $\pm$ 15	Full inverse agonist

All probes (except **16**) were tested in the TR-FRET coactivator recruitment assay and showed varying inverse agonistic behavior as shown in Figure 5.5. A partial inverse agonistic behavior, similar to compound **9**, was observed for probes **15** and **18** (increased size of the substituent), **19** (bis-ortho substitution) and **20** (electron-deficient substituent) (Figure 5.5A). Compound **17** with a fluoro substituent also demonstrated partial behavior, however inducing a greater decrease in coactivator recruitment than the previous four probes (Figure 5.5B), which might be due to its size being more comparable to **GW9662**. Compound **23** (hydroxyl modification) resulted in full inverse agonism, similar to **GW9662**, and **22** (amine modification) also approached the bottom plateau, but with a lower potency than **23** (Figure 5.5B). Compound **21**, with the larger naphthalene substituent, showed a significantly lower potency than the other probes ( $IC_{50} = 1080 \pm 153$  nM) (Figure 5.5B), not reaching the bottom plateau, which makes it hard to confidently characterize it as a partial or full inverse agonist.



**Figure 5.5 |** Dose-response curves from the TR-FRET coactivator recruitment assays on ROR $\gamma$ t, including an overview of the  $IC_{50}$  values. A/B) Titration of covalent probes **GW9662**, **9**, **15**, **18**, **19** and **20** (partial inverse agonists) (A), and **GW9662**, **9**, **17**, **21**, **22** and **23** (partial/full inverse agonists) (B). Data are representative of two independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD.

The modifications with a similar size to the methyl group in compound **9** (e.g. ethyl, methoxy, bis-methyl, trifluoromethyl) appear to result in a partial inverse agonistic behavior, while smaller or polar substituents (e.g. fluoro, hydroxyl, amine) show a full inverse agonistic character. A structural explanation could be derived from the docking studies, where the partial inverse agonists generally show a docking pose similar to compound **9** (warhead pointing toward the orthosteric site), while for the full inverse agonists a docking pose similar to **GW9662** was observed. Combined, compound **9** and four derivatives were found to be covalent partial inverse agonists and suitable candidates as covalent orthosteric probes.



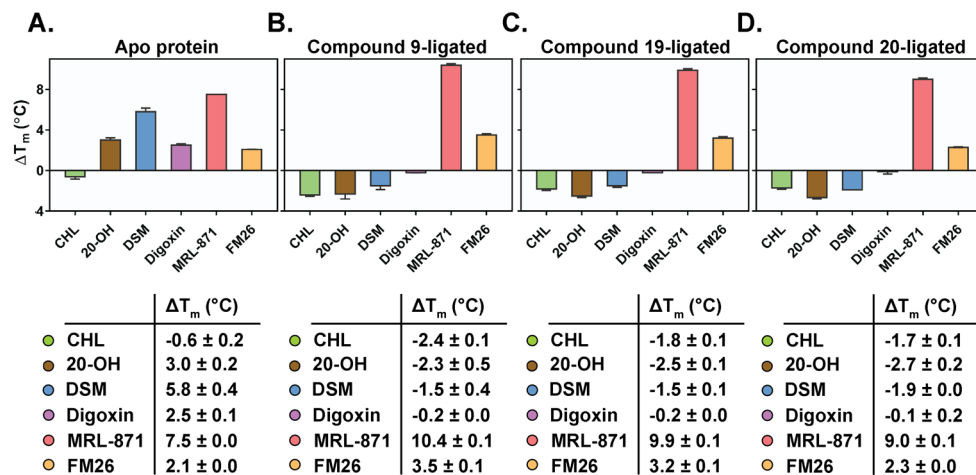
**Figure 5.6 | TR-FRET coactivator recruitment ligand binding assays with ROR $\gamma$ t (ligated) by titration of various orthosteric and allosteric ligands. A-C) Ligand binding assays for compound **9**-ligated protein (A), compound **19**-ligated protein (B) and compound **20**-ligated protein (C). Abbreviations: n.a., not active. Data are representative of two independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD.**

### Covalent partial inverse agonists effectively block orthosteric ligand binding

The TR-FRET ligand binding assay was used to investigate the combination of occlusion of the orthosteric site and potential for allosteric binding. With the partial inverse agonist probes, a screening window is expected to be preserved to detect allosteric inverse agonist binding. Clear occlusion of the orthosteric site for orthosteric agonist binding was observed for RORyt ligated to **9**, **19** and **20** (Figure 5.6A-C), in comparison to the apo protein (Figure 5.4A). In addition, the orthosteric inverse agonist digoxin was also ineffective. In contrast, and most importantly, the allosteric inverse agonists **MRL-871** and **FM26** still induced a clear inverse agonistic response on RORyt (Figure 5.6A-C). Also, the potency of **MRL-871** and **FM26** was not negatively affected by the occlusion of the orthosteric site, revealing that the partial inverse agonism of the covalent probes translates to responsiveness of the allosteric site for ligand binding. The other partial inverse agonistic probes **15** and **18** showed similar occlusion of the orthosteric site while still allowing allosteric binding, albeit with a smaller assay window (Supporting Figure S5.4). Interestingly, in all cases, the  $IC_{50}$  values for the allosteric ligands **MRL-871** and **FM26** were actually decreased to different extents (3-fold and 13-fold, respectively) in comparison to the data for the apo protein (Figure 5.6A-C, Figure 5.4A). This increased potency of the allosteric ligands in the presence of an orthosteric probe is likely caused by a cooperative effect between both binding sites, as observed in previous studies.<sup>18,19,36</sup>

A thermal shift assay (TSA) was used as an orthogonal method to confirm occlusion of orthosteric ligand binding by the most interesting probes (**9**, **19** and **20**) without affecting allosteric modulation. With the apo RORyt protein, all orthosteric ligands showed a significant thermal stabilization ( $\Delta T_m$  between 2.5 and 5.8 °C), except for cholesterol (Figure 5.7A). In contrast, for RORyt ligated to **9**, **19** and **20**, a thermal stabilization effect was not observed anymore for these orthosteric ligands and they even caused a small destabilization effect ( $\Delta T_m$  between -0.1 and -2.7 °C) (Figure 5.7B-D). These results again demonstrate orthosteric site occlusion by the covalent probes. The allosteric ligands **MRL-871** and **FM26** showed a moderate to high thermal stabilization effect for the apo protein ( $\Delta T_m$  of 7.5 and 2.1 °C, respectively) (Figure 5.7A). When RORyt was ligated to a covalent probe, this thermal stabilization effect by the allosteric ligands was preserved and even higher  $\Delta T_m$  values were observed in all cases (Figure 5.7B-D). These results again prove that allosteric binding is still functional with the orthosteric site blocked and that the affinity is even slightly enhanced. The increased stabilization by the allosteric ligands in the presence of the orthosteric covalent probes is consistent with the cooperative behavior observed in the TR-FRET ligand binding

assays. Although the observed cooperative behavior might question the usefulness of the probes at first sight, it will not be an issue from a screening perspective, since it will not lead to false hits but will just enhance an allosteric  $IC_{50}$  value by a limited degree.



**Figure 5.7 | Thermal shift assay (TSA) data for apo and ligated RORyt in the presence of various orthosteric and allosteric ligands.** A) Apo RORyt. B) Compound 9-ligated RORyt. C) Compound 19-ligated RORyt. D) Compound 20-ligated RORyt. The shift in melting temperature ( $\Delta T_m$  in °C), relative to DMSO, is shown. Data are representative of two independent experiments (recorded in triplicate) and are presented as mean ± SD.

## Conclusions

Although allosteric RORyt inverse agonists have high potential for NR drug discovery, the number of examples and chemical diversity have remained limited. An enlargement of the allosteric RORyt ligand library is therefore essential in order to enhance understanding of the SAR and to tune potency and selectivity. Unambiguous screening for allosteric ligands is challenging, since both orthosteric and allosteric ligands will result in an inhibitory response on the protein, and discrimination between them is not trivial.

In this chapter, a method for occlusion of the RORyt orthosteric binding site was introduced, via the ligation of covalent chemical probes to a native cysteine residue. This allows for the unambiguous targeting of the allosteric binding site, which has the potential to facilitate the rapid identification of allosteric inverse agonists. The reference compound **GW9662** showed full ligation to Cys320 of RORyt, but acted as a full inverse agonist and completely inhibited coactivator binding, preventing the detection of allosteric inverse



agonists. From a small library of **GW9662** derivatives, the methyl-substituted compound **9** was identified as a covalent partial inverse agonist. Further SAR studies around compound **9** resulted in the discovery of four additional covalent probes with a partial inverse agonistic character, for which **19** and **20** appeared to be the most promising probes (containing a bis-ortho methyl and trifluoromethyl modification). The partial character of these probes can most probably be explained by an inversed binding conformation compared to **GW9662** as was supported by docking experiments. Co-crystallization attempts of the probes with RORyt were unsuccessful, but could provide more structural evidence. TR-FRET and thermal shift assays revealed complete occlusion of the orthosteric binding site with the covalent probes, while allosteric ligand binding was not inhibited and even occurred with enhanced affinity. This cooperative behavior of the orthosteric inverse agonistic covalent probes with the allosteric ligands is an interesting observation, since these cooperative effects had previously only been observed with orthosteric agonists. The covalent probes are excellent tools that could underpin an assay format that unambiguously screens for allosteric RORyt modulators. Additionally, these covalent orthosteric ligands could be used as inspiration for the development of covalent orthosteric inverse agonists for RORyt, for which future studies could focus on the efficacy and toxicity of covalent RORyt targeting.

## Acknowledgements

Maxime van den Oetelaar performed the major part of the experimental work described in this chapter, for which she is greatly acknowledged. Richard Doveston is acknowledged for the initial work he performed on this research and for the valuable discussions. Ella Sampers is kindly thanked for the synthesis of the first library of compounds and for performing initial biochemical assays.

This chapter was reprinted (adapted) with permission from *ACS Med. Chem. Lett.* 2021, 12, 631-639. Copyright 2021 American Chemical Society. Further permissions related to this material should be directed to the ACS.

## Experimental Section

**Molecular docking studies.** The molecular modeling environment Maestro (version 11.1, Schrödinger LLC) was used with no changes to the default parameters throughout. The crystal structure (PDB: 3LoL) was prepared for the docking simulation using the Protein Preparation Wizard. A receptor grid was generated using the Receptor Grid Generation tool. The examined ligands were drawn in ChemDraw and prepared using the Ligand Preparation tool. The prepared ligands were then docked into the generated receptor grid using the Covalent Docking tool (nucleophilic substitution with Cys320 as reactive residue) in standard precision mode with flexible ligand sampling, to obtain the corresponding docking poses and Glide Scores.

**General chemistry.** All solvents were supplied by Biosolve and used without further purification. Dry solvent was obtained from a MBRAUN Solvent Purification System (MB-SPS-800). Water was purified by a Millipore purification train. Deuterated solvents were obtained from Cambridge Isotope Laboratories. Solvents were removed *in vacuo* using a Büchi rotary evaporator and a diaphragm pump. All reagents were commercially available and were supplied by Sigma-Aldrich, TCI Chemicals, Fluorochem and Iris Biochem GmbH. Proton ( $^1\text{H}$ ) NMR (400 MHz), carbon ( $^{13}\text{C}$ ) NMR (100 MHz) and 2D NMR (400 MHz) were recorded on a Bruker Avance 400 MHz spectrometer. Proton spectra are referenced to tetramethyl silane (TMS). Carbon spectra are referenced to TMS or the solvent peak of the deuterated spectrum. NMR spectra are reported as follows: chemical shift ( $\delta$ ) in parts per million (ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublet, td = triplet of doublets), coupling constant (J) in Hertz (Hz) (if applicable) and integration (proton spectra only). Peak assignments are based on additional 2D NMR techniques (COSY, HMBC, HSQC). Analytical Liquid Chromatography coupled with Mass Spectrometry (LC-MS) was performed on a C<sub>4</sub> Jupiter SuC4300A 150 x 2.0 mm column using ultrapure water with 0.1% formic acid (FA) and acetonitrile with 0.1% FA, in general with a gradient of 5% to 100% acetonitrile over 10 min, connected to a Thermo Fisher LCQ Fleet Ion Trap Mass Spectrometer. The purity of the samples was assessed using a UV detector at 254 nm. Unless otherwise stated all final compounds were >95% pure as judged by HPLC. High resolution mass spectra (HRMS) were recorded using a Waters ACQUITY UPLC I-Class LC system coupled to a Xevo G2 Quadrupole Time of Flight (Q-TOF) mass spectrometer. Manual column chromatography was performed using silica gel with a particle size of 60-200  $\mu\text{m}$  (60 Å). For manual chromatography, solid loading was used. Reaction progress was monitored by thin-layer liquid chromatography (TLC) using Merck TLC silica gel 60 F254 plates. Visualization of the plates was achieved using an ultraviolet lamp ( $\lambda_{\text{max}} = 254 \text{ nm}$ ).

## Synthetic procedures

**General Procedure for Amide Coupling** (all compounds, except **GW9662**, **SB1404**, **3**, **4**, **11**, **13**, **14**, **22** and **23**). The amine (2.95 mmol) was dissolved in dry pyridine (2.5 mL) and cooled to 0 °C under argon. 2-Chloro-5-nitrobenzoyl chloride (2.27 mmol, 500 mg) was added to the reaction mixture. The mixture was stirred overnight at room temperature. Subsequently, the mixture was poured into ice and the resulting solid was filtered, washed with water and freeze dried to obtain the product.<sup>33</sup>

**2-chloro-5-nitro-N-phenylbenzamide (1, GW9662).** 2-Chloro-5-nitrobenzoyl chloride (1.00 g, 5.56 mmol) was dissolved in a solution of CH<sub>2</sub>Cl<sub>2</sub> (25 mL) and triethylamine (0.7 mL, 5.00 mmol) at 0 °C under argon. Aniline (0.44 mL, 4.81 mmol) was added dropwise. The reaction mixture was stirred for 5 min at 0 °C followed by 15 min at room temperature. The solution was diluted with EtOAc and washed with 1 M HCl, H<sub>2</sub>O, 1 M NaHCO<sub>3</sub> and brine. The combined organic phase was dried over MgSO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude mixture was purified by recrystallization from EtOAc, furnishing the product (1105 mg, 85%). <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 8.61 (1H, d), 8.25 (1H, dd), 7.84 (1H, s), 7.75-7.57 (3H, m), 7.41 (2H, t), 7.22 (1H, t); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 162.32, 146.75, 137.68, 137.01, 136.68, 131.73, 129.41, 126.09, 125.68, 125.41, 120.48. LC-MS (ESI): calc. for C<sub>13</sub>H<sub>9</sub>ClN<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 277.04, observed 277.17 (R<sub>t</sub> = 6.29 min). HRMS (ESI): calc. for C<sub>13</sub>H<sub>9</sub>ClN<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 277.0380, observed: 277.0372. Data was consistent with that previously reported.<sup>24</sup>

**2-chloro-5-nitro-N-methylbenzamide (2, SB1404).** 2-Chloro-5-nitrobenzoylchloride (500 mg, 2.27 mmol) was dissolved in a solution of CH<sub>2</sub>Cl<sub>2</sub> (10 mL) and triethylamine (0.79 mL, 5.68 mmol) at 0 °C. A solution of methylamine (2.0 M in THF, 2.95 mmol, 1.5 mL) was added dropwise. The reaction mixture was stirred overnight at room temperature under argon. The reaction mixture was diluted with CH<sub>2</sub>Cl<sub>2</sub> and washed with NH<sub>4</sub>Cl and H<sub>2</sub>O. The combined organic phase was dried over MgSO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude solid was purified by column chromatography, eluting with 100% CH<sub>2</sub>Cl<sub>2</sub> to furnish the product (38 mg, 8%). <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 8.62 (1H, s), 8.25 (2H, d, *J* = 9.5 Hz), 7.82 (1H, d, *J* = 8.5 Hz), 2.80 (3H, d, *J* = 4.6 Hz); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 164.68, 146.00, 137.84, 137.06, 131.29, 125.35, 123.70, 26.03. LC-MS (ESI): calc. for C<sub>8</sub>H<sub>7</sub>ClN<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 215.02, observed 215.00 (R<sub>t</sub> = 3.35 min). HRMS (ESI): calc. for C<sub>8</sub>H<sub>7</sub>ClN<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 215.0223, observed: 215.0221. Data was consistent with that previously reported.<sup>27</sup>

**2-chloro-5-nitro-N-isopropylbenzamide (3).** 2-Chloro-5-nitrobenzoylchloride (250 mg, 1.24 mmol) was dissolved in THF (8 mL) at 0 °C. Isopropylamine (0.061 mL, 1.49 mmol) and triethylamine (0.1 mL, 1.49 mmol) were added dropwise. The reaction mixture was stirred overnight at room temperature under argon. Saturated NaHCO<sub>3</sub> was added to the reaction mixture and it was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x). The combined organic phase was washed with aqueous saturated NaCl, dried over MgSO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude solid was purified by recrystallization from 1:1 water/EtOH, resulting in the product (105 mg, 35%). <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 8.45 (1H, d, *J* = 2.7 Hz), 8.17 (1H, dd, *J* = 8.8, 2.7 Hz), 7.57 (1H, d, *J* = 8.8 Hz), 5.99 (1H, s), 4.37-4.25 (1H, m), 1.30 (6H, d, *J* = 6.6 Hz); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 163.40, 146.54, 137.48, 136.92, 131.33, 125.45, 125.04, 42.71, 29.71, 22.56. LC-MS (ESI): calc. for C<sub>10</sub>H<sub>11</sub>ClN<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 243.06, observed 243.08 (R<sub>t</sub> = 4.73 min). HRMS (ESI): calc. for C<sub>10</sub>H<sub>11</sub>ClN<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 243.0536, observed: 243.0540.<sup>27</sup>

**N-allyl-2-chloro-5-nitrobenzamide (4).** Allylamine (1.61 mmol) was dissolved in dry pyridine (1.4 mL) and the reaction mixture was cooled to 0 °C under argon. 2-Chloro-5-nitrobenzoyl chloride (1.24 mmol, 250 mg) was added to the reaction mixture. The mixture was stirred overnight at room temperature. Subsequently, the mixture was poured into ice. The resulting solid was filtered, washed with water and freeze dried to furnish the product (32 mg, 11%). <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) 8.51 (1H, d, *J* = 2.7 Hz), 8.21 (1H, dd, *J* = 8.8, 2.7 Hz), 7.59 (1H, d, *J* = 8.8 Hz), 6.30 (1H, s), 6.01-5.88 (1H, m), 5.37-5.20

(2H, m), 4.16-4.09 (2H, m);  $^{13}\text{C}$ -NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 164.05, 146.57, 137.51, 136.37, 133.05, 131.47, 125.71, 125.29, 117.52, 42.73. LC-MS (ESI): calc. for  $\text{C}_{10}\text{H}_9\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 241.04, observed 241.00 ( $R_t = 4.55$  min). HRMS (ESI): calc. for  $\text{C}_{10}\text{H}_9\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 241.0380, observed: 241.0378.<sup>33</sup>

**2-chloro-5-nitro-*N*-cyclohexylbenzamide (5).** According to the General Procedure for amide coupling, cyclohexylamine (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the product (69 mg, 11%).  $^1\text{H}$ -NMR (400 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 8.55 (1H, d,  $J = 7.7$  Hz), 8.25 (1H, dd,  $J = 8.8$ , 2.8 Hz), 8.18 (1H, d,  $J = 2.7$  Hz), 7.82 (1H, d,  $J = 8.8$  Hz), 3.80-3.66 (1H, m), 1.90-1.79 (2H, m), 1.78-1.65 (2H, m), 1.62-1.52 (1H, m), 1.40-1.07 (5H, m);  $^{13}\text{C}$ -NMR (100 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 163.79, 146.45, 138.70, 137.53, 131.64, 125.61, 123.93, 48.78, 32.55, 25.63, 24.95. LC-MS (ESI): calc. for  $\text{C}_{13}\text{H}_{15}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 283.09, observed 283.08 ( $R_t = 6.01$  min). HRMS (ESI): calc. for  $\text{C}_{13}\text{H}_{15}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 283.0849, observed: 283.0852.

***N*-benzyl-2-chloro-5-nitrobenzamide (6).** According to the General Procedure for amide coupling, benzylamine (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the product (140 mg, 21%).  $^1\text{H}$ -NMR (400 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 9.22 (1H, t,  $J = 5.8$  Hz), 8.33-8.23 (2H, m), 7.86-7.80 (1H, m), 7.44-7.32 (4H, m), 7.32-7.23 (1H, m), 4.49 (2H, d,  $J = 6.0$  Hz);  $^{13}\text{C}$ -NMR (100 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 164.86, 146.49, 139.09, 138.13, 137.56, 131.80, 128.85, 127.84, 127.47, 125.93, 124.16, 43.15. LC-MS (ESI): calc. for  $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 291.06, observed 291.08 ( $R_t = 6.47$  min). HRMS (ESI): calc. for  $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 291.0536, observed: 291.0523.

**2-chloro-5-nitro-*N*-(*p*-tolyl)benzamide (7).** According to the General Procedure for amide coupling, *p*-toluidine (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the product (580 mg, 88%).  $^1\text{H}$ -NMR (400 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 10.61 (1H, s), 8.44 (1H, d,  $J = 2.6$  Hz), 8.32 (1H, dd,  $J = 8.8$ , 2.6 Hz), 7.89 (1H, d,  $J = 8.8$  Hz), 7.58 (2H, d,  $J = 8.2$  Hz), 7.19 (2H, d,  $J = 8.2$  Hz), 2.29 (3H, s);  $^{13}\text{C}$ -NMR (100 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 163.01, 146.60, 138.30, 137.57, 136.48, 133.71, 131.79, 129.69, 126.08, 124.31, 120.23. LC-MS (ESI): calc. for  $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 291.06, observed 291.08 ( $R_t = 6.22$  min). HRMS (ESI): calc. for  $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 291.0536, observed: 291.0532.

**2-chloro-5-nitro-*N*-(*m*-tolyl)benzamide (8).** According to the General Procedure for amide coupling, *m*-toluidine (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the product (620 mg, 94%).  $^1\text{H}$ -NMR (400 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 10.62 (1H, s), 8.43 (1H, d,  $J = 2.7$  Hz), 8.33 (1H, dd,  $J = 8.8$ , 2.8 Hz), 7.90 (1H, d,  $J = 8.8$  Hz), 7.56 (1H, s), 7.48 (1H, d,  $J = 8.3$  Hz), 7.25 (1H, t,  $J = 7.8$  Hz), 6.98 (1H, d,  $J = 7.5$  Hz), 2.32 (3H, s);  $^{13}\text{C}$ -NMR (100 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 163.18, 146.61, 138.88, 138.59, 138.27, 137.55, 131.80, 129.17, 126.12, 125.39, 124.28, 120.73, 117.45, 21.66. LC-MS (ESI): calc. for  $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 291.06, observed 291.08 ( $R_t = 6.24$  min). HRMS (ESI): calc. for  $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 291.0536, observed: 291.0529.

**2-chloro-5-nitro-*N*-(*o*-tolyl)benzamide (9).** According to the General Procedure for amide coupling, *o*-toluidine (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the product (421 mg, 64%).  $^1\text{H}$ -NMR (400 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 10.18 (1H, s), 8.47 (1H, d,  $J = 2.7$  Hz), 8.33 (1H, dd,  $J = 8.8$ , 2.8 Hz), 7.89 (1H, d,  $J = 8.8$  Hz), 7.50 (1H, d,  $J = 7.9$  Hz), 7.30-7.14 (3H, m), 2.30 (3H, s);  $^{13}\text{C}$ -NMR (100 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  (ppm) 163.58, 146.62, 138.43, 137.53, 135.78, 133.31, 131.78, 130.93, 126.68, 126.55, 126.37, 126.03, 124.31, 39.99, 18.43. LC-MS (ESI): calc. for  $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 291.06, observed 291.08 ( $R_t = 5.82$  min). HRMS (ESI): calc. for  $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 291.0536, observed: 291.0537.

**2-chloro-5-nitro-*N*-(4-methoxyphenyl)benzamide (10).** According to the General Procedure for amide coupling, *p*-anisidine (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the

product (661 mg, 95%).  $^1\text{H-NMR}$  (400 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 10.56 (1H, s), 8.43 (1H, d,  $J$  = 2.6 Hz), 8.34 (1H, dd,  $J$  = 8.8, 2.7 Hz), 7.89 (1H, d,  $J$  = 8.8 Hz), 7.61 (2H, d,  $J$  = 8.9 Hz), 6.96 (2H, d,  $J$  = 8.9 Hz), 3.75 (3H, s);  $^{13}\text{C-NMR}$  (100 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 162.78, 156.35, 146.60, 138.34, 137.59, 132.07, 131.79, 126.05, 124.30, 121.79, 114.45, 55.71. LC-MS (ESI): calc. for  $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 307.05, observed 307.17 ( $R_t$  = 5.78 min). HRMS (ESI): calc. for  $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 307.0486, observed: 307.0478.

**2-chloro-5-nitro-*N*-(thiophen-3-yl)benzamide (11).** Thiophen-3-amine hydrochloride (1.61 mmol) was dissolved in dry pyridine (1.3 mL) and the reaction mixture was cooled to 0°C under argon. 2-Chloro-5-nitrobenzoyl chloride (1.24 mmol, 250 mg) was added to the reaction mixture. The mixture was stirred overnight at room temperature. Subsequently, the mixture was poured into ice. The resulting solid was filtered, washed with water and freeze dried to furnish the product (243 mg, 70%).  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 11.09 (1H, s), 8.44 (1H, d,  $J$  = 2.7 Hz), 8.35 (1H, dd,  $J$  = 8.8 Hz), 7.90 (1H, d,  $J$  = 8.8 Hz), 7.70 (1H, dd,  $J$  = 3.2, 1.3 Hz), 7.56-7.49 (1H, m), 7.18 (1H, dd,  $J$  = 5.2, 1.3 Hz);  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) 162.25, 146.59, 137.73, 137.67, 136.53, 131.86, 126.26, 125.62, 124.40, 122.00, 110.65. LC-MS (ESI): calc. for  $\text{C}_{11}\text{H}_7\text{ClN}_2\text{O}_3\text{S}$   $[\text{M}+\text{H}]^+$ : 283.00, observed 283.00 ( $R_t$  = 5.91 min). HRMS (ESI): calc. for  $\text{C}_{11}\text{H}_7\text{ClN}_2\text{O}_3\text{S}$   $[\text{M}+\text{H}]^+$ : 282.9944, observed: 282.9944.<sup>33</sup>

**5-chloro-2-nitro-*N*-phenylbenzamide (12).** First, 5-chloro-2-nitrobenzoyl chloride was synthesized by dissolving 5-chloro-2-nitrobenzoic acid (457 mg, 2.27 mmol) in thionylchloride (8.3 mL). The reaction mixture was concentrated *in vacuo*, obtaining 5-chloro-2-nitrobenzoyl chloride (100% yield, 500 mg). According to the General Procedure for amide coupling, aniline (2.95 mmol) was reacted with 5-chloro-2-nitrobenzoyl chloride to furnish the product (545 mg, 87%).  $^1\text{H-NMR}$  (400 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 10.71 (1H, s), 8.20 (1H, d,  $J$  = 8.8 Hz), 7.95 (1H, d,  $J$  = 2.3 Hz), 7.86 (1H, dd,  $J$  = 8.8, 2.3 Hz), 7.68-7.62 (2H, m), 7.40-7.33 (2H, m), 7.17-7.09 (1H, m);  $^{13}\text{C-NMR}$  (100 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 162.99, 145.47, 139.22, 139.07, 134.73, 131.22, 129.63, 129.33, 126.82, 124.59, 120.17. LC-MS (ESI): calc. for  $\text{C}_{13}\text{H}_9\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 277.04, observed 277.00 ( $R_t$  = 5.89 min). HRMS (ESI): calc. for  $\text{C}_{13}\text{H}_9\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 277.0380, observed: 277.0368.

**2-chloro-5-nitro-*N*-(4-nitrophenyl)benzamide (13).** 2-Chloro-5-nitro-*N*-(4-nitrophenyl) benzamide was synthesized as described in the patent of Amemiya et al.<sup>34</sup> 4-Nitroaniline (263 mg, 1.90 mmol) was dissolved in DMA (5 mL). 2-Chloro-5-nitrobenzoylchloride (500 mg, 2.27 mmol) was added and the reaction mixture was stirred overnight at room temperature. Ethyl acetate (5 mL) and saturated aqueous sodium bicarbonate (30 mL) were added and the mixture was stirred for 1 h.  $\text{H}_2\text{O}$  (10 mL) was added and the reaction mixture was stirred again for 1 h. The resulting crystals were filtered and freeze dried to furnish the product (502 mg, 82%).  $^1\text{H-NMR}$  (400 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 11.31 (1H, s), 8.56 (1H, d,  $J$  = 2.7 Hz), 8.36 (1H, dd,  $J$  = 8.8, 2.8 Hz), 8.28 (2H, d,  $J$  = 9.2 Hz), 7.99-7.87 (3H, m);  $^{13}\text{C-NMR}$  (100 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 164.04, 146.63, 145.35, 143.31, 137.62, 137.51, 131.89, 126.52, 125.49, 124.57, 120.18. LC-MS (ESI): calc. for  $\text{C}_{13}\text{H}_8\text{ClN}_3\text{O}_5$   $[\text{M}+\text{H}]^+$ : 322.02, observed 322.17 ( $R_t$  = 4.70 min). HRMS (ESI): calc. for  $\text{C}_{13}\text{H}_8\text{ClN}_3\text{O}_5$   $[\text{M}+\text{H}]^+$ : 322.0231, observed: 322.0227.

**(2-chloro-5-nitrophenyl)(1H-pyrrol-1-yl)methanone (14).** (2-Chloro-5-nitrophenyl)(1H-pyrrol-1-yl) methanone was synthesized as described by D'Silva et al.<sup>35</sup> Triethylamine (0.32 mL, 2.30 mmol) and pyrrole (0.21 mL, 3.10 mmol) were dissolved in  $\text{CH}_2\text{Cl}_2$  (1.23 mL, 19.2 mmol). DMAP (27.7 mg, 0.227 mmol) and 2-chloro-5-nitrobenzoyl chloride (500 mg, 2.27 mmol) were then added to the reaction mixture and it was stirred overnight at room temperature. The reaction mixture was diluted in  $\text{CH}_2\text{Cl}_2$  and washed with  $\text{NaHCO}_3$  and  $\text{H}_2\text{O}$ , dried over  $\text{MgSO}_4$ , filtered and concentrated *in vacuo*. The resulting solid was freeze dried to furnish the product (406 mg, 72%).  $^1\text{H-NMR}$  (400 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 8.65 (1H, d,  $J$  = 2.7 Hz), 8.43 (1H, dd,  $J$  = 8.9, 2.7 Hz), 7.97 (1H, d,  $J$  = 8.9 Hz), 7.24 (2H, s), 6.42 (2H, t,

$J = 4.0$  Hz);  $^{13}\text{C}$ -NMR (100 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 163.30, 146.92, 137.20, 134.68, 131.98, 127.42, 125.03, 121.08, 114.98. LC-MS (ESI): calc. for  $\text{C}_{11}\text{H}_7\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 251.02, observed 251.00 ( $R_t = 4.75$  min). HRMS (ESI): calc. for  $\text{C}_{11}\text{H}_7\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 251.0223, observed: 251.0231.

**2-chloro-5-nitro-*N*-(2-ethylphenyl)benzamide (15).** According to the General Procedure for amide coupling, 2-ethylaniline (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the product (573 mg, 83%).  $^1\text{H}$ -NMR (400 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 10.18 (1H, s), 8.46 (1H, d,  $J = 2.8$  Hz), 8.33 (1H, dd,  $J = 8.8, 2.8$  Hz), 7.89 (1H, d,  $J = 8.9$  Hz), 7.51-7.48 (1H, m), 7.33-7.21 (3H, m), 2.68 (2H, q,  $J = 7.5$  Hz), 1.17 (3H, t,  $J = 7.5$  Hz);  $^{13}\text{C}$ -NMR (100 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 164.04, 146.63, 139.48, 138.49, 137.51, 135.07, 131.80, 129.22, 127.29, 127.16, 126.57, 126.02, 124.25, 24.31, 14.99. LC-MS (ESI): calc. for  $\text{C}_{15}\text{H}_{13}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 305.74, observed 305.25 ( $R_t = 6.27$  min). HRMS (ESI): calc. for  $\text{C}_{15}\text{H}_{13}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 305.0693, observed: 305.0680.

**2-chloro-5-nitro-*N*-(2-*tert*-butylphenyl)benzamide (16).** According to the General Procedure for amide coupling, 2-(*tert*-butyl)aniline (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the product (593 mg, 71%).  $^1\text{H}$ -NMR (400 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 10.16 (1H, s), 8.39 (1H, d,  $J = 2.5$  Hz), 8.34 (1H, dd,  $J = 8.7, 2.8$  Hz), 7.90 (1H, d,  $J = 8.8$  Hz), 7.51-7.42 (1H, m), 7.37-7.25 (3H, m), 1.39 (9H, s);  $^{13}\text{C}$ -NMR (100 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 164.86, 147.33, 146.54, 138.44, 137.69, 135.55, 132.16, 131.93, 128.08, 127.30, 127.15, 126.03, 123.86, 35.34, 31.31. LC-MS (ESI): calc. for  $\text{C}_{17}\text{H}_{17}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 333.79, observed 333.08 ( $R_t = 6.63$  min). HRMS (ESI): calc. for  $\text{C}_{17}\text{H}_{17}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 333.1006, observed: 333.0995.

**2-chloro-5-nitro-*N*-(2-fluorophenyl)benzamide (17).** According to the General Procedure for amide coupling, 2-fluoroaniline (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the product after recrystallization from ethyl acetate/hexane (50:50) (158 mg, 46%).  $^1\text{H}$ -NMR (400 MHz, Acetone- $d_6$ ):  $\delta$  (ppm) 9.68 (1H, s), 8.53 (1H, d,  $J = 2.8$  Hz), 8.35 (1H, dd,  $J = 8.8, 2.8$  Hz), 8.31-8.27 (1H, m), 7.84 (1H, d,  $J = 8.8$  Hz), 7.30-7.20 (3H, m);  $^{13}\text{C}$ -NMR (100 MHz, Acetone- $d_6$ ):  $\delta$  (ppm) 164.00, 155.92, 153.47, 147.46, 138.47, 132.19, 126.94, 126.83, 126.77, 126.75, 126.70, 126.67, 126.55, 125.34, 125.07, 124.54, 124.42, 116.29. LC-MS (ESI): calc. for  $\text{C}_{13}\text{H}_8\text{ClFN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 295.68, observed 295.17 ( $R_t = 5.96$  min). HRMS (ESI): calc. for  $\text{C}_{13}\text{H}_8\text{ClFN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 295.0286, observed: 295.0293.

**2-chloro-5-nitro-*N*-(2-methoxyphenyl)benzamide (18).** According to the General Procedure for amide coupling, 2-methoxyaniline (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the product after recrystallization from ethanol (212 mg, 30%).  $^1\text{H}$ -NMR (400 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 10.00 (1H, s), 8.38 (1H, d,  $J = 2.8$  Hz), 8.32 (1H, dd,  $J = 8.8, 2.8$  Hz), 7.98 (1H, dd,  $J = 7.9, 1.6$  Hz), 7.85 (1H, d,  $J = 8.8$  Hz), 7.19 (1H, t,  $J = 7.5$  Hz), 7.10 (1H, d,  $J = 8.3$  Hz), 6.99 (1H, t,  $J = 7.6$  Hz), 3.83 (3H, s);  $^{13}\text{C}$ -NMR (100 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 163.57, 151.22, 146.46, 138.24, 137.65, 131.66, 126.82, 126.29, 125.94, 124.47, 123.74, 120.73, 112.05, 56.22. LC-MS (ESI): calc. for  $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 307.71, observed 307.17 ( $R_t = 6.19$  min). HRMS (ESI): calc. for  $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_4$   $[\text{M}+\text{H}]^+$ : 307.0486, observed: 307.0487.

**2-chloro-5-nitro-*N*-(2,6-dimethylphenyl)benzamide (19).** According to the General Procedure for amide coupling, 2,6-dimethylaniline (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the product (504 mg, 73%).  $^1\text{H}$ -NMR (400 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 10.14 (1H, s), 8.38-8.33 (2H, m), 7.94-7.88 (1H, m), 7.18-7.10 (3H, m), 2.30 (s, 6H);  $^{13}\text{C}$ -NMR (100 MHz, DMSO- $d_6$ ):  $\delta$  (ppm) 162.93, 146.19, 137.90, 137.01, 135.36, 133.94, 131.55, 127.95, 127.10, 125.69, 123.50, 18.28. LC-MS (ESI): calc. for  $\text{C}_{15}\text{H}_{13}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 305.74, observed 305.25 ( $R_t = 6.09$  min). HRMS (ESI): calc. for  $\text{C}_{15}\text{H}_{13}\text{ClN}_2\text{O}_3$   $[\text{M}+\text{H}]^+$ : 305.0693, observed: 305.0682.

**2-chloro-5-nitro-N-(2-(trifluoromethyl)phenyl)benzamide (20).** According to the General Procedure for amide coupling, 2-(trifluoromethyl)aniline (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the product after recrystallization from ethanol (124 mg, 32%). <sup>1</sup>H-NMR (400 MHz, Acetone-*d*<sub>6</sub>):  $\delta$  (ppm) 9.52 (1H, s), 8.50 (1H, d, *J* = 2.7 Hz), 8.36 (1H, dd, *J* = 8.8, 2.7 Hz), 7.98 (1H, d, *J* = 8.0 Hz), 7.86 (1H, d, *J* = 8.9 Hz), 7.84-7.75 (2H, m), 7.55 (1H, t, *J* = 7.5 Hz); <sup>13</sup>C-NMR (100 MHz, Acetone-*d*<sub>6</sub>):  $\delta$  (ppm) 164.75, 147.46, 138.51, 138.32, 135.48, 133.93, 132.37, 130.25, 128.81, 128.00, 127.27, 126.68, 126.10, 125.78, 125.48, 124.87, 123.41. LC-MS (ESI): calc. for C<sub>14</sub>H<sub>8</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 345.68, observed 345.17 (*R*<sub>t</sub> = 6.29 min). HRMS (ESI): calc. for C<sub>14</sub>H<sub>8</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 345.0254, observed: 345.0246.

**2-chloro-5-nitro-N-(naphthalen-1-yl)benzamide (21).** According to the General Procedure for amide coupling, naphthalen-1-amine (2.95 mmol) was reacted with 2-chloro-5-nitrobenzoyl chloride to furnish the product after recrystallization from ethanol (199 mg, 54%). <sup>1</sup>H-NMR (400 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) 10.76 (1H, s), 8.62 (1H, d, *J* = 2.8 Hz), 8.37 (1H, dd, *J* = 8.7, 3.0 Hz), 8.20-8.14 (1H, m), 8.01-7.96 (1H, m), 7.93 (1H, d, *J* = 8.8 Hz), 7.88 (1H, d, *J* = 7.9 Hz), 7.83 (1H, d, *J* = 7.8 Hz), 7.63-7.54 (3H, m); <sup>13</sup>C-NMR (100 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) 163.93, 146.24, 138.03, 137.13, 133.76, 132.65, 131.32, 128.22, 128.17, 126.38, 126.23, 126.20, 125.66, 125.58, 124.04, 122.92, 122.80. LC-MS (ESI): calc. for C<sub>17</sub>H<sub>11</sub>ClN<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 327.75, observed 327.25 (*R*<sub>t</sub> = 6.36 min). HRMS (ESI): calc. for C<sub>17</sub>H<sub>11</sub>ClN<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 327.0536, observed: 327.0524.

**2-chloro-5-nitro-N-(2-aminophenyl)benzamide (22).** *N*-Boc-1,2-phenylenediamine (0.59 mmol) was dissolved in dry pyridine (13 mL) and the reaction mixture was cooled to 0 °C under argon. 2-Chloro-5-nitrobenzoyl chloride (0.45 mmol, 100 mg) was added to the reaction mixture. The mixture was stirred overnight at room temperature. Subsequently, the mixture was poured into ice. The resulting solid was filtered, washed with water and freeze dried to furnish the product which was recrystallized from methanol (9 mg, 11%). <sup>1</sup>H-NMR (400 MHz, Acetone-*d*<sub>6</sub>):  $\delta$  (ppm) 9.23 (1H, s), 8.56 (1H, d, *J* = 3.0 Hz), 8.34 (1H, dd, *J* = 8.8, 2.8 Hz), 7.84 (1H, d, *J* = 8.8 Hz), 7.44-7.38 (1H, m), 7.04 (1H, t, *J* = 7.6 Hz), 6.88 (1H, d, *J* = 8.0 Hz), 6.70 (1H, t, *J* = 7.6 Hz), 4.68 (2H, s); <sup>13</sup>C-NMR (100 MHz, Acetone-*d*<sub>6</sub>):  $\delta$  (ppm) 164.03, 147.51, 143.56, 139.08, 138.50, 132.17, 128.01, 126.83, 126.76, 126.32, 125.03, 123.83, 118.08, 117.55. LC-MS (ESI): calc. for C<sub>13</sub>H<sub>10</sub>ClN<sub>3</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 292.70, observed 292.17 (*R*<sub>t</sub> = 4.61 min). HRMS (ESI): calc. for C<sub>13</sub>H<sub>10</sub>ClN<sub>3</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 292.0489, observed: 292.0483.<sup>33</sup>

**2-chloro-5-nitro-N-(2-hydroxyphenyl)benzamide (23).** 2-Aminophenol (6.82 mmol) was dissolved in dry THF (3 mL) and cooled to 0 °C. 2-Chloro-5-nitrobenzoyl chloride (2.27 mmol, 500 mg) was added dropwise to the solution. The reaction mixture was stirred under argon at 0 °C for 3 h. Additionally it was quenched with 5% aqueous hydrochloric acid, poured into ice and washed with 5% aqueous hydrochloric acid and H<sub>2</sub>O. After trituration in a mixture of ethyl acetate/hexane (50:50), the product was obtained (347 mg, 52%). <sup>1</sup>H-NMR (400 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) 9.95 (1H, s), 9.77 (1H, s), 8.44 (1H, d, *J* = 2.2 Hz), 8.31 (1H, dd, *J* = 8.8, 2.8 Hz), 7.88 – 7.79 (2H, m), 7.05 (1H, t, *J* = 7.7 Hz), 6.93 (1H, d, *J* = 6.6 Hz), 6.84 (1H, t, *J* = 7.7 Hz); <sup>13</sup>C-NMR (100 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  (ppm) 163.02, 148.97, 145.88, 137.62, 137.13, 131.09, 125.78, 125.38, 125.12, 124.04, 123.63, 118.83, 115.64. LC-MS (ESI): calc. for C<sub>13</sub>H<sub>9</sub>ClN<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 293.69, observed 293.17 (*R*<sub>t</sub> = 5.56 min). HRMS (ESI): calc. for C<sub>13</sub>H<sub>9</sub>ClN<sub>2</sub>O<sub>4</sub> [M+H]<sup>+</sup>: 293.0329, observed: 293.0326.<sup>33</sup>

## Protein expression

**ROR $\gamma$ t LBD expression and purification.** His<sub>6</sub>-ROR $\gamma$ t-LBD was expressed and purified as described in Chapter 2.

**Site-directed mutagenesis.** Point mutations (Cys320Ala) were introduced using the QuickChange Lightning Multi Site-Directed Mutagenesis Kit from Agilent Technologies in accordance with the protocols described in the kit manual. The primer with sequence 5'-tgggaaatgtgggaacgtgcgcacatc-

acctgacgg-3' was used to mutate the Cys320 residue into an alanine residue. The QIAquick Purification Kit from Qiagen was used to isolate and purify the mutated DNA.<sup>37</sup>

### Ligation

**Ligation experiment.** Ligation experiments were executed in a buffer with 50 mM TRIS (pH varying from 5.3-5.8), 50 mM NaCl, 5 mM CHAPS and 0.1% Bovine Serum Albumin (BSA). The concentration of RORYt in the buffer was 20  $\mu$ M and the compound concentration varied from 5 to 10 equivalents (Supporting Table S5.1). Total volume was 50  $\mu$ L which was incubated (continuous mixing) at 4 °C or room temperature overnight.

**Q-TOF MS analysis.** Purity and exact mass of the ligated RORYt protein were determined using a High-Resolution LC-MS system consisting of a Waters ACQUITY UPLC I-Class system coupled to a Xevo G2 Quadrupole Time of Flight (Q-TOF). The system was comprised of a Binary Solvent Manager and a Sample Manager with Fixed-Loop (SM-FL). A Polaris C18A reverse phase column (2.0 x 100 mm, Agilent) was used (0.3 mL/min) with a gradient of 15-75% acetonitrile in water supplemented with 0.1% v/v formic acid, before analysis in positive mode in the mass spectrometer. Deconvolution of the m/z spectra was done using the MaxENT1 algorithm in the Masslynx v4.1 (SCN862) software. The percentage ligation was determined from the ratio between the different mass peaks in the deconvoluted spectra.

### Biophysical assays

**Ligation experiment for the TR-FRET coactivator recruitment assay – Dose Response.** Ligation was conducted using 20  $\mu$ M His<sub>6</sub>-RORYt-LBD in buffer containing 50 mM TRIS (pH mentioned in Supporting Table S5.2), 50 mM NaCl, 5 mM CHAPS and 0.1% Bovine Serum Albumin (BSA). A mix of the probe (10 eq, 200  $\mu$ M at the maximum concentration) and His<sub>6</sub>-RORYt-LBD (20  $\mu$ M) were titrated using a 1.4 x dilution series of the probe in Corning white low volume, low binding, 384-well plates at a final volume of 5.71  $\mu$ L. The final DMSO concentration was 1% v/v throughout. The plate was incubated at 4 °C or room temperature overnight (shown in Supporting Table S5.1).

**TR-FRET coactivator recruitment assay – Dose Response.** TR-FRET assays and data analysis were performed as described in Chapter 2, using 100 nM ligated His<sub>6</sub>-RORYt-LBD, without the titration of ligands. Data were recorded in triplicate; error shown is standard deviation from the mean; curves are representative of two independent experiments.

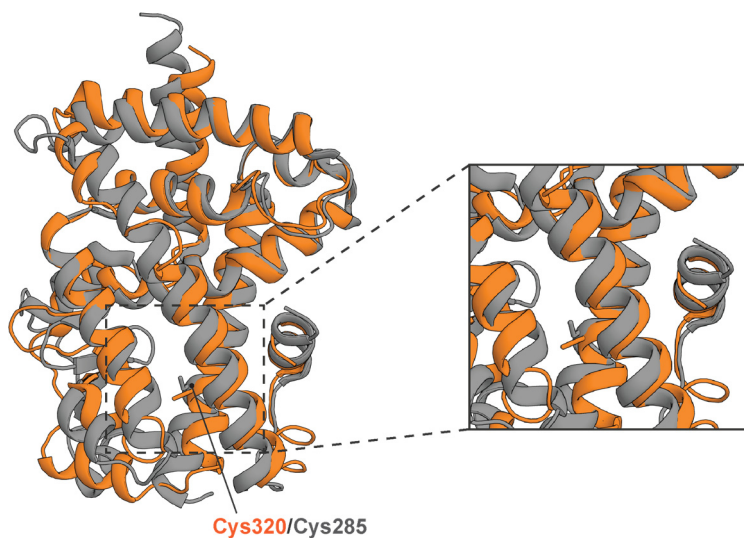
**Ligation experiment for the TR-FRET coactivator recruitment assay – Ligand Binding.** Ligation was conducted using 20  $\mu$ M His<sub>6</sub>-RORYt-LBD in buffer containing 50 mM TRIS (pH mentioned in Supporting Table S5.2), 50 mM NaCl, 5 mM CHAPS and 0.1% Bovine Serum Albumin (BSA). A mix of the probe (5 eq, 100  $\mu$ M) and His<sub>6</sub>-RORYt-LBD (20  $\mu$ M) was incubated (continuous mixing) at 4 °C or room temperature (shown in Supporting Table S5.1) overnight, in a total volume of 50  $\mu$ L. The mixture was rebuffed to TR-FRET assay buffer (10 mM HEPES, 150 mM NaCl) using a PD Spin Trap G-25 column and excess compound was removed.

**TR-FRET coactivator recruitment assay – Ligand Binding.** TR-FRET assays and data analysis were performed as described in Chapter 2, using 20 nM (ligated) His<sub>6</sub>-RORYt-LBD. The gain was set manually to 141 and 163 for Tb and d2, respectively. Data were recorded in triplicate; error shown is standard deviation from the mean; curves are representative of two independent experiments.

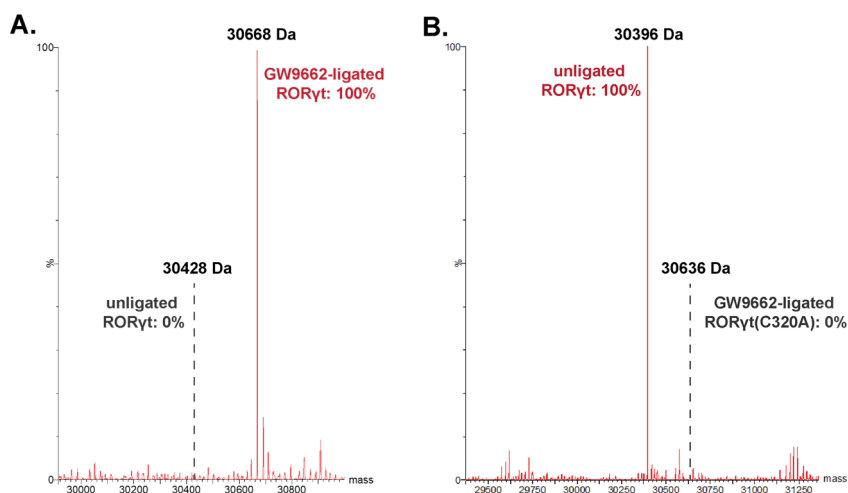
**Thermal shift assays.** Thermal shift assays and data analysis were performed as described in Chapter 3, using 5  $\mu$ M (ligated) RORYt LBD and 10  $\mu$ M compound. Data were recorded in triplicate; error shown is standard deviation from the mean; data are representative of two independent experiments.



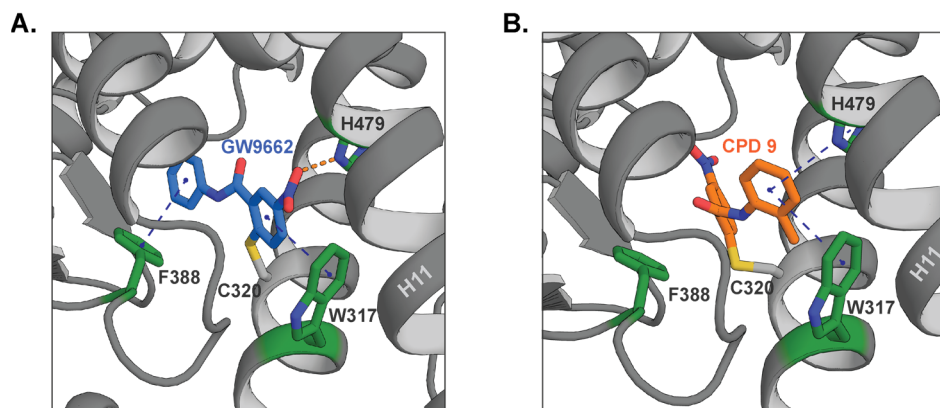
## Supporting Information



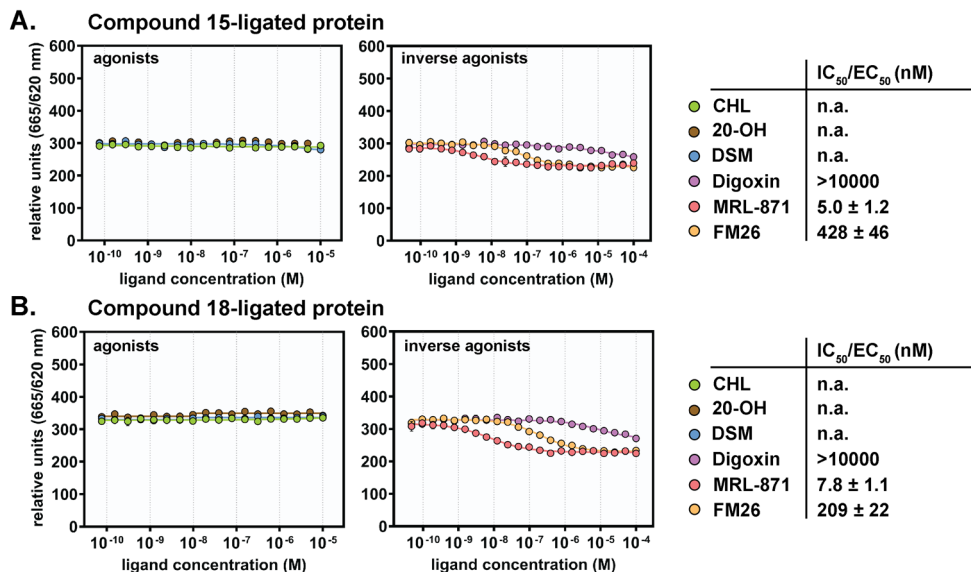
**Figure S5.1** | Overlay of the crystal structures of the ROR $\gamma$  LBD (orange, PDB: 3LoL) and PPAR $\gamma$  LBD (grey, PDB: 3BoR). The overlay demonstrates a similar positioning of the cysteine residues Cys320 (ROR $\gamma$ ) and Cys285 (PPAR $\gamma$ ).



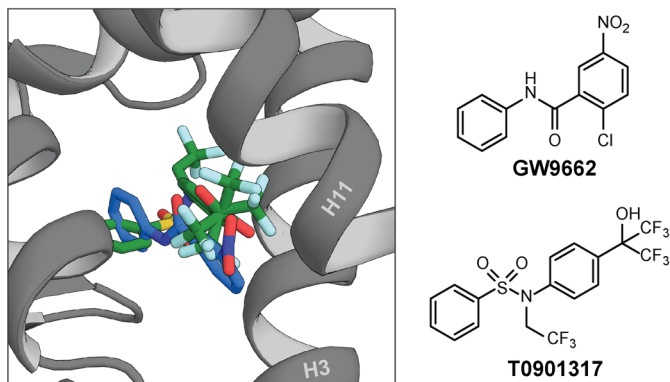
**Figure S5.2** | Deconvoluted Q-TOF spectra for the ligation of GW9662 to ROR $\gamma$ . **A)** ROR $\gamma$  ligated to GW9662: 100% ligation is observed (mass: 30668 Da), while 0% unligated protein is visible (mass: 30428 Da). **B)** ROR $\gamma$  containing a Cys320Ala mutation, ligated to GW9662: 0% ligation is observed (mass: 30636 Da), while 100% unligated protein is visible (mass: 30396 Da).



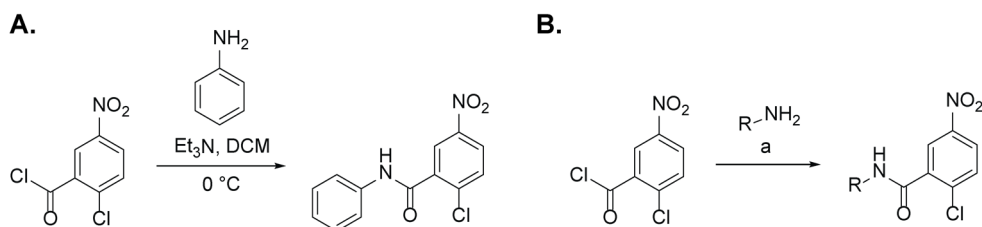
**Figure S5.3** | *In silico* modelled docking poses of GW9662 and compound 9 in complex with the ROR $\gamma$ t LBD (PDB: 3LoL). **A)** Docking pose of GW9662 (blue sticks) (best docking score), making  $\pi$ - $\pi$  interactions (blue dashed lines) with residues Trp317 and Phe388, and a hydrogen bond interaction with His479 (orange dashed line). The warhead is positioned towards H11. **B)** Docking pose of compound 9 (orange sticks) (best docking score), making  $\pi$ - $\pi$  interactions (blue dashed lines) with residues Trp317 and His479. The warhead is positioned into the orthosteric pocket.



**Figure S5.4** | TR-FRET coactivator recruitment ligand binding assays with ROR $\gamma$ t (ligated) by titration of various orthosteric and allosteric ligands. **A/B)** Ligand binding assays for compound 15-ligated protein (A) and compound 18-ligated protein (B). Abbreviations: n.a., not active.



**Figure S5.5 |** Overlay of the co-crystal structure of ROR $\gamma$ t in complex with inverse agonist To901317 (green sticks) (PDB: 4NB6) and the docking pose of GW9662 (blue sticks) (best docking score). The nitro moiety of GW9662 points towards helix 11 (H11), similar to the CF<sub>3</sub>-groups of To901317.



**Scheme S5.1 | A)** Synthesis route for compound GW9662 via an amide coupling reaction. **B)** General synthesis route for all compounds. Reagents and conditions: (a) amine, triethylamine, CH<sub>2</sub>Cl<sub>2</sub>, 0 °C, 18 h, or: pyridine, 0 °C, 18 h, or: DMA, rt, 18 h, or: Et<sub>3</sub>N, DMAP, CH<sub>2</sub>Cl<sub>2</sub>, rt, 18 h, or: THF, rt, 3 h. The reaction conditions for each compound are specified in the experimental section.

**Table S5.1 |** Optimal ligation conditions for all chemical probes.

Compound	Ligation conditions
1 (GW9662), 11, 15, 16 & 17	pH=5.8, 5 eq, 4 °C
2 (SB1404), 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14, 18, 19, 20 & 23	pH=5.8, 5 eq, RT
21	pH=5.8, 10 eq, 4 °C
22	pH=5.3, 10 eq, RT

## References

1. Ivanov, I. I. *et al.* The Orphan Nuclear Receptor ROR $\gamma$ t Directs the Differentiation Program of Proinflammatory IL-17+ T Helper Cells. *Cell* **126**, 1121–1133 (2006).
2. Solt, L. A. & Burris, T. P. Action of RORs and their ligands in (patho)physiology. *Trends Endocrinol. Metab.* **23**, 619–627 (2012).
3. Manel, N., Unutmaz, D. & Littman, D. R. The differentiation of human TH-17 cells requires TGF- $\beta$  and induction of the nuclear receptor ROR $\gamma$ t. *Nat. Immunol.* **9**, 641–649 (2008).
4. Miossec, P. & Kolls, J. K. Targeting IL-17 and TH17 cells in chronic inflammation. *Nat. Rev. Drug Discov.* **11**, 763–776 (2012).
5. Yang, J., Sundrud, M. S., Skepner, J. & Yamagata, T. Targeting Th17 cells in autoimmune diseases. *Trends Pharmacol. Sci.* **35**, 493–500 (2014).
6. Fauber, B. P. & Magnuson, S. Modulators of the nuclear receptor retinoic acid receptor-related orphan receptor-gamma (RORgamma or RORc). *J. Med. Chem.* **57**, 5871–5892 (2014).
7. Cyr, P., Bronner, S. M. & Crawford, J. J. Recent progress on nuclear receptor ROR $\gamma$  modulators. *Bioorganic Med. Chem. Lett.* **26**, 4387–4393 (2016).
8. Pandya, V. B., Kumar, S., Sachchidanand, Sharma, R. & Desai, R. C. Combating Autoimmune Diseases With ROR $\gamma$  Inhibitors: Hits and Misses. *J. Med. Chem.* **61**, 10976–10995 (2018).
9. Bronner, S. M., Zbieg, J. R. & Crawford, J. J. ROR $\gamma$  antagonists and inverse agonists: a patent review. *Expert Opin. Ther. Pat.* **27**, 101–112 (2017).
10. Sun, N., Guo, H. & Wang, Y. Retinoic acid receptor-related orphan receptor gamma-t (ROR $\gamma$ t) inhibitors in clinical development for the treatment of autoimmune diseases: a patent review (2016-present). *Expert Opin. Ther. Pat.* **29**, 663–674 (2019).
11. Li, X. *et al.* Structural studies unravel the active conformation of apo ROR $\gamma$ t NR and a common inverse agonism of two diverse classes of ROR $\gamma$ t inhibitors. *J. Biol. Chem.* **292**, 11618–11630 (2017).
12. Hu, X. *et al.* Sterol metabolism controls TH17 differentiation by generating endogenous ROR $\gamma$  agonists. *Nat. Chem. Biol.* **11**, 141–147 (2015).
13. Huh, J. R. *et al.* Digoxin and its derivatives suppress TH17 cell differentiation by antagonizing ROR $\gamma$ t activity. *Nature* **472**, 486–490 (2011).
14. Meijer, F. A., Leijten-van de Gevel, I. A., de Vries, R. M. J. M. & Brunsvel, L. Allosteric small molecule modulators of nuclear receptors. *Mol. Cell. Endocrinol.* **485**, 20–34 (2019).
15. Scheepstra, M. *et al.* Identification of an allosteric binding site for ROR $\gamma$ t inhibition. *Nat. Commun.* **6**, e8833 (2015).
16. Karstens, W. F. J. *et al.* RORgammaT Inhibitors. *PCT Int. Appl.* WO2012/106995 (2012).
17. Chaudari, S. S. *et al.* Bicyclic Heterocyclic Compounds as ROR gamma Modulators. *PCT Int. Appl.* WO2015/008234 (2015).
18. de Vries, R. M. J. M., Meijer, F. A., Doveston, R. G. & Brunsvel, L. Elucidation of an Allosteric Mode of Action for a Thienopyrazole ROR $\gamma$ t Inverse Agonist. *ChemMedChem* **15**, 561–565 (2020).
19. Meijer, F. A. *et al.* Ligand-Based Design of Allosteric Retinoic Acid Receptor-Related Orphan Receptor  $\gamma$ t (ROR $\gamma$ t) Inverse Agonists. *J. Med. Chem.* **63**, 241–259 (2020).
20. Fauber, B. P. *et al.* Discovery of imidazo[1,5-a]pyridines and -pyrimidines as potent and selective RORc inverse agonists. *Bioorganic Med. Chem. Lett.* **25**, 2907–2912 (2015).
21. Ouvry, G. *et al.* Discovery of phenoxyindazoles and phenylthioindazoles as ROR $\gamma$  inverse agonists. *Bioorganic Med. Chem. Lett.* **26**, 5802–5808 (2016).
22. Jiang, X. *et al.* A novel series of cysteine-dependent, allosteric inverse agonists of the nuclear receptor ROR $\gamma$ t. *Bioorg. Med. Chem. Lett.* **30**, e126967 (2020).
23. Brust, R. *et al.* Modification of the Orthosteric PPAR $\gamma$  Covalent Antagonist Scaffold Yields an Improved Dual-Site Allosteric Inhibitor. *ACS Chem. Biol.* **12**, 969–978 (2017).
24. Leesnitzer, L. M. *et al.* Functional Consequences of Cysteine Modification in the Ligand Binding Sites of Peroxisome Proliferator Activated Receptors by GW9662. *Biochemistry* **41**, 6640–6650 (2002).
25. Ohtera, A. *et al.* Identification of a New Type of Covalent PPAR $\gamma$  Agonist using a Ligand-Linking Strategy. *ACS Chem. Biol.* **10**, 2794–2804 (2015).
26. Seargent, J. M., Yates, E. A. & Gill, J. H. GW9662, a potent antagonist of PPARgamma, inhibits growth of breast tumour cells and promotes the anticancer effects of the PPARgamma agonist rosiglitazone, independently of PPARgamma activation. *Br. J. Pharmacol.* **143**, 933–937 (2004).
27. Bae, H. *et al.* Mechanistic elucidation guided by covalent inhibitors for the development of anti-diabetic PPAR $\gamma$  ligands. *Chem. Sci.* **7**, 5523–5529 (2016).
28. Gehringer, M. & Laufer, S. A. Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **62**, 5673–5724 (2019).

29. Degorce, F. HTRF: A Technology Tailored for Drug Discovery - A Review of Theoretical Aspects and Recent Applications. *Curr. Chem. Genomics* **3**, 22–32 (2009).
30. Kallen, J. *et al.* Structural States of ROR $\gamma$ mat: X-ray Elucidation of Molecular Mechanisms and Binding Interactions for Natural and Synthetic Compounds. *ChemMedChem* **12**, 1014–1021 (2017).
31. Kumar, N. *et al.* The benzenesulfoamide T0901317 is a novel retinoic acid receptor-related orphan receptor- $\alpha$ / $\gamma$  inverse agonist. *Mol. Pharmacol.* **77**, 228–236 (2010).
32. Fauber, B. P. *et al.* Structure-based design of substituted hexafluoroisopropanol-arylsulfonamides as modulators of ROR $\gamma$ . *Bioorg. Med. Chem. Lett.* **23**, 6604–6609 (2013).
33. Henke, A. & Srogl, J. Thioimides: New Reagents for Effective Synthesis of Thioesters from Carboxylic Acids. *J. Org. Chem.* **73**, 7783–7784 (2008).
34. Amemiya, Y.; Wakabayashi, K.; Takaishi, S. & Fukuda, C. PPAR $\gamma$  modulators. CA2407587A1 (2001).
35. D'Silva, C. & Iqbal, R. A New Method to N-Arylmethylenepyrroles from N-Acylpyrroles. *Synthesis (Stuttg.)*. **1996**, 457–458 (1996).
36. de Vries, R. M. J. M., Meijer, F. A., Doveston, R. G., Leijten-van de Gevel, I. A. & Brunsveld, L. Cooperativity between the Orthosteric and Allosteric Ligand Binding Sites of ROR $\gamma$ t. *Proc. Natl. Acad. Sci.* **118**, e2021287118 (2020).
37. Agilent Technologies. QuickChange Lightning Site-Directed Mutagenesis Kit. Instruction Manual (Vol. Revision D). (2011).



# Chapter 6

## Dual Targeting of the Orthosteric and Allosteric Binding Sites of ROR $\gamma$ t with a Bitopic Ligand

### Abstract

ROR $\gamma$ t is a unique NR in that it contains both a canonical, orthosteric and a second, allosteric ligand binding site in its ligand binding domain (LBD). Hence, dual targeting of both binding pockets constitutes an attractive alternative molecular entry for pharmacological modulation. Here, we report an innovative approach to develop a bitopic ligand for ROR $\gamma$ t, enabling concomitant engagement of both binding pockets. Three candidate bitopic ligands, **Bit-L15**, **Bit-L9**, and **Bit-L4**, comprising of an orthosteric and allosteric ROR $\gamma$ t pharmacophore linked via a polyethylene glycol (PEG) linker, were designed, synthesized, and evaluated to examine the influence of linker length on the ROR $\gamma$ t binding mode. **Bit-L15** and **Bit-L9** showed convincing evidence of concomitant engagement of both ROR $\gamma$ t binding pockets, while the shorter **Bit-L4** did not, as was anticipated during the ligand design. As the most potent bitopic ROR $\gamma$ t ligand, **Bit-L15** antagonized ROR $\gamma$ t function in a potent manner in both a biochemical and cellular context. Furthermore, **Bit-L15** displayed an increased selectivity for ROR $\gamma$ t over ROR $\alpha$  and PPAR $\gamma$  compared to the purely orthosteric and allosteric parent compounds. Combined, these results highlight potential advantages of bitopic NR modulation over monovalent targeting strategies.

**This chapter has been published as:** F.A. Meijer<sup>#</sup>, G.J.M. Oerlemans<sup>#</sup> & L. Brunsveld. Orthosteric and Allosteric Dual Targeting of the Nuclear Receptor ROR $\gamma$ t with a Bitopic Ligand. *ACS Chem. Biol.* 16, 510-519 (2021).

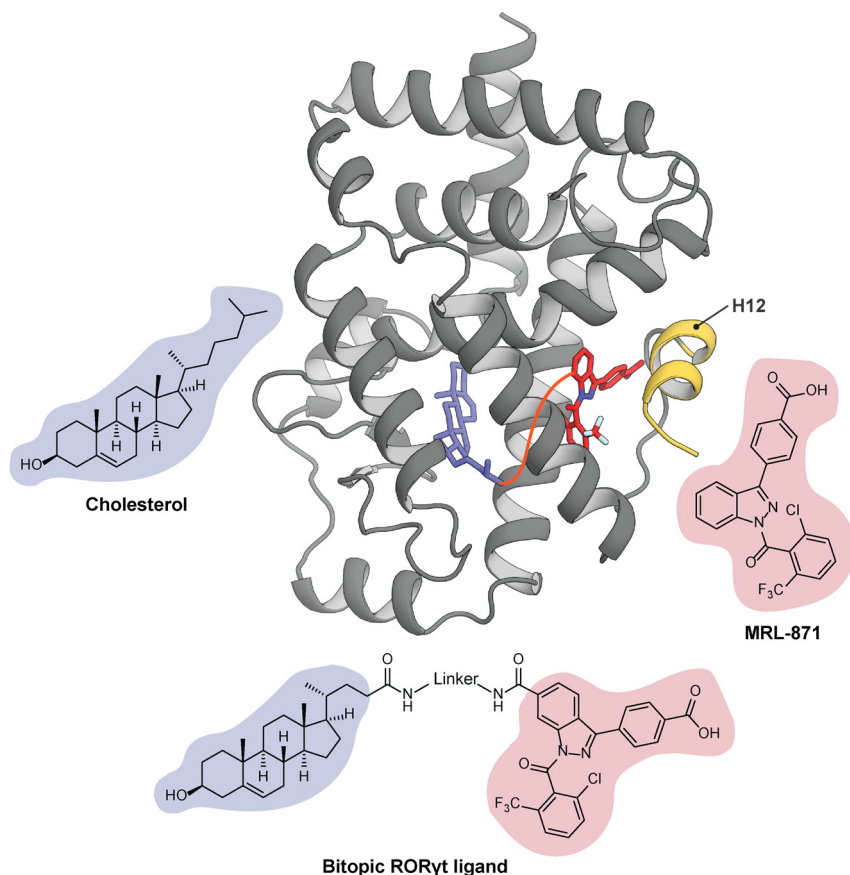
<sup>#</sup> These authors contributed equally

## Introduction

ROR $\gamma$ t is a NR that plays an important regulatory role in the immune system.<sup>1-3</sup> ROR $\gamma$ t expression is limited to the lymphoid system, where it is essential for the differentiation of naïve CD4<sup>+</sup> T cells into Th17 cells, and the production of the pro-inflammatory cytokine IL-17a.<sup>1-3</sup> Elevated IL-17a levels are highly associated with the pathogenesis of autoimmune diseases, including multiple sclerosis, rheumatoid arthritis and psoriasis.<sup>4-7</sup> Disrupting the Th17/IL-17a pathway could therefore potentially be an effective strategy for the treatment of these diseases.<sup>4</sup> The clinical successes of FDA-approved monoclonal antibodies targeting IL-17a or Th17 cell development have already validated the potential of Th17 pathway inhibition as a successful therapeutic strategy.<sup>8</sup> However, inhibition of ROR $\gamma$ t with small molecules might be an attractive alternative strategy to decrease IL-17a production in the treatment of these autoimmune diseases, which has been the focus of many research efforts over the past decades, resulting in the development of several synthetic ROR $\gamma$ t inverse agonists.<sup>9-14</sup>

Typically, NR ligands bind to a highly conserved hydrophobic binding pocket, termed orthosteric site, located within the ligand binding domain (LBD) of ROR $\gamma$ t.<sup>2,9</sup> ROR $\gamma$ t features some level of background transcriptional activity since Helix 12 (H12/AF-2) is already positioned in a conformation that enables coactivator recruitment in the apo structure.<sup>15</sup> Regardless, ROR $\gamma$ t is responsive to ligand binding, with cholesterol (Figure 6.1) and its derivatives acting as agonists for ROR $\gamma$ t,<sup>2,16</sup> stabilizing H12 in an active conformation, which results in an increased recruitment of coactivators. Conversely, inverse agonist binding destabilizes the active conformation of H12, disrupting the coactivator binding groove and thus decreasing the transcriptional activity. Recently, a novel class of ROR $\gamma$ t inverse agonists has been identified, typified by **MRL-871**, which bind to a topographically distinct, allosteric site of the ROR $\gamma$ t LBD, formed by helices 3, 4, 11 and reoriented H12 (Figure 6.1).<sup>17-20</sup> The interactions in this allosteric pocket are predominantly hydrophobic, in addition to the hydrogen bonds between the carboxylic acid moiety of **MRL-871** and the backbone hydrogen atoms of Ala497 and Phe498 as well as the side chain of residue Gln329.<sup>17</sup> These allosteric ligands decrease the transcriptional activity of ROR $\gamma$ t by repositioning H12 in a conformation incompatible with coactivator binding and thus directly affect the activity of ROR $\gamma$ t.<sup>17</sup> Interestingly, these ligands show a high potency (low nanomolar IC<sub>50</sub> values) and potentially possess beneficial properties over orthosteric ligands.<sup>17,21</sup> Therefore, such allosteric ligands are of high relevance in drug discovery.<sup>21-23</sup>





**Figure 6.1** | Crystal structure of ROR $\gamma$ t with MRL-871 in the allosteric site (red sticks) and cholesterol in the orthosteric site (blue sticks), H12 is shown in yellow (PDB: 6T4I).<sup>24</sup> The chemical structures of the orthosteric agonist cholesterol (blue), allosteric inverse agonist MRL-871 (red) and the general design of the bitopic ROR $\gamma$ t ligand (orthosteric cholesterol pharmacophore in blue, allosteric MRL-871 pharmacophore in red) are shown as well. The envisioned path of the linker connecting the orthosteric and allosteric pharmacophores is shown as an orange line in the crystal structure.

Recent studies with orthosteric and allosteric ligands have demonstrated the capability of ROR $\gamma$ t to bind both types of ligands simultaneously, even in a cooperative fashion.<sup>17,24–26</sup> These insights have inspired us to develop ligands that comprise of a covalently linked orthosteric and allosteric pharmacophore to enable simultaneous targeting of both sites, also known as bitopic ligands.<sup>27,28</sup> Bitopic ligands were pioneered for G protein-coupled receptors (GPCRs).<sup>27–30</sup> Recently, the field of bitopic ligands has been expanded to other protein classes including kinases, e.g. mTor<sup>31</sup> and PKC $\alpha$ <sup>32</sup>, and a merged bitopic ligand for the nuclear receptor PPAR $\gamma$ <sup>33</sup>. These chemical biology studies have demonstrated that a dual targeting

strategy is associated with several advantages over monovalent targeting strategies, including an increased affinity<sup>30,34</sup> or selectivity,<sup>30,34–36</sup> a bias in signaling pathway activation,<sup>30,37</sup> and reduced therapeutic resistance<sup>31</sup>.

This chapter describes the design, synthesis and biochemical evaluation of three candidate bitopic ligands that comprise of a covalently linked orthosteric and allosteric pharmacophore for ROR $\gamma$ t (Figure 6.1), as the first linked bitopic ligands for NRs. Biochemical evaluation revealed that both **Bit-L15** and **Bit-L9** (linking both pharmacophores via a biamine linker with fifteen and nine PEG units) showed bitopic ROR $\gamma$ t binding characteristics, while **Bit-L4** (containing a shorter linker with four PEG units) did not, as anticipated by design. Most promisingly, **Bit-L15** had a significantly increased overall efficacy compared to its monovalent counterparts in both a biochemical and cellular context, approaching the activity of **MRL-871**. In addition, **Bit-L15** displayed increased selectivity for ROR $\gamma$ t over ROR $\alpha$  and PPAR $\gamma$ , compared to a cholesterol derivative and **MRL-871**, respectively. Combined, this study shows that bitopic modulation of ROR $\gamma$ t might enable advantageous properties over classical, monovalent NR targeting strategies, providing a framework for future studies investigating bitopic NR modulation.

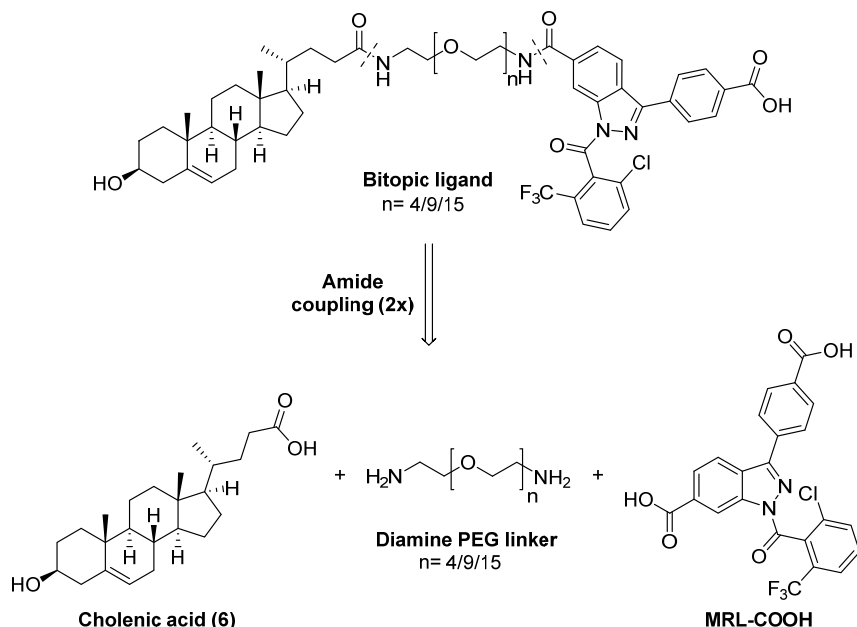
## Design of the bitopic ROR $\gamma$ t ligands

The first step in the design of a bitopic ROR $\gamma$ t ligand was the identification of a suitable pharmacophore pair that could be used for linkage. Because of the concomitant binding observed for the orthosteric agonist cholesterol and allosteric inverse agonist **MRL-871** to ROR $\gamma$ t<sup>24</sup>, these two pharmacophores were chosen for the bitopic ligand design. The crystal structure of ROR $\gamma$ t with cholesterol and **MRL-871** (PDB: 6T41)<sup>24</sup> (Figure 6.1) was used to devise a suitable strategy to link both ligands. The most promising linking strategy, in terms of space and retaining the key pharmacophore interactions with the LBD, was envisioned to be the coupling of the acyclic alkyl chain of cholesterol to the indazole core of **MRL-871**, yielding a bitopic ligand with the general structure shown in Figure 6.1.

Previous structure activity relationship (SAR) studies around **MRL-871** have shown that modifications at the C-6 position of the indazole scaffold are tolerated<sup>17,18</sup>, since this part of the molecule protrudes into an open channel in the co-crystal structure (Supporting Figure S6.1A). In previous studies, the C-6 position of **MRL-871** was functionalized with a carboxylic acid moiety (**MRL-COOH**, Scheme 6.1) and various PEG linkers were attached to this handle via amide coupling chemistry (Supporting Figure S6.1B).<sup>17</sup> These modifications resulted in

an affinity decrease of up to 30-fold relative to **MRL-871**, but the derivatives were still able to bind to the LBD of ROR $\gamma$ t with IC<sub>50</sub> values of 250 nM or lower.<sup>17</sup> Therefore, **MRL-COOH** was used as entry for the attachment of a linker to the allosteric pharmacophore.

To keep the linking pathway between both sites as short as possible, the ideal position for linker attachment to the orthosteric pharmacophore is the alkyl tail of cholesterol (Figure 6.1), which is problematic due to the lack of a reactive handle at this position. Recently, Kallen and colleagues have published the crystal structure of cholenic acid (see chemical structure in Scheme 6.1), extended at its carboxylic acid position (Supporting Figure S6.1D).<sup>38</sup> Their work demonstrates that extended derivatives of cholenic acid maintain the ability to bind ROR $\gamma$ t, and that the receptor is highly flexible in the HII region (Supporting Figure S6.1C).<sup>38</sup> Although the extension induces a protein conformation that is incompatible with allosteric pocket formation due to displacement of HII, a less bulky and less rigid extension of cholenic acid is expected to disturb the agonistic protein conformation to a lesser extent, enabling the formation of the allosteric pocket. Therefore, cholenic acid was selected as entry for the attachment of a linker to the orthosteric pharmacophore.



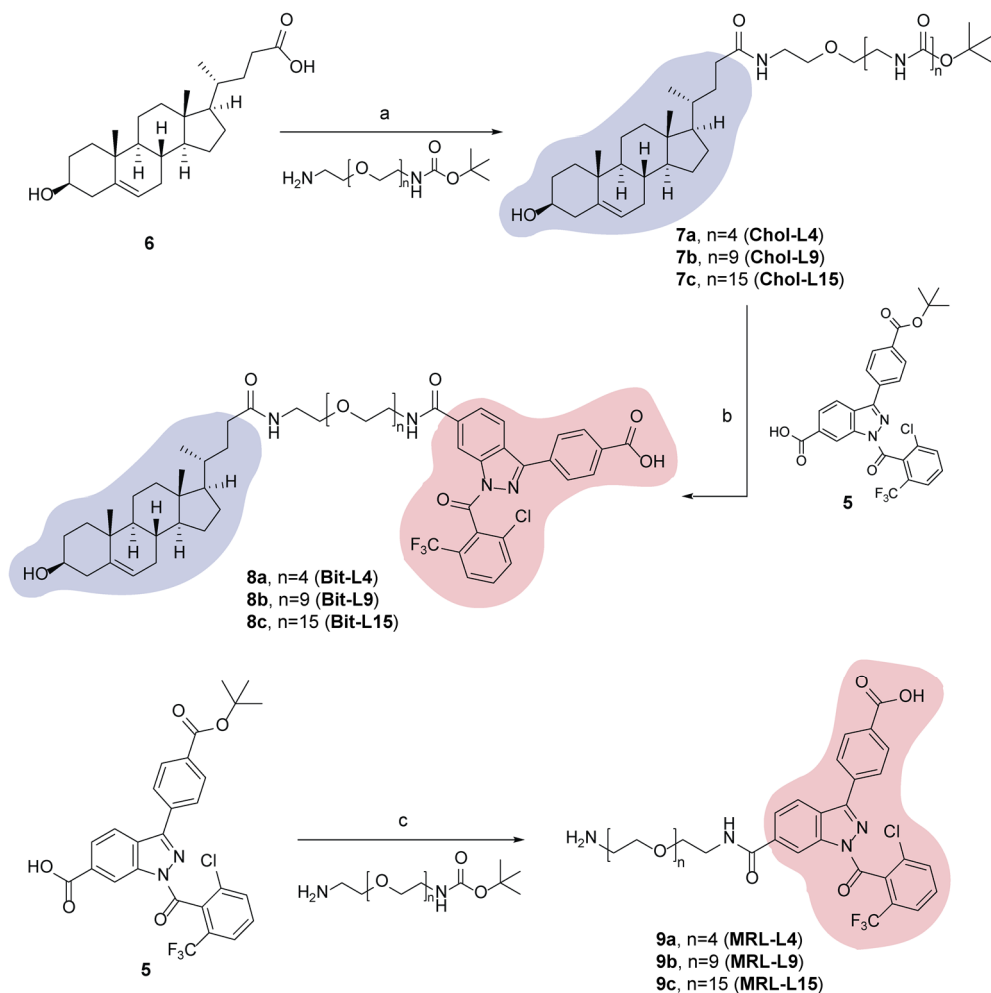
**Scheme 6.1 | Retrosynthesis of the designed bitopic ligands.** The synthesis is established via two amide coupling reactions between the carboxylic acid functionalities of cholenic acid and **MRL-COOH**, and a diamine PEG linker (n= 4/9/15).

The carboxylic acid moieties of **MRL-COOH** and cholenic acid allow the connection of both pharmacophores with a diamine linker via amide coupling chemistry to yield the desired bitopic ligands (Scheme 6.1). In order to avoid nonspecific protein binding and to maintain flexibility and solubility, a polyethylene glycol (PEG) linker was used.<sup>39</sup> This type of linker is expected to maintain a high degree of conformational freedom upon bitopic binding to ROR $\gamma$ t, which is beneficial from an entropic perspective. Additionally, PEG linkers have also been applied successfully in other bitopic ligands.<sup>31</sup> The distance between cholenic acid and the **MRL-871** derivative, following the linker path illustrated in Figure 6.1, was estimated via *in silico* measurements in the crystal structure to be 27.4 Å (Supporting Figure S6.2). A linker consisting of nine PEG units (**Bit-L9**) (MM2-minimalized maximum nitrogen to nitrogen distance of the linker is 35.2 Å) was hypothesized to be just of adequate length to enable engagement with both sites. To verify this hypothesized mode, a linker consisting of four PEG units (**Bit-L4**), which is too short to span the distance around the protein, was also investigated (MM2-minimalized maximum nitrogen to nitrogen distance of the linker is 17.9 Å), as well as a ligand with a longer linker of fifteen PEG units (**Bit-L15**) (MM2-minimalized maximum nitrogen to nitrogen distance of the linker is 56.9 Å).

## Synthesis of the ligands

The synthesis of the bitopic ligands was established via two amide coupling reactions with the three building blocks (Scheme 6.1): 1) the orthosteric ligand cholenic acid, 2) the diamine PEG linker (both commercially available), and 3) the allosteric ligand **MRL-COOH**. In order to prevent chemoselectivity issues during the synthesis, *tert*-butyloxycarbonyl (*t*-Boc) monoprotected biamine PEG linkers were used, and the **MRL-871** derivative was synthesized containing a *tert*-Butyl protected benzoic acid moiety.

The protected **MRL-COOH** derivative **5** was synthesized as described in literature,<sup>17,18</sup> with an overall yield of 39% (Supporting Scheme S6.1). The three bitopic ligands (**8a**, **8b** and **8c**) were synthesized via two amide coupling reactions, to couple the linker to both pharmacophores (Scheme 6.2). While two strategies were tested for pharmacophore attachment (the orthosteric pharmacophore coupled first to the linker, followed by the allosteric pharmacophore, or *vice versa*), the strategy shown in Scheme 6.2 was deemed optimal, since greater ease of purification resulted in overall higher yields.



**Scheme 6.2 | Synthesis of the bitopic and monovalent ligands.** Bitopic ligands: **Bit-L4** (**8a**), **Bit-L9** (**8b**) and **Bit-L15** (**8c**), monovalent ligands: **Chol-L4** (**7a**), **Chol-L9** (**7b**) and **Chol-L15** (**7c**), and **MRL-L4** (**9a**), **MRL-L9** (**9b**) and **MRL-L15** (**9c**). Reagents and conditions: (a) *t*-Boc mono-protected biamine PEG linker, DIPEA, HATU, DMF, rt, 3 h, 95% (**7a**), 86% (**7b**), 85% (**7c**); (b) i) CH<sub>2</sub>Cl<sub>2</sub>/TFA/H<sub>2</sub>O (65:30:5), rt, 3 h; ii) MeOH, 80 °C, 18 h; iii) **5**, DIPEA, HATU, DMF, rt, 3 h; iv) CH<sub>2</sub>Cl<sub>2</sub>/TFA/H<sub>2</sub>O (65:30:5), rt, 3 h; v) MeOH, 80 °C, 18 h, 49% (**8a**), 56% (**8b**), 45% (**8c**); (c) i) *t*-Boc mono-protected biamine PEG linker, DIPEA, HATU, DMF, rt, 3 h; ii) CH<sub>2</sub>Cl<sub>2</sub>/TFA/H<sub>2</sub>O (65:30:5), rt, 3 h, 32% (**9a**), 51% (**9b**), 58% (**9c**).

First, the *t*-Boc monoprotected biamine PEG linkers were coupled to cholenic acid **6** via an amide coupling with DIPEA as base and HATU as coupling reagent, as described by Kallen *et al.*<sup>38</sup> The monovalent cholenic acid derivatives **7a** (**Chol-L4**), **7b** (**Chol-L9**) and **7c** (**Chol-L15**) were obtained in high yields. Subsequently, the linker was deprotected in a mixture of dichloromethane (DCM):trifluoroacetic acid (TFA):water, resulting in the TFA-ester of the compounds (esterified at the alcohol moiety of cholenic acid). These were refluxed in methanol, hydrolyzing the TFA-ester in quantitative yields. Subsequently, these compounds were coupled to **5** via an amide coupling reaction. The suboptimal yields at this stage are believed to be due to the formation of two unidentified side products, suspected to be related to compound **5**. After deprotection of the *tert*-Butyl protected acid of the **MRL-871** pharmacophore in quantitative yields, the desired bitopic ligands **8a**, **8b** and **8c** were obtained, termed **Bit-L4**, **Bit-L9** and **Bit-L15**.

In addition to the bitopic ligands, their monovalent counterparts were also synthesized to be used as a reference in biochemical evaluation. The monovalent orthosteric derivatives (**7a**, **7b** and **7c**) were already obtained in the synthesis route toward the bitopic ligands (Scheme 6.2). The monovalent allosteric derivatives **9a** (**MRL-L4**), **9b** (**MRL-L9**), and **9c** (**MRL-L15**) were synthesized from the **MRL-871** derivative **5** (Scheme 6.2) via a similar amide coupling and deprotection strategy as described for the other ligands.

### Biochemical evaluation of the binding mode of the bitopic ligands

Various types of time-resolved FRET (TR-FRET) binding assays<sup>40</sup> were used to investigate the potency and binding mode of the bitopic ligands and monovalent counterparts (an overview of the chemical structures is shown in Figure 6.2). The coactivator recruitment TR-FRET assays are based on fluorescence emission occurring upon the FRET pairing of a d2-labelled coactivator with a terbium cryptate-labelled ROR $\gamma$ t LBD (see Supporting Figure S6.3, left). In an orthogonal TR-FRET AlexaFluor-MRL recruitment assay (see Supporting Figure S6.3, right), an AlexaFluor647-labelled MRL-871 probe is used instead of the d2-labelled coactivator, enabling direct probing of allosteric site binding.

### The binding characteristics of both orthosteric and allosteric pharmacophores are retained

The binding behavior of the monovalent counterparts **MRL-L15** and **Chol-L15** (Figure 6.2) was evaluated in a TR-FRET coactivator recruitment assay, to ensure that the attachment of a linker to either pharmacophore is not detrimental to ligand binding. The allosteric monovalent counterpart **MRL-L15** showed a dose-dependent inverse agonistic behavior with

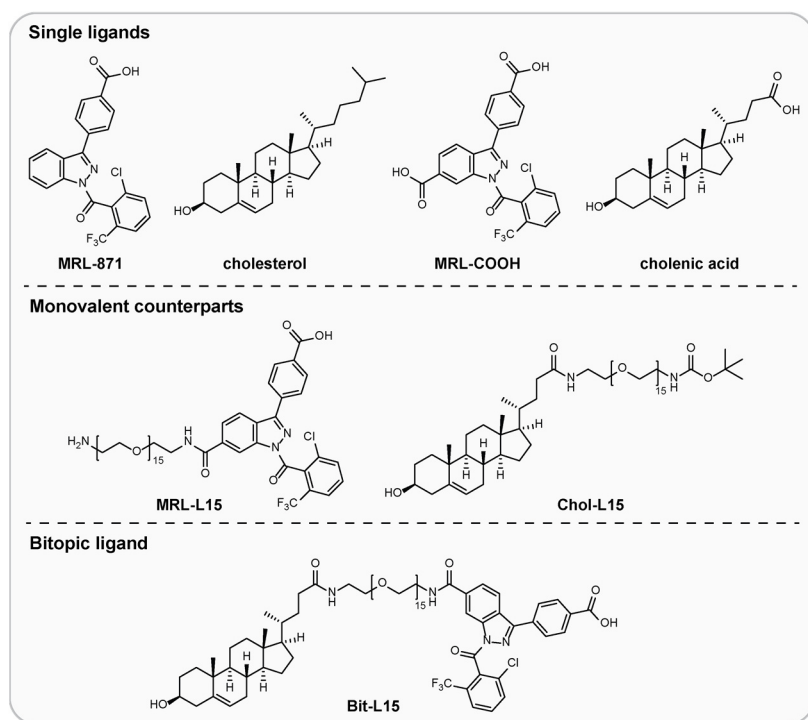
an IC<sub>50</sub> value of  $0.26 \pm 0.02 \mu\text{M}$  (Supporting Figure S6.4B), which is in the same range as **MRL-871**-based probes containing a carboxamide modification at this position.<sup>17,18</sup> In the presence of cholesterol, no decrease in inhibitory potency of **MRL-L15** was observed, indicating an allosteric mode of binding. In fact, even an increase in potency was observed for **MRL-L15** in the presence of cholesterol, indicating a cooperative behavior between both binding sites as has been observed previously.<sup>17,25,26</sup> In contrast to the orthosteric agonist cholesterol<sup>2,16</sup>, the extended **Chol-L15** derivative was not compatible with coactivator recruitment and is thus an inverse agonist with an IC<sub>50</sub> value of  $0.54 \pm 0.08 \mu\text{M}$  (Supporting Figure S6.4A), similar to the previously described extended cholenic acid derivative.<sup>38</sup> **Chol-L15** showed increasing IC<sub>50</sub> values in the presence of cholesterol, which indicates competition between the two ligands, verifying an orthosteric mode of binding.

Investigation of the binding mode of the bitopic ligands was performed via four different TR-FRET assay formats, each probing a different aspect of binding (Figure 6.3). Three different binding modes can be considered<sup>29</sup>: 1) A true bitopic mode of binding, concomitantly occupying both the orthosteric and allosteric site of the protein, 2) A flip-flop mode of binding, switching between a purely allosteric or purely orthosteric mode of binding since both pockets cannot be occupied simultaneously, and 3) A mode of binding in which one bitopic ligand binds orthosterically and a second bitopic ligand binds allosterically, termed 2:1 binding (ligand:protein stoichiometry). Although a true bitopic and a flip-flop mode of binding can be difficult to distinguish experimentally, the latter is not expected in this case since this mode is only worthwhile to consider in case of small size pharmacophores with low binding affinities, instead of voluminous high affinity ligands that have a fixed binding topography, as is the case here.<sup>29,41</sup> Based on their design, a true bitopic binding mode is expected for **Bit-L9** and **Bit-L15**, while a monovalent or 2:1 mode of binding is expected for **Bit-L4**.

### *Linking both pharmacophores increases potency of Bit-L15 for ROR $\gamma$ t compared to monovalent counterparts*

The three bitopic ligands and their monovalent counterparts were examined in a TR-FRET coactivator recruitment assay to investigate the effect of the different linkages of both pharmacophores on the inhibition of coactivator recruitment (Figure 6.3A). The results for the PEG-15 ligands demonstrate that the monovalent counterparts **MRL-L15** and **Chol-L15** showed IC<sub>50</sub> values of  $0.42 \pm 0.05 \mu\text{M}$  and  $0.67 \pm 0.11 \mu\text{M}$ , respectively, whereas the bitopic

ligand **Bit-L15** showed a significantly higher potency with an  $IC_{50}$  value of  $0.0059 \pm 0.0007$   $\mu$ M, comparable to the highly potent allosteric ligand **MRL-871** (Figure 6.3A). These results demonstrate that the linkage of the allosteric **MRL-871** and the orthosteric cholenic acid pharmacophore via a PEG-15 linker is beneficial for the overall potency of the ligand. For the second bitopic ligand with a PEG-9 linker (**Bit-L9**), a similar behavior was observed, although the overall potency of this bitopic ligand was slightly lower than for **Bit-L15** (Supporting Figure S6.5A). In stark contrast, the bitopic ligand **Bit-L4** with the shorter PEG-4 linker did not show a significant increase in potency relative to the monovalent counterparts. Instead, it showed an  $IC_{50}$  value comparable to **MRL-L4** (the highest affinity monovalent counterpart) (Supporting Figure S6.5B), providing evidence that the bitopic binding mode cannot be established with this shorter linker length.



**Figure 6.2 | Chemical structures of the ligands used in TR-FRET assays.** Single ligands (**MRL-871**, cholesterol, **MRL-COOH** and cholenic acid), monovalent counterparts (**MRL-L15** and **Chol-L15**) and bitopic ligand (**Bit-L15**).

In order to further validate these results, the TR-FRET coactivator recruitment assay was performed with a modified ROR $\gamma$ t LBD in which the orthosteric site was blocked via ligation of a chemical probe (compound **20**, Chapter 5) to a native cysteine residue (Cys320) in the



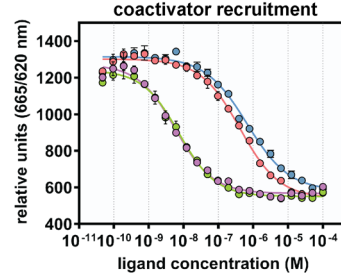
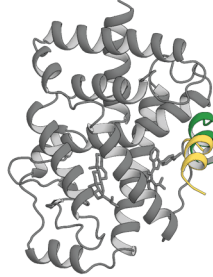
orthosteric ligand binding pocket (see Chapter 5).<sup>42</sup> This probe prevents orthosteric ligands from binding to ROR $\gamma$ t, while the allosteric binding site remains accessible for binding of allosteric ligands. As expected, when the orthosteric site is not available for binding, **Bit-L15** and **Bit-L9** showed IC<sub>50</sub> values in the same ballpark as their monovalent allosteric counterparts **MRL-L15** and **MRL-L9** (Supporting Figure S6.6A,B). The absolute IC<sub>50</sub> values in general were lower than in the regular coactivator recruitment assay, due to cooperativity between the covalent orthosteric probe and the allosteric binding ligands.<sup>17,25,26</sup> In contrast, **Bit-L4** showed a lower potency than its monovalent allosteric counterpart **MRL-L4** (Supporting Figure S6.6C). This lower potency is presumably due to unfavorable interactions or steric clashes between the protein and the unbound orthosteric cholenic acid moiety upon binding of **Bit-L4** to the allosteric site (due to the shorter linker), decreasing the allosteric site affinity relative to **MRL-L4**. Combined, these results demonstrate that the increase in overall potency of **Bit-L15** and **Bit-L9** relative to **MRL-L15** and **MRL-L9** in the coactivator recruitment assay with the native ROR $\gamma$ t LBD is due to concomitant engagement of both binding sites.

#### ***Bit-L15 competes with cholesterol for orthosteric site binding***

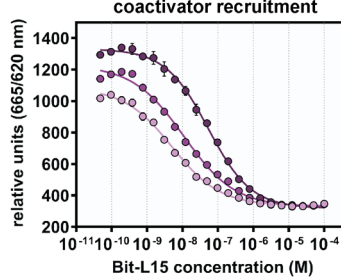
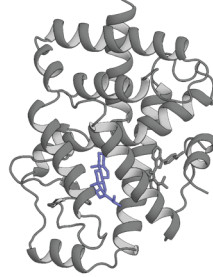
In order to probe the importance of the orthosteric site in binding of the bitopic ligands in more detail, the TR-FRET coactivator recruitment assay was performed in the absence and presence of the orthosteric ligand cholesterol (CHL). The results in Figure 6.3B show that in the absence of cholesterol, **Bit-L15** exhibits a dose-dependent inverse agonistic character, in agreement with the previous assay. However, when the same titration was performed in the presence of a fixed concentration of cholesterol, the IC<sub>50</sub> values increased (shift of the curves to the right) with increasing cholesterol concentration. This shift in IC<sub>50</sub> values demonstrates a competitive character between **Bit-L15** and cholesterol, indicating once more that orthosteric binding is involved in the mode of action of **Bit-L15**. A similar increase in IC<sub>50</sub> values was observed for **Bit-L9** and **Bit-L4** (Supporting Figure S6.7A,B), confirming that these bitopic ligands feature an orthosteric component in their binding mode as well.

Upon closer examination, a trend can be observed between the linker length of the bitopic ligands and the degree of competition with cholesterol. As can be seen in Supporting Figure S6.7C, the bitopic ligands with a longer linker length show a higher fold decrease in potency in the presence of cholesterol compared to ligands with a shorter linker length, indicating that a bitopic ligand with a longer linker becomes relatively more susceptible to competition with cholesterol. This suggests that **Bit-L15** gains more of its overall potency from the orthosteric site compared to **Bit-L9** and **Bit-L4**, i.e. especially **Bit-L4** binds mainly via the allosteric pocket.

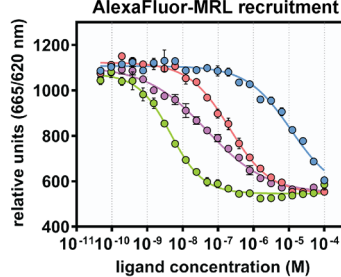
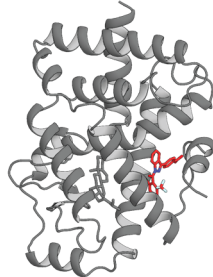
### A. Overall potency



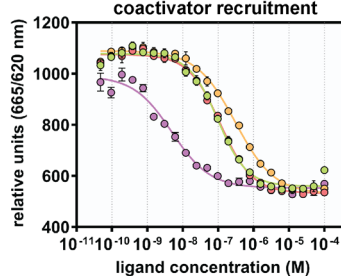
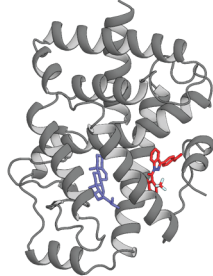
### B. Orthosteric binding



### C. Allosteric binding



### D. Multivalent binding



**Figure 6.3 | Overview of the different TR-FRET assay formats used to investigate the binding mode of the bitopic ligands. A-D)** Overall potency (A), orthosteric binding (B), allosteric binding (C) and multivalent binding (D). Next to each assay schematic, the dose-response curves and an overview of the IC<sub>50</sub> values are shown for the titration of single ligands, monovalent counterparts and bitopic ligands to RORYt (coactivator recruitment-based assays (A,B,C) and AlexaFluor-MRL recruitment-based assay (D)). Abbreviations: CHL, cholesterol; chol. acid, cholenic acid. Data are representative of two independent experiments (recorded in triplicate) and are presented as mean ± SD.

***Bit-L15 shows increased competition with an allosteric probe compared to MRL-L15***

The orthogonal TR-FRET AlexaFluor-MRL recruitment assay was used to investigate the allosteric binding behavior of the bitopic ligands (Figure 6.3C). **MRL-L15** demonstrated a clear dose-response curve, displacing the AlexaFluor-MRL-871 probe with an IC<sub>50</sub> value of  $0.21 \pm 0.02 \mu\text{M}$  (Figure 6.3C). In contrast, **Chol-L15** showed the typical behavior for an orthosteric ligand with an IC<sub>50</sub> value  $>10 \mu\text{M}$ .<sup>25</sup> **Bit-L15** displaced the allosteric probe with an IC<sub>50</sub> value of  $0.047 \pm 0.008 \mu\text{M}$ , confirming allosteric site binding. Relative to **MRL-L15**, the potency was increased 4.5-fold, which can be explained by an enhanced local concentration of the allosteric component due to concomitant binding of **Bit-L15** to the orthosteric site (tethering effect), again validating the bitopic binding mode of **Bit-L15**.

**Bit-L9** and its monovalent counterparts showed a comparable behavior to **Bit-L15** (Supporting Figure S6.8A), with **Bit-L9** featuring a 2.7-fold increase in potency compared to the monovalent **MRL-L9** (IC<sub>50</sub> =  $0.058 \pm 0.007 \mu\text{M}$  vs.  $0.16 \pm 0.02 \mu\text{M}$ , respectively). For **Bit-L4**, allosteric site binding was observed as well, however without a tethering effect (Supporting Figure S6.8B). In contrast to **Bit-L15** and **Bit-L9**, **Bit-L4** was approximately 8-fold less potent than its monovalent allosteric counterpart **MRL-L4** (Supporting Figure S6.8B). This demonstrates that the coupling of the orthosteric pharmacophore to **MRL-L4** weakens the affinity for the allosteric site, presumably due to a steric clash (as was discussed for the coactivator recruitment assay in the presence of the covalent probe), and is in agreement with the hypothesis that **Bit-L4** cannot bind both sites simultaneously. Combined, **Bit-L15** and **Bit-L9** demonstrate an increased potency relative to their allosteric monovalent counterparts, ascribed to a tethering effect from the orthosteric binding pharmacophore.

***Bit-L15 exhibits increased potency relative to co-incubated monovalent counterparts***

With the binding to both the orthosteric and allosteric site confirmed, the affinity of the bitopic ligands was compared to their simultaneously incubated monovalent counterparts in a TR-FRET coactivator recruitment assay (Figure 6.3D) to probe the presence of a multivalent effect (i.e. an increased affinity compared to equimolar amounts of co-incubated unlinked counterparts).<sup>36</sup>

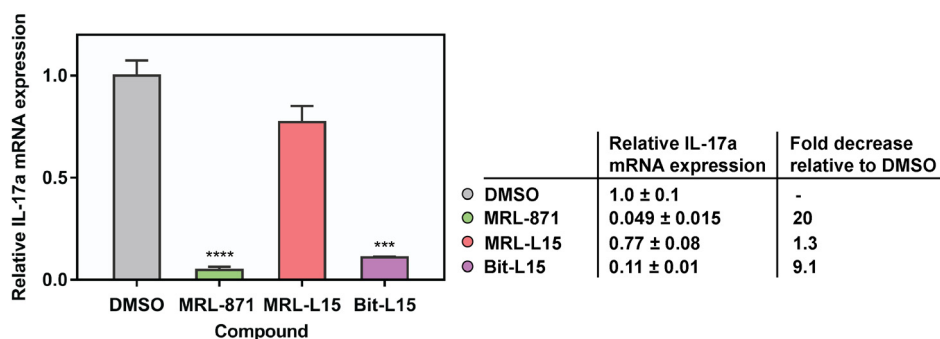
As shown in Figure 6.3D, **Bit-L15** had a higher overall affinity in the coactivator recruitment assay compared to equimolar amounts of co-incubated monovalent counterparts (different combinations of co-incubated monovalent counterparts were examined). Depending on the combination of co-incubated ligands, a 23- to 63-fold higher affinity can be observed for **Bit-L15**. This multivalent effect provides convincing evidence for a true bitopic

binding mode to the ROR $\gamma$ t LBD.<sup>36</sup> A potential 2:1 binding mode can be excluded based on these results, since in this binding scenario, **Bit-L15** would have been equally potent as the co-incubated monovalent ligands with linker.

Similarly, **Bit-L9** also showed a multivalent effect (Supporting Figure S6.9A), although with a slightly lower magnitude (5- to 11-fold increase in potency compared to the co-incubated ligands). In contrast, the potency of **Bit-L4** was similar to the co-incubated monovalent counterparts and therefore lacks a multivalent effect, in agreement with the hypothesis that this ligand cannot bind both pockets simultaneously due to an inadequate linker length (Supporting Figure S6.9B). The results for **Bit-L4** suggest a 2:1 binding stoichiometry, with one **Bit-L4** ligand binding orthosterically and another one binding allosterically.

### **Bit-L15 inhibits IL-17a expression in EL4 cells**

ROR $\gamma$ t is the master transcription factor in Th17 cell differentiation and promotes IL-17a production. Therefore, the cellular activity of **Bit-L15** was determined by measuring the reduction of IL-17a mRNA expression levels by quantitative reverse transcriptase PCR (RT-PCR), to provide a first indication on the efficacy of **Bit-L15** in a cellular context. The inhibition of IL-17a mRNA expression was measured in EL4 cells, a murine lymphoblast cell line that constitutively expresses ROR $\gamma$ t. The EL4 cells were treated with 10  $\mu$ M of **MRL-871**, **MRL-L15** and **Bit-L15** for 24 h before the mRNA levels were measured (Figure 6.4). Both **MRL-871** and **Bit-L15** were active and potent in a cellular context. **MRL-871** significantly reduced IL-17a mRNA expression 20-fold, in line with previous reports.<sup>25</sup> **Bit-L15** led to a significant decrease in IL-17a expression as well (9.1-fold), demonstrating the desired effect not only in a direct biochemical assay, but also in a cellular context despite its non-drug-like chemical structure (high molecular weight, long linker and hydrophobic cholenic acid moiety). In contrast, **MRL-L15** showed only a minor reduction of IL-17a mRNA expression (1.3-fold) compared to **Bit-L15**, which shows that the coupling of the cholesterol pharmacophore to the allosteric pharmacophore results in a significantly higher response than for the monovalent allosteric counterpart alone. The results are in agreement with the results from the TR-FRET coactivator recruitment assays, where **MRL-871** and **Bit-L15** show a similar overall affinity for ROR $\gamma$ t, while **MRL-L15** is less potent ( $IC_{50}$  value of 0.0059  $\mu$ M for **Bit-L15** vs. 0.42  $\mu$ M for **MRL-L15**, Figure 6.3A). However, the increase in cellular efficacy of **Bit-L15** relative to **MRL-L15** might also be caused by active cellular uptake facilitated by the attachment of the cholesterol pharmacophore to **MRL-871**.<sup>43</sup>



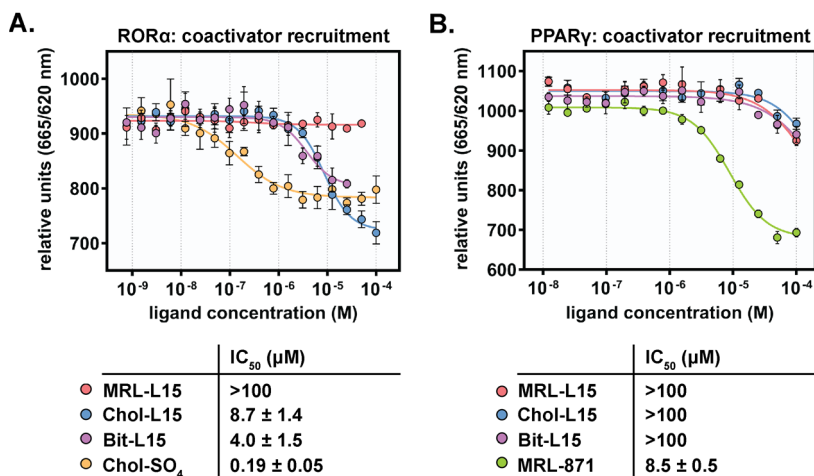
**Figure 6.4 | IL-17a mRNA expression in EL4 cells treated with ligands MRL-871, MRL-L15, Bit-L15 (10  $\mu$ M, 24 h) or DMSO and fold decrease of IL-17a expression relative to DMSO.** The level of IL-17a expression was normalized to that of GAPDH expression. Data are representative of two independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD. The relative gene expression was calculated by the  $2^{-\Delta\Delta C_t}$  (Livak) method using the DMSO control as a calibrator. Statistical analysis was performed using a one-way analysis of variance (ANOVA) compared against the DMSO control following Dunnett's *post hoc* test; \*\*\*  $P < 0.001$  and \*\*\*\*  $P < 0.0001$ .

### Bit-L15 is selective for ROR $\gamma$ t over ROR $\alpha$ and PPAR $\gamma$

In addition to an increased overall potency, the second major feature of a bitopic ligand over its monovalent orthosteric and allosteric ligands is an increased selectivity for its target by concomitant engagement of two sites.<sup>29,41</sup> Cholesterol and its derivatives are known to not only bind to ROR $\gamma$ t, but also to have cross-reactivity towards ROR $\alpha$  with high affinities.<sup>2</sup> In order to investigate the cross-reactivity of **Bit-L15** on ROR $\alpha$ , a TR-FRET coactivator recruitment assay was performed. Whereas sulfated cholesterol (Chol-SO<sub>4</sub>) showed a clear activity toward ROR $\alpha$  with an IC<sub>50</sub> value of 0.19  $\pm$  0.05  $\mu$ M, **Bit-L15** was more than 20 times less active in recruiting coactivator (IC<sub>50</sub> value of 4.0  $\pm$  1.5  $\mu$ M) (Figure 6.5A). These results show that **Bit-L15** exhibits some off-target activity on ROR $\alpha$ , but with a significant decrease in potency compared to sulfated cholesterol.

Furthermore, it has been demonstrated that **MRL-871** and its derivatives possess off-target effects against the peroxisome proliferator-activated receptor  $\gamma$  (PPAR $\gamma$ ).<sup>17,20</sup> In order to probe the cross-reactivity of the bitopic ligands on PPAR $\gamma$ , a TR-FRET assay was performed with **Bit-L15**, in the presence of the orthosteric PPAR $\gamma$  ligand rosiglitazone<sup>44</sup> (Figure 6.5B). In agreement with literature<sup>17,25</sup>, **MRL-871** showed an IC<sub>50</sub> value of 8.5  $\pm$  0.5  $\mu$ M (vs. 0.0074  $\pm$  0.0009  $\mu$ M for ROR $\gamma$ t), whereas **Bit-L15** showed an IC<sub>50</sub> value >100  $\mu$ M (vs. 0.0059  $\pm$  0.0007  $\mu$ M for ROR $\gamma$ t), demonstrating that **Bit-L15** shows a considerably lower cross-reactivity with PPAR $\gamma$  than **MRL-871**. Additionally, a competition experiment was performed with the

orthosteric PPAR $\gamma$  ligand tesaglitazar and increasing concentrations of the bitopic and monovalent ligands.<sup>45</sup> The data show that the competition between **MRL-871** and tesaglitazar is considerably higher than between **Bit-L15** and tesaglitazar (84 vs. 1.3 times maximal increase in IC<sub>50</sub> values, respectively), again indicating a lower activity of **Bit-L15** on PPAR $\gamma$  compared to **MRL-871** (Supporting Figure S6.10). These data clearly demonstrate how bitopic targeting can enhance the selectivity for a target.



**Figure 6.5 | B.** Dose-response curves from the TR-FRET coactivator recruitment assays on ROR $\alpha$  and PPAR $\gamma$ . A/B) Titration of **MRL-L15**, **Chol-L15**, **Bit-L15** and **Chol-SO<sub>4</sub>**/**MRL-871** to ROR $\alpha$  (A) and PPAR $\gamma$  (B), including an overview of the IC<sub>50</sub> values. (The last data points for **Bit-L15** and **MRL-L15** are not shown because of solubility issues at high concentrations.) Data are representative of two independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD.

## Conclusions

The recent discovery of the simultaneous binding of an orthosteric and allosteric ligand to the LBD of ROR $\gamma$ t inspired the design of bitopic ROR $\gamma$ t ligands that in theory could concomitantly occupy both the orthosteric and allosteric site of the protein. Compared to monovalent targeting strategies, successful bitopic targeting of various classes of proteins has been associated with advantages including an increased overall affinity and a higher selectivity profile. Bitopic targeting of ROR $\gamma$ t might therefore yield desirable molecular pharmacological properties. We report the design, synthesis and biochemical and cellular evaluation of three candidate bitopic ligands **Bit-L15**, **Bit-L9** and **Bit-L4**, connecting an orthosteric and allosteric ROR $\gamma$ t pharmacophore via a PEG linker that varies in linker length from four to fifteen PEG units.

A combination of TR-FRET assays was performed to probe different aspects of the mode of binding for all three bitopic ligands. A strong dependence of the overall affinity on linker length was observed, with both **Bit-L15** and **Bit-L9** showing an increase in potency relative to their individual monovalent counterparts, while **Bit-L4** showed a comparable potency as the allosteric monovalent counterpart. **Bit-L15** was the most potent of the bitopic ligands, matching the low nanomolar affinity of the allosteric inverse agonist **MRL-871**. Several follow-up assays confirmed that **Bit-L15** and **Bit-L9** bind both in the orthosteric and allosteric site simultaneously. Comparison of the bitopic ligands with equimolar amounts of co-incubated monovalent counterparts demonstrated that for **Bit-L15** and **Bit-L9**, a more than additive effect of both monovalent counterparts exists. For **Bit-L4** such a multivalent effect was, as expected, not observed. Combined, the TR-FRET data for **Bit-L15** reveal a bitopic binding mode and a multivalent character, illustrated by its higher potency compared to the (co-incubated) monovalent pharmacophores. **Bit-L9** also binds bitopically, but with a lower overall potency and less multivalent character. This could be related to an entropic penalty caused by tension in the shorter linker, upon bitopic binding. **Bit-L4** lacks the typical characteristics of bitopic binding and most probably binds via a 2:1 binding mode.

The most potent bitopic ligand, **Bit-L15**, was also evaluated in a cellular setting to explore functional efficacy. Indeed, **Bit-L15** showed a clear reduction of IL-17a levels, approaching the activity of **MRL-871**. Finally, the selectivity of **Bit-L15** was investigated in TR-FRET assays with PPAR $\gamma$  and ROR $\alpha$ . **Bit-L15** was found to be more than 20 times less active on ROR $\alpha$  than a sulfated cholesterol derivative, and displayed hardly any activity on PPAR $\gamma$  (IC<sub>50</sub> >100  $\mu$ M) compared to **MRL-871**, clearly demonstrating the enhancement of selectivity for ROR $\gamma$ t via a dual targeting strategy.

In conclusion, we have rationally designed three candidate bitopic ROR $\gamma$ t ligands; **Bit-L15**, **Bit-L9** and **Bit-L4**. Biochemical evaluation via various TR-FRET assays provides strong evidence that **Bit-L15** and **Bit-L9** bind to ROR $\gamma$ t in a bitopic manner, with **Bit-L15** showing the most promising characteristics. Furthermore, bitopic targeting results in an increased target selectivity, while retaining overall efficacy in both a biochemical and cellular context. Bitopic NR modulation thus positions itself as a highly promising alternative to monovalent strategies, as chemical biology tool compounds, or maybe even toward alternative NR targeting strategies. Future studies focusing on the drug-likeness of the bitopic ROR $\gamma$ t modulators (e.g. pharmacokinetic properties, cytotoxicity studies and elaboration on the selectivity profile) are required to determine the relevance of these ligands beyond a chemical biology point of view.

## Acknowledgements

Guido Oerlemans performed the major part of the experimental work described in this chapter, for which he is gratefully acknowledged.

This chapter was reprinted (adapted) with permission from *ACS Chem. Biol.* 2021, 16, 510-519. Copyright 2021 American Chemical Society. Further permissions related to this material should be directed to the ACS.

## Experimental Section

**In silico measurements.** In order to estimate the linker length required to bridge the distance between the orthosteric and allosteric site via the protein exterior of the ROR $\gamma$ t LBD, the molecular modeling environment Maestro (version 12.5, Schrödinger LLC). The protein crystal structure with PDB ID 6T4I was imported and prepared using the Protein Preparation Wizard. The chemical structure of both ligands was modified to incorporate the -CONH<sub>2</sub> moiety using the Build function. The measurement tool was used to estimate the length of the pathway from the nitrogen atom of the orthosteric ligand to the nitrogen atom of the allosteric ligand via H11 residues Arg482 and Glu481.

**General chemistry.** All solvents were supplied by Biosolve and used without further purification. Dry solvent was obtained from a MBRAUN Solvent Purification System (MB-SPS-800). Water was purified by a Millipore purification train. Deuterated solvents were obtained from Cambridge Isotope Laboratories. Solvents were removed *in vacuo* using a Büchi rotary evaporator and a diaphragm pump. All reagents were commercially available and were supplied by Sigma-Aldrich, Fluorochem and Iris Biochem GmbH. Proton (<sup>1</sup>H) NMR (400 MHz), carbon (<sup>13</sup>C) NMR (100 MHz), fluor (<sup>19</sup>F) NMR (400 MHz) and 2D NMR (400 MHz) were recorded on a Bruker Avance 400 MHz spectrometer. Proton spectra are referenced to tetramethyl silane (TMS). Carbon spectra are referenced to TMS or the solvent peak of the deuterated spectrum. NMR spectra are reported as follows: chemical shift ( $\delta$ ) in parts per million (ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublet, td = triplet of doublets), coupling constant (*J*) in Hertz (Hz) (if applicable) and integration (proton spectra only). Peak assignments are based on additional 2D NMR techniques (COSY, HMBC, HSQC). Analytical Liquid Chromatography coupled with Mass Spectrometry (LC-MS) was performed on a C<sub>4</sub> Jupiter SuC4300A 150 x 2.0 mm column using ultrapure water with 0.1% formic acid (FA) and acetonitrile with 0.1% FA, in general with a gradient of 5% to 100% acetonitrile over 10 min, connected to a Thermo Fisher LCQ Fleet Ion Trap Mass Spectrometer. The purity of the samples was assessed using a UV detector at 254 nm, or the TIC MS signal in case of low UV signal. Unless otherwise stated all final compounds were >95% pure as judged by HPLC. High resolution mass spectra (HRMS) were recorded using a Waters ACQUITY UPLC I-Class LC system coupled to a Xevo G2 Quadrupole Time of Flight (Q-TOF) mass spectrometer. Flash column chromatography was performed on a Grace Reveleris X2 machine using Grace pre-packed silica columns with 40  $\mu$ m particle size. For flash column chromatography, dry loading was performed. Manual column chromatography was performed using silica gel with a particle size of 60-200  $\mu$ m (60 Å). For manual chromatography, liquid loading was used. Reaction progress was monitored by thin-layer liquid chromatography (TLC) using Merck TLC silica gel 60 F254 plates. Visualization of the plates was achieved using an ultraviolet lamp ( $\lambda_{\text{max}}$  = 254 nm).

### Synthetic procedures

**Methyl 1-(2-chloro-6-(trifluoromethyl)benzoyl)-3-iodo-1H-indazole-6-carboxylate (3).** Thionyl chloride (4.92 g, 3.00 mL, 41.4 mmol) was added to a round bottom flask containing 2-chloro-(6-trifluoromethyl)-benzoic acid (185.9 mg, 0.83 mmol) under an inert atmosphere (argon). The reaction



mixture was heated to 60 °C and stirred for 2 h and the excess thionyl chloride was carefully removed *in vacuo*. The resulting product was dissolved in anhydrous CH<sub>2</sub>Cl<sub>2</sub> and the solution was added to a solution of 4-dimethylaminopyridine (103.0 mg, 0.84 mmol) and methyl 3-iodo-1H-indazole-6-carboxylate (250.9 mg, 0.83 mmol) in anhydrous CH<sub>2</sub>Cl<sub>2</sub> (2.5 mL). The mixture was stirred at room temperature for 5 min under positive argon pressure, followed by the dropwise addition of NEt<sub>3</sub> (0.23 mL, 1.66 mmol, 2.0 eq) under positive argon pressure. The mixture was stirred at room temperature overnight and partitioned between CH<sub>2</sub>Cl<sub>2</sub> and water (each 20 mL). The aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x) and the resulting organic phase was washed with brine, dried over anhydrous sodium sulfate, filtered and concentrated *in vacuo*. The crude residue was purified by column chromatography, eluting with 40% CH<sub>2</sub>Cl<sub>2</sub> in n-heptane, to furnish **3** as a white solid (305.3 mg, 86%). <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 9.32 (1H, s), 8.19 (1H, dd, *J* = 8.4, 1.6 Hz), 7.71-7.66 (2H, m), 7.61-7.56 (2H, m), 4.01 (3H, s); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ 166.22, 163.76, 138.49, 133.57, 133.03, 132.69, 132.65, 132.33 (q, *J* = 2.0 Hz), 131.16, 129.53 (q, *J* = 32.6 Hz), 126.85, 124.82 (q, *J* = 4.2 Hz), 122.82 (q, *J* = 273.0 Hz), 122.04, 117.04, 104.63, 52.70. LC-MS (ESI): calc. for C<sub>17</sub>H<sub>9</sub>ClF<sub>3</sub>IN<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>: 508.93, observed: 508.83 (*R*<sub>t</sub> = 7.78 min).

**Methyl 3-(4-(tert-butoxycarbonyl)phenyl)-1-(2-chloro-6-(trifluoromethyl)benzoyl)-1H-indazole-6-carboxylate (4).** Methyl 1-(2-chloro-6-(trifluoromethyl)benzoyl)-3-iodo-1H-indazole-6-carboxylate **3** (305.3 mg, 0.60 mmol), (4-(tert-butoxycarbonyl)phenyl)boronic acid (200.3 mg, 0.90 mmol), KOAc (177.0 mg, 1.80 mmol) and Pd(dppf)Cl<sub>2</sub> (70.0 mg, 0.096 mmol) were combined in an oven dried Schlenk flask. The Schlenk tube was evacuated and purged with argon gas three times. Subsequently, a mixture of 5:1 dioxane/water (13 mL) was added under positive argon pressure. The reaction mixture was stirred for 3 h at 90 °C. The reaction was cooled down to room temperature and partitioned between CH<sub>2</sub>Cl<sub>2</sub> and water. The aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x). The combined organic phase was washed with brine (3 x) and subsequently dried over MgSO<sub>4</sub>. The mixture was filtered and concentrated *in vacuo*. The crude residue was purified via flash column chromatography, eluting with 5-20% EtOAc in n-heptane to furnish **4** as a yellow solid (150.6 mg, 45%). <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 9.34 (1H, s), 8.22 (1H, d, *J* = 8.4 Hz), 8.12-8.05 (3H, m), 7.88 (2H, d, *J* = 8.3 Hz), 7.74-7.68 (2H, m), 7.61 (1H, app. t, *J* = 8.0 Hz), 4.03 (3H, s), 1.61 (9H, s); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ 166.42, 165.07, 164.66, 150.25, 140.16, 134.58, 133.21, 132.95, 132.91, 132.71, 131.64, 130.95, 129.98, 129.52 (q, *J* = 33.5 Hz), 127.97, 127.73, 126.87, 124.79 (q, *J* = 4.3 Hz), 122.93 (q, *J* = 273.9 Hz), 121.27, 117.46, 81.48, 52.63, 28.17. LC-MS (ESI): calc. for C<sub>28</sub>H<sub>22</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 559.12, observed: 559.08 (*R*<sub>t</sub> = 9.10 min).

**3-(4-(tert-butoxycarbonyl)phenyl)-1-(2-chloro-6-(trifluoromethyl)benzoyl)-1H-indazole-6-carboxylic acid (5).** Methyl 3-(4-(tert-butoxycarbonyl)phenyl)-1-(2-chloro-6-(trifluoromethyl)benzoyl)-1H-indazole-6-carboxylate **4** (150.6 mg, 0.27 mmol) was dissolved in 2.5 mL THF/water (1:1) and the solution was added to a round bottom flask containing LiOH.H<sub>2</sub>O (56.5 mg, 1.35 mmol). The mixture was stirred for 4 h at room temperature, after which the reaction mixture was partitioned between CH<sub>2</sub>Cl<sub>2</sub> and water. The resulting aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x) and the combined organic phase was washed with brine, dried over MgSO<sub>4</sub> and concentrated *in vacuo* to furnish **5** as a white solid (139.2 mg, 95%). <sup>1</sup>H-NMR (400 MHz, DMSO-*d*<sub>6</sub>): δ 9.18 (1H, s), 8.31 (1H, d, *J* = 6.2 Hz), 8.16 (1H, d, *J* = 7.6 Hz), 7.99-7.97 (3H, m), 7.92-7.90 (3H, m), 7.83 (1H, app. t, *J* = 7.7 Hz), 1.49 (9H, s); <sup>13</sup>C-NMR (100 MHz, DMSO-*d*<sub>6</sub>): δ 169.13, 164.28, 164.18, 150.02, 140.25, 139.82, 134.23, 133.69, 132.50, 132.30, 131.49, 129.77, 127.91, 127.72 (q, *J* = 32.5 Hz), 125.40 (q, *J* = 4.2 Hz), 125.24, 122.96 (q, *J* = 273.2 Hz), 120.91, 115.77, 81.00, 27.66. LC-MS (ESI): calc. for C<sub>27</sub>H<sub>20</sub>ClF<sub>3</sub>N<sub>2</sub>O<sub>5</sub> [M+H]<sup>+</sup>: 545.10, observed: 545.08 (*R*<sub>t</sub> = 8.07 min).

**Tert-butyl ((R)-19-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-16-oxo-3,6,9,12-tetraoxa-15-azaicosyl) carbamate (7a).** 5-Cholenic acid-3β-ol **6** (50.2 mg, 0.13 mmol) was dissolved in DMF (0.4 mL) in a round bottom flask. Subsequently, di-isopropylethylamine (DIPEA) (0.1 mL, 0.57 mmol) and a solution of 1-

[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU) (76.2 mg, 0.20 mmol) in DMF (0.5 mL) was added. The mixture was stirred for 10 min at room temperature, followed by the addition of a solution of *tert*-butyl (14-amino-3,6,9,12-tetraoxatetradecyl)carbamate (63.2 mg, 0.19 mmol) in DMF (0.7 mL). The reaction mixture was stirred for 4 h at room temperature, after which the mixture was partitioned between CH<sub>2</sub>Cl<sub>2</sub> and water. The resulting aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x) and the combined organic phase was washed with water (2 x) and brine, dried over anhydrous sodium sulfate and concentrated *in vacuo*. The crude residue was purified via column chromatography, eluting with 0-3% MeOH in CH<sub>2</sub>Cl<sub>2</sub> to furnish **7a** as a colourless and viscous oil (88.6 mg, 95%). <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 6.32 (1H, s), 5.36-5.32 (1H, m), 5.14 (1H, s), 3.70-3.59 (12H, m), 3.57-3.52 (4H, m), 3.52-3.47 (1H, m), 3.47-3.41 (2H, m), 3.35-3.28 (2H, m), 2.33-2.20 (3 H, m), 2.12-1.74 (7H, m), 1.62-1.23 (19H, m), 1.19-1.02 (4H, m), 1.00 (3H, s), 0.97-0.88 (4H, m), 0.68 (1H, s); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ 174.21, 156.58, 140.87, 121.54, 79.33, 71.63, 70.69, 70.43, 70.37, 70.23, 70.01, 69.93, 56.72, 55.87, 50.09, 42.36, 42.27, 40.28, 39.75, 39.08, 37.28, 36.49, 35.54, 33.43, 31.88, 31.75, 31.62, 28.40, 28.15, 24.27, 21.07, 19.40, 18.41, 11.88. LC-MS (ESI): calc. for C<sub>39</sub>H<sub>68</sub>N<sub>2</sub>O<sub>8</sub> [M+H]<sup>+</sup>: 693.50, observed 693.33 (R<sub>t</sub> = 7.34 min). HRMS (ESI): calc. for C<sub>39</sub>H<sub>68</sub>N<sub>2</sub>O<sub>8</sub> [M+H]<sup>+</sup>: 693.5054, observed: 693.5032.

***Tert*-butyl ((R)-34-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-31-oxo-3,6,9,12,15,18,21,24,27-nonaoxa-30-azapentatriacontyl)carbamate (7b).** 5-Cholenic acid-3β-ol **6** (50 mg, 0.13 mmol) was dissolved in DMF (0.4 mL) in a round bottom flask. Subsequently, DIPEA (0.1 mL, 0.57 mmol) and a solution of HATU (76.2 mg, 0.20 mmol) in DMF (0.5 mL) was added. The mixture was stirred for 10 min at room temperature, followed by the addition of a solution of *tert*-butyl (29-amino-3,6,9,12,15,18,21,24,27-nonaoxanonacosyl)carbamate (106.3 mg, 0.19 mmol) in DMF (0.5 mL). The reaction mixture was stirred for 4 h at room temperature, after which the mixture was partitioned between CH<sub>2</sub>Cl<sub>2</sub> and water. The resulting aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x) and the combined organic phase was washed with water (2 x) and brine, dried over MgSO<sub>4</sub> and concentrated *in vacuo*. The crude residue was purified via column chromatography, eluting with 2.5-10% MeOH in CH<sub>2</sub>Cl<sub>2</sub> to furnish **7b** as a colourless and viscous oil (104.6 mg, 86%). <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 6.31 (1H, s), 5.36-5.31 (1H, m), 5.17 (1H, s), 3.77-3.57 (32H, m), 3.59-3.51 (4H, m), 3.51-3.45 (1H, m), 3.45-3.39 (2H, m), 3.34-3.26 (2H, m), 2.34-2.15 (3H, m), 2.14-1.72 (7H, m), 1.62-1.28 (19H, m), 1.19-1.03 (4H, m), 1.00 (3H, s), 0.96-0.87 (4H, m), 0.67 (1H, s); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>): δ 173.66, 156.02, 140.84, 121.57, 79.13, 71.56, 70.56, 70.51, 70.24, 70.20, 69.99, 56.74, 55.88, 50.10, 42.36, 42.28, 40.37, 39.77, 39.15, 37.27, 36.49, 35.54, 33.51, 31.88, 31.75, 31.62, 28.44, 28.16, 24.26, 21.07, 19.40, 18.43, 11.89. LC-MS (ESI): calc. for C<sub>49</sub>H<sub>88</sub>N<sub>2</sub>O<sub>13</sub> [M+H]<sup>+</sup>: 913.63, observed 913.58 (R<sub>t</sub> = 7.30 min). HRMS (ESI): calc. for C<sub>49</sub>H<sub>88</sub>N<sub>2</sub>O<sub>13</sub> [M+H]<sup>+</sup>: 913.6365, observed: 913.6339.

***Tert*-butyl ((R)-52-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-49-oxo-3,6,9,12,15,18,21,24,27,30,33,36,39,42,45-pentadeca-48-azatripentacontyl) carbamate (7c).** 5-Cholenic acid-3β-ol **6** (32.1 mg, 0.086 mmol) was dissolved in DMF (0.3 mL) in a round bottom flask. Subsequently, DIPEA (0.06 mL, 0.34 mmol) and a solution of HATU (49.5 mg, 0.13 mmol) in DMF (0.5 mL) was added. The mixture was stirred for 10 min at room temperature, followed by the addition of a solution of *tert*-butyl (47-amino-3,6,9,12,15,18,21,24,27,30,33,36,39,42,45-pentadeca-48-heptatetracontyl)carbamate (70 mg, 0.085 mmol) in DMF (0.7 mL). The reaction mixture was stirred for 4 h at room temperature, after which the mixture was partitioned between CH<sub>2</sub>Cl<sub>2</sub> and water. The resulting aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x) and the combined organic phase was washed with water (2 x) and brine, dried over MgSO<sub>4</sub> and concentrated *in vacuo*. The crude residue was purified via column chromatography, eluting with 2% MeOH in CH<sub>2</sub>Cl<sub>2</sub> to furnish **7c** as a colourless and viscous oil (85.5 mg, 85%). <sup>1</sup>H-NMR (400 MHz,

CDCl<sub>3</sub>):  $\delta$  6.17 (1H, s), 5.36–5.42 (1H, m), 5.10 (1H, s), 3.84–3.57 (56H, m), 3.57–3.52 (4H, m), 3.52–3.46 (1H, m), 3.46–3.40 (2H, m), 3.33–3.26 (2H, m), 2.32–2.17 (3H, m), 2.16–1.78 (7H, m), 1.64–1.19 (19H, m), 1.19–1.02 (4H, m), 1.01 (3H, s), 0.96–0.85 (4H, m), 0.68 (1H, s); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  173.63, 156.01, 140.85, 121.15, 79.11, 71.61, 70.55, 70.51, 70.21, 69.97, 56.73, 55.88, 50.09, 42.36, 42.28, 40.36, 39.76, 39.15, 37.27, 36.49, 35.53, 33.51, 31.87, 31.74, 31.62, 28.43, 28.16, 24.26, 21.06, 19.40, 18.43, 11.89. LC-MS (ESI): calc. for C<sub>61</sub>H<sub>112</sub>N<sub>2</sub>O<sub>19</sub> [M+H]<sup>+</sup>: 1177.79, observed 1177.50 (R<sub>t</sub> = 7.19 min). HRMS (ESI): calc. for C<sub>61</sub>H<sub>112</sub>N<sub>2</sub>O<sub>19</sub> [M+H]<sup>+</sup>: 1177.7938, observed: 1177.7914.

**4-(1-(2-chloro-6-(trifluoromethyl)benzoyl)-6-(((R)-19-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H cyclopenta[a]phenanthren-17-yl)-16-oxo-3,6,9,12-tetraoxa-15-azaicosyl)carbamoyl)-1H-indazol-3-yl) benzoic acid (8a).** In a round bottom flask, **7a** (32.6 mg, 0.047 mmol) was dissolved in a mixture of CH<sub>2</sub>Cl<sub>2</sub>, TFA and water (65:30:5, 1 mL total volume) (*N*-Boc deprotection). The mixture was stirred for 3 h at room temperature, concentrated *in vacuo* and refluxed overnight in MeOH to yield the intermediate compound with free amine linker (TFA salt) (31.8 mg, quant.). Compound **5** (19.0 mg, 0.035 mmol) was then dissolved in DMF in a round bottom flask. Subsequently, DIPEA (0.024 mL, 0.24 mmol) and a solution of HATU (19.9 mg, 0.052 mmol) in DMF (0.5 mL) were added. The mixture was stirred for 10 min at room temperature, followed by the addition of the cholenic acid intermediate (20.65 mg, 0.035 mmol) in DMF (0.5 mL). The resulting reaction mixture was stirred for 5 h at room temperature. Subsequently, the reaction mixture was partitioned between CH<sub>2</sub>Cl<sub>2</sub> and water. The resulting aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x) and the combined organic phase was washed with water and with brine, dried over MgSO<sub>4</sub> and concentrated *in vacuo*. The crude residue was purified by column chromatography, eluting with 1–3% MeOH in CH<sub>2</sub>Cl<sub>2</sub> to furnish the intermediate bitopic ligand (19.0 mg, 49%). The intermediate (19.0 mg, 0.017 mmol) was dissolved in a mixture of CH<sub>2</sub>Cl<sub>2</sub>, TFA and water (65:30:5, 1 mL total volume) (deprotection of the *tert*-Butyl protected acid of the allosteric pharmacophore). The mixture was stirred for 3 h at room temperature, concentrated *in vacuo* and refluxed in MeOH overnight to yield the TFA salt of **8a** as a light yellow and viscous oil (17.9 mg, quant.). <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  9.04 (1H, s), 8.13–7.99 (4H, m), 7.86 (2H, d, *J* = 8.3 Hz), 7.75–7.69 (2H, m), 7.62 (1H, app. t, *J* = 8.0 Hz), 7.52–7.50 (1H, m), 6.31–6.25 (1H, m), 5.34–5.29 (1H, m), 3.80–3.37 (21H, m), 2.30–2.17 (3H, m), 2.09–1.73 (7H, m), 1.62–1.29 (10H, m), 1.14–1.00 (4H, m), 0.98 (3H, s), 0.90–0.85 (4H, m), 0.64 (3H, s); <sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  174.11, 168.74, 167.03, 164.79, 150.09, 140.74, 140.15, 136.57, 135.36, 133.02, 132.67, 131.02, 130.52, 129.44 (*q*, *J* = 32.5 Hz), 128.01, 126.57, 125.87, 124.85 (*q*, *J* = 4.7 Hz), 122.98 (*q*, *J* = 273.5 Hz), 121.63, 121.52, 114.13, 71.77, 70.59, 70.47, 70.42, 70.38, 70.28, 70.08, 70.07, 69.97, 56.70, 55.81, 50.07, 42.33, 42.25, 40.24, 39.73, 39.31, 37.25, 36.46, 35.52, 33.46, 31.85, 31.76, 31.60, 28.12, 24.24, 21.04, 19.38, 18.38, 11.85. LC-MS (ESI): calc. for C<sub>57</sub>H<sub>70</sub>ClF<sub>3</sub>N<sub>4</sub>O<sub>10</sub> [M+H]<sup>+</sup>: 1063.47, observed 1063.58 (R<sub>t</sub> = 7.73 min). HRMS (ESI): calc. for C<sub>57</sub>H<sub>70</sub>ClF<sub>3</sub>N<sub>4</sub>O<sub>10</sub> [M+H]<sup>+</sup>: 1063.4811, observed: 1063.4845.

**4-(1-(2-chloro-6-(trifluoromethyl)benzoyl)-6-(((R)-34-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H cyclopenta[a]phenanthren-17-yl)-31-oxo-3,6,9,12,15,18,21,24,27-nonaaxa-30 azapentatriacontyl)carbamoyl)-1H-indazol-3-yl)benzoic acid (8b).** In a round bottom flask, **7b** (45.4 mg, 0.050 mmol) was dissolved in a mixture of CH<sub>2</sub>Cl<sub>2</sub>, TFA and water (65:30:5, 1 mL total volume) (*N*-Boc deprotection). The mixture was stirred for 3 h at room temperature and concentrated *in vacuo* to yield the intermediate compound with free amine linker (TFA salt) (44.5 mg, quant.). Compound **5** (14.9 mg, 0.027 mmol) was then dissolved in DMF (0.3 mL) in a round bottom flask. Subsequently, DIPEA (0.02 mL, 0.11 mmol) and a solution of HATU (15.6 mg, 0.041 mmol) in DMF (0.3 mL) were added. The mixture was stirred for 10 min at room temperature, followed by the addition of the cholenic acid intermediate (31.2 mg, 0.033 mmol) in DMF (0.3 mL). The resulting reaction mixture was stirred for 3 h at room temperature. Subsequently, the reaction mixture was partitioned between CH<sub>2</sub>Cl<sub>2</sub> and water. The resulting aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x) and

the combined organic phase was washed with water and brine, dried over  $\text{MgSO}_4$  and concentrated *in vacuo*. The crude residue was purified by column chromatography, eluting with 1-5% MeOH in  $\text{CH}_2\text{Cl}_2$  to furnish the intermediate bitopic ligand (20.3 mg, 56%). The intermediate (31.3 mg, 0.024 mmol) was dissolved in a mixture of  $\text{CH}_2\text{Cl}_2$ , TFA and water (65:30:5, 1 mL total volume) (deprotection of the *tert*-Butyl protected acid of the allosteric pharmacophore). The mixture was stirred for 3 h at room temperature, concentrated *in vacuo* and refluxed in MeOH overnight to yield the TFA salt of **8b** as a light yellow and viscous oil (28.8 mg, quant.).  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  9.07 (1H, s), 8.14 (2H, d,  $J$  = 8.0 Hz), 8.05 (2H, s), 7.88 (2H, d,  $J$  = 7.9 Hz), 7.75-7.69 (2H, m), 7.62 (1H, app. t,  $J$  = 9.3 Hz), 7.51-7.45 (1H, m), 6.37-6.29 (1H, m), 5.34-5.30 (1H, m), 3.77-3.52 (3H, m), 3.50-3.47 (1H, m), 3.47-3.43 (2H, m), 2.31-2.18 (3H, m), 2.15-1.67 (7H, m), 1.61-1.30 (10H, m), 1.17-1.00 (4H, m), 0.99 (3H, s), 0.97-0.90 (4H, m), 0.65 (3H, s);  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  174.12, 168.8, 167.00, 164.71, 150.16, 140.77, 140.18, 136.63, 135.12, 133.00, 132.68, 130.98, 130.51, 129.44 (q,  $J$  = 32.5 Hz), 128.02, 126.59, 125.63, 124.82 (q,  $J$  = 4.7 Hz), 122.96 (q,  $J$  = 273.5 Hz), 121.60, 121.50, 114.40, 71.73, 70.59, 70.55, 70.47, 70.45, 70.43, 70.39, 70.28, 70.15, 70.01, 69.85, 56.72, 55.80, 50.09, 42.34, 42.24, 40.22, 39.75, 39.21, 37.26, 36.46, 35.51, 33.43, 31.85, 31.78, 31.58, 28.14, 24.25, 21.05, 19.39, 18.41, 11.87. LC-MS (ESI): calc. for  $\text{C}_{67}\text{H}_{90}\text{ClF}_3\text{N}_4\text{O}_{15}$   $[\text{M}+\text{H}]^+$ : 1283.60, observed 1283.75 ( $R_t$  = 7.68 min). HRMS (ESI): calc. for  $\text{C}_{67}\text{H}_{90}\text{ClF}_3\text{N}_4\text{O}_{15}$   $[\text{M}+\text{H}]^+$ : 1283.6122, observed: 1283.6069.

4-(1-(2-chloro-6-(trifluoromethyl)benzoyl)-6-(((R)-52-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H cyclopenta[a]phenanthren-17-yl)-49-oxo-3,6,9,12,15,18,21,24,27,30,33,36,39,42,45-pentadeca-4,8-azatripentacontyl)carbamoyl)-1H-indazol-3-yl)benzoic acid (**8c**). In a round bottom flask, **7c** (60 mg, 0.051 mmol) was dissolved in a mixture of  $\text{CH}_2\text{Cl}_2$ , TFA and water (65:30:5, 1 mL total volume) (*N*-Boc deprotection). The mixture was stirred for 3 h at room temperature and concentrated *in vacuo* to yield the intermediate compound with free amine linker (TFA salt) (57.2 mg, quant.). Compound **5** (40.2 mg, 0.07 mmol) was then dissolved in DMF (0.3 mL) in a round bottom flask. Subsequently, DIPEA (0.05 mL, 0.29 mmol) and a solution of HATU (42.1 mg, 0.11 mmol) in DMF (0.3 mL) were added. The mixture was stirred for 10 min at room temperature, followed by the addition of the cholenic acid intermediate (79.4 mg, 0.067 mmol) in DMF (0.3 mL). The resulting reaction mixture was stirred for 4 h at room temperature. Subsequently, the reaction mixture was partitioned between  $\text{CH}_2\text{Cl}_2$  and water. The resulting aqueous phase was extracted with  $\text{CH}_2\text{Cl}_2$  (2 x) and the combined organic phase was washed with water and brine, dried over  $\text{MgSO}_4$  and concentrated *in vacuo*. The crude residue was purified by column chromatography, eluting with 1-5% MeOH in  $\text{CH}_2\text{Cl}_2$  to furnish the intermediate bitopic ligand (53.2 mg, 45%). The intermediate (55.1 mg, 0.034 mmol) was dissolved in a mixture of  $\text{CH}_2\text{Cl}_2$ , TFA and water (65:30:5, 1 mL total volume) (deprotection of the *tert*-Butyl protected acid of the allosteric pharmacophore). The mixture was stirred for 3 h at room temperature, concentrated *in vacuo* and refluxed in MeOH overnight to yield the TFA salt of **8c** as a light yellow and viscous oil (50 mg, quant.).  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  9.06 (1H, s), 8.14 (2H, d,  $J$  = 8.4 Hz), 8.10-8.03 (2H, m), 7.87 (2H, d,  $J$  = 8.4 Hz), 7.75-7.68 (2H, m), 7.62 (1H, app. t,  $J$  = 8.0 Hz), 7.54-7.46 (1H, m), 6.36-6.28 (1H, m), 5.36-5.28 (1H, m), 3.80-3.52 (6H, m), 3.52-3.48 (1H, m), 3.48-3.43 (2H, m), 2.31-2.13 (3H, m), 2.12-1.73 (7H, m), 1.60-1.28 (10H, m), 1.18-1.04 (4H, m), 0.99 (3H, s), 0.97-0.88 (4H, m), 0.66 (3H, s);  $^{13}\text{C-NMR}$  (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  174.05, 168.27, 166.95, 164.72, 150.19, 140.78, 140.18, 136.63, 135.15, 133.01, 132.68, 130.99, 130.48, 129.44 (q,  $J$  = 32.5 Hz), 128.04, 126.60, 125.70, 124.81 (q,  $J$  = 4.7 Hz), 122.96 (q,  $J$  = 273.5 Hz), 121.62, 121.55, 114.35, 71.73, 70.61, 70.50, 70.46, 70.43, 70.41, 70.39, 70.31, 70.18, 69.96, 69.84, 56.73, 55.83, 50.09, 42.36, 42.25, 40.22, 39.75, 39.21, 37.26, 36.48, 35.54, 33.48, 31.86, 31.79, 31.59, 28.15, 24.26, 21.06, 19.39, 18.43, 11.88. LC-MS (ESI): calc. for  $\text{C}_{79}\text{H}_{114}\text{ClF}_3\text{N}_4\text{O}_{21}$   $[\text{M}+\text{H}]^+$ : 1547.76, observed 1547.83 ( $R_t$  = 7.68 min). HRMS (ESI): calc. for  $\text{C}_{79}\text{H}_{114}\text{ClF}_3\text{N}_4\text{O}_{21}$   $[\text{M}+\text{H}]^+$ : 1547.7694, observed: 1547.7714.

**4-(6-((14-amino-3,6,9,12-tetraoxatetradecyl)carbamoyl)-1-(2-chloro-6-(trifluoro-methyl) benzoyl)-1H-indazol-3-yl)benzoic acid (9a).** Compound **5** (40.0 mg, 0.073 mmol) was dissolved in DMF (0.3 mL) in a round bottom flask. Subsequently, DIPEA (0.06 mL, 0.34 mmol) and a solution of HATU (42.5 mg, 0.11 mmol) in DMF (0.3 mL) were added. The mixture was stirred for 10 min at room temperature, followed by the addition of a solution of *tert*-butyl (14-amino-3,6,9,12-tetraoxatetradecyl)carbamate (35.2 mg, 0.10 mmol) in DMF (0.3 mL). The resulting reaction mixture was stirred for 4 h at room temperature. Subsequently, the reaction mixture was partitioned between CH<sub>2</sub>Cl<sub>2</sub> and water. The resulting aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x) and the combined organic phase was washed with water (2 x) and brine, dried over MgSO<sub>4</sub> and concentrated in vacuo. The crude residue was purified by column chromatography, eluting with 0-2% MeOH in CH<sub>2</sub>Cl<sub>2</sub> to furnish the intermediate **MRL-871** derivative (20.6 mg, 32%). The intermediate (20.6 mg, 0.024 mmol) was dissolved in a mixture of CH<sub>2</sub>Cl<sub>2</sub>, TFA and water (65:30:5, 1 mL total volume) (*N*-Boc deprotection). The mixture was stirred for 3 h at room temperature and concentrated *in vacuo* to yield the TFA salt of **9a** as a light yellow and viscous oil (19.3 mg, quant.). <sup>1</sup>H-NMR (400 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  9.07 (1H, t, *J* = 5.3 Hz), 9.03 (1H, s), 8.37 (1H, d, *J* = 8.5 Hz), 8.14-8.09 (3H, m), 8.07 (1H, d, *J* = 7.8 Hz), 8.04-7.95 (3H, m), 7.90 (1H, app. t, *J* = 8.2 Hz), 7.82 (2H, br. s), 3.65-3.50 (18H, m), 2.98 (2H, q, *J* = 5.3 Hz); <sup>13</sup>C-NMR (100 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  166.71, 165.51, 164.29, 150.12, 139.46, 136.71, 133.84, 133.78, 132.57, 132.27, 131.97 (q, *J* = 2.1 Hz), 131.43, 130.17, 128.06, 127.66 (q, *J* = 32.1 Hz), 125.76, 125.61, 125.57 (q, *J* = 4.1 Hz), 122.94 (q, *J* = 273.5 Hz), 122.26, 113.89, 69.76, 69.73, 69.65, 69.61, 69.59, 68.78, 66.65, 39.91, 38.60. LC-MS (ESI): calc. for C<sub>33</sub>H<sub>34</sub>ClF<sub>3</sub>N<sub>4</sub>O<sub>8</sub> [M+H]<sup>+</sup>: 707.20, observed: 707.42 (*R*<sub>t</sub> = 4.72 min). HRMS (ESI): calc. for C<sub>33</sub>H<sub>34</sub>ClF<sub>3</sub>N<sub>4</sub>O<sub>8</sub> [M+H]<sup>+</sup>: 707.2093, observed: 707.2106.

**4-(6-((29-amino-3,6,9,12,15,18,21,24,27-nonaoxanonacosyl)carbamoyl)-1-(2-chloro-6-(trifluoro-methyl)benzoyl)-1H-indazol-3-yl)benzoic acid (9b).** Compound **5** (25.2 mg, 0.046 mmol) was dissolved in DMF (0.3 mL) in a round bottom flask. Subsequently, DIPEA (0.035 mL, 0.20 mmol) and a solution of HATU (28.0 mg, 0.074 mmol) in DMF (0.3 mL) were added. The mixture was stirred for 10 min at room temperature, followed by the addition of a solution of *tert*-butyl (29-amino-3,6,9,12,15,18,21,24,27-nonaoxanonacosyl)carbamate (36.6 mg, 0.066 mmol) in DMF (0.3 mL). The resulting reaction mixture was stirred for 4 h at room temperature. Subsequently, the reaction mixture was partitioned between CH<sub>2</sub>Cl<sub>2</sub> and water. The resulting aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x) and the combined organic phase was washed with water (2 x) and brine, dried over MgSO<sub>4</sub> and concentrated in vacuo. The crude residue was purified by column chromatography, eluting with 1-5% MeOH in CH<sub>2</sub>Cl<sub>2</sub> to furnish the intermediate **MRL-871** derivative (25.3 mg, 51%). The intermediate (25.3 mg, 0.023 mmol) was dissolved in a mixture of CH<sub>2</sub>Cl<sub>2</sub>, TFA and water (65:30:5, 1 mL total volume) (*N*-Boc deprotection). The mixture was stirred for 3 h at room temperature and concentrated in vacuo to yield the TFA salt of **9b** as a light yellow and viscous oil (23.9 mg, quant.). <sup>1</sup>H-NMR (400 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  9.06 (t, *J* = 5.4 Hz, 1H), 9.04 (1H, s), 8.37 (1H, d, *J* = 8.6 Hz), 8.15-8.09 (3H, m), 8.07 (1H, d, *J* = 7.6 Hz), 8.03-7.95 (3H, m), 7.90 (1H, app. t, *J* = 8.0 Hz), 7.79 (2H, br. s), 3.65-3.47 (38H, m), 3.00 (2H, q, *J* = 5.4 Hz); <sup>13</sup>C-NMR (100 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  166.75, 165.50, 164.30, 150.13, 139.49, 136.74, 133.86, 133.83, 132.58, 132.28, 132.01 (q, *J* = 2.0 Hz), 131.47, 130.20, 128.08, 127.69 (q, *J* = 32.3 Hz), 125.78, 125.65, 125.58 (q, *J* = 4.3 Hz), 122.96 (q, *J* = 273.8 Hz), 122.26, 113.94, 69.81, 69.75, 69.67, 69.63, 68.81, 66.68, 39.91, 38.65. LC-MS (ESI): calc. for C<sub>43</sub>H<sub>54</sub>ClF<sub>3</sub>N<sub>4</sub>O<sub>13</sub> [M+H]<sup>+</sup>: 927.33, observed: 927.58 (*R*<sub>t</sub> = 4.87 min). HRMS (ESI): calc. for C<sub>43</sub>H<sub>54</sub>ClF<sub>3</sub>N<sub>4</sub>O<sub>13</sub> [M+H]<sup>+</sup>: 927.3406, observed: 927.3395.

**4-(6-((47-amino-3,6,9,12,15,18,21,24,27,30,33,36,39,42,45-pentadecaaxaheptatetracontyl)carbamoyl)-1-(2-chloro-6-(trifluoromethyl)benzoyl)-1H-indazol-3-yl)benzoic acid (9c).** Compound **5** (20 mg, 0.037 mmol) was dissolved in DMF (0.3 mL) in a round bottom flask. Subsequently, DIPEA (0.03 mL, 0.17 mmol) and a solution of HATU (21 mg, 0.055 mmol) in DMF (0.3 mL) were added. The mixture was stirred for 10 min at room temperature, followed by the addition of a solution of *tert*-butyl(47-amino-

3,6,9,12,15,18,21,24,27,30,33,36,39,42,45-pentadeca-oxaheptatetra-contyl)carbamate (30 mg, 0.037 mmol) in DMF (0.3 mL). The reaction mixture was stirred for 4 h at room temperature. Subsequently, the reaction mixture was partitioned between CH<sub>2</sub>Cl<sub>2</sub> and water. The resulting aqueous phase was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x) and the combined organic phase was washed with water (2 x) and brine, dried over MgSO<sub>4</sub> and concentrated in vacuo. The crude residue was purified by column chromatography, eluting with 2-5% MeOH in CH<sub>2</sub>Cl<sub>2</sub> to furnish the intermediate **MRL-871** derivative (28.5 mg, 58%). In a round bottom flask, the intermediate (28.5 mg, 0.021 mmol) was dissolved in a mixture of CH<sub>2</sub>Cl<sub>2</sub>, TFA and water (65:30:5, 1 mL total volume) (*N*-Boc deprotection). The mixture was stirred for 3 h at room temperature and concentrated in vacuo to yield the TFA salt of **9c** as a light yellow and viscous oil (27.3 mg, quant.). <sup>1</sup>H-NMR (400 MHz, DMSO-d<sub>6</sub>): δ 9.06 (1H, t, *J* = 5.4 Hz), 9.04 (1H, s), 8.37 (1H, d, *J* = 8.7 Hz), 8.14-8.08 (3H, m), 8.07 (1H, d, *J* = 7.7 Hz), 8.04-7.96 (3H, m), 7.90 (1H, app. t, *J* = 8.0 Hz), 7.79 (2H, br. s), 3.64-3.45 (62H, m), 3.00 (2H, q, *J* = 5.4 Hz); <sup>13</sup>C-NMR (100 MHz, DMSO-d<sub>6</sub>): δ 166.67, 165.50, 164.31, 150.14, 139.50, 136.75, 133.88, 133.84, 132.60, 132.29, 132.03 (q, *J* = 2.0 Hz), 131.47, 130.20, 128.10, 127.69 (q, *J* = 32.3 Hz), 125.78, 125.66, 125.58 (q, *J* = 4.3 Hz), 122.96 (q, *J* = 273.8 Hz), 122.27, 113.94, 69.76, 69.67, 69.63, 68.81, 66.69, 39.52, 38.65. LC-MS (ESI): calc. for C<sub>55</sub>H<sub>74</sub>ClF<sub>3</sub>N<sub>4</sub>O<sub>19</sub> [M+H]<sup>+</sup>: 1191.49, observed: 1191.67 (*R*<sub>t</sub> = 5.00 min). HRMS (ESI): calc. for C<sub>55</sub>H<sub>74</sub>ClF<sub>3</sub>N<sub>4</sub>O<sub>19</sub> [M+H]<sup>+</sup>: 1191.4979, observed: 1191.4946.

### Biophysical assays

**ROR $\gamma$ t LBD expression and purification.** His<sub>6</sub>-ROR $\gamma$ t-LBD was expressed and purified as described in Chapter 2.

**TR-FRET coactivator recruitment assays, competition assays and AlexaFluor-MRL-871 recruitment assays on ROR $\gamma$ t.** TR-FRET assays and data analysis were performed as described in Chapter 2. Data were recorded in triplicate; error shown is standard deviation from the mean; curves are representative of > 2 independent experiments.

**Ligation of ROR $\gamma$ t with covalent probe 20.** Ligation was conducted using 20  $\mu$ M His<sub>6</sub>-ROR $\gamma$ t-LBD in buffer containing 50 mM TRIS (pH 5.8), 50 mM NaCl, 5 mM CHAPS and 0.1% Bovine Serum Albumin (BSA). A mix of the covalent probe **20** (2-chloro-5-nitro-*N*-(2-(trifluoromethyl)phenyl)-benzamide, Chapter 5) (5 eq, 100  $\mu$ M) and His<sub>6</sub>-ROR $\gamma$ t-LBD (20  $\mu$ M) was incubated (continuous mixing) at room temperature overnight, in a total volume of 50  $\mu$ L. Ligation was verified with Q-TOF MS (see Chapter 5).

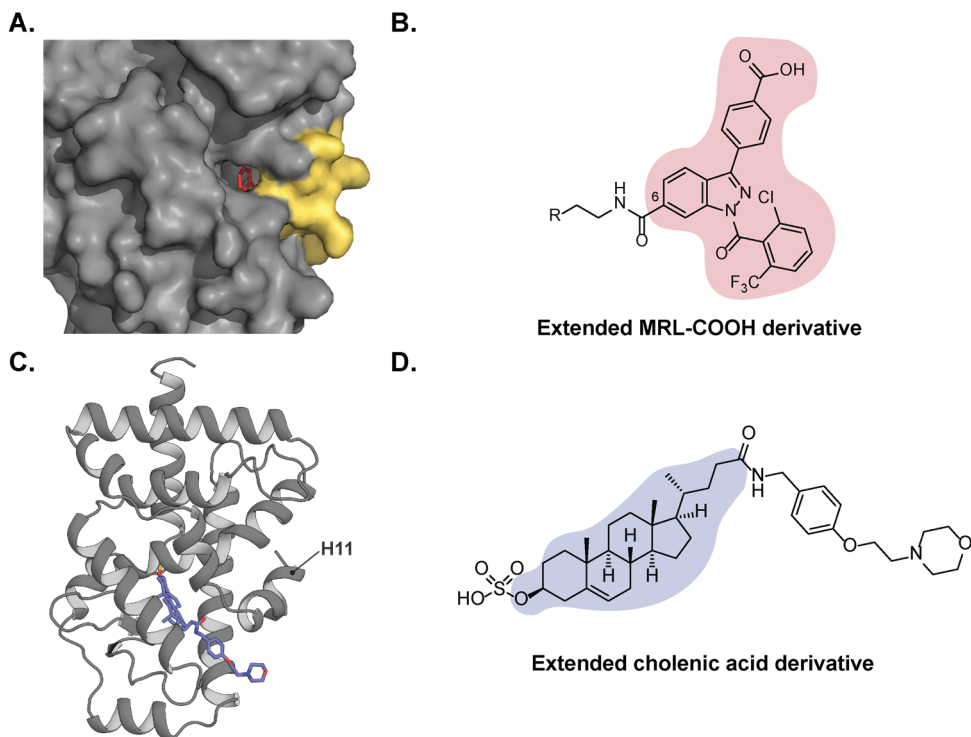
**Quantitative IL-17a mRNA RT-PCR assay.** Cell culture, RT-PCR experiments and data analysis were performed as described in Chapter 2. Statistical analysis was performed using a one-way analysis of variance (ANOVA) comparing against the DMSO control following Dunnett's *post hoc* test (GraphPad Prism 7.0 software). A *P* value < 0.05 was considered statistically significant. Data were recorded in triplicate; error shown is standard deviation from the mean; data are representative of two independent experiments.

### Selectivity TR-FRET coactivator recruitment assays

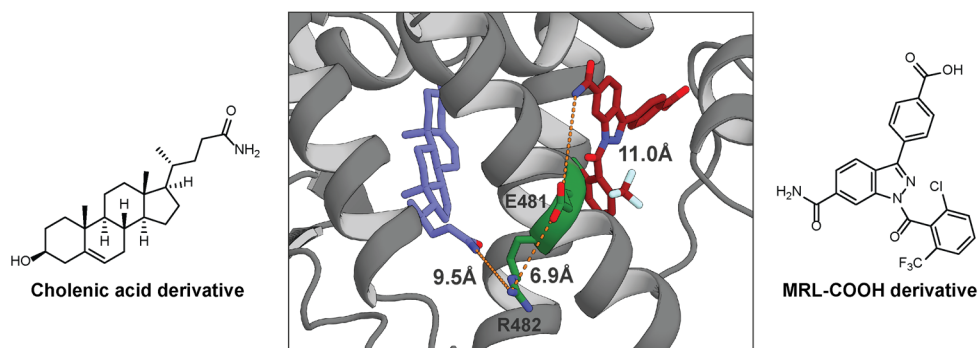
**TR-FRET coactivator recruitment assay on PPAR $\gamma$ .** TR-FRET assays were performed in an analogous fashion to that described above, only using 100 nM His<sub>6</sub>-PPAR $\gamma$ -LBD instead of His<sub>6</sub>-ROR $\gamma$ t-LBD. The assay with rosiglitazone was performed in the presence of 1  $\mu$ M rosiglitazone (in order to initially activate PPAR $\gamma$ ). The assay with tesaglitazar was performed in the presence of 0, 1, 10 or 50  $\mu$ M **MRL-871**, **MRL-L15** or **Bit-L15**, titrating tesaglitazar.

**TR-FRET coactivator recruitment assay on ROR $\alpha$ .** TR-FRET assays were performed in an analogous fashion to that described above, using 25 nM His<sub>6</sub>-ROR $\alpha$ -LBD instead of 20 nM His<sub>6</sub>-ROR $\gamma$ t-LBD.

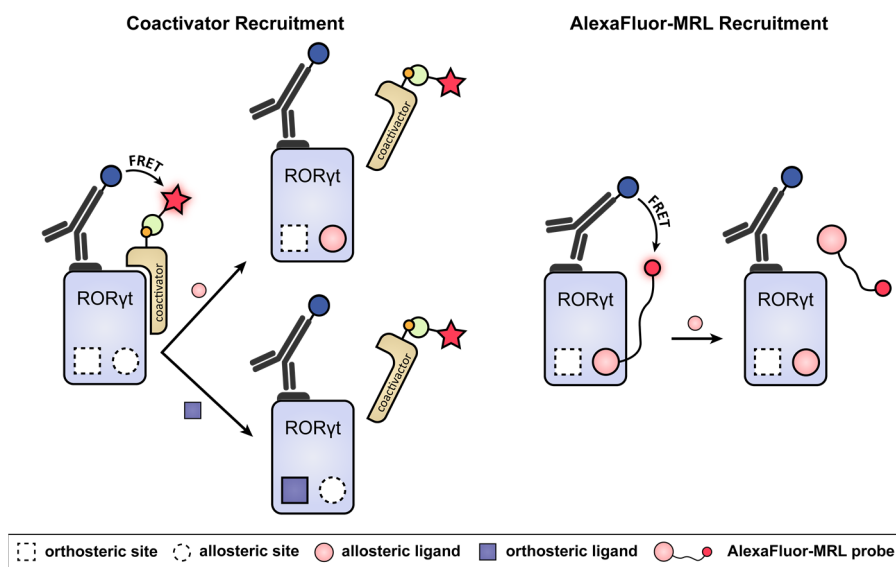
## Supporting Information



**Figure S6.1 | Detailed representation of the pharmacophores used in the design of the bitopic ligands.** **A)** Surface representation of ROR $\gamma$ t in complex with both **MRL-871** (red) and cholesterol (not visible) (PDB: 6T4I), showing the surface accessible C-6 position of the indazole ring of **MRL-871**. **B)** **MRL-COOH** extended with a linker at the carboxylic acid moiety, via amide coupling chemistry. **C)** Crystal structure of the extended cholenic acid derivative (purple) in complex with ROR $\gamma$ t (PDB: 5NTK) showing how H11 is displaced to the position where normally the allosteric pocket is formed. **D)** Chemical structure of the extended cholenic acid derivative as shown in the crystal structure in C.



**Figure S6.2 | Estimation of the distance between the orthosteric and allosteric RORyt pharmacophores.** Crystal structure of RORyt in complex with cholesterol and **MRL-871** (PDB: 6T4I) where cholenic acid and **MRL-COOH** derivatives are built in (see chemical structures). The shortest path that could be followed by a linker over the surface of the protein, via residues Glu481 and Arg482 (shown in green), has been indicated (using Maestro 12.5). Distances are shown as orange dashed lines and the numbers indicate the distances in angstrom.



**Figure S6.3 | Schematic representation of the TR-FRET assays.** Left: TR-FRET coactivator recruitment assay. When RORyt is in its apo or agonist-bound state, the LBD binds to the coactivator, resulting in FRET pairing of an anti-His terbium cryptate donor with the d2-labelled streptavidin acceptor. Inverse agonist binding (orthosteric or allosteric) results in coactivator displacement thus a lower FRET pairing. Right: TR-FRET AlexaFluor-MRL recruitment assay. When the probe binds to the RORyt LBD, there is FRET pairing between the anti-His terbium cryptate donor and the AlexaFluor647-labelled MRL-871 probe. Allosteric inverse agonist binding results in probe displacement thus a lower FRET pairing.



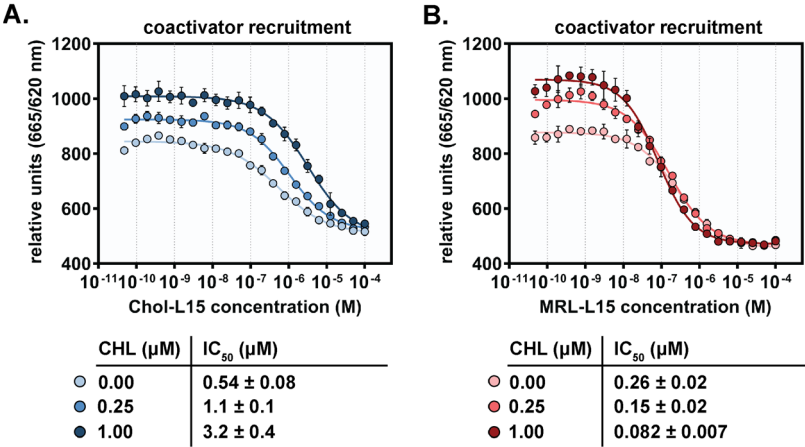


Figure S6.4 | Dose-response curves from the competitive TR-FRET coactivator recruitment assays on ROR $\gamma$ t. A/B) Titration of Chol-L15 (A) and MRL-L15 (B) in the presence of fixed concentrations of cholesterol (CHL) (0.00  $\mu$ M, 0.25  $\mu$ M and 1.00  $\mu$ M).

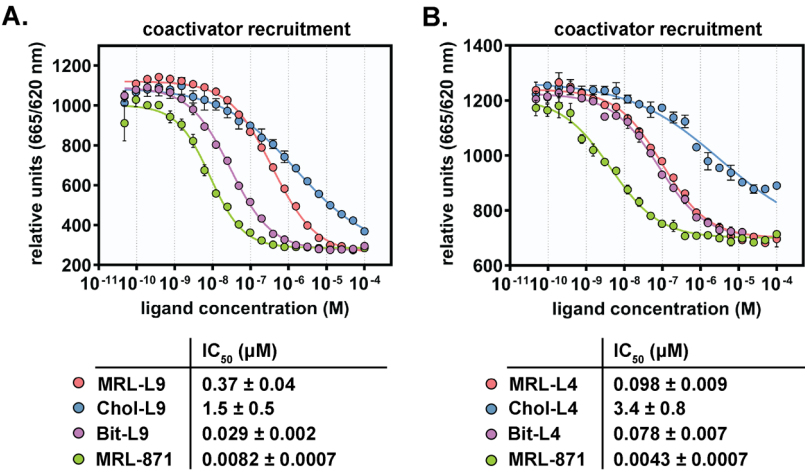


Figure S6.5 | Dose-response curves from the TR-FRET coactivator recruitment assays on ROR $\gamma$ t. A/B) Titration of MRL-871, MRL-L9, Chol-L9 and Bit-L9 (A), and MRL-871, MRL-L4, Chol-L4 and Bit-L4 (B). (The last three data points for Bit-L4 are not shown because of solubility issues at high concentrations.)

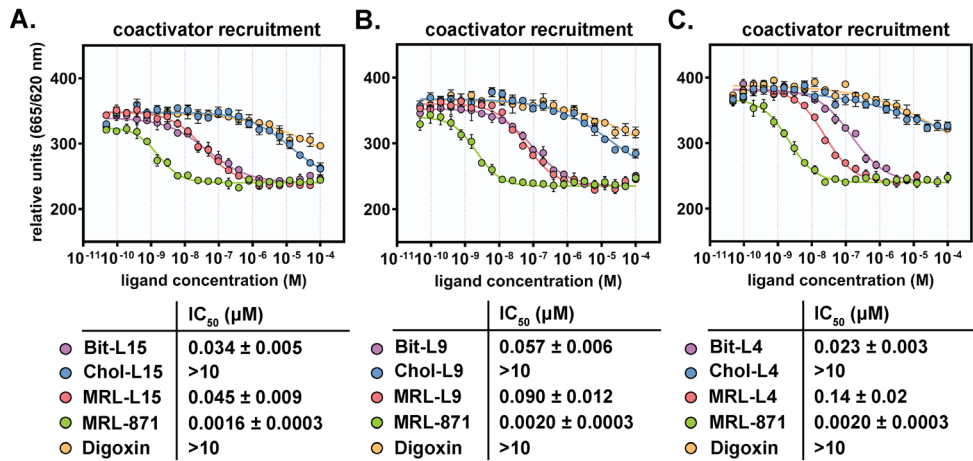


Figure S6.6 | Dose-response curves for all bitopic and monovalent ligands and digoxin in TR-FRET coactivator recruitment assays on orthosterically blocked ROR $\gamma$ t. (The last three data points for Bit-L4 are not shown because of solubility issues at high concentrations.)

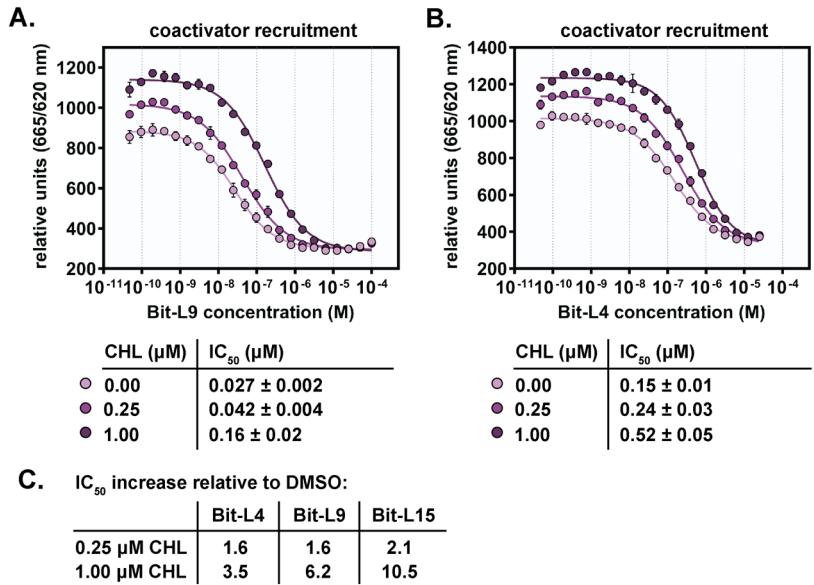
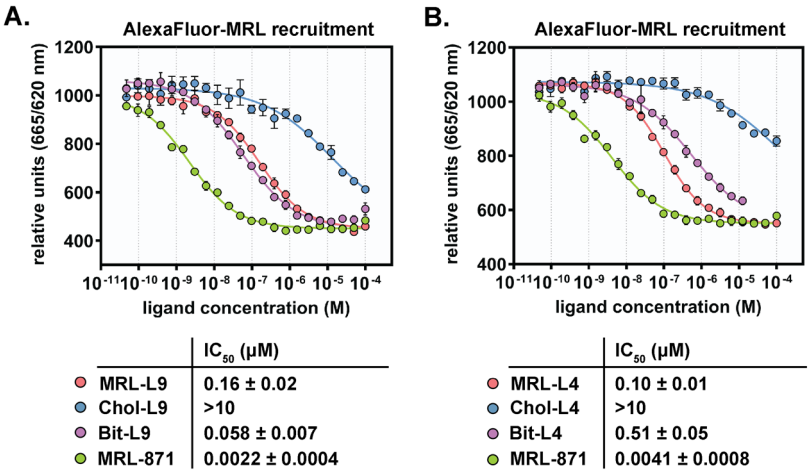
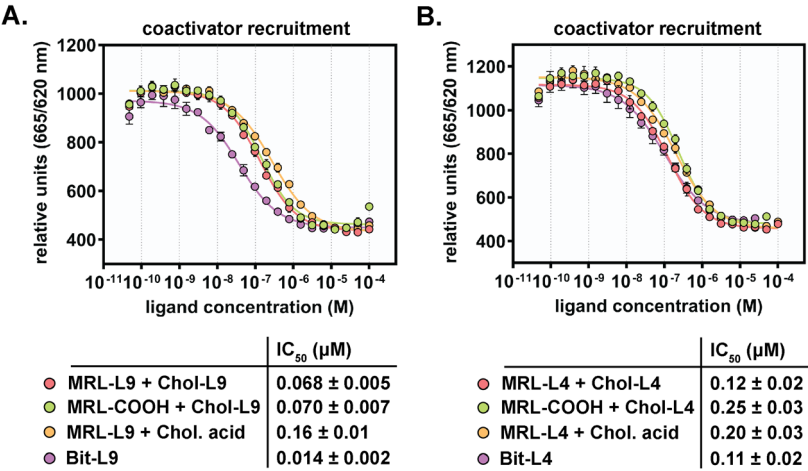


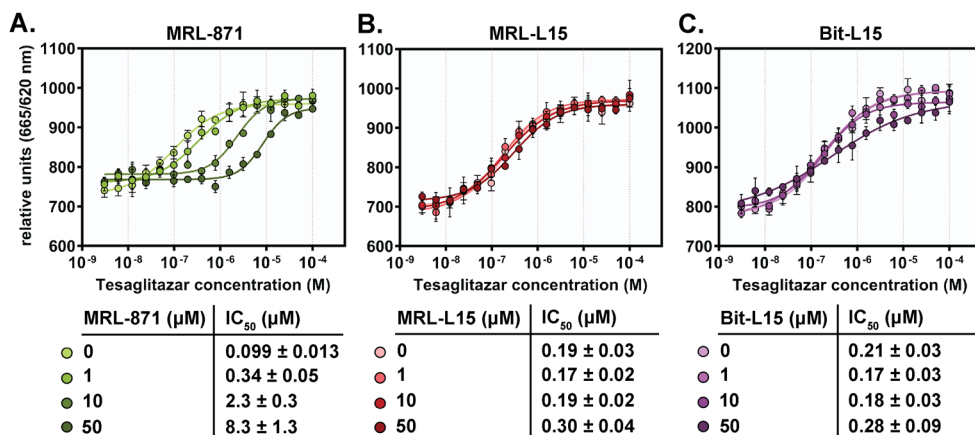
Figure S6.7 | Dose-response curves from the competitive TR-FRET coactivator recruitment assays on ROR $\gamma$ t. A/B) Titration of Bit-L9 (A) and Bit-L4 (B) in the presence of a fixed concentration of cholesterol (CHL) (0.00 μM, 0.25 μM and 1.00 μM). (The last two data points for Bit-L4 are not shown because of solubility issues at high concentrations.) C) The increase of the IC<sub>50</sub> values in the presence of cholesterol (relative to DMSO) is shown in a table for all bitopic ligands.



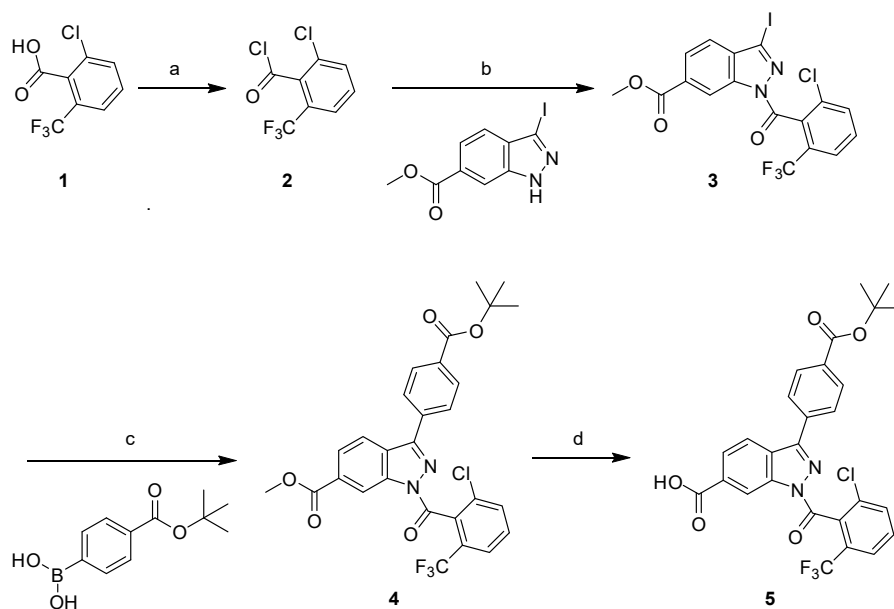
**Figure S6.8 |** Dose-response curves from the TR-FRET AlexaFluor-MRL recruitment assays on ROR $\gamma$ t. A/B) Titration of MRL-871, MRL-L9, Chol-L9 and Bit-L9 (A), and MRL-871, MRL-L4, Chol-L4 and Bit-L4 (B). (The last three data points for Bit-L4 are not shown because of solubility issues at high concentrations.)



**Figure S6.9 |** Dose-response curves from the TR-FRET coactivator recruitment assays on ROR $\gamma$ t, to probe the presence of a multivalent effect. A/B) Titration of Bit-L9 (A) and Bit-L4 (B) and equimolar amounts of co-incubated monovalent counterparts. Abbreviations: chol. acid, cholenic acid. (The last two data points for Bit-L4 and the last data point for MRL-COOH + Chol-L4 are not shown because of solubility issues at high concentrations.)



**Figure S6.10 |** Dose-response curves from the competitive TR-FRET coactivator recruitment assays on PPAR $\gamma$ . A-C) Titration of tesaglitazar in the presence of 0, 1, 10 and 50  $\mu\text{M}$  MRL-871 (A), MRL-L15 (B) and Bit-L15 (C).



**Scheme S6.1 |** Synthesis of MRL-871 derivative 5. Reagents and conditions: (a)  $\text{SOCl}_2$ ,  $60^\circ\text{C}$ , 3 h; (b) indazole, DMAP,  $\text{Et}_3\text{N}$ ,  $\text{CH}_2\text{Cl}_2$ , rt, 18 h, 86%; (c) boronic acid,  $\text{Pd(dppf)Cl}_2$ , KOAc, dioxane/ $\text{H}_2\text{O}$  (5:1),  $90^\circ\text{C}$ , 3 h, 45%; (d)  $\text{LiOH}\cdot\text{H}_2\text{O}$ ,  $\text{THF}/\text{H}_2\text{O}$  (1:1), rt, 4 h, quant.

## References

- Ivanov, I. I. *et al.* The Orphan Nuclear Receptor ROR $\gamma$ t Directs the Differentiation Program of Proinflammatory IL-17+ T Helper Cells. *Cell* **126**, 1121–1133 (2006).
- Solt, L. A. & Burris, T. P. Action of RORs and their ligands in (patho)physiology. *Trends Endocrinol. Metab.* **23**, 619–627 (2012).
- Manel, N., Unutmaz, D. & Littman, D. R. The differentiation of human TH-17 cells requires TGF- $\beta$  and induction of the nuclear receptor ROR $\gamma$ t. *Nat. Immunol.* **9**, 641–649 (2008).
- Miossec, P. & Kolls, J. K. Targeting IL-17 and TH17 cells in chronic inflammation. *Nat. Rev. Drug Discov.* **11**, 763–776 (2012).
- Yang, J., Sundrud, M. S., Skepner, J. & Yamagata, T. Targeting Th17 cells in autoimmune diseases. *Trends Pharmacol. Sci.* **35**, 493–500 (2014).
- Burkett, P. R. & Kuchroo, V. K. IL-17 Blockade in Psoriasis. *Cell* **167**, 1669 (2016).
- Lock, C. *et al.* Gene-microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis. *Nat. Med.* **8**, 500–508 (2002).
- Isono, F., Fujita-Sato, S. & Ito, S. Inhibiting ROR $\gamma$ t/Th17 axis for autoimmune disorders. *Drug Discov. Today* **19**, 1205–1211 (2014).
- Fauber, B. P. & Magnuson, S. Modulators of the nuclear receptor retinoic acid receptor-related orphan receptor- $\gamma$  (ROR $\gamma$  or ROR $\gamma$ c). *J. Med. Chem.* **57**, 5871–5792 (2014).
- Bronner, S. M., Zbieg, J. R. & Crawford, J. J. ROR $\gamma$  antagonists and inverse agonists: a patent review. *Expert Opin. Ther. Pat.* **27**, 101–112 (2017).
- Cyr, P., Bronner, S. M. & Crawford, J. J. Recent progress on nuclear receptor ROR $\gamma$  modulators. *Bioorganic Med. Chem. Lett.* **26**, 4387–4393 (2016).
- Huh, J. R. *et al.* Identification of Potent and Selective Diphenylpropanamide ROR $\gamma$  Inhibitors. *ACS Med. Chem. Lett.* **4**, 79–84 (2013).
- Kotoku, M. *et al.* Discovery of Second Generation ROR $\gamma$  Inhibitors Composed of an Azole Scaffold. *J. Med. Chem.* **62**, 2837–2842 (2019).
- Duan, J. J. W. *et al.* Structure-based Discovery of Phenyl (3-Phenylpyrrolidin-3-yl)sulfones as Selective, Orally Active ROR $\gamma$ t Inverse Agonists. *ACS Med. Chem. Lett.* **10**, 367–373 (2019).
- Li, X. *et al.* Structural studies unravel the active conformation of apo ROR $\gamma$ t NR and a common inverse agonism of two diverse classes of ROR $\gamma$ t inhibitors. *J. Biol. Chem.* **292**, 11618–11630 (2017).
- Hu, X. *et al.* Sterol metabolism controls TH17 differentiation by generating endogenous ROR $\gamma$  agonists. *Nat. Chem. Biol.* **11**, 141–147 (2015).
- Scheepstra, M. *et al.* Identification of an allosteric binding site for ROR $\gamma$ t inhibition. *Nat. Commun.* **6**, e8833 (2015).
- Karstens, W. F. J. *et al.* ROR $\gamma$  Inhibitors. *PCT Int. Appl.* WO2012/106995 (2012).
- Ouvry, G. *et al.* Discovery of phenoxyindazoles and phenylthioindazoles as ROR $\gamma$  inverse agonists. *Bioorganic Med. Chem. Lett.* **26**, 5802–5808 (2016).
- Fauber, B. P. *et al.* Discovery of imidazo[1,5-a]pyridines and -pyrimidines as potent and selective ROR $\gamma$  inverse agonists. *Bioorganic Med. Chem. Lett.* **25**, 2907–2912 (2015).
- Meijer, F. A., Leijten-van de Gevel, I. A., de Vries, R. M. J. M. & Brunsveld, L. Allosteric small molecule modulators of nuclear receptors. *Mol. Cell. Endocrinol.* **485**, 20–34 (2019).
- Tice, C. M. & Zheng, Y.-J. Non-canonical modulators of nuclear receptors. *Bioorg. Med. Chem. Lett.* **26**, 4157–4164 (2016).
- Moore, T. W., Mayne, C. G. & Katzenellenbogen, J. a. Minireview: Not picking pockets: nuclear receptor alternate-site modulators (NRAMs). *Mol. Endocrinol.* **24**, 683–695 (2010).
- de Vries, R. M. J. M., Meijer, F. A., Doveston, R. G., Leijten-van de Gevel, I. A. & Brunsveld, L. Cooperativity between the Orthosteric and Allosteric Ligand Binding Sites of ROR $\gamma$ t. *Proc. Natl. Acad. Sci.* **118**, e2021287118 (2020).
- Meijer, F. A. *et al.* Ligand-Based Design of Allosteric Retinoic Acid Receptor-Related Orphan Receptor  $\gamma$ t (ROR $\gamma$ t) Inverse Agonists. *J. Med. Chem.* **63**, 241–259 (2020).
- de Vries, R. M. J. M., Meijer, F. A., Doveston, R. G. & Brunsveld, L. Elucidation of an Allosteric Mode of Action for a Thienopyrazole ROR $\gamma$ t Inverse Agonist. *ChemMedChem* **15**, 561–565 (2020).
- Fronik, P., Gaiser, B. I. & Sejer Pedersen, D. Bitopic Ligands and Metastable Binding Sites: Opportunities for G Protein-Coupled Receptor (GPCR) Medicinal Chemistry. *J. Med. Chem.* **60**, 4126–4134 (2017).
- Newman, A. H., Battiti, F. O. & Bonifazi, A. 2016 Philip S. Portoghese Medicinal Chemistry Lectureship: Designing Bivalent or Bitopic Molecules for G-Protein Coupled Receptors. The Whole Is Greater Than the Sum of Its Parts. *J. Med. Chem.* **63**, 1779–1797 (2020).

29. Lane, J. R., Sexton, P. M. & Christopoulos, A. Bridging the gap: bitopic ligands of G-protein-coupled receptors. *Trends Pharmacol. Sci.* **34**, 59–66 (2013).
30. Valant, C., Sexton, P. M. & Christopoulos, A. Orthosteric/allosteric bitopic ligands: going hybrid at GPCRs. *Mol. Interv.* **9**, 125–135 (2009).
31. Rodrik-Outmezguine, V. S. *et al.* Overcoming mTOR resistance mutations with a new-generation mTOR inhibitor. *Nature* **534**, 272–276 (2016).
32. Ma, N. *et al.* Bitopic Inhibition of ATP and Substrate Binding in Ser/Thr Kinases through a Conserved Allosteric Mechanism. *Biochemistry* **57**, 6387–6390 (2018).
33. Brust, R. *et al.* Modification of the Orthosteric PPAR $\gamma$  Covalent Antagonist Scaffold Yields an Improved Dual-Site Allosteric Inhibitor. *ACS Chem. Biol.* **12**, 969–978 (2017).
34. Battiti, F. O. *et al.* The Significance of Chirality in Drug Design and Synthesis of Bitopic Ligands as D<sub>3</sub> Receptor (D<sub>3</sub>R) Selective Agonists. *J. Med. Chem.* **62**, 6287–6314 (2019).
35. Morales, P. *et al.* Discovery of Homobivalent Bitopic Ligands of the Cannabinoid CB<sub>2</sub> Receptor. *Chem. – A Eur. J.* **26**, 15839–15842 (2020).
36. Steinfeld, T., Mammen, M., Smith, J. A. M., Wilson, R. D. & Jasper, J. R. A Novel Multivalent Ligand That Bridges the Allosteric and Orthosteric Binding Sites of the M<sub>2</sub> Muscarinic Receptor. *Mol. Pharmacol.* **72**, 291–302 (2007).
37. Egyed, A. *et al.* Controlling receptor function from the extracellular vestibule of G-protein coupled receptors. *Chem. Commun. (Camb)*. **56**, 14167–14170 (2020).
38. Kallen, J. *et al.* Structural States of ROR $\gamma$ mat: X-ray Elucidation of Molecular Mechanisms and Binding Interactions for Natural and Synthetic Compounds. *ChemMedChem* **12**, 1014–1021 (2017).
39. Bastings, M. M. C., de Greef, T. F. A., van Dongen, J. L. J., Merckx, M. & Meijer, E. W. Macrocyclization of enzyme-based supramolecular polymers. *Chem. Sci.* **1**, 79–88 (2010).
40. Degorce, F. HTRF: A Technology Tailored for Drug Discovery - A Review of Theoretical Aspects and Recent Applications. *Curr. Chem. Genomics* **3**, 22–32 (2009).
41. Mohr, K. *et al.* Rational design of dualsteric GPCR ligands: quests and promise. *Br. J. Pharmacol.* **159**, 997–1008 (2010).
42. Leesnitzer, L. M. *et al.* Functional Consequences of Cysteine Modification in the Ligand Binding Sites of Peroxisome Proliferator Activated Receptors by GW9662. *Biochemistry* **41**, 6640–6650 (2002).
43. Petrova, N. S. *et al.* Carrier-free cellular uptake and the gene-silencing activity of the lipophilic siRNAs is strongly affected by the length of the linker between siRNA and lipophilic group. *Nucleic Acids Res.* **40**, 2330–2344 (2012).
44. Cuzzocrea, S. *et al.* Rosiglitazone, a ligand of the peroxisome proliferator-activated receptor- $\gamma$ , reduces acute inflammation. *Eur. J. Pharmacol.* **483**, 79–93 (2004).
45. Chira, E. C. *et al.* Tesaglitazar, a dual peroxisome proliferator-activated receptor  $\alpha/\gamma$  agonist, reduces atherosclerosis in female low density lipoprotein receptor deficient mice. *Atherosclerosis* **195**, 100–109 (2007).

# Chapter 7

## Epilogue

### **Abstract**

Nuclear receptors (NRs) are generally considered as highly attractive drug targets because of their essential role in disease processes. The modulation of NRs via allosteric mechanisms, orthogonal to the classical orthosteric site, is a promising novel approach for future drug discovery strategies. The work described in this thesis demonstrates the potential benefits and possibilities of allosteric NR targeting, exemplified by the modulation of ROR $\gamma$ t via a unique, secondary binding pocket. This study provided insights into the mode of action of allosteric ROR $\gamma$ t ligands, as well as the investigation of a cooperative behavior with the classical binding site. Additionally, ROR $\gamma$ t was utilized as a learning platform for the exploration of novel techniques associated with the allosteric pocket, including dual targeting and the development of a screening assay for the unambiguous identification of allosteric ligands. This final chapter discusses future perspectives on the allosteric modulation of ROR $\gamma$ t, and how certain techniques could be applied to the nuclear receptor field in general.

## Introduction

Allosteric modulation of nuclear receptors (NRs) has recently gained significant attention as a promising alternative targeting approach, in contrast to classical modulation via the orthosteric binding pocket.<sup>1,2</sup> Several drugs with allosteric modes of action are already on the market for other protein classes, e.g. GPCRs and kinases, and multiple other allosteric ligands are undergoing clinical trials.<sup>3,4</sup> The first allosteric ligands for NRs were shown to have rather low binding affinities towards their target.<sup>2</sup> However, recently, potent and selective allosteric site modulators have been discovered for several NRs, bringing allosteric modulation of NRs also to the center of attention.<sup>1,2</sup>

The ligand binding domain (LBD) is presently the most promising NR domain for allosteric targeting, due to the high affinities and functional effects of the allosteric compounds that have thus far been identified.<sup>1,5</sup> In addition, there is considerable structural information available for this domain, making *de novo* drug development, computational guided studies, and medicinal chemistry more straightforward. Targeting other NR domains could be an interesting alternative approach, but the scarcity of full-length structures limits a thorough understanding of allosteric effects on the multi-domain protein level.<sup>5</sup>

ROR $\gamma$ t is a unique example of a NR that features such an allosteric pocket in the LBD, next to the canonical, orthosteric binding site.<sup>6</sup> The existence of this allosteric site provides an attractive learning platform for gaining insight into the possibilities of allosteric NR modulation. Besides, ROR $\gamma$ t in itself is an extremely interesting target for (allosteric) inhibition from a drug discovery perspective. In this thesis, the modulation of ROR $\gamma$ t via the allosteric site has been further explored. We extended the available allosteric chemotypes for ROR $\gamma$ t by identifying a novel class of allosteric ROR $\gamma$ t ligands, as described in Chapter 2 and 3. Additionally, the existence of a second, allosteric site sparked the interest of investigating the possibility of dual ligand binding. The research described in Chapter 4 provided evidence that an orthosteric and allosteric ligand can indeed bind to ROR $\gamma$ t simultaneously, even in a cooperative fashion, which progressed into the development of a bitopic ligand in Chapter 6. Furthermore, in Chapter 5, we occluded the orthosteric site with a covalent probe, showing potential for the unambiguous identification of novel allosteric ligands.

In this final chapter, additional ideas on ROR $\gamma$ t allosteric targeting will be described, via traditional approaches but also via more innovative techniques, including possible therapeutic implications. Furthermore, this chapter will provide a future perspective on allosteric modulation of NRs in general, particularly focused on approaches to identify novel allosteric pockets and allosteric ligands, which appears to be the main challenge in the field.

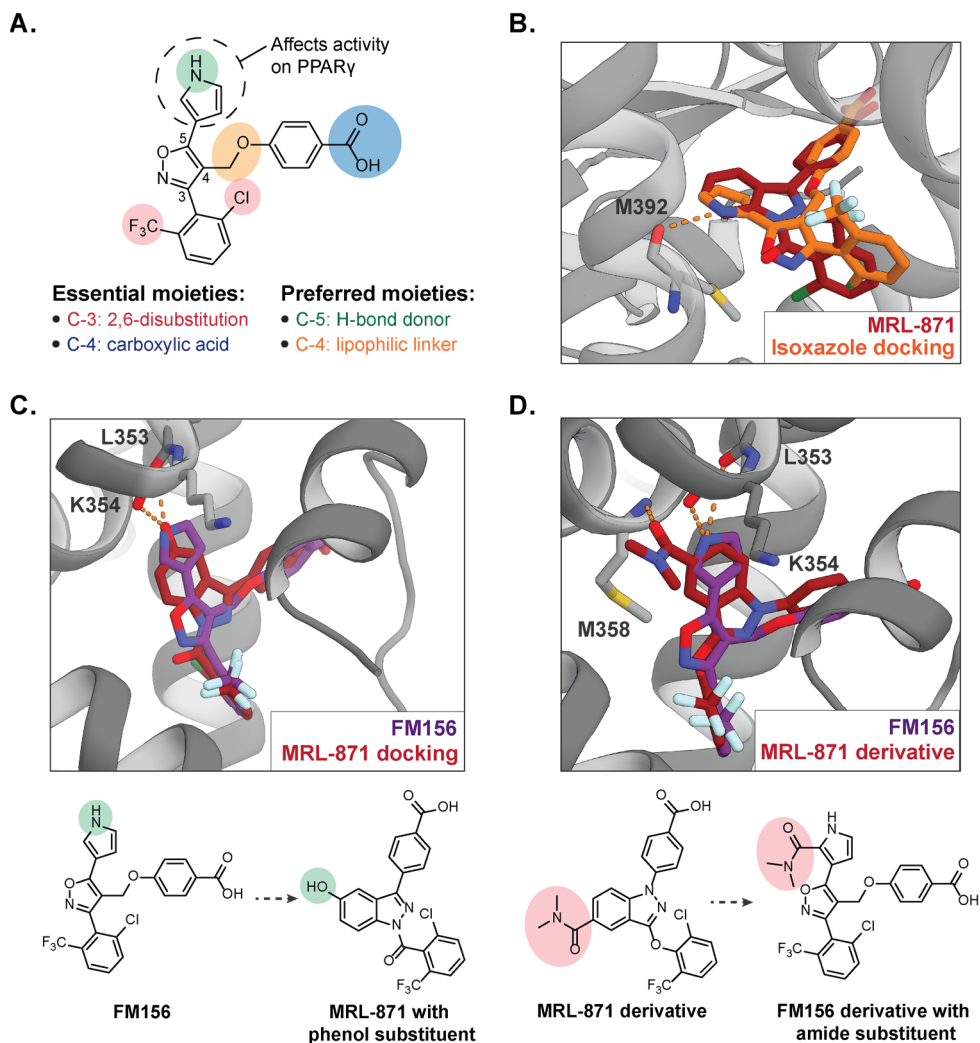


## Future perspective on the allosteric modulation of ROR $\gamma$ t

### *Allosteric ROR $\gamma$ t ligands*

The novel class of allosteric ROR $\gamma$ t ligands discovered in Chapter 2 and 3, represented by the trisubstituted isoxazoles, shows promising properties for the allosteric modulation of ROR $\gamma$ t, with a high potency and good selectivity profile. Additionally, the research described in this thesis provided essential insights into the structure activity relationship (SAR) of this isoxazole class and the structural features that appear to be favorable for efficient allosteric binding (i.e. a lipophilic C-4 linker and a polar C-5 moiety) (Figure 7.1A). Future work should focus on improving the physicochemical properties of the isoxazole series further (e.g. logP, metabolic stability etc.) while maintaining the potency, and subsequently moving into preclinical studies for the development of drugs against autoimmune diseases. The versatility of the trisubstituted isoxazole scaffold, containing three chemical handles, provides many potential avenues to build in beneficial pharmacokinetic (PK) properties. The focus of the PK optimization should lie on adjusting the pyrrole moiety which can be prone to oxidation in drug metabolism<sup>7</sup> (e.g. a pyrazole moiety could be further investigated, since it showed a significantly different PK profile, as was shown in Chapter 3). Also the carboxylic acid moiety might need to be changed because of its unfavorable drug-like properties.<sup>8</sup> The developed chemistry allows for easy modification of the ligands to incorporate favorable PK features, such as pyrrole derivatives or carboxylic acid bioisosteres<sup>8</sup> (although this proved to be difficult for the **MRL-871** series). Additionally, the alkene C-4 linker results in the highest potency for the isoxazole series, but oxidation of an alkene bond is a well-known problem in metabolism (e.g. CYP450-catalyzed epoxidation).<sup>9</sup> Therefore, cyclopropanation of the alkene moiety or synthesis of a cyclized compound (connecting the linker to the pyrrole) could improve the PK character while maintaining the good potency. Another strategy could be to transform the compounds into a prodrug with improved ADME properties, which can be converted to the pharmacologically active ligand after metabolism in the human body.<sup>10</sup>

Furthermore, the SAR knowledge on the isoxazole series provides entry points for further optimization of other known allosteric ROR $\gamma$ t ligands, such as **MRL-871**. Most interestingly, the characteristic hydrogen bond donor moiety at the C-5 position of the isoxazole series could be incorporated into **MRL-871** to obtain a higher polarity, basically mimicking the pyrrole moiety. Transforming the phenyl of the indazole scaffold into a phenol (Figure 7.1C), or an appropriate bioisostere thereof, would be a good starting point, as was supported by an explorative *in silico* docking study, where the OH of the phenol appears to be positioned at the same position in the pocket as the NH of the pyrrole in the isoxazole series, with a promising



**Figure 7.1 | Additional suggestions for allosteric ROR $\gamma$ t ligands.** **A)** Chemical structure of isoxazole ligand **FM156**, with essential and preferred moieties indicated. **B)** Overlay of the co-crystal structure of PPAR $\gamma$  in complex with **MRL-871** (red) (PDB: 6TDC) and the docking pose of an isoxazole derivative with a 2-pyrrole at the isoxazole C-5 position (orange). A potential hydrogen bond between the NH of the pyrrole and the backbone of Met392 is indicated with orange dashes. **C)** Overlay of the co-crystal structure of ROR $\gamma$ t in complex with isoxazole ligand **FM156** (purple) where the pyrrole moiety forms an essential hydrogen bond with the main chain of residues Leu353 and Lys354 (PDB: 7NPC) and the docking pose of an **MRL-871** derivative, containing a phenol moiety (red). Incorporation of a phenol in **MRL-871** could be explored to form an additional hydrogen bond interaction with the protein. **D)** Overlay of the co-crystal structures of ROR $\gamma$ t in complex with an **MRL-871** derivative containing an amide substituent (red) that forms a hydrogen bond interaction with the backbone of residue Met358 (PDB: 5LWP) and ROR $\gamma$ t in complex with isoxazole ligand **FM156** (purple) (PDB: 7NPC). An amide moiety could be incorporated at the C-2 position of the pyrrole moiety to establish additional interactions in the allosteric site.

docking score (-14.91 for the phenol derivative vs. -14.99 for **MRL-871**) (Figure 7.1C). *Vice versa*, a dimethylamide at the 6-position of the indazole scaffold of an **MRL-871** derivative results in an increased potency by forming a hydrogen bond interaction between its carbonyl and the backbone NH of Met358 (Figure 7.1D).<sup>11</sup> In an overlay of the co-crystal structures of ROR $\gamma$ t in complex with this **MRL-871** derivative and with isoxazole **FM156**, the 6-position of the indazole is at the same location as the 2-position of the pyrrole, which would indicate that incorporation of an amide at the 2-position of the pyrrole could have a beneficial effect (Figure 7.1D).

Another interesting topic which could be further explored is the apparent cross-reactivity of some allosteric ROR $\gamma$ t ligands with PPAR $\gamma$ , as was shown for the **MRL-871** series but also for some isoxazoles (Chapter 3). **MRL-871** and some derivatives act as partial agonists on PPAR $\gamma$ , which could be an interesting entry point in the treatment of type II diabetes, showing reduced side effects in contrast to full agonists.<sup>12</sup> The knowledge obtained in this thesis, combined with the knowledge about the activity of **MRL-871** on PPAR $\gamma$ , including its co-crystal structure (Figure 7.1B)<sup>13</sup>, could guide the investigation of potent and selective PPAR $\gamma$  ligands. Interestingly, the isoxazole compounds containing an additional hydrophobic (methyl) group at the pyrrole ring in particular showed a significantly higher cross-reactivity with PPAR $\gamma$  which will likely be due to additional hydrophobic interactions. Additionally, the compound containing a 2-pyrrole at the C-5 position improved PPAR $\gamma$  binding compared to the other isoxazoles, which could be caused by a hydrogen bond interaction with the backbone of Met392 according to the docking pose (Figure 7.1B). Future studies could thus focus on the rational design of PPAR $\gamma$  ligands, guided by structure-based design and docking studies.

Lastly, it would be interesting to diversify the chemotypes for allosteric ROR $\gamma$ t inverse agonists further, since currently there is still only a limited number of different scaffolds known. Possible strategies could rely on the performance of a pharmacophore search around the novel isoxazole class, or a tethering approach as will be discussed later in more detail. Special attention could be focused on scaffolds containing a higher sp<sup>3</sup> hybridized character (instead of flat ring systems) with a more flexible molecular shape and greater three-dimensionality, which can better exploit the chemical space in the allosteric pocket.<sup>14</sup>

### ***Challenges in targeting ROR $\gamma$ t from a therapeutic point of view***

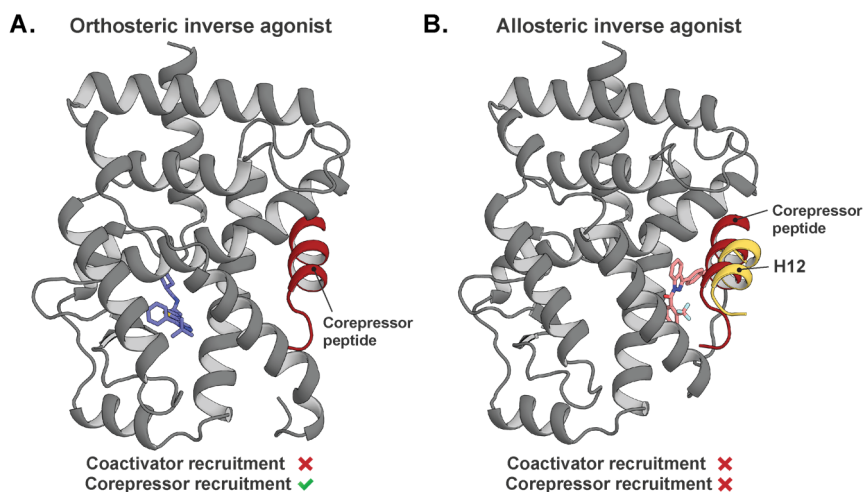
ROR $\gamma$ t has been identified as a very promising target in autoimmune diseases, because of its key regulating role in the development of Th17 cell differentiation and proliferation, and the production of pro-inflammatory cytokines including IL-17a.<sup>15</sup> Inhibition of the Th17/

IL-17a pathway has already proven to be successful in clinical applications for psoriasis, with several anti-IL-17a monoclonal antibodies that are currently approved by the FDA.<sup>16</sup> However, treatment of other autoimmune diseases (e.g. Crohn's disease) with IL-17a antibodies, appeared to be inefficient, which might be due to the fact that several other cytokines produced by Th17 cells may also have critical roles in inflammation.<sup>17</sup> Therefore, RORyt inhibition can be more efficient as it regulates the expression of multiple effector cytokines and thus will cause the inhibition of a broader panel of these. Additionally, inhibition of RORyt with small molecules could have significant advantages over antibody-based therapies. However, the therapeutic effect of inhibiting RORyt function is not fully understood yet. Recent studies have shown that mice deficient in RORyt or treated with RORyt inhibitors develop significant thymic aberrations (e.g. downregulation of anti-apoptotic molecules), suggesting a possible risk of developing T cell lymphoma.<sup>18–21</sup> However, other studies have pointed out that no adverse reactions were observed in mice treated with RORyt inverse agonists.<sup>22–24</sup> Ongoing preclinical and clinical studies should thus reveal more knowledge on the pharmacological effect and apparent risks of RORyt inhibition, since the broader implications of modulating RORyt are not apparent yet.

The eventual side effects observed in the mentioned studies are all based on experiments with orthosteric inverse agonists for RORyt. Interestingly, we found that allosteric inverse agonists inhibit RORyt via a different mode of action than orthosteric ligands, directly interacting with H12. Orthosteric and allosteric inverse agonists both lead to inhibition of coactivator recruitment, however most orthosteric inverse agonists do create a groove for corepressor binding, by destabilization of H12 (Figure 7.2A).<sup>25,26</sup> In contrast, the protein conformation in the presence of an allosteric inverse agonist, with H12 folding back over the allosteric ligand, will most likely also block corepressor recruitment (thus repressing both coactivator and corepressor binding) (Figure 7.2B). Because of the distinct binding mode and potential different effect on cofactor recruitment, the physiological effect of allosteric ligands could be different from orthosteric ligands and also the reported side effects might be circumvented. Additionally, partial (orthosteric or allosteric) inverse agonists may have a better safety profile as well.<sup>27</sup> Future studies could focus on the effect of these different types of RORyt inverse agonists both in a biochemical setting (e.g. TR-FRET assays with corepressor peptides) and in a therapeutic environment (i.e. cellular assays and preclinical studies), which could be extremely relevant for the development of RORyt drugs.

Another aspect that should be considered from a therapeutic point of view is the cooperative behavior between the orthosteric and allosteric binding site, as was discovered in

Chapter 4. The potency of allosteric ligands for ROR $\gamma$ t could be influenced by the binding of endogenous orthosteric ligands and could thus affect the pharmacological effect of allosteric inverse agonists. Therefore, further studies could focus on how this cooperativity effect, which has currently only been studied in biochemical assays, influences cellular responses.

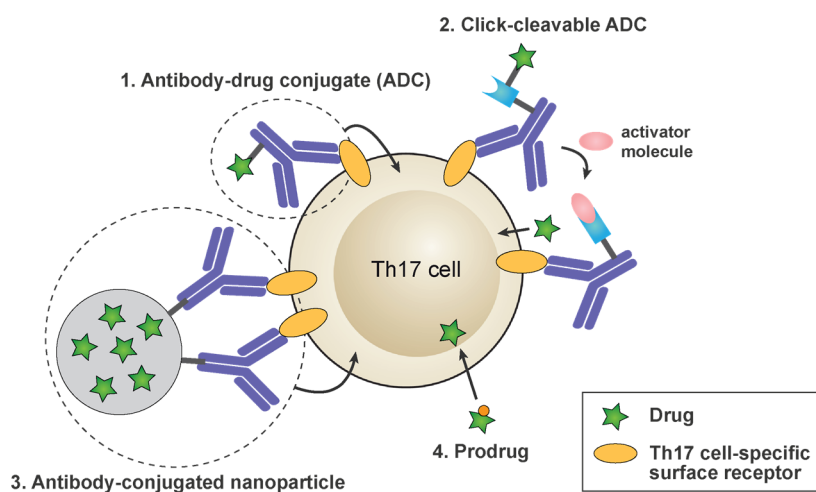


**Figure 7.2 | Difference in cofactor recruitment for orthosteric and allosteric ROR $\gamma$ t inverse agonists. A)** Binding of an orthosteric inverse agonist (blue sticks) results in destabilization of H12 (not visible), disrupting coactivator binding, but inducing the binding of a corepressor peptide (red) in most cases (PDB: 6A22). **B)** Binding of an allosteric inverse agonist (pink sticks) positions H12 (yellow) in a conformation that folds back over the allosteric ligand, repressing coactivator binding, but most likely also blocking corepressor binding (red) (PDB: 4YPQ, in overlay with a corepressor peptide (PDB: 6A22)).

Furthermore, the two ROR $\gamma$  isoforms (ROR $\gamma$ 1 and ROR $\gamma$ t) have identical ligand binding domains, so it is likely that ROR $\gamma$ t inhibitors will show cross-reactivity with ROR $\gamma$ 1 in nonimmune cells, which could lead to potential side effects (and in part may also be the reason for some of the mentioned side effects discussed).<sup>17,27</sup> The selective inhibition of ROR $\gamma$ t will thus be an extremely challenging task, but the fact that ROR $\gamma$ t expression is limited to the thymus might provide an interesting entry point. As an ultimate aim in ROR $\gamma$ t drug development, future studies could therefore focus on selective drug delivery to Th17 cells in the thymus in order to selectively target ROR $\gamma$ t.

A potential approach for the selective release of a ROR $\gamma$ t inverse agonist to these Th17 cells, could be based on combining a drug with a delivery system that can target a specific receptor on the surface of Th17 cells. Recently, a set of surface markers have been identified that are exclusively expressed by Th17 cells and are therefore extremely interesting to target

in this approach.<sup>28</sup> A first strategy could be the development of an antibody-drug conjugate (ADC), consisting of a monoclonal antibody that recognizes the specific surface receptor, tethered to a drug via a chemical linker (Figure 7.3, example 1).<sup>29</sup> Upon binding of the ADC to a receptor protein, the complex is internalized via endocytosis and the drug can be released intracellularly.<sup>29</sup> To date, five ADCs have been approved by the FDA (mostly for cancer chemotherapy), and over 100 are being studied in clinical trials, highlighting the potential of the technique.<sup>29</sup> Closely related to this technique are the recently developed click-cleavable ADCs (Figure 7.3, example 2), which are especially interesting in case of non-internalizing receptor proteins.<sup>30</sup> The structure of a click-cleavable ADC is similar to the original ADCs, apart from the linker that contains a moiety which can react with an activator molecule via an elimination click-reaction, resulting in extracellular release of the drug.<sup>30</sup> A second technology for target drug delivery, which builds on the potential of ADCs, is the use of antibody-conjugated nanoparticles (with the drug encapsulated), that can again target Th17-specific surface receptors (Figure 7.3, example 3).<sup>31,32</sup> The main difference compared to ADCs are the higher drug to antibody ratio and the potentially easier drug release (since the drug is not covalently attached as is the case for ADCs).<sup>31</sup> Finally, a different approach for selective ROR $\gamma$ t targeting, independent of surface receptor targeting, is the design of a prodrug that could be converted to the active drug by Th17 cell-specific enzymes (Figure 7.3, example 4), although these are not known in the literature yet.

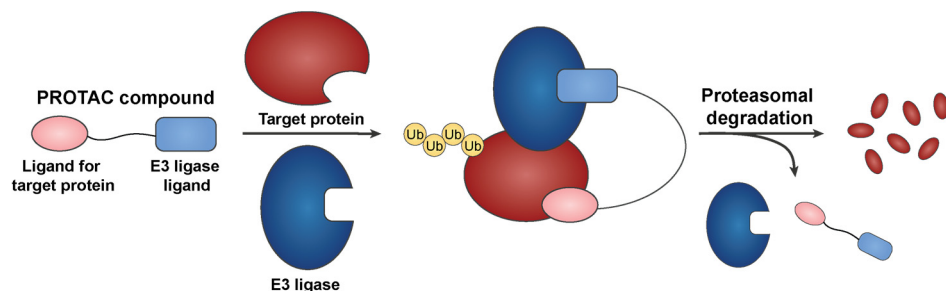


**Figure 7.3 | Potential strategies for selective ROR $\gamma$ t drug delivery to Th17 cells.** 1) Via antibody-drug conjugates (ADCs) targeting Th17 cell-specific surface receptors, 2) By using click-cleavable ADCs, 3) Via the formation of antibody-conjugated nanoparticles, 4) By delivery of a prodrug that is converted into the active species by Th17 cell-specific enzymes.

## Targeted protein degradation

The inhibition of ROR $\gamma$ t activity is a promising strategy in the treatment of autoimmune diseases. Although ROR $\gamma$ t modulation with inverse agonists shows many possibilities as described in this thesis, a promising alternative could be the targeted degradation of ROR $\gamma$ t by using proteolysis targeting chimeras (PROTACs).<sup>33–35</sup>

PROTACs have recently received significant attention as a new therapeutic strategy in drug development.<sup>33–36</sup> The technology is based on the development of bivalent ligands that can bring together an E3 ligase protein (typically Von Hippel Lindau (VHL) or Cereblon (CRBN)) and a target protein that is to be degraded (Figure 7.4).<sup>33–35</sup> PROTAC ligands thus usually consist of three chemical elements: a ligand that targets an E3 ligase, a ligand that binds to the target protein, and a defined linker connecting the two ligands. Upon binding of the PROTAC ligand, the proteins are brought into close proximity, which allows the transfer of ubiquitin to the target protein. The poly-ubiquitinated target protein is then recognized by the proteasome, which leads to proteasomal degradation.



**Figure 7.4 | Schematic representation of the mode of action of a PROTAC compound.** The bivalent PROTAC compound consists of a ligand for the target protein, a flexible linker and an E3 ligase ligand. Upon formation of the ternary complex (E3 ligase, PROTAC and target protein), the E3 ligase and target protein are brought into close proximity, resulting in poly-ubiquitination (Ub) of the target protein, which triggers proteasomal degradation. The PROTAC compound is released and can exert its function again. Figure adapted from Scheepstra *et al.*<sup>35</sup>

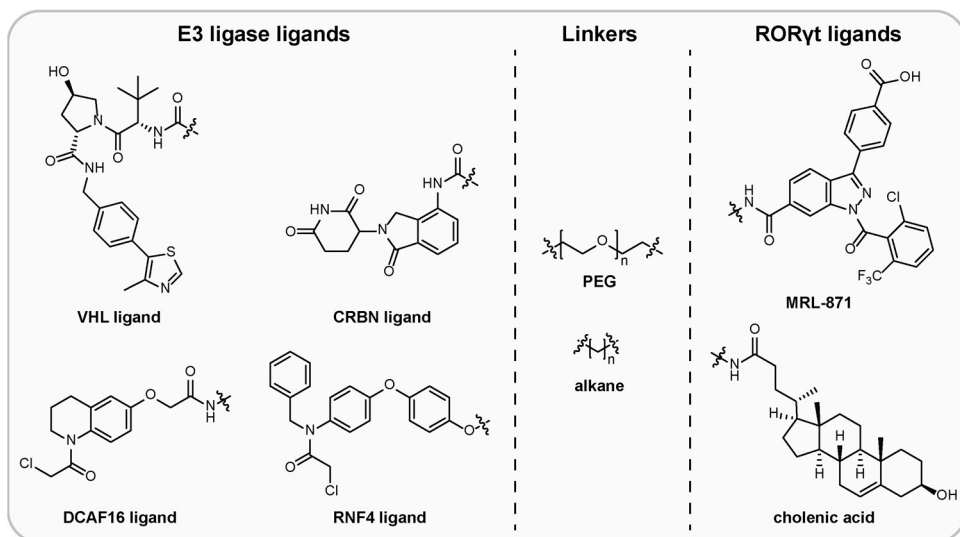
The main difference between protein degradation strategies compared to traditional protein inhibition approaches is the fact that PROTACs follow an event-driven instead of an occupancy-driven approach (i.e. eliminating the target protein rather than to inhibiting it), which has several key advantages.<sup>33–36</sup> First, PROTACs are catalytic in their mode of action, since after degradation of the target protein, they can repeat their action on a next target protein.<sup>33–36</sup> Consequently, a lower concentration of a PROTAC can be used to obtain full

protein degradation, compared to the concentration that would be needed for traditional inhibitors. Second, the ligand does not need to have a functional activity towards the target protein in order to exert its effect as a degrader; ligands that simply bind to the target, even with lower affinities, will be effective as well, since the ligand is only used to recruit the E3 ligase to the target protein.<sup>33–36</sup> This also means that in case of drug resistance, degrader molecules can offer a good alternative. Third, PROTACs can be more effective in decreasing protein activity than a small molecule inhibitor, because they will establish a more significant and long lasting effect by degrading the protein.<sup>33–36</sup> Finally, several examples have been reported where a non-selective ligand results in selective target degradation.<sup>37–38</sup> Although the PROTAC ligand forms a binary complex with multiple proteins in these cases, ternary complex formation is preferred for one certain protein, which is mostly due to stabilizing interactions between that specific protein and the E3 ligase (positive cooperativity), while other proteins show charge repulsion or steric clashes with the E3 ligase.<sup>33,35</sup>

However, several challenges still remain to be addressed in the development of PROTACs for therapeutic applications, especially because of the large size of the molecules, which generally leads to suboptimal PK properties (i.e. solubility, cell permeability and metabolic stability).<sup>33–36</sup> Additionally, the formation of a ternary complex by binding of a PROTAC is still no guarantee for degradation of the target protein (e.g. because the E3 ligase is potentially not positioned at the right distance or in the right orientation to induce ubiquitination) and optimization via rational design is difficult in most cases.<sup>33–36</sup> Despite these challenges, many research groups focus on implementing the technique in therapeutic applications, with two PROTAC compounds, **ARV-110**<sup>39</sup> and **ARV-471**<sup>40</sup> (both developed by the company Arvinas), currently studied in clinical trials for the treatment of prostate cancer and breast cancer (by targeting the nuclear receptors AR and ER $\alpha$ ).<sup>41,42</sup>

Since the targeted degradation of ROR $\gamma$ t could offer a promising alternative to small molecule inhibition, we explored the possibility of degrading ROR $\gamma$ t using a PROTAC strategy. Especially combining the PROTAC technology with allosteric targeting of ROR $\gamma$ t could have double benefits. Therefore, we designed and evaluated a set of PROTAC ligands, initially targeting the allosteric binding site of ROR $\gamma$ t. A library of 16 PROTACs was designed (compounds **1-16**, Table 7.1), consisting of an E3 ligase ligand (VHL or CRBN ligand), a polyethylene glycol (PEG-) or alkane-based linker (with various lengths) and a ROR $\gamma$ t allosteric ligand (**MRL-871**, with the linker attached to the 6-position, similar to the bitopic ligand design in Chapter 6) (Figure 7.5).

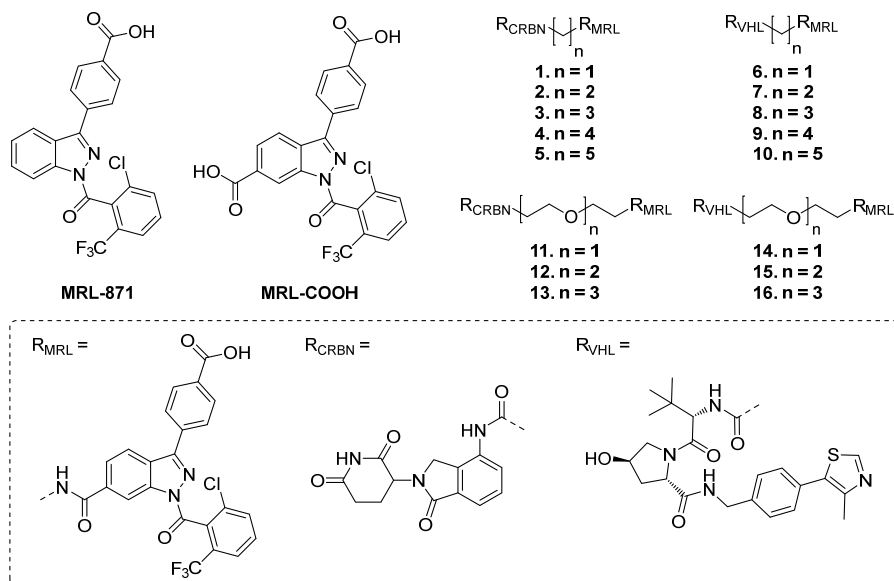




**Figure 7.5 | PROTAC design for RORγt.** Combinations of E3 ligase ligands (VHL, CRBN, DCAF16 or RNF4 ligands), linkers (PEG or alkane) and RORγt allosteric or orthosteric ligands (**MRL-871** and cholenic acid, respectively).

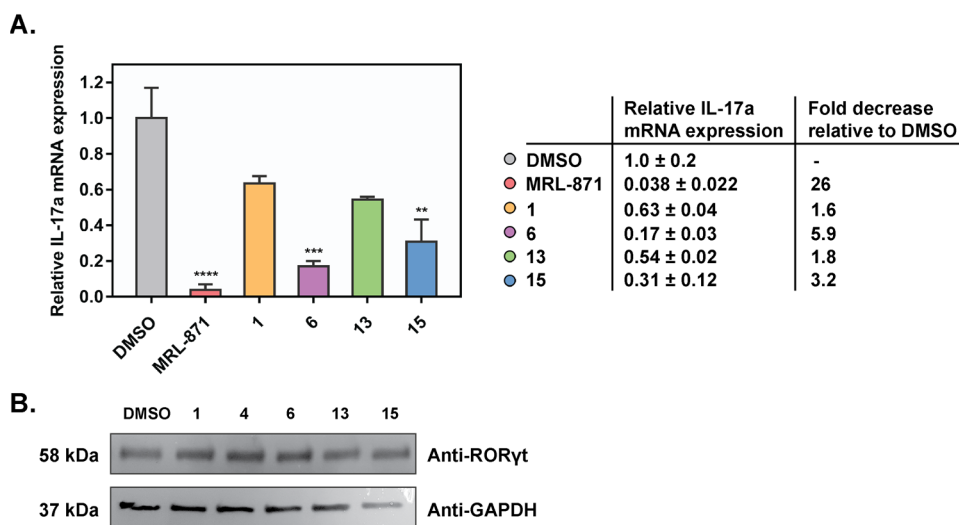
The PROTACs were synthesized in collaboration with the company Symeres (Nijmegen, the Netherlands) (synthesis route not shown) and the binding to RORγt was evaluated in TR-FRET assays, for which the results are shown in Table 7.1. All PROTAC ligands **1-16** were potent inverse agonists for RORγt with  $IC_{50}$  values in the range of 50-350 nM, comparable to reference compound **MRL-COOH** ( $IC_{50} = 220 \pm 20$  nM), while **MRL-871** showed an  $IC_{50}$  value of  $13 \pm 1$  nM, as known from literature.<sup>6</sup> Overall, the PROTACs containing a shorter linker appear to have a slightly higher potency, although this correlation does not hold in all cases and the differences are rather small. Additionally, the conformation and solubility of the linkers could play a role as well in affecting potency. The PROTACs were then also evaluated in a TR-FRET AlexaFluor-MRL recruitment assay, where the compounds showed a significant displacement of the allosteric probe (with similar potency to **MRL-COOH**), and the  $IC_{50}$  values clearly correlate with the  $IC_{50}$  values from the coactivator recruitment assay.

**Table 7.1 | Biochemical assays with PROTAC ligands 1-16 on ROR $\gamma$ t.** TR-FRET IC<sub>50</sub> values ( $\mu$ M) from coactivator recruitment and AlexaFluor-MRL recruitment assays, and fold decrease in IL-17a expression levels relative to DMSO (RT-PCR). TR-FRET and RT-PCR data are representative of two independent experiments (recorded in triplicate) and TR-FRET data are presented as mean  $\pm$  SD.



Compound	IC <sub>50</sub> ( $\mu$ M) TR-FRET Coactivator	IC <sub>50</sub> ( $\mu$ M) TR-FRET AlexaFluor-MRL	Fold decrease in IL-17a expression
MRL-871	0.013 $\pm$ 0.001	0.0079 $\pm$ 0.0005	26
MRL-COOH	0.22 $\pm$ 0.02	0.11 $\pm$ 0.01	n.d.
1	0.16 $\pm$ 0.02	0.10 $\pm$ 0.01	1.6
2	0.27 $\pm$ 0.05	0.16 $\pm$ 0.07	n.d.
3	0.23 $\pm$ 0.03	0.16 $\pm$ 0.08	n.d.
4	0.13 $\pm$ 0.01	0.073 $\pm$ 0.008	n.d.
5	0.36 $\pm$ 0.04	0.24 $\pm$ 0.03	n.d.
6	0.087 $\pm$ 0.008	0.081 $\pm$ 0.008	5.9
7	0.28 $\pm$ 0.03	0.16 $\pm$ 0.02	n.d.
8	0.30 $\pm$ 0.04	0.18 $\pm$ 0.03	n.d.
9	0.30 $\pm$ 0.04	0.17 $\pm$ 0.02	n.d.
10	0.36 $\pm$ 0.05	0.26 $\pm$ 0.04	n.d.
11	0.048 $\pm$ 0.004	0.042 $\pm$ 0.003	n.d.
12	0.21 $\pm$ 0.02	0.097 $\pm$ 0.010	n.d.
13	0.13 $\pm$ 0.01	0.083 $\pm$ 0.007	1.8
14	0.12 $\pm$ 0.02	0.040 $\pm$ 0.003	n.d.
15	0.21 $\pm$ 0.04	0.055 $\pm$ 0.003	3.2
16	0.30 $\pm$ 0.03	0.15 $\pm$ 0.01	n.d.

Subsequently, a subset of PROTACs (**1**, **6**, **13** and **15**, see Table 7.1) were evaluated in a RT-PCR assay to investigate their functional effect on ROR $\gamma$ t in a cellular environment. EL<sub>4</sub> cells were treated with 10  $\mu$ M of PROTAC ligand or **MRL-871** for 24 h, before the IL-17a mRNA levels were measured (Figure 7.6A, Table 7.1). **MRL-871** showed a significant inhibition of IL-17a mRNA levels, with a 26-fold reduction compared to DMSO, as shown previously.<sup>6</sup> The PROTACs consisting of **MRL-871** coupled to a VHL ligand (**6** and **15**) resulted in a 5.9- and 3.2-fold reduction of IL-17a, which was shown to be significant (Figure 7.7), but they were clearly not as active as **MRL-871**. The PROTACs consisting of **MRL-871** coupled to a CRBN ligand (**1** and **13**) showed an even lower effect, with a 1.6- and 1.8-fold decrease in IL-17a levels, respectively, which was not shown to be significant. These results demonstrate that the PROTACs (especially the VHL ones) appear to be cell-permeable and are active on full length ROR $\gamma$ t (reduction of IL-17a mRNA expression), however to a lower extent than **MRL-871**, which might be due to a lower potency, but also because of a lower cell permeability. However,



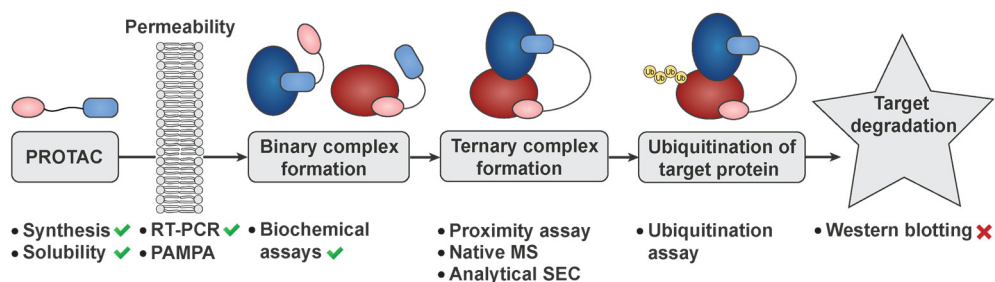
**Figure 7.6 | Investigation of the cellular activity of the PROTAC ligands and ROR $\gamma$ t degradation.**  
**A)** IL-17a mRNA expression levels in EL<sub>4</sub> cells treated with **MRL-871** and PROTAC ligands **1**, **6**, **13** and **15** (10  $\mu$ M, 24 h) or DMSO, and fold decrease of IL-17a expression relative to DMSO. The level of IL-17a expression was normalized to that of GAPDH expression. Data are representative of two independent experiments (recorded in triplicate) and are presented as mean  $\pm$  SD. The relative gene expression was calculated by the  $2^{-\Delta\Delta C_t}$  (Livak) method using the DMSO control as calibrator. Statistical analysis was performed using a one-way analysis of variance (ANOVA) compared against the DMSO control following Dunnett's *post hoc* test; \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$  and \*\*\*\*  $P < 0.0001$ .  
**B)** Western blots showing ROR $\gamma$ t and GAPDH levels in EL<sub>4</sub> cells, treated with 10  $\mu$ M of PROTAC ligand (**1**, **4**, **6**, **13** or **15**) or DMSO (24 h incubation).

from these data it could not yet be concluded if these PROTACs show degradation of the protein, since their effect could also be caused by inhibition of ROR $\gamma$ t via their allosteric inverse agonistic behavior. In case of degradation, potentially a more significant reduction in IL-17a levels would have been expected for the PROTACs than was observed here.

We next evaluated the ability of the PROTACs to effectively degrade ROR $\gamma$ t. EL4 cells were again treated with a subset of PROTAC compounds (concentration of 10  $\mu$ M) for 24 h. Subsequently, the cells were lysed and the ROR $\gamma$ t expression levels were investigated via Western blotting. Additionally, the expression of GAPDH was evaluated as control. The Western blot data did not show any significant difference in the ROR $\gamma$ t protein levels between the cells treated with PROTAC and the control cells (treated with DMSO), which indicates that the PROTACs were not able to induce ROR $\gamma$ t degradation (Figure 7.6B). The experiment was repeated using lower concentrations of PROTAC ligands (0.01, 0.1 and 1.0  $\mu$ M), since a too high concentration could result in lower degradation effects (hook effect)<sup>35</sup>, however still no change in ROR $\gamma$ t protein levels was observed.

In order to further explore the PROTAC approach on ROR $\gamma$ t, a novel subset of PROTACs was synthesized, by coupling **MRL-871** to the ligands for two different E3 ligase proteins (DCAF16<sup>43</sup> and RNF4<sup>44</sup> (see Figure 7.5)), which both contain an  $\alpha$ -chloro ketone moiety that bind to the ligase proteins via covalent attachment. Additionally, the orthosteric ROR $\gamma$ t ligand cholenic acid was also coupled to the four different E3 ligase ligands (Figure 7.5), in order to explore if PROTACs targeting a different binding site of ROR $\gamma$ t would show a different result. Although all these PROTACs showed binding to ROR $\gamma$ t in the TR-FRET assays and also demonstrated the desired cellular effect in terms of IL-17a expression, again no degradation of ROR $\gamma$ t was observed by Western blot. Additionally, the PROTACs targeting DCAF16 appeared to be toxic, since EL4 cell death was observed, which is likely to be caused by non-specific ligation to other proteins.

This study on the design of ROR $\gamma$ t PROTACs highlights that the task of developing PROTACs for a certain target is not trivial, as was also the case for some examples in the literature.<sup>45</sup> Despite the fact that the PROTACs show good affinity for the target protein (binary complex formation) and appear to be cell permeable, they are not able to induce target degradation, for which the reason could lie at different stages in the PROTAC pathway. Figure 7.7 shows the different steps in the complex journey of a PROTAC ligand, from administration to target degradation, for which the different techniques that could be used to study each stage have been indicated.<sup>46</sup>



**Figure 7.7 | The pathway of a PROTAC ligand, from administration to target degradation.** The chart highlights the various steps involved and the different techniques that could be used to study each stage. A green mark indicates a technique that has been used in the RORYt PROTAC study which showed the desired results, whereas the red cross indicates a technique that has been used which did not show the desired result. Figure adapted from Hughes *et al.*<sup>46</sup>

In order to further explore why the PROTACs do not induce RORYt degradation and try to establish a functional system, future studies could focus on investigating the steps in between binary complex formation and target degradation (Figure 7.7). First, it should be investigated if the ternary complex is formed, since a mismatch between RORYt and the specific E3 ligase (negative cooperativity) could prevent proper formation of the complex.<sup>33,35</sup> Although a mismatch with all four different E3 ligase proteins would be unexpected, additional assays could shed further light on this. A biochemical assay, based on FRET pairing between the two proteins (proximity assay), would be a simple approach to show if indeed proximity is induced in the presence of a PROTAC ligand.<sup>47,48</sup> This assay could also aid in getting more insight into the hook effect, by exploring the concentration range in which ternary complex formation is preferred (since at high concentrations of PROTAC, it will saturate the individual proteins instead of bringing them together).<sup>35</sup> Other methods to explore ternary complex formation would be native mass spectrometry (native MS)<sup>49</sup> or analytical size exclusion chromatography (SEC), by comparing a situation with and without a PROTAC ligand present.

Additionally, the cellular effect of the PROTACs can be further investigated. Although the performed RT-PCR assay already indicated cell permeability, a PAMPA assay (parallel artificial membrane permeability assay) could provide further evidence. Also the functionality of the PROTAC ligands in cells can be tested, by performing an assay to check if ubiquitination of RORYt is happening. Examination of the structure of RORYt shows that a significant number of lysine residues should be available for ubiquitination in the environment of both the allosteric and orthosteric binding site, but the E3 ligase might be oriented in an improper way, not suitable for target ubiquitination.<sup>35</sup>

Apart from these techniques, investigating a broader range of different concentrations and incubation times for all PROTACs might affect the degradation ability. Also a different cell line (e.g. Jurkat cells) could be used, since the presence of the targeted E3 ligases in the EL4 cell line is not proven yet. Another problem could be the fact that RORYt is mostly present in the nucleus of the cell and is not located in proximity of the E3 ligases. However, this is unlikely since DCAF16 and RNF4 reside in the nucleus, similar to RORYt, and VHL and CRBN will also be in proximity of RORYt when the protein is shuttling to the cytoplasm.

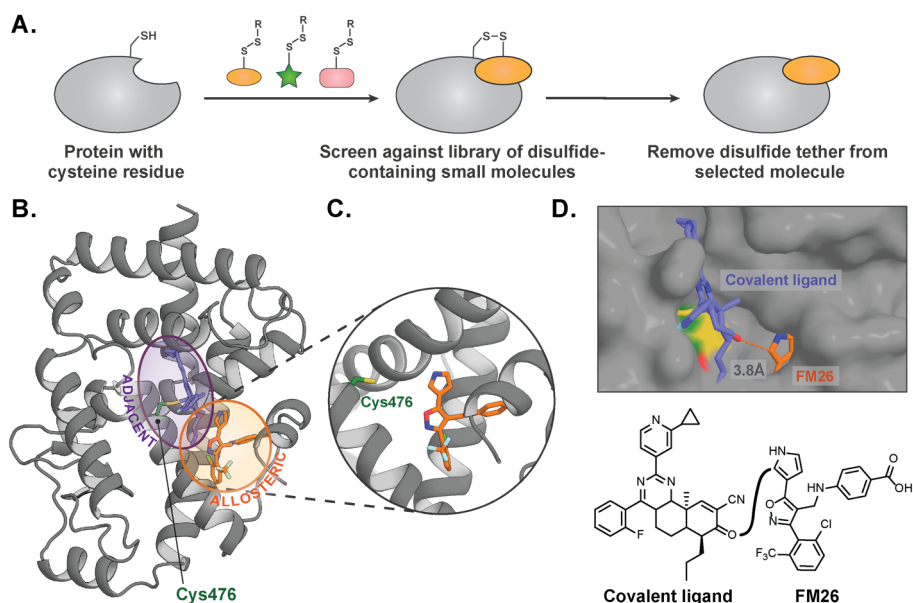
If these follow-up studies would indeed show that ternary complex formation of the different E3 ligases and RORYt is established, but in a non-optimal fashion (i.e. not leading to ubiquitination and degradation), attempts could be made to optimize the intermolecular interactions between the two proteins in the ternary complex, preferentially via structure-based design after obtaining the crystal structure. Optimizing the linker of the PROTAC, in order to establish additional interactions with both proteins, has helped in certain cases to force ternary complex formation.<sup>47,50</sup>

Despite the challenges shown here, the technique of targeted protein degradation generally has high potential for targeting NRs but also other protein classes (including ‘undruggable’ targets). In order to implement the technology as a therapeutic approach, future studies should primarily focus on how to rationally design PROTAC ligands for any specific target. Improving the structural understanding of ternary complex formation (i.e. by X-ray crystallography) could assist in the design of linkers that can induce target-specific interactions in the ternary complex (i.e. improving the cooperativity), or even in the formation of a merged PROTAC ligand that could bring the proteins together like a molecular glue.<sup>33,35</sup> Furthermore, alternative E3 ligases (and their binding motifs) could be identified for targeting to expand the PROTAC toolbox.<sup>33,35</sup> Lastly, the PK properties of PROTACs and their eventual use as drugs should be further explored, as is currently investigated for the first PROTACs that have entered clinical trials.<sup>33,35</sup>

## Techniques for identifying novel allosteric ligands

The growing interest in the allosteric modulation of NRs comes with a concurrent demand for novel screening techniques that could aid in the identification of allosteric pockets and the discovery of small molecules specifically targeting these pockets. Whereas the identification of novel ligands typically relies on high throughput screening (HTS) campaigns, the discovery of allosteric ligands is not trivial by experimental approaches. This is basically caused by the fact that allosteric ligands are not easily distinguished from classical ligands via most screening approaches; a series of complex experiments would be needed afterwards to discriminate allosteric ligands from orthosteric ones. Competition assays with a fluorescently labelled allosteric probe could provide a possible entry, however this approach can only be used for finding new allosteric ligands if an allosteric site has already been established. Additionally, this technique suffers from many problems, including a high false positive rate and difficulties in identifying weaker binding ligands. Therefore, there is a high demand for approaches that can easily and unambiguously identify allosteric ligands, for which several entries will be discussed in this paragraph.

A first strategy is the occlusion of the classical, orthosteric binding site of a NR, as has been done for ROR $\gamma$ t in Chapter 5, but in principle could also be applied for other NRs or proteins. The protein with a blocked orthosteric site can then be used in screening approaches, for which all identified hit compounds should bind to allosteric sites. One option for orthosteric site occlusion, as performed for ROR $\gamma$ t in Chapter 5, is by covalent binding of a chemical probe to a native or engineered cysteine residue (with minimal effect on protein structure and activity). This approach can be easily translated to other NR proteins. A second option is mutagenesis, where one or more small amino acid residues in the orthosteric site are replaced by larger ones, that can occupy the orthosteric site with their amino acid side chains. This approach circumvents the covalent attachment of a chemical probe and optimization of ligation conditions, but just relies on the expression of a protein containing one or more mutated amino acid residues. In preliminary studies, we focused on the mutagenesis of different residues in the orthosteric site of ROR $\gamma$ t (i.e. different combinations of Cys320Phe, Val376Phe and Ile400Phe mutations). We showed that the orthosteric site was blocked, while the allosteric site was still available for ligand binding, demonstrating that the technique is thus a good alternative for covalent probe attachment. These results are highly promising for implementing the technique on ROR $\gamma$ t but also on other NRs.



**Figure 7.8 | Disulfide tethering approaches could be used for the identification of novel allosteric ligands.** **A)** Schematic representation of the disulfide tethering technique, based on screening of a library of disulfide-containing fragments that can be covalently attached to a cysteine residue. Figure adapted from Erlanson *et al.*<sup>51</sup> **B)** The co-crystal structure of RORγt in complex with allosteric ligand FM26 (orange sticks) in an overlay with the predicted binding mode of the covalent ligand (purple sticks) in the adjacent site, where the Cys476 residue is also indicated. **C)** Enlarged view of FM26 in the allosteric site, showing that Cys476 points towards the allosteric site and could be used for tethering approaches. **D)** The covalent ligand and isoxazole FM26 bind close to each other (3.8 Å) and could be linked together, by attaching the carbonyl to the C-2 pyrrole position.

Furthermore, in the case that an allosteric pocket has already been discovered, novel allosteric ligands can be identified via a covalent tethering approach. Different tethering strategies have been developed, for which disulfide tethering is the most well-known.<sup>51-52</sup> The technique involves the formation of a reversible disulfide bond via the linking of a thiol-containing fragment to a cysteine residue (native or engineered) near the allosteric site of a protein (Figure 7.8A). At equilibrium, the mixture will consist predominantly of the protein bound to the fragment with the highest binding affinity, which can be identified via MS/MS technology. By use of the covalent bond formation between the fragment and the protein, the affinity of the fragment for the target protein is amplified, which can be useful for detection of fragments with lower affinities. The advantage of the identification of smaller fragments instead of larger compounds is the fact that the chemical space of a binding pocket can be explored more thoroughly, via the identification of multiple fragments that can also bind to sub-pockets. The identified fragments can be further developed into high affinity ligands by



either merging fragments together or growing of a fragment. Afterwards, either the linkage can be removed or the compound can be used as a covalent allosteric ligand<sup>53,54</sup> (which is especially relevant in case of mutated residues correlated to disease development).

The disulfide tethering technique can be extremely interesting for ROR $\gamma$ t, because of the presence of a native Cys476 residue on H11, pointing towards the allosteric pocket (Figure 7.8B,C). This residue could therefore be an interesting starting point for tethering of disulfide fragments, to further explore the chemical space in the allosteric site and expand the chemical diversity of allosteric ROR $\gamma$ t ligands. Alternatively, other residues around the allosteric site could be mutated to a cysteine residue, if Cys476 is not optimally positioned for targeting the allosteric site. Interestingly, the Cys476 residue is also used for ligation of the covalent ligand **RTA-1701**, an allosteric clinical candidate mentioned in Chapter 1, that partly binds to the allosteric site and partly to an adjacent site as was discovered via docking studies and MD simulations (Figure 7.8B).<sup>55</sup> Since Cys476 is at the interface of the allosteric site and the adjacent site, the chemical space of both binding pockets can be explored via tethering of fragments to this residue. Related to this, future studies could also focus on linking the covalent ligand and an allosteric (isoxazole) ligand together to obtain a larger compound that can address both the allosteric site but also the adjacent pocket (Figure 7.8D). A plausible approach would be the attachment of the carbonyl of the covalent ligand to the pyrrole C-2 position of an isoxazole ligand, which points towards the covalent ligand via an open channel. However, since the acrylamide warhead (that reacts with the protein) is modified in this way, this approach will only succeed by the formation of a non-covalent ligand.

Besides the mentioned techniques, more diverse screening methods, such as those based on computational approaches<sup>56,57</sup>, fragment-based screening by X-ray crystallography<sup>58</sup>, and phage display<sup>59</sup> could be further optimized or developed towards allosteric site targeting.

## Closing thoughts

NRs have been shown to harbor multiple allosteric sites that can potentially be used for drug discovery. A number of promising examples of small molecule allosteric NR targeting strategies have recently been reported, illustrating the high potential of this concept. The work described in this thesis can be used as an inspiration for the further development of allosteric targeting approaches for NRs. Ultimately, combining allosteric targeting with other techniques, such as bitopic ligand design or allosteric PROTAC development, could be used to even more accurately and selectively force a desired biological response. Combined, the future for allosteric small molecule modulators of nuclear receptors is highly promising.

## Acknowledgements

The PROTAC research was performed in collaboration with scientists from the company Symeres (Nijmegen, The Netherlands) (Marcel Scheepstra, Koen Hekking, Luc van Hijfte and Rutger Folmer). Marcel Scheepstra performed the synthesis of the PROTAC ligands, for which he is greatly acknowledged. Iris Leijten-van de Gevel is kindly acknowledged for performing site-directed mutagenesis (mutations at the orthosteric site of ROR $\gamma$ t), together with Justin Houx, Ingrid Kolen and Lisanne Tempelaar during their bachelor projects. Furthermore, Mike Visnick (Reata Pharmaceuticals, Irving, Texas, USA) and Yigitcan Eken (Michigan State University, East Lansing, Michigan, USA) are kindly thanked for sharing the predicted binding mode of the covalent ROR $\gamma$ t ligand (RTA-1701 derivative).

## Experimental Section

### Biophysical assays

**ROR $\gamma$ t LBD expression and purification.** His<sub>6</sub>-ROR $\gamma$ t-LBD was expressed and purified as described in Chapter 2.

**TR-FRET coactivator and Fluor-MRL-871 recruitment assays.** TR-FRET assays and data analysis were performed as described in Chapter 2. Data were recorded in triplicate; error shown is standard deviation from the mean; curves are representative of > 2 independent experiments.

**Quantitative IL-17a mRNA RT-PCR assay.** Cell culture, RT-PCR experiments and data analysis were performed as described in Chapter 2. Data were recorded in triplicate; error shown is standard deviation from the mean; data are representative of two independent experiments.

**Western blotting.** EL4 cells (Sigma-Aldrich) were grown in DMEM (Gibco) supplemented with 10% FBS and 1% penicillin-streptomycin. The cells were seeded onto a 12-well plate (300,000 cells/well) and after 24 h, the cells were treated with PROTAC ligand (10  $\mu$ M) or DMSO. After 24 h incubation, the cells (approx.  $5 \times 10^6$ ) were pelleted in a conical tube by spinning at 2100 rpm for 5 min. The pellet was washed with ice cold PBS and again spinning at 2100 rpm for 5 min was performed. The cells were lysed for 30 min on ice using 0.5 mL lysis buffer (RIPA buffer; 150 mM NaCl, 1.0% IGEPAL® CA-630, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 8.0) containing protease inhibitor cocktail powder. The cell lysates were centrifuged by spinning at 12,000 rpm for 10 min at 4 °C, the supernatant was collected and the protein concentration was determined by absorbance (A<sub>280</sub> on nanodrop). Samples were mixed with 2x Laemmli buffer (12.5% glycerol, 1% SDS, 0.01% bromophenol blue, 0.1 M DTT, 50 mM Tris, pH 6.8) and denatured at 95 °C for 5 min. 2  $\mu$ g of total protein per sample was loaded on a 12% SDS-PAGE gel and run for 1 h at 150 V in TGS buffer (25 mM Tris, 192 mM Glycine, 0.1% SDS, pH 8.3). The proteins were transferred to a nitrocellulose membrane (GE, Whatman) using a standard procedure. After blotting, the membrane was cut in two identical pieces and blocked using 5 mL 5% (w/v) skim milk in PBST (PBS + 0.1% Tween-20) in a 50 mL tube on a tube roller. One membrane was then incubated with a Rabbit Anti-Mouse ROR $\gamma$ t antibody (ab207082, 0.658 mg/mL, 1:2000 dilution), while the other membrane was incubated with a Rabbit Anti-Mouse GAPDH antibody (ab181603, 1.485 mg/mL, 1:15,000 dilution), both in 5 mL 5% (w/v) skim milk in PBST for 1 h. Next, the membranes were washed three times in 5 mL PBST. Both membranes were incubated with a Goat Anti-Rabbit antibody (HRP-conjugated) (ThermoFisher #31466, 1 mg/mL, 1:1000 dilution) in 5 mL 0.5% (w/v) skim milk in PBST for 1 h. The membranes were washed three times in 5 mL PBST, the blots were then dried on tissue and placed in a container. TMB substrate (Sigma Aldrich #To440) was added, covering the entire blot (2.5 mL), and a picture was taken.

## References

1. Caboni, L. & Lloyd, D. G. Beyond the ligand-binding pocket: targeting alternate sites in nuclear receptors. *Med. Res. Rev.* **33**, 1081–1118 (2013).
2. Tice, C. M. & Zheng, Y.-J. Non-canonical modulators of nuclear receptors. *Bioorg. Med. Chem. Lett.* **26**, 4157–4164 (2016).
3. Jensen, A. A. & Bräuner-Osborne, H. Allosteric modulation of the calcium-sensing receptor. *Curr. Neuropharmacol.* **5**, 180–186 (2007).
4. Hassan, A. Q., Sharma, S. V & Warmuth, M. Allosteric inhibition of BCR-ABL. *Cell Cycle* **9**, 3710–3714 (2010).
5. Moore, T. W., Mayne, C. G. & Katzenellenbogen, J. a. Minireview: Not picking pockets: nuclear receptor alternate-site modulators (NRAMs). *Mol. Endocrinol.* **24**, 683–695 (2010).
6. Scheepstra, M. *et al.* Identification of an allosteric binding site for ROR $\gamma$ t inhibition. *Nat. Commun.* **6**, e8833 (2015).
7. Dalvie, D. K., Kalgutkar, A. S., Khojasteh-Bakht, S. C., Obach, R. S. & O'Donnell, J. P. Biotransformation Reactions of Five-Membered Aromatic Heterocyclic Rings. *Chem. Res. Toxicol.* **15**, 269–299 (2002).
8. Carlo, B., Donna, M. H. & Amos, B. Smith, L. Carboxylic Acid (Bio)Isosteres in Drug Design. *ChemMedChem* **8**, 385–395 (2013).
9. Corina, I. & Caira, M. *Drug Metabolism: Current Concepts.* (Springer The Netherlands, 2005). doi:10.1007/1-4020-4142-X
10. Maag, H. Prodrugs of Carboxylic Acids BT - Prodrugs: Challenges and Rewards Part 1. in (eds. Stella, V. J. *et al.*) 703–729 (Springer New York, 2007). doi:10.1007/978-0-387-49785-3\_20
11. Ouvry, G. *et al.* Discovery of phenoxyindazoles and phenylthioindazoles as ROR $\gamma$  inverse agonists. *Bioorganic Med. Chem. Lett.* **26**, 5802–5808 (2016).
12. Kroker, A. J. & Bruning, J. B. Review of the Structural and Dynamic Mechanisms of PPAR $\gamma$  Partial Agonism. *PPAR Res.* **2015**, 816856 (2015).
13. Leijten-van de Gevel, I. A., van Herk, K. H. N., de Vries, R. M. J. M. & Brunsveld, L. PDB ID: 6TDC. PPAR gamma ligand binding domain in complex with MRL-871. (2019). doi:10.2210/pdb6tdc/pdb
14. Lovering, F., Bikker, J. & Humblet, C. Escape from Flatland: Increasing Saturation as an Approach to Improving Clinical Success. *J. Med. Chem.* **52**, 6752–6756 (2009).
15. Ivanov, I. I. *et al.* The Orphan Nuclear Receptor ROR $\gamma$ t Directs the Differentiation Program of Proinflammatory IL-17+ T Helper Cells. *Cell* **126**, 1121–1133 (2006).
16. Isono, F., Fujita-Sato, S. & Ito, S. Inhibiting ROR $\gamma$ t/Th17 axis for autoimmune disorders. *Drug Discov. Today* **19**, 1205–1211 (2014).
17. Zhong, C. & Zhu, J. Small-Molecule ROR $\gamma$ t Antagonists: One Stone Kills Two Birds. *Trends Immunol.* **38**, 229–231 (2017).
18. Ueda, E. *et al.* High incidence of T-cell lymphomas in mice deficient in the retinoid-related orphan receptor RORgamma. *Cancer Res.* **62**, 901–909 (2002).
19. Liljevald, M. *et al.* Retinoid-related orphan receptor  $\gamma$  (ROR $\gamma$ ) adult induced knockout mice develop lymphoblastic lymphoma. *Autoimmun. Rev.* **15**, 1062–1070 (2016).
20. Guntermann, C. *et al.* Retinoic-acid-orphan-receptor-C inhibition suppresses Th17 cells and induces thymic aberrations. *JCI insight* **2**, e91127 (2017).
21. Guo, Y. *et al.* Inhibition of ROR $\gamma$ T Skews TCR $\alpha$  Gene Rearrangement and Limits T Cell Repertoire Diversity. *Cell Rep.* **17**, 3206–3218 (2016).
22. Imura, C. *et al.* A novel ROR $\gamma$ t inhibitor is a potential therapeutic agent for the topical treatment of psoriasis with low risk of thymic aberrations. *J. Dermatol. Sci.* **93**, 176–185 (2019).
23. Vankayalapati, H & Yerramreddy, V. (Arrien Pharmaceuticals LLC). Substituted 2,3-dihydro-1H-inden-1-one retinoic acid-related orphan nuclear receptor antagonists for treating multiple sclerosis. *PCT Int. Appl.* WO2015/038350 (2015).
24. Bronner, S. M., Zbieg, J. R. & Crawford, J. J. ROR $\gamma$  antagonists and inverse agonists: a patent review. *Expert Opin. Ther. Pat.* **27**, 101–112 (2017).
25. Noguchi, M. *et al.* Ternary complex of human ROR $\gamma$  LBD, inverse agonist and SMRT peptide shows a unique mechanism of corepressor recruitment. *Genes to Cells* **22**, 535–551 (2017).
26. Noguchi, M. *et al.* Ternary crystal structure of human ROR $\gamma$  ligand-binding-domain, an inhibitor and corepressor peptide provides a new insight into corepressor interaction. *Sci. Rep.* **8**, 17374 (2018).
27. Sun, N., Guo, H. & Wang, Y. Retinoic acid receptor-related orphan receptor gamma-t (ROR $\gamma$ t) inhibitors in clinical development for the treatment of autoimmune diseases: a patent review (2016-present). *Expert Opin. Ther. Pat.* **29**, 663–674 (2019).

28. Bonnevier, J., Guerra, A. & Goetz, C. Identification of novel cell surface markers on mouse and human TH17 cells (P6354). *J. Immunol.* **190**, 199.3 LP-199.3 (2013).
29. Chau, C. H., Steeg, P. S. & Figg, W. D. Antibody-drug conjugates for cancer. *Lancet (London, England)* **394**, 793–804 (2019).
30. Rossin, R. *et al.* Chemically triggered drug release from an antibody-drug conjugate leads to potent antitumour activity in mice. *Nat. Commun.* **9**, 1484 (2018).
31. Johnston, M. C. & Scott, C. J. Antibody conjugated nanoparticles as a novel form of antibody drug conjugate chemotherapy. *Drug Discov. Today Technol.* **30**, 63–69 (2018).
32. Schmid, D. *et al.* T cell-targeting nanoparticles focus delivery of immunotherapy to improve antitumor immunity. *Nat. Commun.* **8**, 1747 (2017).
33. Lai, A. C. & Crews, C. M. Induced protein degradation: an emerging drug discovery paradigm. *Nat. Rev. Drug Discov.* **16**, 101–114 (2017).
34. Ottis, P. & Crews, C. M. Proteolysis-Targeting Chimeras: Induced Protein Degradation as a Therapeutic Strategy. *ACS Chem. Biol.* **12**, 892–898 (2017).
35. Scheepstra, M., Hekking, K. F. W., van Hijfte, L. & Folmer, R. H. A. Bivalent Ligands for Protein Degradation in Drug Discovery. *Comput. Struct. Biotechnol. J.* **17**, 160–176 (2019).
36. Gao, H., Sun, X. & Rao, Y. PROTAC Technology: Opportunities and Challenges. *ACS Med. Chem. Lett.* **11**, 237–240 (2020).
37. Zengerle, M., Chan, K.-H. & Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chem. Biol.* **10**, 1770–1777 (2015).
38. Yang, K. *et al.* Development of the first small molecule histone deacetylase 6 (HDAC6) degraders. *Bioorg. Med. Chem. Lett.* **28**, 2493–2497 (2018).
39. Neklesa, T. *et al.* ARV-110: An oral androgen receptor PROTAC degrader for prostate cancer. *J. Clin. Oncol.* **37**, 259 (2019).
40. Flanagan, J. J. *et al.* Abstract P5-04-18: ARV-471, an oral estrogen receptor PROTAC degrader for breast cancer. *Cancer Res.* **79**, P5-4-18-P5-04-18 (2019).
41. Mullard, A. First targeted protein degrader hits the clinic. *Nat. Rev. Drug Discov.* **18**, 237–239 (2019).
42. Mullard, A. Arvinas PROTACs pass first safety and PK analysis. *Nat. Rev. Drug Discov.* **18**, 895 (2019).
43. Zhang, X., Crowley, V. M., Wucherpennig, T. G., Dix, M. M. & Cravatt, B. F. Electrophilic PROTACs that degrade nuclear proteins by engaging DCAF16. *Nat. Chem. Biol.* **15**, 737–746 (2019).
44. Ward, C. C. *et al.* Covalent Ligand Screening Uncovers a RNF4 E3 Ligase Recruiter for Targeted Protein Degradation Applications. *ACS Chem. Biol.* **14**, 2430–2440 (2019).
45. Konstantinidou, M. *et al.* The tale of proteolysis targeting chimeras (PROTACs) for Leucine-Rich Repeat Kinase 2 (LRRK2). *ChemMedChem* **15**, 1–8 (2020).
46. Hughes, S. J. & Ciulli, A. Molecular recognition of ternary complexes: a new dimension in the structure-guided design of chemical degraders. *Essays Biochem.* **61**, 505–516 (2017).
47. Gadd, M. S. *et al.* Structural basis of PROTAC cooperative recognition for selective protein degradation. *Nat. Chem. Biol.* **13**, 514–521 (2017).
48. Wurz, R. P. *et al.* A ‘Click Chemistry Platform’ for the Rapid Synthesis of Bispecific Molecules for Inducing Protein Degradation. *J. Med. Chem.* **61**, 453–461 (2018).
49. Beveridge, R. *et al.* Native Mass Spectrometry Can Effectively Predict PROTAC Efficacy. *ACS Cent. Sci.* **6**, 1223–1230 (2020).
50. Farnaby, W. *et al.* BAF complex vulnerabilities in cancer demonstrated via structure-based PROTAC design. *Nat. Chem. Biol.* **15**, 672–680 (2019).
51. Erlanson, D. A., Wells, J. A. & Braisted, A. C. Tethering: Fragment-Based Drug Discovery. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 199–223 (2004).
52. Hardy, J. A. A link means a lot: disulfide tethering in structure-based drug design. *Comput. Struct. Approaches to Drug Discov. Ligand-Protein Interact.* 318–347 (2008).
53. Lu, S. & Zhang, J. Designed covalent allosteric modulators: an emerging paradigm in drug discovery. *Drug Discov. Today* **22**, 447–453 (2017).
54. Bruning, J. M. *et al.* Covalent Modification and Regulation of the Nuclear Receptor Nurrl by a Dopamine Metabolite. *Cell Chem. Biol.* **26**, 674–685.e6 (2019).
55. Jiang, X. *et al.* A novel series of cysteine-dependent, allosteric inverse agonists of the nuclear receptor ROR $\gamma$ t. *Bioorg. Med. Chem. Lett.* **30**, e126967 (2020).
56. Huang, W. *et al.* Allosteric: A method for predicting allosteric sites. *Bioinformatics* **29**, 2357–2359 (2013).
57. Lu, S., Huang, W. & Zhang, J. Recent computational advances in the identification of allosteric sites in proteins. *Drug Discov. Today* **19**, 1595–1600 (2014).
58. Bauman, J. D. *et al.* Detecting Allosteric Sites of HIV-1 Reverse Transcriptase by X-ray Crystallographic Fragment Screening. *J. Med. Chem.* **56**, 2738–2746 (2013).
59. Pande, J., Szewczyk, M. M. & Grover, A. K. Phage display: Concept, innovations, applications and future. *Biotechnol. Adv.* **28**, 849–858 (2010).

## Exploring Allosteric Modulation of the Nuclear Receptor ROR $\gamma$ t from a Drug Discovery Perspective

Nuclear receptors (NRs) are a family of ligand-regulated transcription factors that are highly attractive drug targets because of their central role in several disease processes in the human body. Allosteric modulation of NRs, via the binding of small molecules to non-canonical, allosteric binding sites, has recently gained significant attention as a promising and novel drug discovery strategy. It offers several potential benefits over classical modulation via the canonical, orthosteric ligand binding pocket. However, allosteric sites are often not easily identified and thus discovering allosteric drugs presents a significant challenge.

ROR $\gamma$ t is an illustrative example of a NR that features both a clearly defined orthosteric binding site, but also a highly unique allosteric binding site, near helix 12 in the ligand binding domain. ROR $\gamma$ t plays an essential role in the immune system via the differentiation of T helper 17 cells and the expression of the inflammatory cytokine IL-17a. Inhibition of ROR $\gamma$ t with small molecules shows high potential in the treatment of autoimmune disorders by reducing the inflammatory response. In this thesis, the inhibition of ROR $\gamma$ t via the allosteric binding site has been further explored, with the aim to gain new insights into allosteric ROR $\gamma$ t modulation from a drug discovery perspective, but also to provide a learning platform for the translation of the knowledge to other NRs.

The majority of allosteric ROR $\gamma$ t modulators identified thus far are limited to one series of closely related indazoles, typified by **MRL-871**. Chapter 2 describes a highly efficient *in silico*-guided approach that led to the identification of a novel class of allosteric ROR $\gamma$ t inverse agonists, featuring a trisubstituted isoxazole core. *In silico* docking studies guided the optimization of an early hit compound, delivering **FM26** as the most potent lead compound with sub-micromolar inhibition in a coactivator recruitment assay and effective inhibition of cellular IL-17a expression levels. Crystallography studies proved the designed allosteric mode of action, and showed that the presence of a hydrogen bond donating moiety at the C-5 isoxazole position significantly increases the potency towards ROR $\gamma$ t by formation of an additional hydrogen bond interaction with the backbone of the protein. In Chapter 3, the structure activity relationship (SAR) profile of the isoxazoles is further explored. Structural and *in silico*-guided studies revealed that the potency is significantly increased via the incorporation of a lipophilic C-4 linker, and also highlighted the importance of the polar character of the C-5 moiety. The trisubstituted isoxazoles proved to be valuable leads as allosteric ROR $\gamma$ t inverse agonists, with low nanomolar potency, an increased selectivity profile towards ROR $\gamma$ t and promising ADME properties. Thus, this novel class of allosteric ROR $\gamma$ t ligands shows high potential for future development in drug discovery programs against autoimmune diseases.

In Chapter 4, biochemical studies revealed a cooperative dual ligand binding behavior between the orthosteric and allosteric binding sites of ROR $\gamma$ t, where the potency of an allosteric ligand is enhanced in the presence of an orthosteric ligand. The co-crystal structures of ROR $\gamma$ t in complex with all different combinations of orthosteric and allosteric ligands were elucidated, demonstrating a conformational change within the protein upon binding of an orthosteric ligand, which results in a clamping motion of the allosteric pocket. Molecular dynamics simulations shed light on the underlying molecular mechanism of this contraction behavior, highlighting the ability of Ala355 to switch between helix 4 and 5. Orthosteric ROR $\gamma$ t ligands promote the helix 5 conformation of this residue, thereby shifting helix 4 towards the allosteric pocket and enhancing the potency of the allosteric inverse agonists.

Although allosteric modulation of NRs shows high potential, the targeted discovery of novel allosteric ligands is highly challenging via the currently available methods. Chapter 5 describes how the occlusion of the orthosteric site of ROR $\gamma$ t, using covalent chemical probes, can underpin a dedicated screening methodology for the identification of novel allosteric ROR $\gamma$ t modulators. After SAR optimization of an initial hit compound, a set of chemical probes were identified that selectively ligate to a native cysteine residue in the orthosteric site of ROR $\gamma$ t and act as partial inverse agonists. In several binding assays, efficient inhibition of orthosteric ligand binding was observed, while the allosteric site was still capable of binding ligands. These covalent probes that occlude the orthosteric site of ROR $\gamma$ t are thus excellent tool compounds for the development of a screening method that could be used for the unambiguous and rapid identification of novel allosteric ROR $\gamma$ t ligands.

Chapter 6 describes an innovative approach for the dual targeting of the orthosteric and the allosteric binding sites of ROR $\gamma$ t by use of a bitopic ligand, that comprises of a covalently linked orthosteric and allosteric pharmacophore. The concomitant engagement of both binding pockets was validated by various biochemical assays, showing proof of a bitopic mode of action, with an improved inhibitory efficacy compared to the monovalent counterparts in both a biochemical and cellular context. Additionally, the bitopic ligand demonstrated increased selectivity for ROR $\gamma$ t over a selection of other NRs. These results highlight the potential advantages of bitopic NR modulation over monovalent targeting strategies.

Finally, the Epilogue (Chapter 7) presents a future perspective on the allosteric modulation of NRs, and ROR $\gamma$ t in particular, including potential challenges. Additional approaches for the identification of novel allosteric ligands and innovative techniques related to allosteric modulation are described as well.

Together, the data described in this thesis show a variety of successful applications for the allosteric modulation of ROR $\gamma$ t, ranging from the development of novel allosteric modulators and dual targeting strategies, to screening approaches for new allosteric ligands. The insights obtained in this research could be used as an inspiration for the further development of allosteric targeting approaches for NRs in drug discovery strategies.

## Onderzoek naar de Allosterische Modulatie van de Kernreceptor ROR $\gamma$ t vanuit het Perspectief van Medicijnontwikkeling

Kernreceptoren, ofwel hormoon receptoren, zijn een familie van eiwitten in het menselijk lichaam die zeer aantrekkelijke doelwitten voor geneesmiddelen zijn vanwege hun centrale rol in verschillende ziekteprocessen. De allosterische modulatie van kernreceptoren, via de binding van kleine moleculen aan alternatieve bindingsplaatsen, lijkt een veelbelovende strategie te zijn voor de ontwikkeling van nieuwe medicijnen. Het biedt een aantal mogelijke voordelen ten opzichte van de reguliere modulatie van kernreceptoren via de standaard bindingsplaats. Allosterische bindingsplaatsen zijn echter vaak niet gemakkelijk te vinden en daarom is het ontdekken van allosterische geneesmiddelen een grote uitdaging.

ROR $\gamma$ t is een voorbeeld van een kernreceptor die zowel een reguliere, orthostere bindingsplaats heeft, maar ook een unieke, allosterische bindingsplaats. ROR $\gamma$ t speelt een essentiële rol in het immuunsysteem via het opwekken van de signaalstof IL-17a die een ontstekingsreactie op gang brengt. Het blokkeren van ROR $\gamma$ t met kleine moleculen lijkt veelbelovend te zijn bij de behandeling van auto-immuunziekten, doordat dit de ontstekingsreactie zal verminderen. In dit proefschrift is het blokkeren van ROR $\gamma$ t via de allosterische bindingsplaats verder onderzocht, met als doel nieuwe inzichten te krijgen in de allosterische modulatie van ROR $\gamma$ t vanuit het perspectief van medicijnontwikkeling, maar ook als leerplatform voor de vertaling van de kennis naar andere kernreceptoren.

Het grootste deel van de tot nu toe geïdentificeerde allosterische moleculen (ook wel liganden genoemd) voor ROR $\gamma$ t is beperkt tot één reeks vergelijkbare moleculen, gekenmerkt door het ligand **MRL-871**. Hoofdstuk 2 beschrijft de identificatie van een nieuwe klasse van allosterische ROR $\gamma$ t liganden, met een tri-gesubstitueerde isoxazoolkern. Met behulp van *in silico* studies werd **FM26** geïdentificeerd als meest potente ligand, zowel in het remmen van coactivator binding in biochemische metingen, als in het remmen van cellulaire IL-17a expressie niveaus. Door middel van kristallografie studies werd het allosterische werkingsmechanisme bewezen en werd aangetoond dat de aanwezigheid van een waterstofbrug-donerende groep op de C-5 isoxazool positie de affiniteit sterk verhoogt door de vorming van een extra waterstofbrug-interactie met het eiwit. In hoofdstuk 3 wordt deze nieuwe klasse van isoxazool liganden verder onderzocht. Uitgebreide *in silico* en kristallografie studies lieten zien dat de affiniteit van de liganden significant hoger wordt in het geval van een lipofiele C-4 linker, en het belang van het polaire karakter van de C-5 groep wordt verder benadrukt. De tri-gesubstitueerde isoxazolen blijken zeer interessante allosterische liganden te zijn voor ROR $\gamma$ t, met een zeer hoge affiniteit (lage nanomolair), een verbeterd selectiviteitsprofiel voor ROR $\gamma$ t ten opzichte van andere kernreceptoren en veelbelovende farmacokinetische eigenschappen. Deze nieuwe klasse van allosterische ROR $\gamma$ t liganden biedt veel mogelijkheden voor de toekomstige ontwikkeling van geneesmiddelen tegen auto-immuunziekten.

In hoofdstuk 4 werd een coöperatief gedrag aangetoond tussen de orthostere en allosterie bindingsplaats van RORyt, waarbij de affiniteit van een allosteer ligand wordt versterkt in aanwezigheid van een orthosteer ligand. De kristalstructuren van RORyt, gebonden aan alle verschillende combinaties van orthostere en allosterie liganden, lieten een conformationele verandering van het eiwit zien na binding van een orthosteer ligand, wat resulteert in een 'klemmende' beweging rond de allosterie bindingsplaats. Moleculaire dynamica simulaties gaven meer inzicht in het onderliggende moleculaire mechanisme van deze contractie, waarbij het aminozuur Ala355 kan schakelen tussen helix 4 en 5. De orthostere RORyt liganden bevorderen de helix 5 conformatie van dit residu, waardoor helix 4 naar de allosterie bindingsplaats beweegt en de affiniteit van de allosterie liganden hoger wordt.

Hoewel de allosterie modulatie van kernreceptoren erg interessant is, blijft het ontdekken van nieuwe allosterie liganden een grote uitdaging via de methodes die momenteel beschikbaar zijn. Hoofdstuk 5 beschrijft het blokkeren van de orthostere bindingsplaats van RORyt met behulp van covalente moleculen, wat gebruikt kan worden als een methode om nieuwe allosterie liganden te kunnen identificeren. Er werden een aantal liganden geïdentificeerd die selectief aan het juiste cysteine residu in de orthostere bindingsplaats ligeren, en functioneren als partiële remmers. In verschillende bindingstesten werd een efficiënte remming van de binding van een orthosteer ligand waargenomen, terwijl de allosterie bindingsplaats nog steeds beschikbaar was voor het binden van liganden. Deze covalente liganden zouden dus gebruikt kunnen worden voor de ontwikkeling van een methode voor de snelle identificatie van nieuwe allosterie RORyt liganden.

Hoofdstuk 6 beschrijft de ontwikkeling van een bitopisch ligand, dat bestaat uit een covalent gekoppeld orthosteer en allosteer ligand, om op die manier beide bindingsplaatsen van RORyt te kunnen bezetten. Het tegelijkertijd binden in beide bindingsplaatsen werd bewezen door middel van verschillende biochemische metingen, waarbij het bitopische ligand leidt tot een sterkere inhibitie van het eiwit dan de afzonderlijke monovalente liganden, in zowel een biochemische als cellulaire context. Bovendien vertoont het bitopische ligand een verhoogde selectiviteit voor RORyt. Deze resultaten laten de potentiële voordelen zien van bitopische modulatie ten opzichte van monovalente strategieën.

Ten slotte beschrijft de Epiloog (hoofdstuk 7) een toekomstperspectief voor de allosterie modulatie van kernreceptoren, en RORyt in het bijzonder. Ook worden een aantal innovatieve technieken beschreven voor de identificatie van nieuwe allosterie liganden.

Samenvattend beschrijft de data in dit proefschrift een aantal succesvolle toepassingen voor de allosterie modulatie van RORyt, variërend van de ontwikkeling van een nieuwe klasse van allosterie liganden en bitopische liganden, tot het screenen van nieuwe allosterie liganden. De inzichten die in dit onderzoek zijn verkregen, kunnen worden gebruikt als inspiratie voor het verder onderzoeken van allosterie liganden voor kernreceptoren als strategie voor het ontwikkelen van nieuwe geneesmiddelen.



## Curriculum Vitae



Femke Meijer was born on September 30<sup>th</sup>, 1993 in Breda (the Netherlands). After finishing her pre-university education at the Stedelijk Gymnasium in Breda, she started studying Biomedical Engineering at the Eindhoven University of Technology in 2011. She obtained her Bachelor of Science (BSc) degree *cum laude* in 2014 and continued with the Master's track of Chemical Biology at the same university. She also participated in a 1-year honors program of the Netherlands Research School of Chemical Biology (NRSCB). During her graduation project, she worked on the structure-based design and synthesis of peptide-small molecule hybrids for the modulation of 14-3-3 protein-protein interactions under the guidance of prof. Luc Brunsveld and dr. Christian Ottmann. Afterwards, she performed a research internship in the group of dr. Anna Barnard and prof. Ed Tate at Imperial College in London on the chemical synthesis of functionalized  $\alpha$ -helix mimetics. After obtaining her Master of Science (MSc) degree *cum laude* in 2017, she started her PhD project under supervision of prof. Luc Brunsveld in the Chemical Biology group at the Eindhoven University of Technology. The goal of her research was to explore the modulation of the nuclear receptor ROR $\gamma$ t via an allosteric binding site. The most important results of this research are described in this thesis.

Femke Meijer werd geboren op 30 september 1993 te Breda. Na het behalen van haar gymnasium diploma aan het Stedelijk Gymnasium te Breda begon zij de studie Biomedische Technologie aan de Technische Universiteit Eindhoven. Ze behaalde haar Bachelor of Science (BSc) graad *cum laude* in 2014 en studeerde door in de Mastertrack Chemische Biologie aan dezelfde universiteit. Daarnaast heeft ze deelgenomen aan een eenjarig honors programma van de 'Netherlands Research School of Chemical Biology' (NRSCB). Tijdens haar afstudeerproject heeft ze, onder begeleiding van prof. Luc Brunsveld en dr. Christian Ottmann, gewerkt aan het ontwerpen en synthetiseren van peptiden, gekoppeld aan kleine moleculen, met als doel het moduleren van 14-3-3 eiwit-eiwit interacties. Vervolgens heeft ze stage gelopen in de groep van dr. Anna Barnard en prof. Ed Tate aan Imperial College in Londen, waar ze gewerkt heeft aan de chemische synthese van gefunctionaliseerde  $\alpha$ -helix mimetica. Na het *cum laude* behalen van haar Master of Science (MSc) graad in 2017, begon ze haar promotieonderzoek onder begeleiding van prof. Luc Brunsveld in de Chemische Biologie groep aan de Technische Universiteit Eindhoven. Het doel van haar onderzoek was het bestuderen van de modulatie van de kernreceptor ROR $\gamma$ t via een allosterische bindingsplaats. De belangrijkste resultaten van dit promotieonderzoek staan beschreven in dit proefschrift.



## List of Publications

**F.A. Meijer**, A.O.W.M. Saris, R.G. Doveston, G.J.M. Oerlemans, R.M.J.M. de Vries, B.A. Somsen, A. Unger, B. Klebl, C. Ottmann, P.J. Cossar & L. Brunsveld. Structure Activity Relationship Studies of Trisubstituted Isoxazoles as Selective Allosteric Ligands for ROR $\gamma$ t. *J. Med. Chem.* In press (2021).

**F.A. Meijer**<sup>#</sup>, M.C.M. van den Oetelaar<sup>#</sup>, R.G. Doveston, E.N.R. Sampers & L. Brunsveld. Covalent Occlusion of the ROR $\gamma$ t Ligand Binding Pocket Allows Unambiguous Targeting of an Allosteric Site. *ACS Med. Chem. Lett.* **12**, 631-639 (2021).

**F.A. Meijer**<sup>#</sup>, G.J.M. Oerlemans<sup>#</sup> & L. Brunsveld. Orthosteric and Allosteric Dual Targeting of the Nuclear Receptor ROR $\gamma$ t with a Bitopic Ligand. *ACS Chem. Biol.* **16**, 510-519 (2021).

**F.A. Meijer**<sup>#</sup>, R.M.J.M. de Vries<sup>#</sup>, R.G. Doveston, I.A. Leijten-van de Gevel & L. Brunsveld. Cooperativity between the Orthosteric and Allosteric Ligand Binding Sites of ROR $\gamma$ t. *Proc. Natl. Acad. Sci.* **118**, e2021287118 (2021).

L. Brunsveld, R.G. Doveston, S. Leysen, L.-G. Milroy, **F.A. Meijer**, M. Scheepstra & C. Ottmann. Substituted Heterocyclic Compounds and their Use as Retinoic Acid Receptor-Related Orphan Receptor (ROR) gamma-t Inhibitors. *PCT Int. Appl.* WO2020/149740 (2020).

R.M.J.M. de Vries, R.G. Doveston, **F.A. Meijer** & L. Brunsveld. Elucidation of an Allosteric Mode of Action for a Thienopyrazole ROR $\gamma$ t Inverse Agonist. *ChemMedChem* **15**, 561-565 (2020).

**F.A. Meijer**<sup>#</sup>, R.G. Doveston<sup>#</sup>, R.M.J.M. de Vries, G.M. Vos, A.A.A. Vos, S. Leysen, M. Scheepstra, C. Ottmann, L.-G. Milroy & L. Brunsveld. Ligand-Based Design of Allosteric ROR $\gamma$ t Inverse Agonists. *J. Med. Chem.* **63**, 241-259 (2020).

**F.A. Meijer**<sup>#</sup>, I.A. Leijten-van de Gevel<sup>#</sup>, R.M.J.M. de Vries<sup>#</sup> & L. Brunsveld. Allosteric Small Molecule Modulators of Nuclear Receptors. *Mol. Cell. Endocrinol.* **485**, 20-34 (2019).

**F.A. Meijer**<sup>#</sup>, S.A. Andrei<sup>#</sup>, J. Filipe Neves, L. Brunsveld, I. Landrieu, C. Ottmann & L.-G. Milroy. Inhibition of 14-3-3/Tau by Hybrid Small-Molecule Peptides operating via Two Different Binding Modes. *ACS Chem. Neurosci.* **9**, 2639-2654 (2018).

J.W. Prokop, N.C. Yeo, C. Ottmann, S.B. Chhetri, K.L. Florus, E.J. Ross, N. Sosonkina, B.A. Link, B.I. Freedman, C.J. Coppola, C. McDermott-Roe, S. Leysen, L.-G. Milroy, **F.A. Meijer**, A.M. Geurts, F.J. Rauscher 3rd, R. Ramaker, M.J. Flister, H.J. Jacob, E.M. Mendenhall & J. Lazar. Characterization of Coding/Noncoding Variants for SHROOM3 in Patients with CDK. *J. Am. Soc. Nephrol.* **29**, 1525-1535 (2018).

M. Scheepstra, S.A. Andrei, R.M.J.M. de Vries, **F.A. Meijer**, J.-N. Ma, E.S. Burstein, R. Olsson, C. Ottmann, L.-G. Milroy & L. Brunsveld. Ligand Dependent Switch from RXR Homo- to RXR-NURR1 Heterodimerization. *ACS Chem. Neurosci.* **8**, 2065-2077 (2017).

L.M. Stevers, C.V. Lam, S.F.R. Leysen, **F.A. Meijer**, D.S. van Scheppingen, R.M.J.M. de Vries, G.W. Carlile, L.-G. Milroy, D.Y. Thomas, L. Brunsveld & C. Ottmann. Characterization and Small-Molecule Stabilization of the Multisite Tandem Binding between 14-3-3 and the R domain of CFTR *Proc. Natl. Acad. Sci.* **113**, e11152-e11161 (2016).

# These authors contributed equally

## Acknowledgements

En dan komt mijn tijd op de TU/e toch echt tot een einde: in 2011 begon ik als een onbevangen eerstejaars en nu, bijna 10 jaar later, sluit ik m'n tijd hier af als een (bijna!) doctor. Wat was het een mooi avontuur en wat heb ik veel mogen leren op zowel wetenschappelijk als persoonlijk vlak. De afgelopen 4 jaar promotieonderzoek zijn voorbij gevlogen, en dan is nu 'opeens' je proefschrift af en ben je bij het laatste onderdeel aangekomen: het dankwoord! In de afgelopen jaren heb ik één ding zeker geleerd: samen kom je zo veel verder dan alleen. Dit promotieonderzoek had ik dan ook nooit kunnen uitvoeren zonder de directe en indirecte hulp van heel veel geweldige mensen, die ik via deze weg ontzettend wil bedanken.

Allereerst wil ik mijn promotor Luc Brunsveld enorm bedanken. Luc, tijdens de studie had ik je al leren kennen als een zeer bevlogen docent tijdens de colleges biochemie en moleculaire celbiologie; je bracht je enthousiasme dan ook snel op mij over! Nadat ik zowel m'n bachelor eindproject als afstudeerproject in de Chemische Biologie groep had gedaan, had ik nog lang geen genoeg van het drug discovery onderzoek, sterker nog: ik was er juist alleen maar enthousiaster over geworden! Daarom was ik ook erg blij toen je me een promotieplek aanbood in je groep, waarbij het onderzoek gericht zou zijn op de allosterische modulatie van kernreceptoren vanuit een drug discovery invalshoek. Nogmaals bedankt dat je me de kans hebt gegeven om aan zo'n fantastisch mooi en uitdagend project te kunnen werken. Ik heb veel bewondering voor de manier waarop je zo'n grote onderzoeksgroep draaiende weet te houden, waarbij je iedereen veel vrijheid geeft binnen het onderzoek, maar toch ook op de juiste momenten persoonlijke begeleiding weet te bieden. Ik heb ontzettend veel van je mogen leren; van het kritisch analyseren van resultaten, tot het ontwikkelen van een innovatieve mindset, maar ook het beknopt beschrijven van de data (waarschijnlijk was deze thesis anders nog 200 pagina's dikker geweest, haha!). Nogmaals bedankt voor je betrokkenheid, je inspirerende manier van begeleiden en je altijd oprechte interesse, zowel in de ontwikkelingen van mijn onderzoek als in mij als persoon. Ik wens je het allerbeste!

The next person I would like to thank is my co-promotor, Richard Doveston. Richard, when I started my PhD four years ago, you already worked on the RORyt research as a postdoc. In the first months of my project, you took the time to teach me a lot of techniques, but apart from that, I also learned a lot from you as a researcher in general, including how to have efficient meetings (you are really good at that!). I am very happy that after you left the TU/e, we always kept in contact and that you have always been so involved and interested in the research. I could always ask for any tips, suggestions and advice on experiments and manuscripts, which was very helpful and from which I learned a lot! Richard, thanks a lot for your tremendous support! I think I can say that without your initial results, valuable input and sharp observations, the research would have never been at this level. I wish you all the best and I hope to keep in touch!

Of course, I would also like to thank the members of my defense committee for their time and effort to assess the work presented in this thesis. Christian Ottmann, dankjewel voor je betrokkenheid bij mijn onderzoek en je nuttige feedback tijdens meetings. Je vrolijke en enthousiaste instelling werken zeer aanstekelijk! Prof. Bert Meijer, ik weet nu nog dat u bij de start van m'n PhD tijdens de blauwe lijst meeting vertelde hoe bevoorrecht we waren dat we in deze werkomgeving mochten werken, met alle faciliteiten en kennis in huis, en dat we die optimaal moesten benutten. Ik heb dit altijd in m'n achterhoofd gehouden en heb m'n best gedaan om er het beste uit te halen. Bedankt voor alle interessante colleges tijdens de opleiding, voor uw input en scherpe vragen over mijn onderzoek tijdens colloquia en voor uw bereidheid om deel te nemen aan m'n promotiecommissie. Anna Barnard, many thanks for being part of my defense committee and for critically reading my thesis. I have great memories of my time in London, where I undertook my internship project under your supervision. You are a great researcher and I really learned a lot from you. I hope to visit you guys again soon in London! Prof. Mario van der Stelt, tijdens het NRSCB honors programma kwam ik erachter hoe sterk het type onderzoek dat jullie in Leiden doen overeenkomt met het onderzoek in onze groep, en in combinatie met uw achtergrond in het RORyt onderzoek past u naar mijn idee perfect in m'n commissie. I wil u hartelijk bedanken voor uw interesse in mijn werk en voor uw deelname aan mijn promotiecommissie. Lastly, prof. Ulrike Holzgrabe, thanks a lot for being part of my committee and I am looking forward to a nice discussion during my defense. Verder wil ik de voorzitter, Maarten Merks, bedanken voor het leiden van de promotiezitting.

Dan komen we bij mijn twee paranimfen: Rens & Iris! Ik voel me heel bevoorrecht dat ik zo veel met jullie heb mogen samenwerken de afgelopen jaren; dat heb ik altijd als heel positief ervaren. Met z'n drietjes waren we een perfect team waarin ieder z'n eigen expertise had en waarin we elkaar heel mooi konden aanvullen en uitdagen. Daarnaast heb ik de conferenties waar we samen heen zijn geweest altijd heel gezellig gevonden, we hebben wat afgelachen! Rens, het aantal publicaties waar onze beide namen op staan, verraadt eigenlijk al aan hoeveel onderzoeksprojecten wij hebben samengewerkt in de afgelopen jaren, waarbij we elkaar altijd hebben gestimuleerd er het maximale uit te halen. Bedankt voor alle gezelligheid, je droge humor, je positieve instelling en je hulp met alles wat je je maar kunt bedenken (figuren maken, experimenten uitdenken, computerproblemen oplossen, gewoon even m'n verhaal aanhoren, en ik kan hier nog wel 100 andere dingen noemen). In het bijzonder wil ik je bedanken voor het werkend krijgen van de kristallografie van RORyt; ik durf wel te zeggen dat we dit zonder jou nooit voor elkaar hadden gekregen en ik bewonder je doorzettingsvermogen echt enorm! Iris, we hebben elkaar leren kennen tijdens NRSCB en zijn sindsdien bevriend geraakt en daarna werden we ook nog collega's! Onze nieuwsgierigheid (ik weet het, ik ben nóg iets erger...) leverde vaak grappige situaties op en samen hebben we heel wat avonturen beleefd; onder andere de reis naar Heidelberg zullen

we niet snel meer vergeten, haha! Ik ben blij dat we elkaar altijd zo mooi konden aanvullen in het onderzoek, in combinatie met heel veel gezelligheid. Ik vind het ontzettend knap hoe je je werk weet te combineren met jullie gezinnetje en ik wens jullie heel veel geluk samen.

Natuurlijk kan de BitterbalBBQ/lunch/koffie&thee-groep niet ontbreken in dit dankwoord. Rens, Iris, Simone, Lenne, Maaïke, Daan, Suzanne en Wiggert, super bedankt voor alle gezelligheid tijdens de lunches, borrels, en koffie/theepauzes (waarbij jullie er meestal voor zorgden dat ik er tijd voor maakte). Simone, jij staat altijd voor iedereen klaar en we konden altijd ons hart even bij elkaar luchten, bedankt daarvoor. Ik was altijd blij als ik het biolab binnen kwam en mijn bench-buurvrouw er ook was, dat maakte het pipetteren een stuk gezelliger! Dankjewel ook voor al die lekkere baksels en hopelijk mogen we weer snel helpen tappen bij Guus Meeuwis! Lenne, jouw eeuwige optimisme, fantastische humor én inspirerende lunchmaaltijden ga ik niet snel vergeten, dankjewel voor al je gezelligheid en heel veel geluk samen met Maaïke! Maaïke, ook al was je niet altijd in Helix, als je er was dan had je ook altijd tijd voor een gezellig praatje of natuurlijk voor het uitproberen van een nieuwe snack in de kantine. Veel succes met het afronden van je onderzoek, maar met jouw enthousiasme en nuchtere instelling komt dat vast helemaal goed! Daan, wie had dat gedacht toen wij 10 jaar geleden samen begonnen in hetzelfde introgroepje; dat we nu bijna tegelijkertijd ons promotieonderzoek zouden afronden. Ik moet nog lachen als ik terugdenk aan je zang- en karaoke-kunsten, je mooie dansmoves en je altijd terugkerende tanks-grappen. Suzanne, als het niet over formule 1 ging tijdens de lunch, dan ging het wel over welk nieuw plantje je wilde kopen voor op kantoor; een green office, maar dan nóg iets greener! Door jouw altijd rijke fantasie en je mooie, sarcastische opmerkingen was iedere lunch een feestje. Ik ben blij om te zien dat je weer vol enthousiasme met je project aan de gang bent en ik wens je heel veel succes met het afronden van je PhD! Wiggert, ik kan me weinig momenten herinneren dat ik je niet met een big smile op je gezicht heb gezien; jij bent altijd de vrolijkheid zelf. Met die positieve houding en al je creatieve en innovatieve ideeën ga jij het denk ik nog heel ver schoppen in de wetenschap! Dank jullie wel allemaal voor de fantastische tijd en veel succes in de toekomst!

Verder wil ik uiteraard alle collega's en oud-collega's binnen MST bedanken voor de fijne werksfeer en de gezellige activiteiten (borrels, retreats, sportdagen en groepsuitjes), maar natuurlijk ook voor de nuttige input en ideeën tijdens meetings. Een paar mensen wil ik graag in het bijzonder bedanken.

Lech, I would like to thank you for all your help and ideas during the first year of my project, before you started your career in industry. Your passion for science and drug discovery has always inspired me and I enjoyed working under your guidance. Additionally, you always try to get the best out of people in a nice and positive way, which I really like. Thanks again and I wish you and your family all the best!

Loes, alweer 7 jaar geleden begon ik met m'n bachelor eind project onder jouw begeleiding. Je hebt me toen al heel enthousiast gemaakt voor het doen van onderzoek en ik heb ontzettend veel van je mogen leren. Jij hebt er ook voor gezorgd dat ik me heel zelfverzekerd ging voelen in het lab, want wat vond ik het toen allemaal nog spannend! Dankjewel voor je positieve en enthousiaste begeleiding en later natuurlijk voor het zijn van een hele gezellige collega.

Sebastian, ik kende je eerst vooral als de broer van Paul, maar al snel bleek dat je een ontzettend gaaf promotieonderzoek deed en ik heel graag m'n afstudeeronderzoek bij jou wilde gaan doen. Hoewel onze manier van werken echt compleet verschillend is, heb ik heel veel van je geleerd. Jouw creativiteit en relaxte manier van werken hebben me altijd geïnspireerd. Ik ben heel bij dat we twee jaar later ook nog collega's werden en zelfs bij elkaar op kantoor kwamen te zitten; daarover later meer!

Marcel, in het eerste jaar van m'n master wilde ik graag nog wat extra ervaring in het lab opdoen en koos ik voor een onderzoeksproject in de 'vrije ruimte'. Ik weet nog dat Luc vertelde over jouw project en de combinatie van Nuclear Receptor onderzoek en organische synthese klonk meteen interessant (toen al! :)). Jij hebt me er tijdens het project extra enthousiast voor gemaakt én me de kneepjes van het vak geleerd in het chemisch lab, dankjewel daarvoor!

Dan nog wat kortere bedankjes, maar niet minder gemeend! Pim, jouw creatieve ideeën en kritische blik op onderzoek hebben me altijd geïnspireerd en uitgedaagd, en iedere meeting had je wel weer een aantal goede suggesties of tips. Dankjewel daarvoor en veel succes en plezier met je postdoc! Peter, I was very happy when you started a few years ago in our group as a postdoc with a lot of experience in organic synthesis. Thanks for being such a nice and social colleague, for always being so interested, and for giving valuable input and suggestions. Bente, bedankt voor de gezellig sfeer die je bracht in de groep en super bedankt voor het uitwerken van de laatste RORyt kristalstructuren samen met Guido; zonder jullie hulp hadden we dit zeker niet zo snel voor elkaar gekregen. Emira, respect voor hoe hard jij werkt en hoe jij je kunt verdelen tussen twee onderzoeksgroepen, dat lijkt me zeker niet makkelijk! Ook jou wil ik bedanken voor al je gezelligheid, onder andere tijdens de Dutch Medicine Days in Leiden, samen met Bente en Iris. Sylvia, jou heb ik (lang, lang geleden) nog mogen begeleiden als BEP student en ik vind het heel leuk dat je ervoor hebt gekozen om een PhD te gaan doen en dat we uiteindelijk collega's zijn geworden. Het is super om te zien hoe goed je nu bezig bent in je PhD, veel succes en natuurlijk ook bedankt voor het meten van alle Q-TOF samples! Auke, ook jou heb ik in een ver verleden nog als BEP student mogen begeleiden. Daarom vind ik het extra leuk dat we collega's zijn geworden en zo fijn konden sparren over het RORyt onderzoek en de synthese. Hopelijk kun je dit mooie onderzoek nog een beetje levende houden; er valt nog heel wat te halen! Eline, ik weet nog dat ik ooit twijfelde of ik een PhD zou gaan doen en jij (na een middagje samen kristallografie condities pipetteren) alle twijfels bij me had weggenomen door al je enthousiaste en positieve verhalen.



Thanks voor al je fijne tips en adviezen, ook tijdens de afronding van m'n project. Heel veel succes bij AmbAgon, maar dat komt vast goed met al jouw creatieve ideeën, en veel geluk samen met Sam en jullie gezinnetje. Roy, hoe wordt synthese op het lab nóg leuker? Met een gezellige zuurkast-buurman natuurlijk! Bedankt voor alle gezellige lab-uurtjes en voor de nodige keren dat we onze frustraties even bij elkaar kwijt konden.

Of course I would like to thank my office mates for the nice working atmosphere! Annelies, weet je nog dat ik bij jou op kantoor kwam te zitten en Luc zei: 'maar het moet niet té gezellig worden, hè dames ;)'. Gezellig was het zeker, maar dat bevorderde de werksfeer zullen we maar zeggen! Bedankt voor alle discussies over onderzoek, goede gesprekken en alle grappige momenten. Sebastian, jij was de meest chille kantoorgenoot die je je maar kunt voorstellen; muziekje aan, beentjes omhoog en gaan. Natuurlijk konden we goed sparren over onderzoek (én muziek), maar soms was het ook gewoon even tijd voor ...: het 'roddel' uurtje! Ik mis die gezellige tijd en leuke dinertjes nog steeds, kom maar snel weer eens langs! Zoals jij het zou zeggen: 'dat was toch goud!'. Heel veel succes met je postdoc in Londen en het ga je goed! Teresa, you are such a hard worker, that's incredible; always busy, but also always kind and happy. Thank you for the nice times and for the good atmosphere in our office. Dave, it was a pleasure to share an office with you; you are such a social guy and always interested. I hope I have taught you about some typical Dutch traditions and food - it was a pleasure. I wish you all the best and good luck with your PhD, I'm sure you will do great.

Tijdens mijn promotietraject heb ik een heel aantal talentvolle studenten mogen begeleiden die allemaal een hele belangrijke bijdrage hebben geleverd aan dit proefschrift. Ella, toen Richard z'n tijd op de TU afsloot mocht ik jou als eerste master student gaan begeleiden. Het was een heel uitdagend onderzoek waarin je sterk bent gegroeid, en ondanks dat het op het einde wat tegenzat, heb je het project mooi kunnen afsluiten en de fundering gelegd voor het werk in hoofdstuk 5. Heel veel succes en plezier met je avontuur in Zweden! Annet, jouw inzet en enthousiasme hebben ontzettend veel bijgedragen aan de optimalisatie van de isoxazool synthese, en mede dankzij jou zijn we die lastige synthese de baas geworden. Het was een plezier om je te mogen begeleiden en hoe leuk is het dat we binnenkort collega's worden bij Symeres! Maxime, als er iemand een doorzetter is, dan ben jij het! Hoewel het project in het begin tegen zat en lastig reproduceerbaar was, heb jij ervoor gezorgd dat we de ligatie van de probes uiteindelijk werkend hebben gekregen, wat geleid heeft tot een hele mooie publicatie. Je mag trots zijn op jezelf en ik vind het mooi om te zien hoe zelfstandig en enthousiast je nu bent begonnen in je PhD onderzoek, zet 'm op! Guido, ik weet nog hoe spannend je het vond om je afstudeeronderzoek op een volledig nieuw onderwerp te gaan doen, maar wat heb je er een mooi project van gemaakt! Je hebt het heel zelfstandig opgepakt, iedere stap perfect uitgedacht, en alle data tot in het kleinste detail geanalyseerd. Je bent een topper en ik ben ontzettend trots op de publicatie die uit dit project is gerold. Heel veel succes met je PhD, maar ik ben ervan overtuigd dat je dat super gaat doen! Naast deze vier master

studenten heb ik ook nog een aantal bachelor studenten mogen begeleiden. Kim, Thom, Michelle, Guido en Maxime, het was ontzettend leuk om jullie te mogen begeleiden tijdens jullie projecten in het chemisch lab, en ook best uitdagend omdat ik zelf ook nog maar net begonnen was met m'n PhD. Justin, Lisanne en Ingrid; ook al konden jullie zelf de experimenten niet afmaken i.v.m. de lockdown, jullie hebben gelukkig toch heel veel kunnen leren in het lab en de mutant eiwitten lijken hele interessante resultaten te laten zien! Wie weet wat daar nog voor mooi onderzoek uit gaat rollen. Bedankt allemaal voor jullie enthousiasme!

Dat het reilen en zeilen in een grote groep als MST zo goed geregeld kan zijn, heeft me altijd positief verrast. Dit is te danken aan de inzet van de MST stafleden die ik hiervoor heel erg wil bedanken. Tanja en Marjo, ontzettend bedankt voor het regelen van alle administratieve zaken, en dat waren er stiekem heel wat! Met al die nieuwe systemen was het soms nog best een uitdaging, maar jullie wisten het altijd weer snel en netjes op te lossen. Peggy, als er iemand geschikt is als lab-manager dan ben jij het wel! Hoe doe je dat toch in het Biolab: altijd alles op orde, bestellingen worden geregeld, het lab ziet er piekfijn uit, tussendoor nog onderwijs verzorgen voor honderden studenten, én je bent (bijna ;) ) altijd de vrolijkheid zelve. En dan maakte je vaak ook nog tijd om even mee te lunchen of koffie te drinken. Chapeau en bedankt! Ralf, dankjewel voor het georganiseerd houden van alle zaken in het chemisch lab. Hans, ik wil je ontzettend bedanken voor het altijd snel en netjes afhandelen van alle bestellingen; zonder jou zou dat nooit zo soepel kunnen lopen! Het analytisch team; super bedankt voor het draaiende houden van alle apparatuur in het analytisch lab en voor jullie eindeloze geduld als de LC-MS voor de zoveelste keer kapot was. Joost, jou wil ik ook nog bedanken voor alle HR-MS metingen en het helpen uitwerken van de data. Jolanda, mede dankzij jou heb ik een solide basis kunnen leggen voor de organische synthese, toen ik 8 jaar geleden samen met Annelies bij jou op het lab kwam werken tijdens een summer internship. Ik had toen ook al direct door dat jouw rol binnen MST er niet zomaar een was. Dankjewel voor het creëren van een fijne én veilige werkomgeving en voor alle random vragen waarvoor ik bij je kon aankloppen. Christine, jou wil ik bedanken voor het schoon houden van alle kantoren en de labs, en voor al je mooie verhalen.

Tijdens mijn project heb ik ook met verschillende mensen van buiten de TU mogen samenwerken. Marcel, het idee van het PROTAC project was super gaaf en ik heb met veel plezier aan het project gewerkt. Hoewel de experimenten niet het verwachte resultaat lieten zien, blijft het een mooi concept en wie weet dat het in de toekomst nog eens opgepakt gaat worden. Ik wil Marcel, en ook Luc van Hijfte, Rutger Folmer en Koen Hekking van Symeres (Mercachem, Nijmegen) bedanken voor de prettige samenwerking. I would also like to thank Bert Klebl, Anke Unger and Matthias Baumann from the Lead Discovery Centre in Dortmund for performing ADME measurements and for critically assessing the data. The data were a very valuable addition to the publications and have lifted the isoxazole research to a higher

level. I'm curious to see if the additional pharmacokinetic and *in vivo* experiments can bring the research one step further.

Hoewel ik het maken van figuren aardig onder de knie heb gekregen tijdens m'n PhD, zou ik de fantastische cover van mijn proefschrift nooit zelf zo gaaf hebben kunnen ontwerpen. Koen en Milan van de ICMS animatie studio: super bedankt hiervoor, het is echt een kunstwerkje geworden en precies zoals ik het voor ogen had!

Gelukkig was er naast m'n onderzoek nog genoeg tijd over voor ontspanning. Ik wil m'n vrienden dan ook bedanken voor alle gezellige avondjes, bbq's, uitjes, theetjes en wijntjes, met een paar bijzondere bedankjes. Natuurlijk de TU/e chickies; vanaf het 1<sup>e</sup> jaar van de studie kennen we elkaar al en ondanks dat iedereen z'n eigen weg is gegaan, hebben we gelukkig nog veel contact. Hopelijk kunnen we snel weer met z'n allen een goed feestje bouwen, thanks voor alle gezelligheid! Nina, wie had dat gedacht op de middelbare school, dat we allebei dr. zouden worden. Wat ben ik blij dat we elkaar nog steeds regelmatig zien om gezellig bij te kletsen. Anniek en Anouk, onze vakgebieden zijn denk ik de drie meest uitersten van elkaar die je je maar kunt bedenken, maar dat maakt ons dan misschien ook zo'n mooi clubje. Onze gezellige avondjes geven me altijd weer nieuwe energie, dankjewel! Maaïke, onze lunch wandelingen waren altijd weer een goede oppepper voor ons beiden denk ik! Ooit waren we bestuurs-genootjes bij Quadrivium, en nu allebei bijna dr. én een stuk burgerlijker ;). De Zythologie-groep uit Knegsel, naast het samen muziek maken wil ik jullie bedanken voor de gezellige feestjes, etentjes en uitstapjes, hopelijk snel weer met z'n allen! Iedereen van het studentenorkest, en in het bijzonder Maureen, Marjolein (sax matties!), Maaïke, Emma, Marieke en Marieke, bedankt voor alle fijne repetities als afleiding na een drukke werkdag en we houden contact! Sanne, alweer een paar jaar geleden tijdens NRSCB leerden we elkaar kennen en daarna kwamen we zelfs bij dezelfde afstudeerbegeleider terecht. Wie weet heb je nog wat leuke inspiratie op kunnen doen in dit proefschrift om je leerlingen nóg iets enthousiaster te maken tijdens de lessen scheikunde ;).

M'n schoonfamilie, de Heezemansjes: ja, ja, vanaf september ga ik daar ook officieel bij horen! Hopelijk zijn de termen 'eiwit', 'molecuul' en 'kernreceptor' jullie ondertussen iets meer bekend; ik vind het mooi hoe geïnteresseerd jullie altijd zijn in m'n onderzoek, hoe ver het ook van jullie eigen veld af staat. Ans en Jos, chapeau dat jullie de A2-route nog steeds op een dag af rijden om alle kids in een dag te kunnen zien; Amsterdam, Utrecht, Den Bosch en Eindhoven: we hadden met z'n allen niet mooier in een lijn kunnen gaan wonen (al had je ons allemaal toch het liefste in Knegsel gehad, hè Ans). Leonie en Lennard, respect voor wat een harde werkers jullie zijn met jullie opleidingen naast een druk sociaal leven en nu met de kleine Olivier erbij. Pieter en Sabine (en de andere twee lieve neefjes), bedankt voor alle tips en inspiratie voor het solliciteren, dat heeft me erg geholpen! En Rogier en Manon, blijven handelen in aandelen van farmaceuten hè! ;) Dank voor al jullie gezelligheid!

Pap en mam, dankjewel voor al het vertrouwen dat jullie altijd in me hebben gehad en hoe jullie me altijd hebben gestimuleerd alles uit jezelf te halen, maar vooral ook om plezier te hebben in wat je doet. Pap, hoewel jouw werk in de klinische chemie toch nog wel een stuk af staat van de biochemie, kon ik je altijd redelijk goed uitleggen wat nou eigenlijk toch het hele doel was van mijn onderzoek, wat vaak nog leidde tot grappige discussies. Ik wil jou en Lianne bedanken voor jullie betrokkenheid en voor de gezellige avondjes in Cromvoirt. Mam, zo'n zorgzame en betrokken moeder als jij bent, zo ken ik er maar weinig. Altijd oprecht geïnteresseerd, altijd klaarstaan voor anderen, en altijd zorgen voor gezelligheid. Dankjewel voor hoe je ons hebt grootgebracht en voor hoe je nog steeds altijd voor ons klaarstaat. Natuurlijk ook bedankt voor de leuke uitstapjes en voor onze gezellige telefoontjes, die de lange dagen schrijven gelukkig wat minder saai maakten. Maaïke, zussie, wat is het toch grappig dat we in hetzelfde vakgebied terecht zijn gekomen. Zo kon je zus je nog eens helpen met die lastige en saaie (nee, leuke!) biochemie vakken en kijken onze mannen ons soms zeer vragend aan als we weer eens een discussie hebben over allerlei lab-kwesties. Maaïke en Jordy, dank jullie wel voor jullie betrokkenheid en ik wens jullie heel veel geluk samen. Oma, u kan ik zeker niet vergeten in dit dankwoord! Als u het nog zou kunnen, zou u dit proefschrift van A tot Z uitlezen en daarna nog alle details opzoeken op internet. Hopelijk heeft het lekenpraatje ook voldaan ;). We weten in ieder geval dat ik mijn nieuwsgierigheid en interesse in onderzoek niet van vreemden heb, en ik ben heel blij dat u dit nog mag meemaken.

Lieve Jeroen, jij bent van iedereen de grootste steun en toeverlaat geweest tijdens m'n PhD en je motiveert me altijd om het beste uit mezelf te halen. Jij zorgde ervoor dat ik na lange dagen schrijven ook de tijd nam voor ontspanning, je maakte me altijd weer aan het lachen en je bood een luisterend oor als er iets niet mee zat of de frustratie er gewoon even uit moest. Maar je vierde natuurlijk ook de momenten van enthousiasme en succes met me mee! Dankjewel voor je altijd oprechte interesse in m'n onderzoek en hoewel de slingers (lees:  $\alpha$ -helices in de kristalstructuur) je in het begin niet veel zeiden, weet je nu haarfijn aan anderen uit te leggen dat dat een eiwit is waar een molecuul aan kan binden in de allosterische bindingsplaats, je wordt nog eens expert! Moleculen maken was eerst ook maar iets abstracts, maar zodra ik dan die kleine bouwstenen gingen vergelijken met de grote bouwstenen in de bouw werd het opeens interessant voor meneer de constructeur ;). Dankjewel voor al je liefde, interesse en steun, en voor je geduld als ik er even doorheen zat. Jij bent de liefste en ik kijk heel erg uit naar september, wanneer ik je dan eindelijk echt m'n man mag gaan noemen. Eén geluk van het uitstellen van de bruiloft: je mag nu wel met een doctor gaan trouwen ;).

Het lab is opgeruimd, mijn bureau is leeg, en dit zijn de allerlaatste regels van m'n proefschrift. Met een voldaan gevoel kijk ik terug op vier mooie jaren, maar het is nu tijd voor de volgende stap en een mooie, nieuwe uitdaging. Nogmaals bedankt iedereen voor alles en tot snel!

*Femke*

