

# Discovery of novel inhibitors to investigate diacylglycerol lipases and $\alpha/\beta$ hydrolase domain 16A

Janssen, F.J.

### Citation

Janssen, F. J. (2016, December 1). *Discovery of novel inhibitors to investigate diacylglycerol lipases and α/β hydrolase domain 16A*. Retrieved from https://hdl.handle.net/1887/44705

Version:	Not Applicable (or Unknown)				
License:	<u>Licence agreement concerning inclusion of doctoral thesis in the</u> <u>Institutional Repository of the University of Leiden</u>				
Downloaded from:	https://hdl.handle.net/1887/44705				

Note: To cite this publication please use the final published version (if applicable).

Cover Page



# Universiteit Leiden



The handle <u>http://hdl.handle.net/1887/44705</u> holds various files of this Leiden University dissertation

Author: Janssen, Freek J. Title: Discovery of novel inhibitors to investigate diacylglycerol lipases and  $\alpha/\beta$  hydrolase domain 16A Issue Date: 2016-12-01

## Potent *sn*-1 diacylglycerol lipase $\alpha$ inhibitor discovered by *in silico* screening<sup>\*</sup>

#### Introduction

Sn-1 specific diacylglycerol lipase  $\alpha$  (DAGL $\alpha$ ) is an intracellular, multidomain protein responsible for the formation of the endocannabinoid 2-arachidonovlglycerol (2-AG) in the central nervous system.<sup>1,2</sup> 2-AG is an endogenous signaling lipid that interacts with the cannabinoid CB1 and CB2 receptors.<sup>3</sup> However, little is known about the regulation of its biosynthetic pathway and it is largely unclear to what extent 2-AG is responsible for distinct CB1 receptor mediated biological processes. Selective inhibitors of DAGLa may contribute to a more fundamental understanding of the physiological role of 2-AG and may serve as potential drug candidates for the treatment of obesity and neurodegenerative diseases.<sup>4,5</sup> Currently there are no selective inhibitors available to study the function of DAGL $\alpha$ .<sup>6-12</sup> The identification of selective DAGL $\alpha$  inhibitors is hampered by a lack of structural knowledge of the target. No crystal structures are available and no homology models have been reported to aid hit identification and to guide optimization of inhibitors. Here a knowledge-based in silico screening approach on DAGL $\alpha$  is presented, followed by screening of a focused library of lipase inhibitors. This resulted in the identification of  $\alpha$ -keto heterocycle LEI104. A potential binding mode of LEI104 is postulated, based on covalent docking and molecular dynamics optimization. The binding mode is supported by preliminary structure-activity relationships (SAR).

<sup>&</sup>lt;sup>\*</sup> Published as part of: Baggelaar, M.P., Janssen F.J., van Esbroeck, A.C.M.; den Dulk, H., Allarà, M., Hoogendoorn, S., McGuire, R., Florea, B.I., Meeuwenoord, N., van den Elst, H., van der Marel, G A., Brouwer, J., di Marzo, V., Overkleeft, H.S., van der Stelt, M. Development of an activity-based probe and *in silico* design reveal highly selective inhibitors for diacylglycerol lipase  $\alpha$  in brain. *Angew. Chemie Int. Ed.* **2013**, *52*, 12081– 12085.

#### **Results and discussion**

To identify novel DAGL $\alpha$  inhibitors, a pharmacophore model based on tetrahydrolipstatin (THL), a known non-selective DAGL $\alpha$  inhibitor, was constructed using Discovery Studio Software Suite from Accelrys. Since THL can assume many different conformations, the protein crystallographic database was examined for crystal structures with a bioactive conformation for THL. A co-crystal structure of THL with fatty acid synthase (PDB code: 2PX6) was identified (Figure 1A)<sup>13</sup> that contains the same Ser-His-Asp catalytic triad and typical  $\alpha/\beta$  hydrolase fold motif as DAGL $\alpha$ . In this co-crystal structure, the nucleophilic Ser of the enzyme is covalently attached to the carbonyl moiety. Reconstitution of the ester formed the  $\beta$ -lactone to recover the active warhead of THL. After optimization of the geometry of the lactone, the resulting conformation was used to generate two pharmacophore models (Figure 1B,C). The essential features of both models are 1) a hydrogen bond acceptor mimicking the carbonyl from the  $\beta$ -lactone; 2) hydrophobic hot spots corresponding to the lipophilic tails of THL; 3) a hydrogen-bond acceptor positioned at the sn-2 ester functionality and 4) exclusion volumes representing the space occupied by the nucleophilic Ser and the backbone oxyanion hole residues in the active site of DAGL $\alpha$ . Model 2 contained an additional hydrogen-bond donor feature derived from the leucinyl formamide moiety of THL (Figure 2C). Using these two models, a set of commercially available lipase inhibitors were screened, which were mainly selected for their reactivity towards enzymes involved in endocannabinoid signaling (Table 1).



**Figure 1.** Development of THL based DAGL $\alpha$  pharmacophore models 1 and 2. **A)** Bioactive conformation of THL as retrieved from its known fatty acid synthase (FAS) co-crystal structure (PDB code 2PX6).<sup>13</sup> **B)** Pharmacophore model 1, consisting of 3 hydrogen bond acceptors (HBA), 1 hydrogen bond donors (HBD), 3 hydrophobic spheres and exclusion volumes as depicted. **C)** Pharmacophore model 2, consisting of 3 HBA, 3 hydrophobic spheres and exclusion volumes as depicted. Arrows represent the interaction vector.

Table 1. Structures and known targets of a set of commercially available lipase inhibitors, which were mainly
selected for their reactivity towards enzymes involved in endocannabinoid signaling. Inhibitors were bought at
Cayman Chemicals, Sigma Aldrich or Thermo Fisher.

Entry/Code	Structure	Target	Entry/ Code	Structure	Target
<b>1</b> (LEI103) CAY10499		HSL	<b>9</b> NO-1886		LPL
<b>2</b> CAY10590	Concernence of the second seco	PLA <sub>2</sub>	<b>10</b> Chlorpro- mazine		5-HT receptor
<b>3</b> CAY10594		PLD <sub>2</sub>	<b>11</b> URB597	°→NH₂ C→C→C→T <sup>H</sup> →C→	FAAH
<b>4</b> CAY10566		SCD	<b>12</b> JZL 184		MAGL
<b>5</b> (LEI104) OL-100	Constant of the second	FAAH	<b>13</b> URB602	N <sup>N</sup> <sup>N</sup>	MAGL
<b>6</b> JZL 195	$(\mathbf{x}_{\mathbf{x}},\mathbf{y},\mathbf{y},\mathbf{y},\mathbf{y},\mathbf{y},\mathbf{y},\mathbf{y},$	FAAH / MAGL	<b>14</b> PF-3845	<sup>N</sup> → <sup>L</sup> <sup>1</sup> N→ <sup>L</sup> <sub>P</sub>	FAAH
<b>7</b> WWL70	NJ () () () () () () () () () () () () ()	ABHD6	<b>15</b> TOFA	Contraction of the officer of	FAS
<b>8</b> FIPI		PLD	<b>16</b> RHC80267	Owali Contraction	Several lipases

Analysis of the docking results revealed that two compounds ranked in the top five of both models, LEI103 (1) and LEI104 (5; Figure 2, Table 2) based on fit value. The fit value is a quantitative measure of how well the pharmacophore overlaps with the compound chemical features. Expectantly also 15, a FAS inhibitor showed binding poses in both models. No binding mode was identified for compounds 3, 4, 7–11, 13, 14, and 16 (Table 2) in either one or both models. This result demonstrates that the pharmacophore models were capable of discriminating between related structures of lipase inhibitors.



Figure 2. Docking results in pharmacophore models 1 and 2: Highest ranked binding pose of A) LEI103 (5 features, model 1), B) LEI104 (4 features, model 1), C) LEI103 (5 features, model 2) and D LEI104 (4 features, model 2).

 Table 2. Ranking order of hits in both pharmacophore models, as determined by pharmacophore screening. Fit values for model 2 are lower because it is more stringent due to more features.

Rank	Model 1	Fit value	Model 2	Fit value
1	2	3.51	5 (LEI104)	2.25
2	<b>1</b> (LEI103)	3.32	12	2.17
3	8	2.84	11	1.95
4	5 (LEI104)	2.38	<b>1</b> (LEI103)	1.15
5	14	2.11	15	0.98
6	12	2.09	6	0.60
7	3	2.07	2	0.42
8	6	1.95	3	-
9	15	1.88	8	-
10	11	-	14	-
11	4	-	4	-
12	7	-	7	-
13	9	-	9	-
14	10	-	10	-
15	13	-	13	-
16	16	-	16	-

Compound **1**, LEI103 is an oxadiazolone known to inhibit hormone-sensitive lipase,<sup>14</sup> and **5**, LEI104 (OL-100) is an  $\alpha$ -ketoheterocycle that has been reported to inhibit fatty acid amide hydrolase (FAAH).<sup>15</sup> LEI104 has been shown to be active in *in vivo* models of antinociception through inhibition of FAAH activity.<sup>16–19</sup> Both hits represent new chemotypes that have not previously been shown to display DAGL $\alpha$  inhibitory activity. To validate the *in silico* hits, a colorimetric biochemical DAGL $\alpha$  activity assay was set up, which makes use of the hydrolysis of *para*-nitrophenyl (PNP) butyrate by membrane preparations from HEK293T cells

transiently transfected with human DAGL $\alpha$ , as previously reported.<sup>20</sup> Screening of the compound library against hDAGL $\alpha$  (Figure 3A) confirmed LEI103 (1) and LEI104 (5) as the only compounds to inhibit hDAGL $\alpha$  enzymatic activity by over 50% at 10  $\mu$ M. Determination of the concentration–response curves resulted in an IC<sub>50</sub> of 37 nM (n = 4) for LEI104, thus making it a hundred-fold more potent than LEI103 (IC<sub>50</sub> = 3.8  $\mu$ M, n = 4; Figure 3B). Of note, the reported DAGL $\alpha$  inhibitor RHC80267 (16) showed no inhibitory activity at 10  $\mu$ M in the biochemical assay. In addition, the activity of the hits in a radiometric assay using 1-[<sup>14</sup>C]oleoyl-2-arachidonoylglycerol as natural-like substrate was investigated.<sup>1</sup> This assay confirmed that LEI104 is a more potent inhibitor of DAGL $\alpha$  (IC<sub>50</sub> = 2.9 ± 0.1  $\mu$ M) than LEI103 (37% inhibition at 10  $\mu$ M).



**Figure 3.** Characterization of novel inhibitors for DAGLa. **A)** Screen of the targeted library using the colorimetric biochemical assay. Normalized residual activity was measured against hDAGLa in HEK293T cell membranes. **B)** Dose–response curves of LEI103 (black) and LEI104 (gray) against hDAGLa as determined with the colorimetric assay. LEI103:  $IC_{50} = 3.8 \mu$ M; LEI104:  $IC_{50} = 37$  nM. **C)** Structures of LEI103 and LEI104.

LEI104 was resynthesized according to Scheme 1 and retested in the colorimetric hDAGLa activity assay, displaying equipotent activity. The alcohol precursor of LEI104 (18) showed no activity, a result in line with the assumption that the ketone functions as a reversible electrophilic trap for the catalytic Ser472. Replacement of the oxazolopyridine heterocycle with a benzoxazole (17, Scheme 1) led to a 100-fold loss in activity, indicating that the pyridine nitrogen could form a potentially important interaction with the active site of the enzyme.<sup>15</sup> To understand the interaction of LEI104 with hDAGL $\alpha$  on a molecular level, a homology model was developed of DAGLa from a representative construct using the automated YASARA procedures.<sup>21-24</sup> The sequence used to represent DAGL $\alpha$  was constructed by editing the full sequence to remove the N terminal residues 1-287, the regulatory domain (residues 555- 623) and the C terminal region (residues 663-1042). These regions are not observed in crystal structures of lipases and the remaining residues form the routinely observed core  $\alpha/\beta$  hydrolase fold. Deleting the regulatory domain necessitated selection of the most appropriate residues and thus optimizes alignment. This was done manually with reference to the three dimensional structure of the template. DAGL $\alpha$  contains a large insert compared to other lipases, the cysteine rich loop, not present in any available lipase structure. As such, residues Tyr308- Phe358 were modeled as the cysteine rich loop insert (Figure 4).

The crystal structure of the S146A mutant of Thermomyces (Humicola) Lanuginosa Lipase in complex with oleic acid (PDB code 1GT6) is the highest scoring template of an enzyme:ligand complex.<sup>25</sup> The alignment had 230 of 306 target residues (75.2%) aligned to template residues and 18.7% sequence identity. Given the very low sequence homology, its unmatched cysteine rich loop and the artificial construct used for the post regulatory domain, the homology model quality is good. The poor 3D-packing however, reflects the remaining uncertainties in the model (Z-score of -2.440), however the calculated overall model quality (Z-score of -1.850) together with visual analysis of the catalytic core and the residues lining the binding cavity rate the homology model acceptable. Of note, the model does not include a potential regulatory cap present in DAGL $\alpha$ ,<sup>2</sup> as this sequence was not similar to other published structures.<sup>26</sup>

The model represents the typical  $\alpha/\beta$  hydrolase fold and has the catalytic triad (Ser472, His650, and Asp524) appropriately aligned in the binding cavity (Figure 5). The tetrahedral transition state of LEI104, which is formed through the nucleophilic attack of Ser472 on the  $\alpha$ -carbonyl, was minimized and subjected to molecular dynamics refinement. According to the model, the oxyanion intermediate is stabilized by the backbone N–H of the residue adjacent to the catalytic serine, Leu473 (71% of the snapshots taken in the molecular dynamics (MD) simulation showed this interaction). In addition, both the side chain O–H and the backbone N–H of Thr400 are observed to make hydrogen bonds with the oxyanion (93% in MD simulation). The oxazole nitrogen of LEI104 formed hydrogen-bond interactions with His650 (51% in MD simulation) and the pyridine nitrogen showed hydrogen-bond interactions with His471 (76% in MD simulation), both of which could further stabilize the transition state, while the hydrophobic pocket lined with aliphatic amino acids accommodated the flexible acyl chain of LEI104 (Figure 6). This proposed binding mode is consistent with the observed structure–activity relationships in this study.



**Scheme 1.** Synthesis of LEI104 (**5**) and α-keto benzoxazole **17**. *i*) DMSO, oxalyl chloride, DCM, Et<sub>3</sub>N, -78 °C, 99% *ii*) KCN, THF:H<sub>2</sub>O (1:1), 64% *iii*) Acetyl chloride, CHCl<sub>3</sub>, EtOH; 2-amino-3-hydroxy-pyridine/2-aminophenol, EtOH, reflux, 3.5% (**18**), 57% (**21**). *iv*) Dess-Martin periodinane, CH<sub>2</sub>Cl<sub>2</sub>, 44% (**5**), 98% (**17**).

#### >DAGLalpha ----RYKEVCYYMLFALAAYGWPM-----YLMRKPACGLCQLARSCSCCLCPARPRFA PGVTIEEDNCCGCNAIAIRRHFLDENMTAVDIVYTSC----HDAVYETPFYVAVDHDKKKVVISIRGTL SPKDALTDLTGDAERLPVEGHHGTWLGHKGMVLSAEYIKKKLEQEMVLSQAFG-RDLGRGTKHYGLIVVG HSLGAGTAAILSFLLRPQYPTLKCFAYSPPGGLLSEDAMEYSK---EFVTAVVLGKDLVPRIGLSQLEGF RRQLLDVLQRSTKPKWREPTYFAIWGDNKAFNEVTISPAMLHEHLPYVVMEGLNKV >1GT6-A: EVSQDLFNQFNLFAQYSAAAYCGKNNDAPAGTNITCTGNAC-----DVTGFLALDNTNKLIVLSFRGSR SIENWIGNLNFDLKEINDIC--SGCRGHDGFTSSWRSVADTLRQKVEDAVREHP------DYRVVFTG HALGGALATVAGADLRGNGYDIDVFSYGAPRVCNRAFAEFLTVQTGGTLYRITHTNDIVPRLPPREFGYS HSSPEYWIKSGTLVPVTRNDIVKIEGIDATGGNNOPNIPDIPAHLWYFGLIGTCL-

**Figure 4.** Manually adjusted alignment of DAGLα construct and 1GT6 Cysteine-Rich Loop highlighted in yellow, post-regulatory domain region in green, catalytic residues in bold.



**Figure 5. A)** Homology model of the DAGLα construct. The sequence used to represent DAGLα was constructed by editing the full sequence to remove the N terminal residues 1-287, the regulatory domain (residues 555-623) and the C terminal region (residues 663-1042). Cysteine-rich loop highlighted in yellow, post-regulatory domain region highlighted in green. **B)** The homology model overlaid with the original template crystal structure PDB code 1GT6 (magenta).<sup>25</sup> Graphical representation produced using PyMol (The PyMOL Molecular Graphics System, Version 1.4.1 Schrödinger, LLC).



**Figure 6.** Binding pose of LEI104 in a homology model of hDAGLα. Docking was performed using Yasara docking procedures which implements AutoDock Vina.<sup>27</sup> Molecular dynamics were performed using the AMBER03 forcefield.<sup>28</sup> The MD trajectories were studied using Yasara analysis tools and showed that in 71% of the snapshots a hydrogen bond was observed between the oxyanion intermediate and the backbone NH of Leu473. The Thr400 residue was observed to continually interact with the oxyanion intermediate throughout the simulation with 93% of the snapshots via both the backbone NH and sidechain OH. A total of 76% of the snapshots exhibited hydrogen bonds between His471 and the pyridine nitrogen of LEI104, while H-bond formation between the catalytic His650 and the oxazole nitrogen of LEI104 was also observed in 51% of the snapshots.

#### Conclusions

The development of a THL based pharmacophore model and screening of a focused library rapidly identified the  $\alpha$ -ketoheterocycle LEI104 as a DAGL $\alpha$  inhibitor. The development of a homology model of DAGL $\alpha$  was used to investigate the binding mode of LEI104, indicating a crucial role for the 4N-oxazolopyridine nitrogen. As LEI104 was originally reported as a potent FAAH inhibitor,<sup>15</sup> optimization of selectivity over this target is required. The homology model provides a clear view of the opportunities to improve the potency and selectivity over FAAH. As such, it is anticipated that the  $\alpha$ -ketoheterocycle class will provide an excellent lead series to investigate 2-AG mediated cannabinoid CB1R signaling and for the development of *in vivo* active and selective DAGL $\alpha$  inhibitors, because these compounds 1) have a clearly defined scaffold with good physicochemical properties; 2) are not based on the natural substrate and do not contain known toxicophores (for example fluorophosponate);<sup>6</sup> 3) are plasma membrane permeable; 4) are highly selective; 5) do not form irreversible covalent bonds,<sup>16–19</sup> which could lead to problems with immunogenicity; and 6) have been shown to be bioavailable and active in animal models.<sup>16–19</sup> The structural insights provided by the DAGL $\alpha$  homology model, may serve as a basis for the development of new therapeutics that can be used to study and treat diseases such as obesity and neurodegeneration.

## Experimental

#### Experimental procedures in silico

#### Development of the DAGLa pharmacophore models

The pharmacophore models are based on the bioactive conformation of THL co-crystalized in fatty acid synthase (FAS, pdb code 2PX6).<sup>13</sup> The  $\beta$ -lactone of the bioactive conformation was closed and the geometry of the reconstituted β-lactone was optimized using Discovery Software Suite 3.5 from Accelrys. The bioactive conformation was separated from the crystal structure, and pharmacophore features were assigned using the manual pharmacophore construction protocol. The essential features that were assigned to pharmacophore model 1 were: i) an H-bond acceptor mimicking the carbonyl from the  $\beta$ -lactone warhead, ii) hydrophobic hot spots corresponding to the lipophilic tails of THL, iii) an H-bond acceptor positioned at the sn-2 ester functionality and iv) exclusion volumes representing the space occupied by the nucleophilic serine and the backbone oxyanion hole residues in the active site of DAGLa. An additional donor feature derived from the leucinyl formamide moiety of THL was assigned to Model 2. The automated Screen Library protocol (Catalyst/FAST, Accelrys) available from Discovery Studio 3.5 Software was used for docking and conformation generation. This protocol uses the CHARMm force field for energy calculations and a poling mechanism.<sup>29,30</sup> The pharmacophore model was considered rigid, and the inhibitors flexible (255 conformations for each compound). Slight flexibility of the bioactive conformation is accounted for by the sphere radii. The remaining parameters were left on default settings. Each template was visually judged on plausible positioning towards the active site serine. Conformations that did not fit were discarded, while well positioned conformations were documented. This resulted in 36 verified conformations for Model 1 and 21 for Model 2. All conformations were ranked according to feature hits and fit value (Catalyst/FAST, Accelrys fit value calculation procedure).

#### Development of the DAGLa homology model

The sequence used to represent DAGL $\alpha$  was constructed by editing the full sequence to remove the N terminal residues 1-287, the regulatory domain (residues 555-623) and the C terminal region (residues 663-1042). The alignment) has 230 of 306 target residues (75.2%) aligned to template residues, sequence identity is 18.7% and similarity 37.4% ('similar' means blocks substitution matrix score, BLOSUM62 (> 0). Homology modeling was performed using the automated YASARA procedures.<sup>21-24</sup> These use PSI-Blast<sup>31</sup> to build a position specific scoring matrix (PSSM) from related UniProt sequences which is used to search the PDB for potential templates. Templates are ranked by alignment score and structural quality.<sup>32</sup> Alignments are generated using sequencebased profiles of target and template from UniProt sequences, optionally augmented with structure-based profiles from related template structures. The alignment also considers structural information contained in the template (avoiding gaps in secondary structure elements, keeping polar residues exposed etc.), as well as the predicted target secondary structure.<sup>33</sup> This structure based alignment correction is partly based on SSALN scoring matrices.<sup>34</sup> Loops arising from insertions and deletions are modeled using an indexed version of the PDB to determine optimal loop anchoring points and potential loop conformations. Prefered side chain rotamers are detected, the hydrogen bonding network is optimized and the entire system is subjected to an unrestrained high-resolution refinement with explicit solvent molecules, using a knowledge based forcefield.<sup>24</sup> The homology model of DAGLa was built using the crystal structure of the S146A Mutant Of Thermomyces (Humicola) Lanuginosa Lipase in Complex With Oleic Acid, PDB code 1GT6, the highest scoring template of an enzyme:ligand complex.<sup>25</sup> Blast E-value 3-e38, Align score 47.0, 76% coverage, Total score 19.93. The resulting homology model is shown in Figure 4, overlaid on the template 1GT6.<sup>25</sup> Yasara Model Quality checks compare model features with those observed in high quality X-ray structures. These checks show good quality dihedrals (Z-score 0.027) and satisfactory 1D packing (Z-score -1.844). The poor 3D-packing z-score of -2.440 reflects the remaining uncertainties in the model, however the calculated overall model quality Z-score of -1.850 (rated as

satisfactory) together with visual analysis of the catalytic core and the residues lining the binding cavity suggest the homology model is acceptable.

#### Docking of LEI104 in the DAGLa homology model

Docking was performed using Yasara docking procedures which implement AutoDock Vina.<sup>27</sup> The protein model was kept rigid while the ligand was treated flexibly. The top 12 scoring poses were retained and examined visually. Figure 5 shows the top scoring pose for LEI104. Energy minimization followed by short Molecular Dynamics simulations were run in order to refine the structure of the docked pose shown in Figure 5. This pose was modified in Yasara to covalently bind the LEI104 structure to the catalytic serine (Ser185 in the model). The geometry around the tetrahydral intermediate was updated and the structure was minimized in Yasara then subjected to a brief Molecular Dynamics procedure. The MD comprised solvating the structure in water in a simulation cell extending 10 Å around all atoms, and subsequent use of default Yasara parameters, including automatic placement of counter ions (Na<sup>+</sup> and Cl<sup>-</sup>) to neutralize the simulation cell at pH 7.4, performing an initial energy minimization before reassigning velocities according to a Boltzmann distribution and then running molecular dynamics using the AMBER03 forcefield,<sup>28</sup> at 298 K, with periodic boundary conditions, Van der Waals interaction cut-off 7.86 Å, Long-range (particle-mesh-Ewald)<sup>35</sup> electrostatic interactions, timestep 2.5 fs. Simulation snapshots were saved every 25 ps over the course of the 8 ns simulation. Analysis was performed after discarding the initial 80 snapshots i.e. allowing a short 2 ns equilibration. Structural alignment of initial homology model and the minimized structure resulting from the MD-run showed an RMSD of 1.77 Å. The MD trajectories were studied using Yasara analysis tools and showed that in 71% of the snapshots a hydrogen bond was observed between the oxyanion intermediate and the backbone NH of Leu473. The Thr400 residue was observed to continually interact with the oxyanion intermediate throughout the simulation, in 93% of the snapshots via both the backbone NH and sidechain OH. 76% of the snapshots exhibited hydrogen bonds between His471 and the pyridine nitrogen of LEI104 while H-bond formation between His650 and the oxazole nitrogen of LEI104 was also observed in 51% of the snapshots. Histidine H-bonds were typically observed with suboptimal geometries but were only counted when calculated to be above the Yasara default energy cut-off of 1.5 kcal mol<sup>-1</sup>.

#### Experimental procedures biochemistry

#### DAGL<sub>a</sub> plasmids

For the preparation of the hDAGLα construct, full length human cDNA was purchased from Biosource and cloned into mammalian expression vector pcDNA3.1, containing genes for ampicillin and neomycin resistance. All plasmids were grown in XL-10 Z-competent cells and prepped (Maxi Prep, Qiagen). The sequences were confirmed by sequence analysis at the Leiden Genome Technology Centre.

#### Cell culture and membrane preparation.

HEK293T cells were grown in DMEM with stable glutamine and phenolred (PAA) with 10% New Born Calf serum, penicillin and streptomycin. Cells were passaged every 2-3 days by resuspending in medium and seeding them to appropriate confluence. Membranes were prepared from transiently transfected HEK293T cells. One day prior to transfection  $10^7$  cells were seeded in a 15 cm petri dish. Cells were transfected by the addition of a 3:1 mixture of polyethyleneimine (60 µg) and plasmid DNA (20 µg) in 2 mL serum free medium. The medium was refreshed after 24 h, and after 72 h the cells were harvested by suspending them in 20 mL medium. The suspension was centrifuged for 10 min at 1000 rpm, and the supernatant was removed. The cell pellet was stored at -80 °C until use. Cell pellets were thawed on ice and suspended in lysis buffer A (20 mM HEPES, 2 mM DTT, 0.25 M sucrose, 1 mM MgCl<sub>2</sub>, 1x cocktail (Roche cOmplete EDTA free), 25 U/µL Benzonase). The suspension was homogenized by polytrone (3 × 7 sec) and incubated for 30 min on ice. The suspension was subjected to ultracentrifugation (93.000 × g, 30 min, 4°C, Beckman Coulter, Type Ti70 rotor) to yield the cytosolic fraction in the supernatant and the membrane fraction as a pellet. The pellet was resuspended in lysis buffer B (20 mM HEPES, 2 mM DTT, 1x Cocktail (Roche cOmplete EDTA free)). The protein concentration was determined with Quick Start Bradford assay (Biorad). The protein fractions were diluted to a total protein concentration of 1 mg/mL and stored in small aliquots at -80°C until use.

#### Biochemical DAGLα activity assay.

The biochemical hDAGL $\alpha$  activity assay is based on the hydrolysis of *para*-nitrophenylbutyrate (PNP-butyrate) by membrane preparations from HEK293T cells transiently transfected with hDAGLa. 200 µL reactions were performed in flat bottom Greiner 96-wells plates in a 50 mM pH 7.2 HEPES buffer. Membrane protein fractions from HEK293T cells transiently transfected with hDAGL $\alpha$  (0.05 µg/µL final concentration) were used as hDAGL $\alpha$ source. Inhibitors were introduced in 5 µL DMSO. The mixtures were incubated for 20-30 minutes before 5.0 µL 12 mM PNP-butyrate (final concentration 0.3 mM) in DMSO was added (final DMSO concentration 5.0%). Kinetics were followed immediately after addition of PNP-butyrate on a plate reader (TECAN GENios microplate reader), by measuring the OD<sub>420</sub> every 60 seconds, for 20 minutes at 37°C. The slope of the linear region from 5-15 minutes was determined, and all experiments were performed at N = 2, n = 2 for experimental measurements and N = 2, n = 4 for controls. Data analysis: Z'-factor of each plate was determined for the validation of each experiment, using the following formula  $Z' = 1-3(\sigma pc+\sigma nc)/(\mu pc - \mu nc)$ . The slope from 5-15 minutes of the positive control (pc: DAGL $\alpha$  DMSO), and the negative control (nc: mock DMSO) was used. Plates were accepted for further analysis when Z' > 0.6. Kinetic measurements were corrected for the average absorption of the negative control (mock DMSO). The slope of the linear region from 5-15 minutes was determined. The average, standard deviation (SD) and standard error of mean (SEM) were calculated and normalized to the corrected positive control. Data were exported to Graphpad Prism 5.0 for the calculation of the IC<sub>50</sub> using a nonlinear dose-response analysis.

#### **Experimental procedures chemistry**

#### **General remarks**

All reactions were performed using oven or flame-dried glassware and dry solvents. Reagents were purchased from Sigma Aldrich, Acros and Merck and used without further purification unless noted otherwise. All moisture sensitive reactions were performed under an argon atmosphere. Traces of water were removed from starting compounds by co-evaporation with toluene. <sup>1</sup>H- and <sup>13</sup>C-NMR spectra were recorded on a Bruker AV 400 MHz spectrometer at 400.2 (<sup>1</sup>H) and 100.6 (<sup>13</sup>C) MHz or a Bruker DMX-600 spectrometer 600 (<sup>1</sup>H) and 150  $(^{13}C)$  MHz using CDCl<sub>3</sub> or CD<sub>3</sub>OD as solvent, unless stated otherwise. Chemical shift values are reported in ppm with tetramethylsilane or solvent resonance as the internal standard (CDCl<sub>3</sub>:  $\delta$  7.26 for <sup>1</sup>H,  $\delta$  77.0 for <sup>13</sup>C, CD<sub>3</sub>OD:  $\delta$  3.31 for <sup>1</sup>H). Data are reported as follows: chemical shifts ( $\delta$ ), multiplicity (s = singlet, d = doublet, dd = double doublet, td = triple doublet, t = triplet, q = quartet, quintet = quint, br = broad, m = multiplet), coupling constants J (Hz), and integration. HPLC purification was performed on a preparative LC-MS system (Agilent 1200serie) with an Agilent 6130 Quadruple MS detector. High-resolution mass spectra (HRMS) were recorded on a Thermo Scientific LTQ Orbitrap XL. Flash chromatography was performed using SiliCycle silica gel type SiliaFlash P60 (230 – 400 mesh). TLC analysis was performed on Merck silica gel 60/Kieselguhr F254, 0.25 mm. Compounds were visualized using either Seebach's reagent (a mixture of phosphomolybdic acid (25 g), cerium (IV) sulfate (7.5 g), H<sub>2</sub>O (500 mL) and H<sub>2</sub>SO<sub>4</sub> (25 mL)) or a KMnO<sub>4</sub> stain (K<sub>2</sub>CO<sub>3</sub> (40 g), KMnO<sub>4</sub> (6 g), H<sub>2</sub>O (600 mL) and 10% NaOH (5 mL)).

#### 6-Phenylhexanal

A solution of DMSO (0.75 g, 10 mmol) in  $CH_2CI_2$  was added dropwise to a solution of oxalyl chloride (0.64 g, 5 mmol) in  $CH_2CI_2$  (5 mL) at -78°C. The mixture was stirred for 1 h before commercially available 6-phenylhexan-1-ol (297 mg, 1.67 mmol) in  $CH_2CI_2$  (1 mL) was added drop-wise. After addition, the mixture was stirred overnight, and  $Et_3N$  (1.7 mL, 16 mmol) was added drop-wise. The mixture was allowed to warm to rt and washed with 1 M HCl. The aqueous layer was extracted with  $CH_2CI_2$ , and the combined organic layers were washed with brine and subsequently dried ( $Na_2SO_4$ ). Volatiles were removed under reduced pressure and the residue was purified by flash chromatography yielding 6-phenylhexanal (289 mg, 1.65 mmol, 99 %) as a colorless oil. <sup>1</sup>H NMR (400 MHz, CDCI<sub>3</sub>)  $\delta$  9.62 (t, *J* = 1.6 Hz, 1H), 7.05-7.17 (m, 5H), 2.51 (t, *J* = 7.6 Hz, 2H), 2.29 (dt, *J* = 1.6, 7.6 Hz, 2H), 1.50-1.56 (m, 4H), 1.24-1.28 (m, 2H); <sup>13</sup>C NMR (101 MHz, CDCI<sub>3</sub>)  $\delta$  202.66, 142.38, 128.37, 128.23, 125.60, 43.79, 35.69, 31.21, 28.73, 21.89. Spectroscopic data are in agreement with those reported.<sup>36</sup>

#### 2-Hydroxy-7-phenylheptanenitrile

To a solution of 6-phenylhexanal (289 mg, 1.6 mmol) in (1:1) THF:H<sub>2</sub>O (100 mL) was added potassium cyanide (1.2 g, 18 mmol). The reaction mixture was stirred at rt for 72 h. H<sub>2</sub>O (10 mL) and Et<sub>2</sub>O (10 mL) were added, and the layers were separated. The aqueous layer was extracted with Et<sub>2</sub>O and the combined organic layers were washed with saturated NaHCO<sub>3</sub> (aq.) and brine before drying (Na<sub>2</sub>SO<sub>4</sub>). Volatiles were removed under reduced pressure and the residue was purified by flash chromatography yielding 2-hydroxy-7-phenylheptanenitrile (213 mg, 1.0 mmol, 64 %) as a colorless oil. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.22-7.36 (m, 5H), 4.45 (dt, *J* = 6.0, 6.0 Hz, 1H), 3.71 (d, *J* = 4.8 Hz, 1H), 2.67 (t, *J* = 7.6 Hz, 2H), 1.85 (q, *J* = 8.4 Hz, 2H), 1.70 (quint, *J* = 7.6 Hz, 2H), 1.56 (quint, *J* = 7.2 Hz, 2H), 1.43 (quint, *J* = 7.2 Hz, 2H); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  142.38, 128.40, 128.32, 125.74, 120.17, 61.07, 35.68, 34.90, 31.12, 28.47, 24.40.

#### 1-(Oxazolo[4,5-b]pyridin-2-yl)-6-phenylhexan-1-ol

A mixture of CHCl<sub>3</sub> (2 mL) and ethanol (2 mL) was cooled to 0°C before acetyl chloride (2 mL) was added dropwise. The mixture was stirred for 30 minutes, and 2-hydroxy-7-phenylheptanenitrile (60 mg, 0.28 mmol) in CHCl<sub>3</sub> (2 mL) was added. The mixture was stirred for another 1.5 h and the solvent was removed under reduced pressure while keeping the temperature below 25 °C. The residue was taken up in ethanol (6.5 mL) ethoxyethanol (1 mL) and 2-amino-3-hydroxypyridine (26 mg, 0.24 mmol) was added and the reaction mixture was heated to reflux for 6 h. The solvent was removed under reduced pressure, and the residue was dissolved in EtOAc (30 mL) and 1 M NaOH (10 mL) was added. The layers were separated, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine and dried (Na<sub>2</sub>SO<sub>4</sub>). Volatiles were removed under reduced pressure and the residue was further purified by flash chromatography yielding 1-(oxazolo[4,5-*b*]pyridin-2-yl)-6-phenylhexan-1-ol (3.0 mg, 0.01 mmol, 3.5 %) as a yellow solid. <sup>1</sup>H NMR (400 MHz, CDCl3)  $\delta$  8.57 (d, *J* = 3.7 Hz, 2H), 7.84 (d, *J* = 8.0 Hz, 2H), 77.34 – 7.22 (m, 3H), 7.18 – 7.14 (m, 2H), 5.00 (t, *J* = 6.2 Hz, 1H), 3.03 (bs, 1H), 2.60 (t, *J* = 7.7 Hz, 2H), 2.15 – 1.90 (m, 2H), 1.70 – 1.36 (m, 6H); <sup>13</sup>C APT NMR (101 MHz, CDCl3)  $\delta$  171.10, 154.90, 146.68, 142.55, 128.41, 128.28, 125.67, 124.81, 120.34, 118.67, 68.21, 35.81, 31.27, 29.73, 28.89, 24.72; Purity > 90 % as determined by LC-MS; mass [M+H]<sup>+</sup> = 297.07 m/z.

#### 1-(Benzo[d]oxazol-2-yl)-6-phenylhexan-1-ol

A mixture of CHCl<sub>3</sub> (2.5 mL) and ethanol (2.5 mL) was cooled to 0°C before acetyl chloride (2.5 mL) was added dropwise. The mixture was stirred for 30 minutes, and 2-hydroxy-7-phenylheptanenitrile (319 mg, 1.57 mmol) in CHCl<sub>3</sub> (2.5 mL) was added. The mixture was stirred for another 1.5 h and the solvent was removed under reduced pressure while keeping the temperature below  $25^{\circ}$ C. The residue was taken up in ethanol (6.5 mL), and 2-aminophenol (170 mg, 1.56 mmol) was added. The mixture was refluxed for 6 h. Volatiles were removed under reduced pressure and the residue was purified by flash chromatography yielding 1-(benzo[d]oxazol-2-yl)-

6-phenylhexan-1-ol (263 mg, 0.89 mmol, 57 %) as a brown solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.71 (m, 1H), 7.28 (m, 1H), 7.34 (m, 2H), 7.29-7.23 (m, 2H), 7.19-7.13 (m, 3H), 4.96 (dd, *J* = 5.2, 2.0 Hz, 1H), 3.15 (bs, 1H), 2.59 (t, *J* = 7.6 Hz, 2H), 2.11 – 1.91 (m, 2H), 1.64 (quint, *J* = 8.0 Hz, 2H), 1.52 (quint, *J* = 7.6 Hz, 2H), 1.44 – 1.36 (m, 2H); <sup>13</sup>C APT NMR (101 MHz, CDCl<sub>3</sub>) δ 167.95, 150.95, 142.70, 140.42, 128.52, 128.38, 125.77, 125.37, 124.73, 120.11, 110.95, 68.22, 35.94, 35.64, 31.38, 29.05, 24.90. HRMS (ESI+) m/z: calculated for C<sub>19</sub>H<sub>21</sub>NO<sub>2</sub> (m + H) 296.165, found 296.164.

#### 1-(Oxazolo[4,5-b]pyridin-2-yl)-6-phenylhexan-1-one

1-(Oxazolo[4,5-*b*]pyridin-2-yl)-6-phenylhexan-1-ol (3.0 mg, 0.01 mmol) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (2 mL), and Dess-Martin periodinane (6.8 mg, 0.016 mmol) was added. The mixture was stirred at rt overnight. Saturated NaHCO<sub>3</sub> (aq.) was added and the mixture was stirred for an additional 15 minutes. The layers were separated and the aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic layers were washed with brine, and dried (Na<sub>2</sub>SO<sub>4</sub>). Volatiles were removed under reduced pressure and the residue was further purified by preparative HPLC, yielding 1-(oxazolo[4,5-*b*]pyridin-2-yl)-6-phenylhexan-1-one (1.3 mg, 4.4 µmol, 44 %) as a yellow solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.76 (dd, *J* = 4.7, 1.1 Hz, 1H), 8.00 (dd, *J* = 8.3, 1.3 Hz, 1H), 7.50 (dd, *J* = 8.3, 4.8 Hz, 1H), 7.31 – 7.24 (m, 2H), 7.22 – 7.12 (m, 3H), 3.28 (t, *J* = 7.4 Hz, 2H), 2.64 (t, *J* = 7.7 Hz, 2H), 1.82 (quint, *J* = 7.6, 2H), 1.70 (quint, *J* = 8.0 Hz, 2H), 1.51 – 1.41 (m, 2H); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  190.24, 158.48, 154.14, 148.70, 143.58, 142.33, 128.35, 128.24, 125.66, 123.11, 120.20, 39.71, 35.66, 31.09, 28.65, 23.64. Purity > 95 % as determined by LC-MS; mass (M+H) = 295.07 m/z. Spectroscopic data are in agreement with those reported.<sup>15</sup>

#### 1-(Benzo[d]oxazol-2-yl)-6-phenylhexan-1-one

1-(Benzo[*d*]oxazol-2-yl)-6-phenylhexan-1-ol (44 mg, 0.15 mmol) was dissolved in CH<sub>2</sub>Cl<sub>2</sub>. Dess-Martin periodinane (96 mg, 0.23 mmol) was added, and the reaction mixture was stirred for 3.5 h. Volatiles were removed under reduced pressure, and the residue was purified by flash chromatography yielding 1-(benzo[*d*]oxazol-2-yl)-6-phenylhexan-1-one (43 mg, 0.15 mmol, 98 %) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.89 (d, *J* = 8.0 Hz, 1H), 7.65 (d, *J* = 8.2 Hz, 1H), 7.53 (t, *J* = 7.7 Hz, 1H), 7.46 (t, *J* = 7.7 Hz, 1H), 7.31 – 7.22 (m, 3H), 7.18 (d, *J* = 6.9 Hz, 3H), 3.22 (t, *J* = 7.4 Hz, 3H), 2.68 – 2.57 (m, 2H), 1.85 (p, *J* = 7.5 Hz, 2H), 1.76 – 1.62 (m, 2H), 1.48 (h, *J* = 7.2, 6.5 Hz, 2H); <sup>13</sup>C APT NMR (101 MHz, CDCl<sub>3</sub>) δ 190.37, 157.25, 150.83, 142.54, 140.62, 128.61, 128.51, 128.39, 125.86, 125.80, 122.3, 112.07, 39.58, 35.84, 31.30, 28.83, 23.81. HRMS (ESI+) m/z: calculated for C<sub>19</sub>H<sub>19</sub>NO<sub>2</sub> (m + H) 294.149, found 294.149.

#### References

- 1. Bisogno, T. *et al.* Cloning of the first *sn*1-DAG lipases points to the spatial and temporal regulation of endocannabinoid signaling in the brain. *J. Cell Biol.* **163**, 463–468 (2003).
- Reisenberg, M., Singh, P. K., Williams, G. & Doherty, P. The diacylglycerol lipases: structure, regulation and roles in and beyond endocannabinoid signaling. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 3264–3275 (2012).
- Katona, I. & Freund, T. F. Multiple functions of endocannabinoid signaling in the brain. Annu. Rev. Neurosci. 35, 529–58 (2012).
- 4. Di Marzo, V. Endocannabinoid signaling in the brain: biosynthetic mechanisms in the limelight. *Nat. Neurosci.* **14**, 9–15 (2011).
- 5. Nomura, D. K. *et al.* Endocannabinoid hydrolysis generates brain prostaglandins that promote neuroinflammation. *Science* **334**, 809–813 (2011).
- 6. Bisogno, T. *et al.* Development of the first potent and specific inhibitors of endocannabinoid biosynthesis. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1761**, 205–212 (2006).
- 7. Bisogno, T. *et al.* Synthesis and pharmacological activity of a potent inhibitor of the biosynthesis of the endocannabinoid 2-arachidonoylglycerol. *ChemMedChem* **4**, 946–950 (2009).
- Johnston, M. et al. Assay and inhibition of diacylglycerol lipase activity. *Bioorganic Med. Chem. Lett.* 22, 4585–4592 (2012).
- Ortar, G. *et al.* Tetrahydrolipstatin analogues as modulators of endocannabinoid 2-arachidonoylglycerol metabolism. *J. Med. Chem.* **51**, 6970–6979 (2008).
- 10. Hsu, K. L. *et al.* DAGLβ inhibition perturbs a lipid network involved in macrophage inflammatory responses. *Nat. Chem. Biol.* **8**, 999–1007 (2012).
- 11. Marrs, W. R. *et al.* The serine hydrolase ABHD6 controls the accumulation and efficacy of 2-AG at cannabinoid receptors. *Nat. Neurosci.* **13**, 951–7 (2010).
- 12. Bisogno, T. *et al.* A novel fluorophosphonate inhibitor of the biosynthesis of the endocannabinoid 2arachidonoylglycerol with potential anti-obesity effects. *Br. J. Pharmacol.* **169**, 784–793 (2013).
- 13. Pemble, C. W., Johnson, L. C., Kridel, S. J. & Lowther, W. T. Crystal structure of the thioesterase domain of human fatty acid synthase inhibited by Orlistat. *Nat. Struct. Mol. Biol.* **14**, 704–9 (2007).
- 14. Ben Ali, Y. *et al.* The molecular mechanism of human hormone-sensitive lipase inhibition by substituted 3-phenyl-5-alkoxy-1,3,4-oxadiazol-2-ones. *Biochimie* **94**, 137–145 (2012).
- Boger, D. L. *et al.* Exceptionally potent inhibitors of fatty acid amide hydrolase: the enzyme responsible for degradation of endogenous oleamide and anandamide. *Proc. Natl. Acad. Sci. U.S.A.* 97, 5044–5049 (2000).
- Lichtman, A. H. Reversible Inhibitors of Fatty Acid Amide Hydrolase That Promote Analgesia: Evidence for an Unprecedented Combination of Potency and Selectivity. J. Pharmacol. Exp. Ther. **311**, 441–448 (2004).
- DeMartino, J. K., Garfunkle, J., Hochstatter, D. G., Cravatt, B. F. & Boger, D. L. Exploration of a fundamental substituent effect of a-ketoheterocycle enzyme inhibitors: Potent and selective inhibitors of fatty acid amide hydrolase. *Bioorganic Med. Chem. Lett.* 18, 5842–5846 (2008).
- Leung, D., Hardouin, C., Boger, D. L. & Cravatt, B. F. Discovering potent and selective reversible inhibitors of enzymes in complex proteomes. *Nat. Biotechnol.* 21, 687–691 (2003).
- Boger, D. L., Miyauchi, H. & Hedrick, M. P. alpha-Keto heterocycle inhibitors of fatty acid amide hydrolase: carbonyl group modification and alpha-substitution. *Bioorg. Med. Chem. Lett.* **11**, 1517–20 (2001).
- Pedicord, D. L. *et al.* Molecular characterization and identification of surrogate substrates for diacylglycerol lipase a. *Biochem. Biophys. Res. Commun.* 411, 809–814 (2011).
- 21. Krieger, E., Koraimann, G. & Vriend, G. Increasing the precision of comparative models with YASARA NOVA A self-parameterizing force field. *Proteins Struct. Funct. Genet.* **47**, 393–402 (2002).
- 22. Krieger, E., Darden, T., Nabuurs, S. B., Finkelstein, A. & Vriend, G. Making optimal use of empirical energy functions: Force-field parameterization in crystal space. *Proteins Struct. Funct. Genet.* **57**, 678–683 (2004).
- 23. Krieger, E., Nielsen, J. E., Spronk, C. A. E. M. & Vriend, G. Fast empirical pKa prediction by Ewald summation. *J. Mol. Graph. Model.* **25**, 481–486 (2006).
- 24. Krieger, E. *et al.* Improving physical realism, stereochemistry and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. **77**, 114–122 (2009).

- Yapoudjian, S. *et al.* Binding of Thermomyces (Humicola) lanuginosa lipase to the mixed micelles of cisparinaric acid/NaTDC: Fluorescence resonance energy transfer and crystallographic study. *Eur. J. Biochem.* 269, 1613–1621 (2002).
- Baggelaar, M. P. *et al.* Development of an activity-based probe and in silico design reveal highly selective inhibitors for diacylglycerol lipase-a in brain. *Angew. Chem. Int. Ed.* 52, 12081–12085 (2013).
- 27. Trott, O. & Olson, A. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, effcient optimization and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
- Duan, Y. *et al.* A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* 24, 1999–2012 (2003).
- Smellie, A., Teig, S. L. & Towbin, P. Poling: promoting conformational variation. J. Comput. ... 16, 171– 187 (1995).
- 30. Brooks, B. R. *et al.* CHARMm: The Biomolecular Simulation Program *J. Comput. Chem.* **30**, 1545–1614 (2009).
- Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402 (1997).
- 32. Hooft, R. W. W., Vriend, G., Sander, C. & Abola, E. E. Errors in protein structures. *Nature* **381**, 272–272 (1996).
- 33. King, R. D. & Sternberg, M. J. Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. *Protein Sci.* **5**, 2298–2310 (1996).
- Qiu, J. & Elber, R. SSALN: An alignment algorithm using structure-dependent substitution matrices and gap penalties learned from structurally aligned protein pairs. *Proteins Struct. Funct. Genet.* 62, 881–891 (2006).
- 35. Essmann, U. et al. A smooth particle mesh Ewald method. J Chem Phys 103, 8577–8593 (1995).
- Murphy, J. A. *et al.* One-carbon extrusion from a tetraazafulvalene. Isolation of aldehydes and a study of their origin. *J. Am. Chem. Soc.* **131**, 6475–6479 (2009).