Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202401393

Chemistry Europe
European Chemical Societies Publishing

www.chemeurj.org

Isolation, Structure Elucidation, and Biological Activity of the Selective TACR2 Antagonist Tumonolide and its Aldehyde from a Marine Cyanobacterium

Sofia Kokkaliari,[a] Laura Grauso,[b] Alfonso Mangoni,[c] Gustavo Seabra,[a] Valerie J. Paul,[d] and Hendrik Luesch*[a]

The macrocyclic tumonolide (**1**) with enamide functionality and the linear tumonolide aldehyde (**2**) are new interconverting natural products from a marine cyanobacterium with a peptide-polyketide skeleton, representing a hybrid of apratoxins and palmyrolides or laingolides. The planar structures were established by NMR and mass spectrometry. The relative configuration of the stereogenically-rich apratoxin-like polyketide portion was determined using **J**-based configuration analysis. The absolute configuration of tumonolide (**1**) was determined by chiral analysis of the amino acid units and computational methods, followed by NMR chemical shift and ECD spectrum prediction, indicating all-**R** configuration for the polyketide portion, as in palmyrolide A and contrary to the all-**S** configuration in apratoxins. Functional screening against a panel of 168 GPCR targets revealed tumonolide (**1**) as a selective antagonist of TACR2 with an IC_{50} of 7.0 μM, closely correlating with binding affinity. Molecular docking studies established the binding mode and rationalized the selectivity for TACR2 over TACR1 and TACR3. RNA sequencing upon treatment of HCT116 colorectal cancer cells demonstrated activation of the pulmonary fibrosis idiopathic signaling pathway and the insulin secretion signaling pathway at 20 μM, indicating its potential to modulate these pathways.

Introduction

Marine cyanobacteria have proven to be a rich source of bioactive natural products that can be used as therapeutic leads.[1,2] Cyanobacterial natural products display a wide range of specific pharmacology, targeting ion channels, enzymes such as proteases, cytoskeletal proteins, GPCRs (G-protein coupled receptors) and many others.[3] For example, palmyrolide A is a *N*-methyl enamide macrolide that acts as an ion channel inhibitor (Figure 1).[4–7] Laingolides are closely related, differing by the macrocycle size, substitution patterns, and configurations of the stereocenters (Figure 1). Palmyrolide A was shown to inhibit calcium oscillations in murine cerebrocortical neurons and block sodium channels.[7] Apratoxins are a family of potent cytotoxic agents that inhibit cotranslational translocation by targeting the Sec61 in the endoplasmic reticulum.[8,9] Structure-activity relationship (SAR) analysis indicated several crucial elements with the C33-C43 chain being one of the key units in the activity of the compounds (Figure 1).[10–14] Furthermore, the *S,S,S,S* configuration of the C33-C43 unit, as well as the C-37 methyl group were proven to be crucial for its activity.[12,15–17] Here we isolated a new macrocyclic natural product, tumonolide (**1**), which contains a similar C9 unit as in apratoxin C but opposite configuration of all stereocenters, like in palmyrolide A (Figure 1).[7,13,18,19] Both compounds feature the *N*-Me enamide functionality, while tumonolide additionally features two amino acids (Figure 1). Functionally, all three compound classes are distinct, having different molecular targets. We have shown that tumonolide selectively targets the GPCR tachykinin receptor 2 (TACR2) and represents a new structural class of TACR2 antagonists.

GPCRs are the largest family of membrane receptors and prime targets for drug discovery due to their involvement in various physiological functions and diseases, being the target of 40% of all approved drugs.[20–22] Currently 134 GPCR's are targets of FDA approved drugs.[21] Although previously underrepresented, there is a growing list of GPCR-targeting marine natural products from cyanobacteria with differential selectivity profiles.[23] Cyanobacterial compounds can be selective GPCR antagonists such as brintonamides (CCR10, OXTR, SSTR3,

[a] *Dr. S. Kokkaliari, Dr. G. Seabra, Prof. H. Luesch*
Department of Medicinal Chemistry and Center for Natural Products, Drug Discovery and Development (CNPD3)
University of Florida
1345 Center Drive, Gainesville, Florida 32610, United States
E-mail: luesch@cop.ufl.edu

[b] *Prof. L. Grauso*
Dipartimento di Agraria
Università degli Studi di Napoli Federico II
80055 Portici, Napoli (Italy)

[c] *Prof. A. Mangoni*
Dipartimento di Farmacia
Università degli Studi di Napoli Federico II
80131 Napoli, Italy

[d] *Dr. V. J. Paul*
Smithsonian Marine Station at Ft. Pierce
701 Seaway Drive, Ft. Pierce, Florida 34949, United States

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202401393

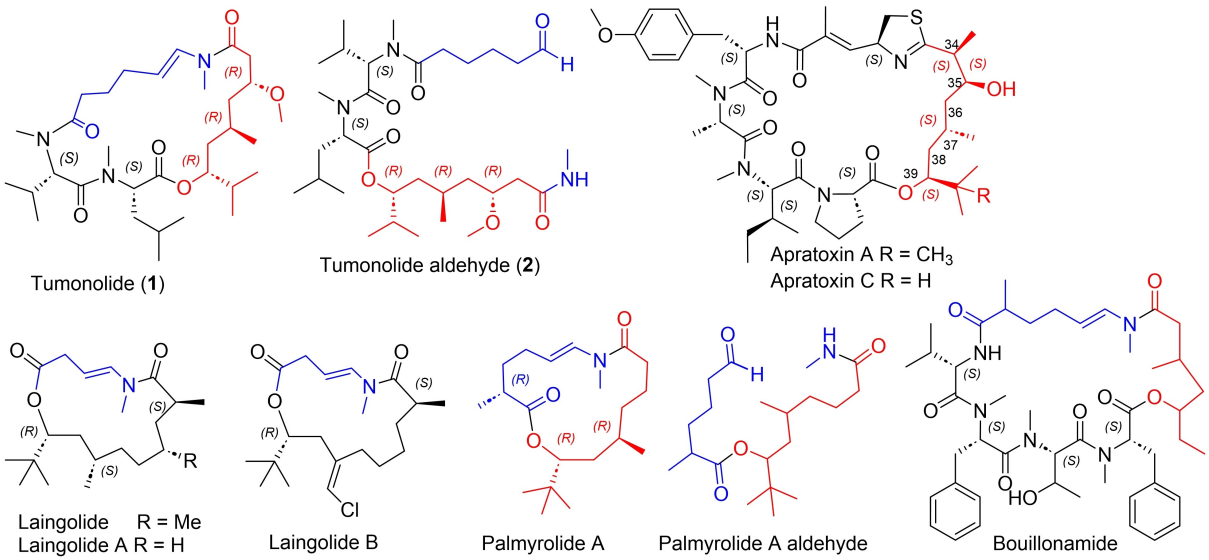

Figure 1. Structure of tumonolide (**1**) and tumonolide aldehyde (**2**), apratoxins A and C, laingolide, laingolide A and B, palmyrolide A and palmyrolide A aldehyde and bouillonamide.

TACR2) or GPCR agonists such as amantamide (CXCR7).[23,24] One of the largest neuropeptide families described in animals, from lower invertebrates to mammals, are the tachykinins which target three GPCRs, the tachykinin receptors (TACR1/NK1R, TACR2/NK2R and TACR3/NK3R).[25–27] The tachykinin receptors are seven-transmembrane-helix receptors belonging to the rhodopsin-like GPCRs (family 1), and all share the same structural unit.[28] The tachykinins share a common C-terminal and show differential affinity, but not selectivity, for the three receptors (Substance P (SP) for TACR1, Neurokinin A (NKA) for TACR2 and Neurokinin B (NKB) for TACR3). The receptors are located in various tissues such as the gastrointestinal tract and smooth muscles of the respiratory system, with TACR3 being located in the peripheral tissues, while the other two are found in both the peripheral and nervous system.[27,29] A TACR1 antagonist, aprepitant, has been approved for chemotherapy induced nausea and vomiting and postoperative nausea and vomiting. TACR2 receptor antagonists are involved in various diseases such as asthma, anxiety and irritable bowel syndrome and have the potential of being used as therapeutics.[30–32] Various TACR2 antagonists have advanced to clinical trials including saredutant, which was evaluated in Phase III studies for major depressive disorders (MDD) and generalized anxiety disorder (GAD). Their development was dropped after completion of the trial and ibodutant, that reached Phase III for irritable bowel syndrome (IBS), was terminated for not showing efficacy in primary and secondary endpoints.[33,34] These compounds showed selectivity for TACR2 compared to TACR1 and TACR3. The widespread distribution of the receptor and its involvement in various diseases indicates the importance of identifying selective antagonists with different scaffolds to further investigate the role of TACR2. In contrast to tumonolide, the structures of the existing selective TACR2 antagonists have common features such as substituted aromatic and indole rings (Figure S1).[33]

We isolated the TACR2 antagonist tumonolide (**1**) along with tumonolide aldehyde (**2**) from a cyanobacterium collected in Tumon Bay, Guam in 2016 and 2023. Tumonolide aldehyde (**2**) is the linearized form. Due to the aldehyde's poor stability, compound **2** could not be subjected to biological testing (Figure S19).

Results and Discussion

Isolation and Planar Structure Elucidation

The marine cyanobacterium VPG16-37 and a recollection of it (VPG23-95) were collected in Tumon Bay, Guam. The freeze-dried sample was extracted with EtOAc:MeOH (1:1), and the extract subjected to solvent-solvent partition, silica gel chromatography and reversed phase HPLC, yielding tumonolide (**1**) and tumonolide aldehyde (**2**).

Tumonolide (**1**) was isolated as a yellow amorphous solid with a molecular formula of $C_{32}H_{57}N_3O_6$ ($[M+H]^+$ 580.4294, calcd. for $C_{32}H_{58}N_3O_6$, 580.4321) as determined by HRESIMS. Prominent features of the 1D NMR showed the presence of three *N*-Me groups ($\delta_{H/C}$ 2.75/30.5, 2.90/30.6, 3.06/30.0), one methoxy group ($\delta_{H/C}$ 3.37/57.1) and a double bond with a disubstituted carbon-carbon double bond ($\delta_{H/C}$ 4.88/110.1, 6.74/130.7). COSY analysis suggested the presence of four spin systems, including two *N*-Me amino acid units, *N*-Me-Val and *N*-Me-Leu (Table 1, Figure 2). For *N*-Me-Leu, H-2 (δ_H 4.97) showed a COSY correlation to H-3 (δ_H 1.88, 1.57), which in turn showed a COSY correlation to H-4 (δ_H 1.54) and that to H-5 (δ_H 1.05) and H-6 (δ_H 0.96) (Figure 2). The adjacent *N*-Me-Val was established from the COSY correlations of H-9 (δ_H 5.09) to H-10 (δ_H 2.41) and of that to H-11 (δ_H 0.86) and H-12 (δ_H 0.76).

The third unit was determined as a tetrasubstituted (dioxygenated and dimethylated) nonanoic acid moiety based

Research Article
doi.org/10.1002/chem.202401393
Chemistry Europe
European Chemical Societies Publishing

Table 1. NMR data for tumonolide (**1**) and tumonolide aldehyde (**2**) in $CDCl_3$.

	Tumonolide A (**1**)					Tumonolide aldehyde (**2**)				
Position	δH, mult., *J* (Hz)[a]	δC[b]	COSY	HMBC	NOESY	δH, mult., *J* (Hz)[a]	δC[b]	COSY	HMBC	NOESY
1		170.7					171.3			
2	4.97, dd (10.0, 4.3)	58.6	3a, 3b	1, 3, 4, 7	3a, 3b, 5	5.32, dd (10.6, 5.4)	54.3	3	1, 3, 4, 7	6
3a	1.88, m	38.3	4	1, 2, 4, 5		1.71, m	36.9	2, 4		
b	1.57, ddd (−14.1, 10.0, 3.9)			1, 2, 4						
4	1.54, m	24.4	3a, 5, 6	3		1.35, m	25.0	3, 5, 6		
5	1.05, d (6.5)	22.2	4	3, 6		0.92, d (6.7)	23.2	4	4, 6	
6	0.96, d (6.6)	23.5	4	3, 4, 5		0.87, d (6.6)	21.1	4	5, 6	
7	2.75, s	30.5		2, 8		2.99, s	31.1		2, 8	
8		170.3					171.2			
9	5.09, d (10.5)	57.5	10	8, 10, 11, 12, 13, 14	4, 5, 11, 12, 13	5.22, d (10.8)	57.8	10	8, 10, 11, 13, 14	7, 11, 12
10	2.41, m	27.3	11, 12	8, 9, 11, 12		2.33, m	27.1	9, 11, 12	9	
11	0.86, d (6.5)	19.8	10	9, 10, 12		0.96, d (6.5)	19.3	10		
12	0.76, d (6.8)	17.7	10	9, 10, 11	9, 10, 13	0.83, d (6.7)	18.2	10		
13	2.90, s	30.6		9, 14		2.93, s	30.0		9, 14	
14		173.3					173.0			
15a	2.24, td (−12.9, 11.6, 5.0)	33.7	16	14, 16, 17		2.33, m	33.2	16	16, 17	
b	2.15, td (−12.9, 12.6, 5.5)		16	14, 16, 17						
16	1.72, m	26.0	15a, 15b	14, 15, 17, 18		1.71, o	24.3	15, 17	14	
17a	2.41, m	30.7		15, 16, 18, 19		1.71, o	21.5	17, 18		
b	1.88, m		16	15, 16, 18, 19						
18	4.88, ddd (13.9, 5.1, 3.8)	110.1	17a, 17b	16, 17, 19		2.47, m	43.6	17, 19	16, 17, 19	
19	6.74, d (13.9)	130.7	18	17, 18, 21	15b, 17a, 17b, 22b	9.78, t (1.5)	202.2	18	17, 18	
20	3.06, s	30.0		19, 21		2.81, d (4.8)	25.5	21-NH	21	
21		169.1					171.6			
22a	2.78, dd (−12.4, 4.6)	37.6	23	21, 23, 24		2.41, dd (−14.9, 5.6)	40.5	23	21, 23, 24	
b	2.69, dd (−12.4, 8.9)		23	21, 23	24a, 31	2.33, dd (−14.9, 5.1)				
23	3.59, dddd (11.0, 8.9, 4.6, 2.5)	76.6	22a, 22b, 24a	22, 31	22a, 24b, 31, 32	3.62, m	75.9	22, 24	21, 31	22a, 31, 32
24a	1.42, ddd (−14.0, 11.0, 3.2)	44.3	23, 24b, 25	23, 32	22b, 27, 31	1.71 m, o	40.4	23, 25		
b	1.06, ddd (−14.0, 10.5, 2.5)		24a, 25		23	1.05, m				
25	1.72, m	25.1	24a, 24b, 26a, 26b, 32			1.71, o	25.6	24, 26, 32		

Table 1. continued

	Tumonolide A (**1**)					Tumonolide aldehyde (**2**)				
Position	δH, mult., *J* (Hz)[a]	δC[b]	COSY	HMBC	NOESY	δH, mult., *J* (Hz)[a]	δC[b]	COSY	HMBC	NOESY
26a	1.34, ddd (−13.3, 10.3, 3.9)	40.4	25, 27	24, 25, 27, 28, 32	15a, 30	1.45, ddd (−14.3, 9.2, 5.1)	38.6	25, 27	24, 25, 27, 32	32
b	1.14, ddd (−13.1, 11.2, 4.8)		25, 27	24, 25, 27, 28, 32	15b, 27, 32	1.35, m				
27	4.89, ddd (10.3, 4.8, 2.2)	77.1	26a, 26b	1, 26, 28, 29, 30		4.87, m	77.0	26, 27	29, 1	28, 29
28	1.88, sept d (6.8, 2.2)	28.6	29, 30	26, 29, 30		1.79, m	31.2	27, 29, 30	26, 27, 29	29, 30
29	0.88, d (6.8)	20.1	28	27, 28, 30		0.87, d (6.8)	17.6	28	27, 28, 30	
30	0.83, d (6.8)	15.2	28	27, 28, 29		0.87, d (6.8)	18.1	28	27, 28, 29	
31	3.37, s	57.1		23		3.34, s	56.6		23	
32	0.82, d (6.5)	17.1		24, 25, 26		0.92, d (6.6)	20.1	25	24, 25, 26	
20-NH	-	-				6.37, brs				20, 22a

[a] Recorded at 600 MHz, [b] Recorded at 150 MHz.

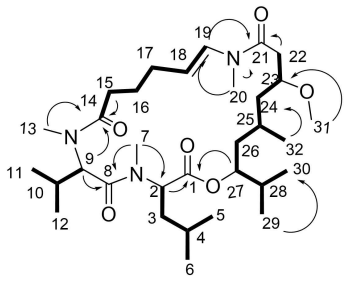

Figure 2. Key COSY (bold) and HMBC (arrows) data for tumonolide (**1**).

on COSY analysis (Table 1, Figure 2). The position of the C-31 methoxy group (δ_H 3.37/δ_C 57.1) was established based on HMBC to the oxygenated carbon C-23 (δ_C 76.6). HMBC correlations from H-22 to C-21, coupled with chemical shift data, indicated the presence of the carbonyl C-21.

The fourth unit was established to be an unsaturated C6 unit with a terminal polarized double bond. COSY correlations were observed for H-15 (δ_H 2.24, 2.15) to H-16 (δ_H 1.72) and H-16 to H-17 (δ_H 2.41, 1.88). The unit was further extended from the COSY correlation of H-17 to H-18 (δ_H 4.88) and of H-18 to H-19 (δ_H 6.74), indicative of a nitrogenated double bond (enamine) with *trans* configuration ($J_{H-18/H-19}$ = 13.9 Hz), accounting for the third nitrogen that had to be present based on the odd molecular weight. The third nitrogen was also methylated (δ_H 3.06/δ_C 30.0), and positioning was supported by HMBC correlations to C-19 (δ_C 130.7). HMBC correlations from H-15 to C-14 established the carbonyl functionality.

The four units were readily connected by HMBC analysis. The amino acid units were linked via an amide bond through the C terminus of *N*-Me-Val based on HMBC correlations of H-7 (δ_H 2.75) to C-2 (δ_C 58.6) and C-8 (δ_C 170.3) and of H-9 to C-8 (Figure 2, Table 1). The carbonyl of the enamine-containing hexenoic acid unit was positioned adjacent to the *N*-Me-Val based on the HMBC correlations from H_3-13 (δ_H 2.90) to C-14 (δ_C 173.3). The enamide was proposed based on HMBC correlations of H-20 (δ_H 3.06) to C-21 (δ_C 169.1) and C-19 (δ_C 130.1), supported by the HMBC from H-22 to C-21. The ring closure was established from the HMBC correlation of H-27 and H-2 to C-1 (δ_C 170.7), fulfilling the molecular formula requirement and completing the planar structure of (**1**).

Tumonolide aldehyde (**2**) was isolated as a colorless oil with a molecular formula of $C_{32}H_{59}N_3O_7$ ([M + H]$^+$598.4410, calcd. for $C_{32}H_{60}N_3O_7$ 598.4431) as determined by HRESIMS, indicating an additional oxygen compared with **1**. The ^{1}H NMR showed similarities to that of **1**, but lacked the presence of the double bond and instead featured aldehyde signals (δ_H 9.78/δ_C 202.2). The compound had to be linear to fulfill the five degrees of unsaturation (Table 1, Figure 3). The two amino acids were determined as *N*-Me-Leu and *N*-Me-Val using COSY and HMBC correlations. The units were connected from the HMBC correlations of H-7 (δ_H 2.99) to C-2 (δ_C 54.3) and C-8 (δ_C 171.2) and of H-9 to C-8 to form the dipeptide as in **1**. The tetrasubstituted nonanoic unit was also present in **1**, supported by COSY and HMBC analysis (Figure 3, Table 1). However, COSY

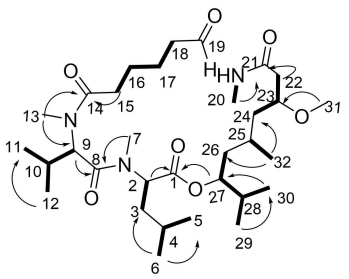

Figure 3. Key COSY (bold) and HMBC (arrows) data for tumonolide aldehyde (**2**).

correlation of H-20 (δ_H 2.81) to the broad singlet of the 20-NH (δ_H 6.34) indicated the presence of a secondary amide in **2**. The fourth unit was established as a carboxylic acid aldehyde based on COSY correlations and HMBC correlations of H-15 to C-16 (δ_C 24.3) and C-17 (δ_C 21.5) due to the overlap of H-16 and H-17 (δ_H 1.71). The unit was further extended from the COSY correlation of H-17 to H-18 (δ_H 2.47) and of H-18 to the aldehyde H-19 (δ_H 9.78). This unit was positioned next to the *N*-Me-Val based on the HMBC correlations of H-13 (δ_H 2.93) and H-16 to C-14 (δ_C 173.0). Compound 2 appeared to be the ring-opened form of **1** (Figures 1 and 3).

Tumonolide aldehyde (**2**) is the linear aldehyde of **1**. There has been one previous report of a similar aldehyde, palmyrolide A aldehyde, which was the synthetic byproduct of the hydrolysis of palmyrolide A, as well as an intermediate in its total synthesis.[7,19] However, we isolated both compounds from the cyanobacterial extracts, suggesting that the aldehyde is also a natural product. Furthermore, monitoring of the ^{1}H NMR of the two-compound mixture, suggested that **2** slowly converts in solution into **1**, indicating that the configurations of all stereogenic centers are the same (Figure S19). Ring opening of **1** can be achieved under basic conditions, after treatment with 1 M LiOH for 72 h (Figure S20).

Relative and Absolute Stereochemical Determination

The configurational assignment for the five stereocenters was made for **1**. The absolute configurations of the two amino acids units were determined using acid hydrolysis followed by chiral HPLC-MS analysis. Both *N*-Me-Val and *N*-Me-Leu were found to possess L configuration (both *S*). As mentioned above, H-18 and H-19 of the enamide unit were determined to be a part of a *trans* double bond system, based on their $J_{H,H}$ coupling constant ($J = 13.9$ Hz). The C-8/N and C-21/N amide bonds were established as *cis*, based on the NOE correlations of H-2 and H-9 and H-19 and H-22 respectively, while the C-14/N amide was *trans* based on the NOE correlations of H-13 to H-15. The relative configuration of the substituted nonanoic acid unit was established using Murata's *J* based analysis of two- and three-bond proton-carbon and vicinal proton-proton coupling constants.[35,36] This type of analysis was initially developed for acyclic compounds but has since been utilized in macrocycles and cyclic compounds.[10,36,37] The values were obtained using ^{1}H NMR and 1D TOCSY for the $^3J_{H,H}$ coupling constants, HETLOC and HSQMBC experiments for the $^{2,3}J_{C,H}$ (Figure 4). Due to the presence of only 1,3-methine systems, all conformers could be distinguished based on the $^{2,3}J_{C,H}$ and $^3J_{H,H}$ values. Two medium values were observed, one for the C22-C23 system and one for the C26-C27 system. As a result, the possibility of rotamers was considered, but was eliminated due to the other values of the same system not being medium as expected from the model (Figure 4A).[36] The 2D NOESY data were in agreement with the proposed conformation (Figure 4B). The relative configuration of the unit was therefore established as either 23*S*,25*S*,27*S* or 23*R*,25*R*,27*R*. Apratoxin A and its analogues have all *S* configuration, which is crucial for the activity of the compound and is conserved among the members of the family.[10] On the contrary, the laingolides and palmyrolide A, which share a similar unit adjacent to the *N*-Me enamide, have different configurations, with the oxygenated methine bearing the *t*-Bu unit being *R*. The methyl-bearing methine in the 1,3-dimethine system is *R* for palmyrolide A and *S* for laingolide A (Figure 1), indicating potentially different specificities within the biosynthetic pathways for the two compounds, as well as for apratoxins.

To determine the absolute configuration of the unit, we attempted to hydrolyze **1** using conventional base hydrolysis methods, such as 1 M NaOH, LiOH and NaOMe/MeOH, in order to use the newly formed 27-OH for Mosher's analysis. The reaction yielded no product with NaOH and LiOH, and resulted in the generation of multiple diastereomers using NaOMe/MeOH by LCMS prior to the opening of the lactone ring. Acid hydrolysis resulted in degradation of the unit.

Therefore, a computational approach was utilized to determine the absolute configuration of the three remaining centers.

A quantum mechanical study of the alternative stereoisomers 23*R*,25*R*,27*R*-**1** and 23*S*,25*S*,27*S*-**1** was performed aimed at predicting their respective NMR, UV and ECD spectra. The most critical part of this computational study was the conformational search, which is described in detail below.

Following an established protocol, the conformational search was performed using molecular dynamics (MD) in the CVFF force field using the InsightII/Discover suite.[38] A 250-ns MD at 600 K generated thousands of different conformers for each stereoisomer, and several hundred of them were within 6 kcal/mol from the lowest energy conformer. Most of these conformers did not match the conformational features described above, i.e. the conformation of the amide bonds and the *gtttg* (*gauche*/*trans*/*trans*/*trans*/*gauche*) conformation of the C-21/C-28 segment determined by application of the Murata method. Moreover, the best conformer that matched NMR-determined conformational features (hereafter "NMR-compliant conformers") was over 3 kcal/mol higher in energy than the lowest-energy conformer. The conformers were then re-optimized in the MMFF94 force field, but again the NMR-compliant conformers were not among the lowest-energy conformers, although we retrospectively found that MMFF94 energies correlated better than CVFF energies with quantum mechanical energies.[39]

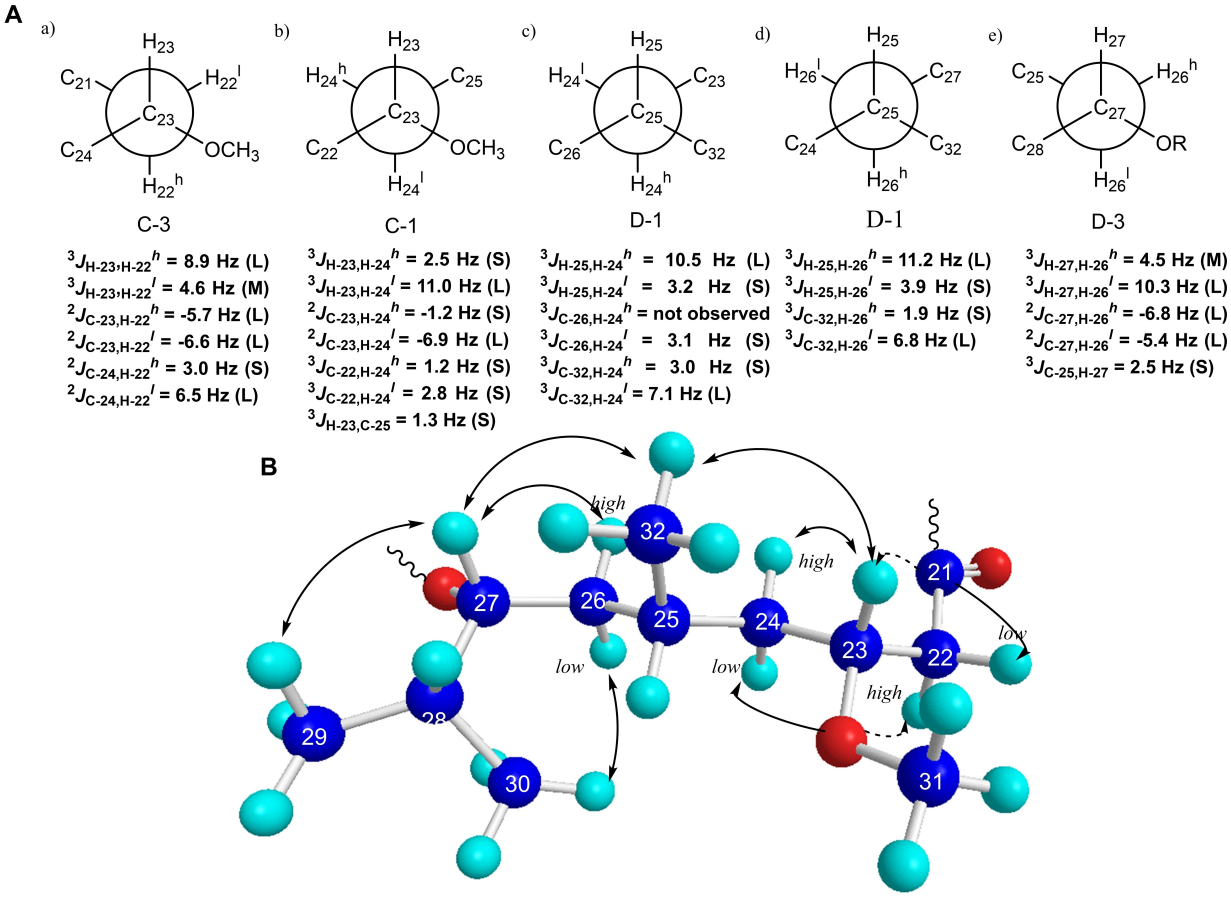

Figure 4. A) $^3J_{H,H}$, $^{2,3}J_{C,H}$ values that resulted in determining the relative configuration of the C21-C32 unit of tumonolide (**1**). The magnitude of the coupling constants was determined as small (S), medium (M) and large (L), B) NOE correlations for the tumonolide (**1**) polyketide unit (C22-C32), indicating the *J*-based analysis matches the NOE data. H^h and H^l indicate the diastereotopic methylene protons observed at higher and lower field, respectively.

Because evaluation of relative energy of conformers in an empirical force field can be highly inaccurate, we decided to re-examine selected low-energy conformers at the quantum mechanical level.[40] Therefore, we selected all the NMR-compliant conformers, the 30 lowest-energy conformers in the CVFF force field, and the 30 lowest-energy conformers in the MMFF94 force field, and optimized all of them using Gaussian 16 at the B3LYP/6-31G(d) level of theory.[41] The energies of the NMR-compliant conformers were still higher than those of other conformers, but to a lesser degree (0.76 Kcal/mol for the best NMR-compliant conformer). Finally, we used the higher B3LYP/6-311G+(d,p) level of theory with the SMD solvation model to re-evaluate the energy of the optimized conformers, and found the lowest-energy conformers were all NMR-compliant conformers.[42]

Therefore, the conformational search protocol was defined as follows: 250-ns dynamics at 600 K in the CVFF force field; re-optimization in the MMFF94 force field; selection of NMR-compliant conformers; re-optimization at the B3LYP/6-31G(d) level; energy evaluation at the B3LYP/6-311G+(d,p)/SMD. To achieve a comprehensive search, we applied this search protocol repeatedly for each stereoisomer, until no new low-energy, NMR-compliant conformers was found. This identified 9 conformers for *RRR*-**1** and 7 conformers for *SSS*-**1** with population >1%, which were used for all the subsequent calculations. The conformers were all similar, with differences limited to the position of the side chain isopropyl and isobutyl groups and in the conformation of the three-methylene chain C-15/C-17 (Figure 5A).

Prediction of ^{1}H and ^{13}C NMR chemical shifts was performed at the mPW1PW91/6-311+G(d,p)/PCM($CHCl_3$) level for application of the DP4+ statistical analysis, and at the mPW1PW91/6-311+G(2d,p)/SMD($CHCl_3$) level to allow the use of the pre-calculated scaling factors available in the literature for chemical shift calculation.[43,44] DP4+ statistical analysis showed a good, but not excellent, overall preference for *RRR*-**1** (89.73%), although some of the terms of the DP4+ statistics (particularly the unscaled ^{1}H chemical shifts) favored *SSS*-**1** (Figure S22). The error of ^{13}C chemical shifts calculated according to ref. [39] was comparable (RMSD was 1.89 ppm for *RRR*-**1** and 1.90 ppm for *SSS*-**1**) while the error of calculated ^{1}H chemical shifts was remarkably better for *RRR*-**1** (RMSD was 0.123 ppm vs. 0.151 ppm), again suggesting *RRR*-**1** as the correct structure (Tables S2–S3 and Figure S23).

Overall, prediction of NMR chemical shift provided strong suggestion, but not clearcut demonstration of the correct stereoisomer. This was not surprising, because comparison of the lowest energy conformation of *RRR*-**1** and *SSS*-**1** showed

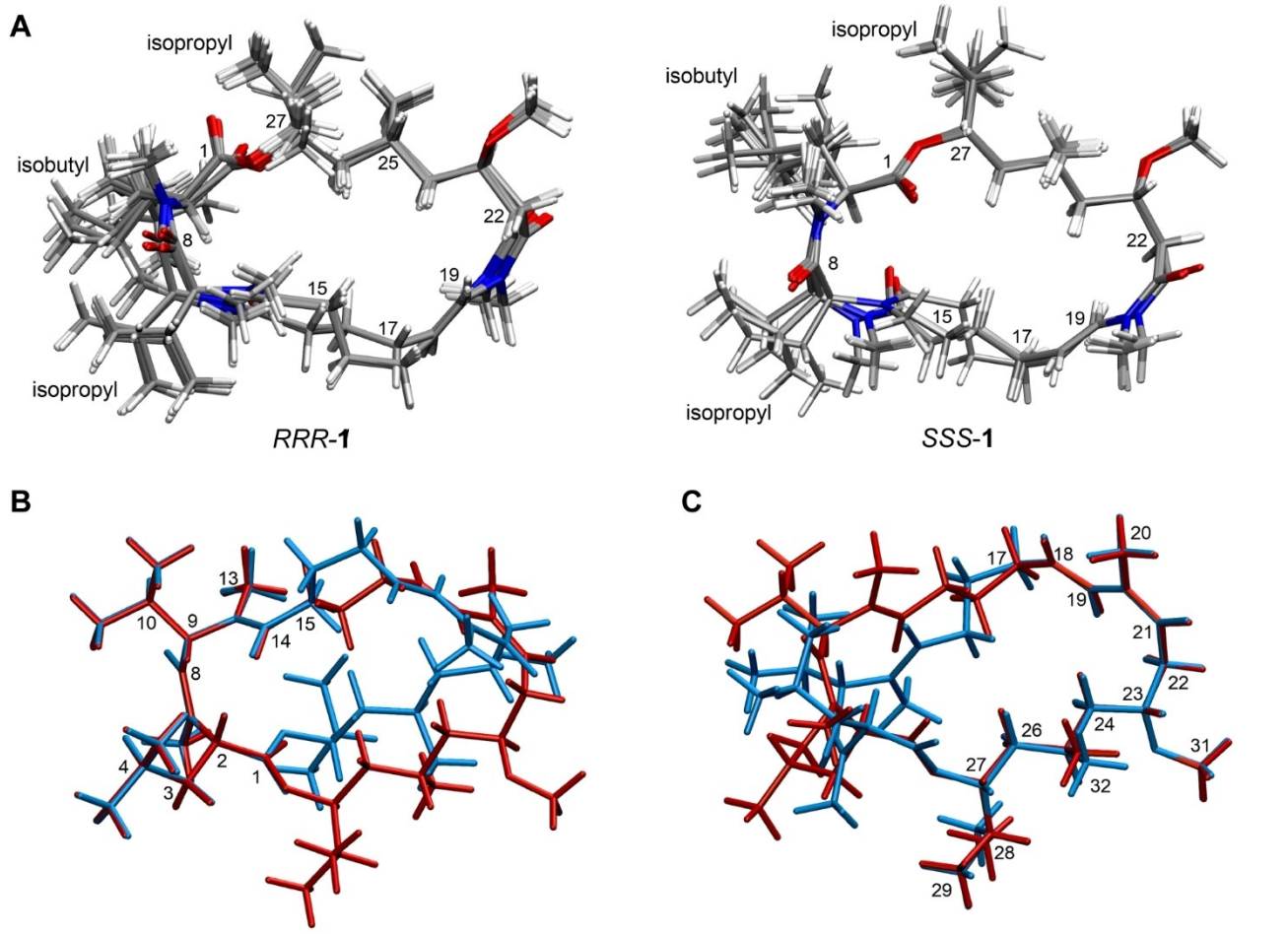

Figure 5. A) Differences between conformers are the positions of the isopropyl and isobutyl side chains and the conformation of the three-methylene chain C-15/C-17. B) Overlay of *RRR*-**1** (blue) and *SSS*-**1** (red) (respective lowest energy conformers). The C-1/C-15 segments of the molecules are almost exactly superimposable. C) Overlay of *RRR*-**1** (blue) and the mirror image of *SSS*-**1** (red) (respective lowest energy conformers). The C-17/C-32 segments are almost exactly superimposable, and therefore they are enantiomeric in *RRR*-**1** and *SSS*-**1**.

that the C-1/C-15 segments of the molecules are almost superimposable, whereas the C-17/C-29 segments are almost enantiomeric (Figure 5B and 5 C). In both segments, therefore, chemical shifts and coupling constants of *RRR*-**1** and *SSS*-**1** are expected to be very similar, making it difficult to discriminate between the two molecules.

Fortunately, clearcut NMR evidence for the 23*R*,25*R*,27*R* configuration of **1** was found in the intense NOESY correlation between H_3-7 (δ_H 2.75) and H_3-29 (δ_H 0.88), which could be interpreted with confidence thanks to the detailed knowledge acquired about the conformational behavior of *RRR*-**1** and *SSS*-**1**. These methyl groups are very close to each other (the distance between the closest protons is <2.5 Å) in 6 out of the 9 conformers of *RRR*-**1**, which together account for 93.7% of population (Figure 6). In contrast, owing to the inverted configuration at C-27, these methyl groups are far from each other (the distance between the closest protons is >4.5 Å) in all the low-energy conformers of *SSS*-**1**.

Prediction of the ECD spectrum of **1** confirmed the relative configuration determined from NMR data as well as the absolute configuration determined by chemical shift prediction

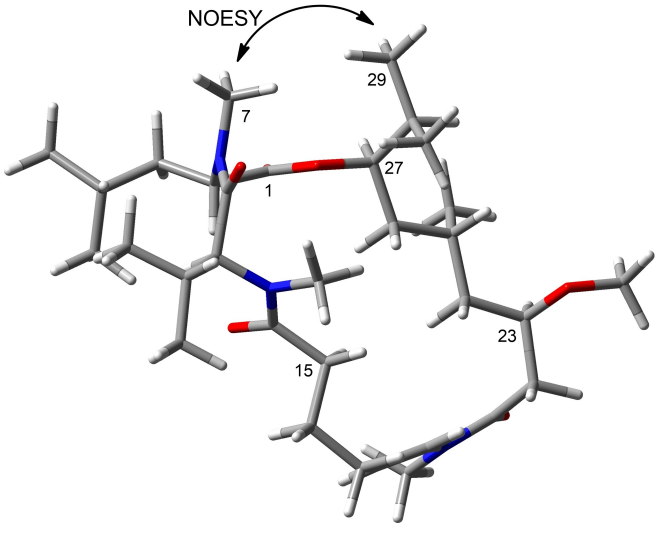

Figure 6. The lowest energy conformer of *RRR*-**1** showing the diagnostic NOESY correlation between H_3-7 and H_3-29, which are too far from each other in any low-energy conformer of SSS-**1**.

and NOE analysis. The UV and ECD transitions of *RRR*-**1** and *SSS*-**1** were calculated at the B3LYP/TZVP/PCM(MeOH) level, and the UV and ECD spectra of individual conformers were obtained using the program SpecDis. Finally, the Boltzmann average of the UV and ECD spectra of individual conformers (Figure S24) yielded the predicted UV and ECD spectra. The predicted ECD spectrum of *RRR*-**1** was very similar to the experimental ECD spectrum of tumonolide (**1**), whereas neither the predicted ECD spectrum of *SSS*-**1** nor its mirror image matched the experimental spectrum (Figure 7).

Biological Evaluation

Due to the large number of drugs targeting the GPCR family, tumonolide (**1**) was profiled against a GPCR panel using the β-arrestin functional assay (Figures 8A,B) at a single concentration (20 μM) in duplicate.[21] Screening against 168 targets revealed three primary hits in antagonist mode (>70% inhibition) and no hit in the agonist mode. The activity of **1** against the three targets, TACR2, CNR2 and NPSR1b, was further evaluated by dose-response analysis. The compound showed an IC_{50} of 7.0 μM against TACR2, but activity against the other two targets was significantly lower, with IC_{50} values of 20.0 μM against NPSR1b and 35.0 μM against CNR2 (Figure 8C, Table 2). Furthermore, from the initial screening it was determined that **1** showed high selectivity for TACR2 compared to the other members of the tachykinin family of receptors showing over 10-fold selectivity over TACR1 and TACR3 (IC_{50} for TACR1 > 100 μM and IC_{50} for TACR3 ≈ 98.6 μM) (Figure 8D), indicating a remarkable selectivity to discriminate among family members.

TACR2 is one of the GPCRs that activate a phosphatidylinositol-calcium second messenger system.[45] Stimulation of the tachykinin receptors leads to elevation of intracellular calcium through different pathways. The natural ligands, neurokinins, can activate phospholipase C by binding to the Gi (decreases cAMP) and Gs (increases cAMP) subunits of the receptors. Inositol (1,4,5)-trisphosphate (IP3), which can increase the cytosolic calcium concentration, is the product of the hydrolysis of phosphatidylinositol 4,5-biphosphate by phospholipase C.[46] We tested **1** in a TACR2 mobilization assay where it showed an IC_{50} of 4.6 μM in the antagonist mode, consistent with the other functional β-arrestin assays (Figure 8E). We then evaluated the binding ability of **1**. A specific binding assay was performed in agonist mode to determine the binding affinity of **1** to TACR2, with the compound showing an IC_{50} of 2.1 μM and a *Ki* of 1.2 μM (Figure 8F), indicating a strong correlation between binding and functional responses across assay systems.

To further examine the activity of **1**, we decided to evaluate the cytotoxicity of **1** in a series of experiments. Tumonolide (**1**) showed no cytotoxicity at the highest tested concentration (100 μM) in HEK293 and RAW264.7 cells, and exhibited an $IC_{50} = 53.5 \pm 7.5$ μM in HCT116 cells (Figure S21). To further probe the activity of the compound on the transcriptional level for validation as well as determination of pleiotropic activities, we decided to proceed with a global approach, using HCT116 colorectal cancer cells. This cell line was chosen because of the high expression of TACR2 in the gastrointestinal tract and colon cells.[29,32] HCT116 cells were then treated with tumonolide (**1**) at 20 μM for 12 h and processed for RNA sequencing. The results showed that the compound up- and down-regulated selected genes and activated seven pathways, as determined by

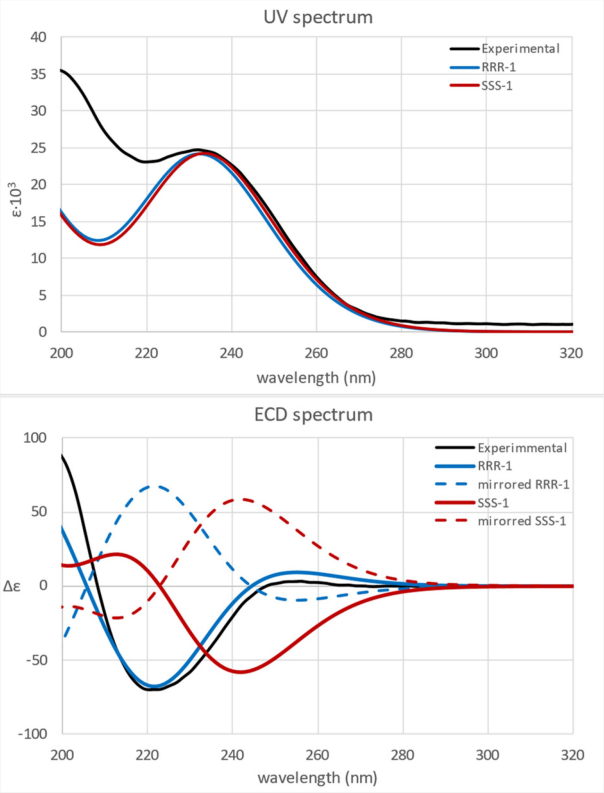

Figure 7. Experimental UV and ECD spectra of tumonolide (**1**) (black line) and predicted UV and ECD spectra of *RRR*-**1** (blue line) and SSS-**1** (red line). Level of theory B3LYP/TZVP/PCM(MeOH); σ = 0.45 eV; UV shift = −5 nm.

Table 2. IC_{50} values for tumonolide (**1**) in the GPCR screening.

Target	IC_{50} (μM) β-arrestin pathway	IC_{50} (μM) Calcium mobilization	IC_{50}/Ki (μM) Binding
TACR1	>100	–	–
TACR2	7.0	4.6	2.1/1.2
TACR3	≈98.6	–	–
CNR2	35.0	–	–
NSPR1b	20.0	–	–

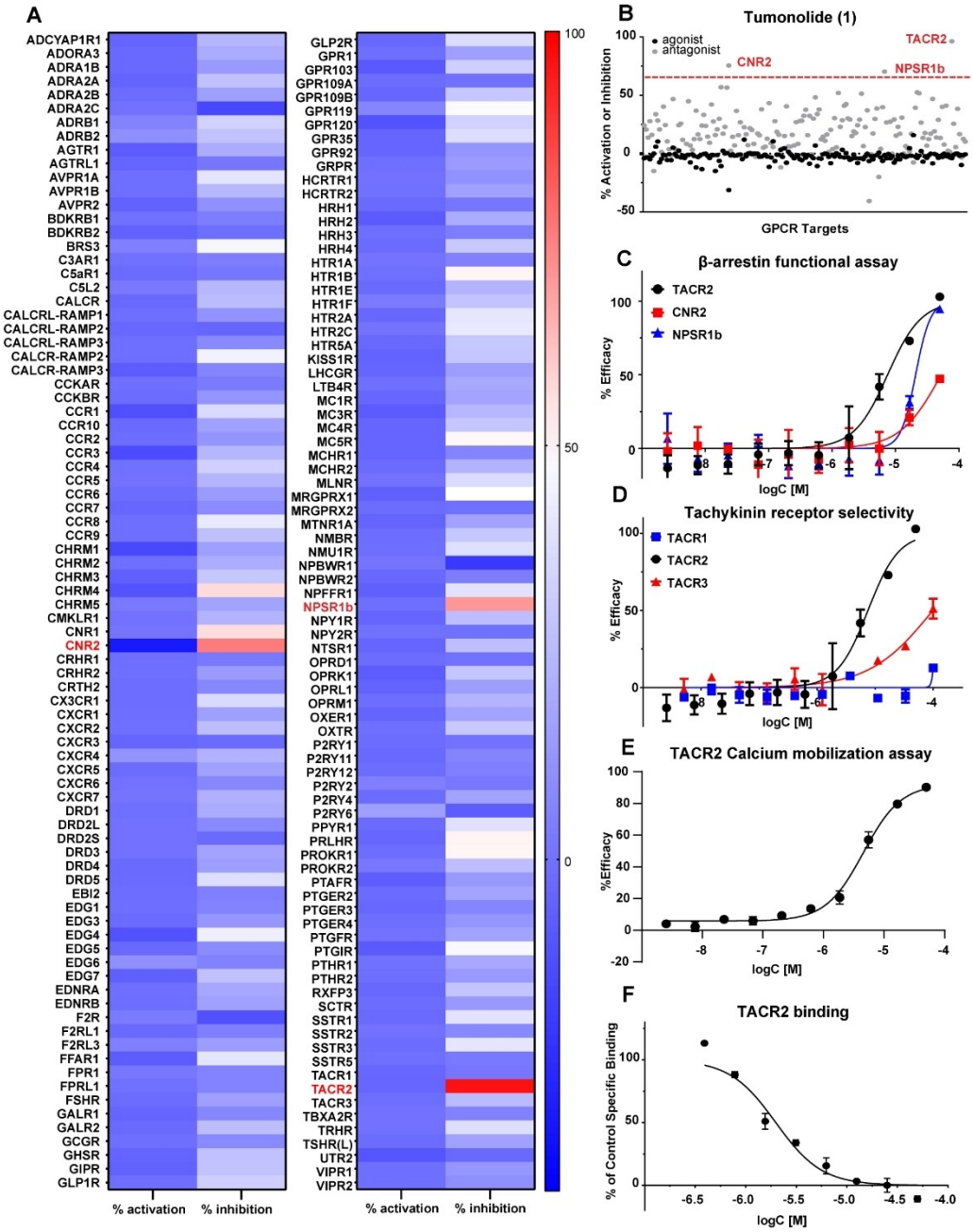

Figure 8. GPCR profiling of tumonolide (**1**). A) Heat map and B) scatter plot showing all GPCR targets at 20 μM, C) Dose response curves for the three targets that showed > 70% inhibition at 20 μM in the primary screen, D) Dose response curves for TACR1, TACR2 and TACR3, E) Calcium mobilization data for **1**, F) specific binding of **1** to TACR2. All experiments were performed in duplicates. Error bars indicate the ± SD.

Ingenuity Pathway Analysis, IPA (Figure 9A, B), including GPCR signaling, the insulin secretion pathway and also the pulmonary fibrosis idiopathic signaling pathway (Figure S20, Table 3). Most of the upregulated pathways could be the result of a different effect not yet discovered. The upregulation of the GPCR signaling pathway could be related to the inhibition of TACR2 through feedback, but the direct relationship of TACR2 inhibition and GPCR signaling pathway upregulation would need to be probed in future experiments.[47–49] The top diseases and functions identified through IPA are the inhibition of the degeneration of the nervous system, inhibition of organismal death and apoptosis and the increase of the size of body and increase of cellular homeostasis.

Molecular Docking

Next, we aimed to identify the binding mode of tumonolide (**1**) to TACR2. In the cryo-EM structure, TACR2 establishes hydrogen bonds (HB) to Neurokinin-A via Asn86, Tyr93, and Tyr289 (Figure S25).[50] The binding pose obtained shows tumonolide (**1**) establishing two of the same interactions with Tyr93 and

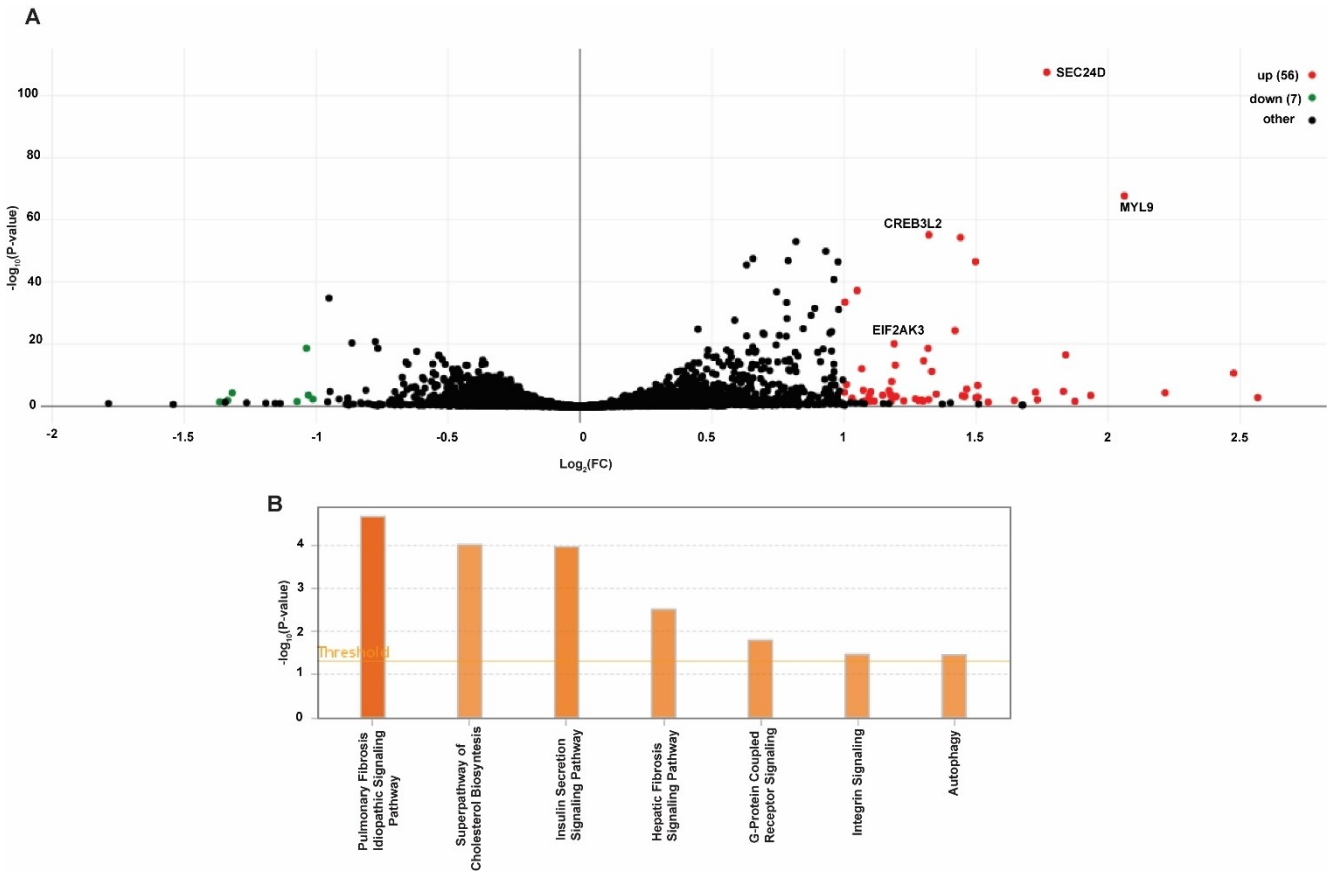

Figure 9. RNA sequencing and pathway analysis (HCT116 cells). A) Volcano plot of up- (red) and down- (green) regulated genes by **1** after 12 h of treatment, with $p < 0.05$ and fold change > 2. The genes shown are common between some of the generated pathways. B) Canonical pathways for **1** with a positive z-score > 2.

Table 3. Top canonical pathways with z-scores and p values.

Canonical Pathways	p-value	z-score	Affected genes	
Pulmonary Fibrosis Idiopathic Signaling Pathway	2.14E–05	3	ACTB, ACTG1, ACVR1C, COL13A1, COL16A1, EDNRA, FOS, FZD2, GLI1, LPAR1, PLAU, SERPINE1	upregulated
Superpathway of Cholesterol Biosynthesis	9.55E–05	2	HMGCS1, MSMO1, SC5D, SQLE	upregulated
Insulin Secretion Signaling Pathway	1.05E–04	2.333	CREB3L2, EIF2AK3, FYN, SEC11C, SEC61A1, SEC61B, SEC61G, SSR1, SSR3, STAT4	upregulated
ABRA Signaling Pathway	1.26E–04	2.449	ACTB, ACTG1, FOS, MYL9, TAGLN, TPM4	upregulated
Hepatic Fibrosis Signaling Pathway	3.16E–03	2.121	ACVR1C, CREB3L2, EDNRA, FOS, FZD2, GLI1, KLF9, MYL9, PTCH1, SERPINE1	upregulated
G-Protein Coupled Receptor Signaling	1.58E–02	2.111	CREB3L2, EDNRA, FICD, FOS, FYN, FZD2, GPR87, LPAR1, MYL9, NFATC4, RASGRP1	upregulated
			LTB4R2	downregulated
Integrin Signaling	3.31E–02	2	ACTB, ACTG1, FYN, MYL9, PPP1R12B	upregulated
Autophagy	3.55E–02	2	CREB3L2, EIF2AK3, FOS, OPTN, WIPI1	upregulated

Tyr289, plus two new hydrogen bonds with Asn90 and Asn110, and this pose was used as input for 100 ns MD simulations (Figure S26).

Figure 10 shows a snapshot of the system at the end of the MD simulation, and Figure 11 shows the most persistent interactions between tumonolide (**1**) and TACR2 during the last 80 ns of simulation. The hydrogen bonds to **1** from Gln166 and

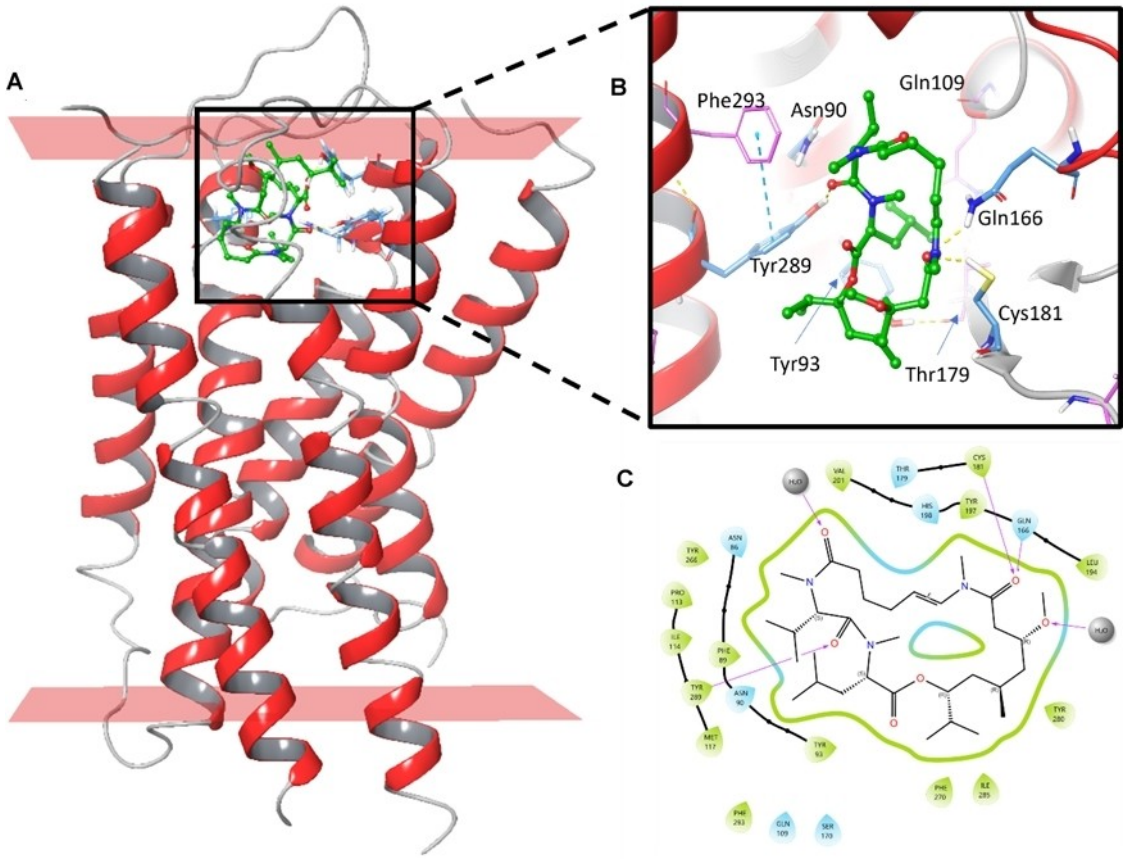

Figure 10. Structure of tumonolide bound to TACR2 obtained at the end of the molecular dynamics. A) The whole complex shows the membrane limits in red slabs, tumonolide in green, and TACR2 residues interacting with tumonolide highlighted in blue. B) A close-up on tumonolide binding site. C) Diagram of interactions between tumonolide and TACR2. Purple arrows indicate hydrogen bonds (donor to acceptor), and grey circles indicate solvent exposure.

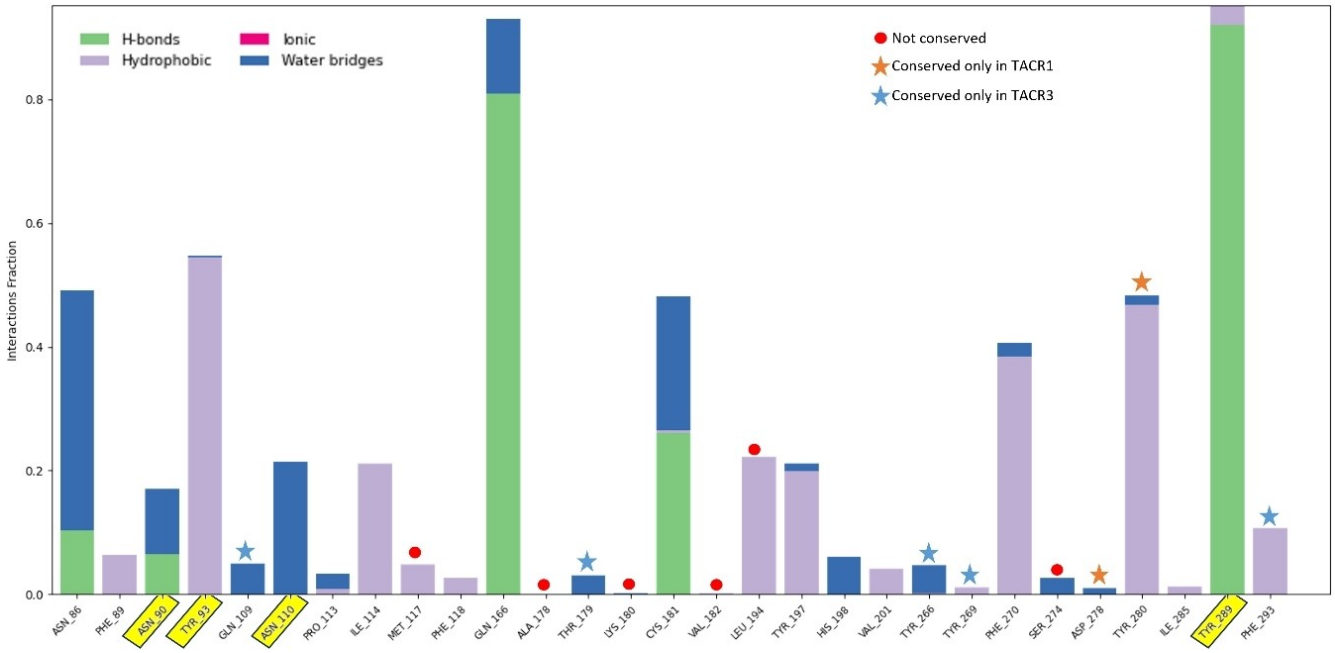

Figure 11. Persistence of the interactions during the last 80 ns of molecular dynamics. Interactions highlighted in yellow were present as hydrogen bonds in the docked structure. Red dots indicate interactions with residues unique to TACR2, orange stars indicate interactions with residues conserved only between TACR2 and TACR1, and blue stars indicate interactions with residues preserved only in TACR3. Interactions with no indicator are conserved in all three TACRs.

Tyr289 are maintained for 80% and 92% of the time, respectively. Gln166 also interacts with **1** through water bridges another 12% of the time. Other important hydrogen bonds are Asn86 and Cys181, which are both present ~50% of the time either as direct bonds or water bridges. The hydrogen bond to Tyr93 is quickly lost and replaced by hydrophobic interactions to the leucine residue in **1** The interactions with Asn90 and Asn110 are still present ~18% and 20% of the time.

To understand the differential binding to TACR1 and TACR3, the sequences for TACR1–3 were obtained from UniProt and aligned using the TACR2 sequence as a reference. The docked pose of tumonolide to TACR2 was used to determine the binding site residues, as all residues with atoms within 6 Å of the ligand. The complete sequences of TACR1 and 3 share a 60% and 62% identity to TACR2, respectively. However, when only the binding site residues are considered (Table 4), TACR1 shares only 65% of its identity with TACR2, while TACR3 shares 84%. Binding site residues that differ from TACR2 are listed in Table 4.

In the binding site, there are 7 residues conserved only between TACR2 and 3, five of them are within 4 Å of the ligand. Gln109, replaced by a histidine in TACR1, makes hydrogen bonds to Asn90 and Thr179, this last one replaced by a valine in TACR1. Tyr266 and Tyr269 are both replaced by a phenylalanine in TACR1. The loss of the hydroxyl group breaks interactions with neighbor helices that hold the binding site structure. Phe293 makes a π stacking interaction with Tyr289, positioning Tyr289 for hydrogen bonding with tumonolide. The replacement of Phe293 by a methionine in TACR1 is likely to weaken this hydrogen bond (Figures S27–28).

Table 4. Binding site residues that differ from TACR2 (in red). The proximity column indicates how close the residues are to the ligand in the docked TACR2 binding pose.

TACR2	Proximity (Å)	TACR1	TACR3
Gln109	<4	His108	Gln161
Leu111	<6	Phe110	Phe163
Met117	<6	Val116	Val169
Ala178	<6	Arg177	Arg230
Thr179	<4	Val178	Thr231
Lys180	<4	Val179	Leu232
Val182	<6	Met181	Phe234
Val183	<6	Ile182	Val235
Leu194	<4	Tyr192	His244
Tyr266	<4	Phe264	Tyr315
Tyr269	<4	Phe267	Tyr318
Ile271	<6	Leu269	Ile320
Gly273	<6	Pro271	Thr322
Ser274	<6	Tyr272	Ala323
Tyr280	<4	Tyr278	Asn328
Cys281	<6	Leu279	Arg330
Phe293	<4	Met291	Phe342

Conclusions

In conclusion, we isolated a selective TACR2 inhibitor, tumonolide (**1**) and its aldehyde (**2**). The two compounds show the most structural similarities to palmyrolide A and the semi-synthetic palmyrolide A aldehyde but also contain an apratoxin signature. The compounds were fully elucidated through various NMR and computational methods to obtain their planar structure and relative and absolute configuration. Tumonolide aldehyde (**2**) was not evaluated for its biological activity because the compound was not stable. Further experiments are needed to determine the effect of tumonolide in other cell types and its potential as a selective TACR2 inhibitor, and to evaluate the effect of it on the pathways identified from RNA-sequencing data. Furthermore, the reactivity of tumonolide aldehyde (**2**) needs to be further examined.

Experimental Section

General Experimental Procedures

Optical rotations were recorded on a Rudolph Research Analytical Autopol III automatic polarimeter. NMR data were collected on a Bruker Avance neo 600 MHz, high resolution 5-mm cryoprobe spectrometer and a Wide Bore 600 MHz spectrometer, both operating at 600 MHz for ^{1}H and 150 MHz for ^{13}C, using residual solvent signal (δ_H 7.27; δ_C 77.16 ppm $CDCl_3$) as internal standard. The edited HSQC and HMBC experiments were optimized for $^{1}J_{CH}$ = 140 Hz $^{n}J_{CH}$ = 8 Hz. HRMS data were obtained using a Q Exactive Focus with electrospray ionization (ESI). Chiral analysis was performed using an Applied Biosystems 3200 QTRAP triple quad/linear trap. Circular Dichroism data were obtained on a Chirascan™ Circular Dichroism Spectrometer (Applied Photophysics, Surrey, UK), using the Pro-Data Chirascan and Pro-Data viewer software 4.7.0.

Biological Material

Samples of this *Moorena* sp. cyanobacterium, VPG16-37 (28.45 g dry weight) and VPG23-95 (70.7 g dry weight), were collected from Tumon Bay in Guam on December 3rd, 2016 and April 6th, 2023, respectively.

Extraction and Isolation

The cyanobacterium (VPG16-37) was lyophilized and extracted using EtOAc:MeOH (1:1) followed by EtOH:H_2O (1:1) to provide 1.69 g and 2.1 g of extract, respectively. The nonpolar extract was partitioned between EtOAc and H_2O, followed by partitioning the polar layer between *n*-BuOH and H_2O. The EtOAc fraction was subjected to a silica column using dichloromethane (DCM) and increasing amounts of *i*-PrOH to give ten fractions. The silica gel fraction eluting with 2% *i*-PrOH in DCM was further purified by HPLC (Synergi hydro C18 column 250×10 mm, 4.0 mL/min; UV detection at 220 and 240 nm) with a H_2O/MeCN gradient (30% for 5 min, followed by 30–100% over 30 min, then 100% for 5 min) from which the fraction at t_R 29.3 min was pure tumonolide (**1**) (7.5 mg). The silica gel fraction eluting with 1% *i*-PrOH in DCM was also purified using a H_2O/MeCN gradient (25% for 3 min, followed by 25–100% over 22 min, then 100% for 15 min) by HPLC (Phenomenex Synergi 4 µ Hydro-RP 80 Å, 250×10 mm, 4 µm; flow rate, 4.0 mL/min; UV detection at 220 nm and 240 nm), from which

the fraction at t_R 24.3 min was pure tumonolide (**1**) (8 mg). Overall yield of the compound was found as **0.03 %** for **1.**

The recollection of this cyanobacterium (VPG23-95) was lyophilized and extracted using EtOAc:MeOH (1 : 1) followed by EtOH:H_2O (1 : 1) to provide 5.5 g and 1.79 g of extract, respectively. The nonpolar extract was partitioned as above and was subjected to silica gel chromatography using an Isolera with DCM and increasing amounts of *i*-PrOH to give fourteen fractions. Using LCMS, the fraction in which **1** was detected was further purified by HPLC (Luna C18 column 250×10 mm, 4.0 mL/min; UV detection at 220 nm) with a H_2O/MeCN gradient (40 % for 5 min, followed by 40–100 % over 30 min, then 100 % for 5 min) from which the fraction at t_R 26.2 min was subjected to a second round of HPLC (Phenomenex Synergi 4 μ Hydro-RP 80 Å, 250×4.6 mm, 4 μm; flow rate, 1.0 mL/min; UV detection at 220 nm using H_2O/MeOH gradient (60 % for 5 min, followed by 60–100 % over 30 min, then 100 % for 5 min) to yield pure tumonolide aldehyde (**2**) t_R 26.2 min (1.3 mg) and **1** (35 mg). Overall yield of the compounds was found as 0.05 % of dry weight for **1** and 0.002 % of dry weight for **2.**

Tumonolide (**1**): yellow amorphous solid, $[\alpha]^{20}_D$ −50 (c 0.77, MeOH), ^{1}H NMR, ^{13}C NMR, COSY and HMBC data in $CDCl_3$, see Table 1, HRESIMS *m/z* ($[M+H]^+$580.4294, calcd. for $C_{32}H_{58}N_3O_6$, 580.4321).

Tumonolide aldehyde (**2**): white colorless oil, $[\alpha]^{20}_D$ −93 (c 0.09, MeOH), ^{1}H NMR, ^{13}C NMR, COSY and HMBC data in $CDCl_3$, see Table 1, HRESIMS *m/z* ($[M+H]^+$598.4410, calcd. for $C_{32}H_{58}N_3O_6$, 598.4431).

Acid Hydrolysis

A portion (50 μg) of **1** or **2** was subjected to acid hydrolysis (6 N HCl, 116 °C, 18 h) and then evaporated to dryness. The sample was reconstituted in 50 uL H_2O and subjected to chiral analysis (Chirobiotic TAG (4.6 mm ×250 mm), Supelco; solvent, MeOH − 10 mM NH_4OAc (40 : 60); flow rate, 0.5 mL/min; detection by ESIMS in positive mode)]. The retention times (t_R, min) were as follows *N*-Me–L-Val (10.9) and *N*-Me–L-Leu (13.2). The authentic standards eluted at *N*-Me–L-Val (10.9), *N*-Me–L-Leu (13.2), *N*-Me–D-Val (30.0) and *N*-Me–D-Leu (70.0).

Base Hydrolysis

Three different routes were utilized. First: A portion (500 μg) of **1** was dissolved in 1 mL of MeOH and treated with 1 M LiOH. The reaction was stirred at room temperature for 72 h and then neutralized using dilute HCl. Extraction with DCM gave a mixture of **1** and **2**. Second: For the Mosher's esterification, a portion (500 μg) of **1** was dissolved in 1 M NaOH or LiOH and stirred at room temperature for 24 h. It was then neutralized using dilute HCl. Extraction with DCM gave starting material only. Third: A portion (1 mg) of **1** was dissolved in 500 uL of MeOH and 50 uL of 0.1 M NaOMe were added. The reaction was stirred at room temperature for 24 h before neutralizing with dilute HCl NaOMe. Multiple peaks were observed by LCMS indicating diastereomers.

Computational Methods

The following computational chemistry software was used: Biovia Insight II/Discover to build the models and perform molecular dynamics (MD) simulations, Pcmodel (v. 10.075000) for molecular mechanics optimizations, Gaussian 16 (Revision C.01) for quantum mechanical calculations, OpenBabel (v. 3.0.0), GaussView (v. 6.0.16), and VMD (v. 1.9.3) for format conversion, data analysis, and visualization. Docking and molecular dynamics of the TACR2-Tumonolide system used the Schrödinger Suite (v. 2023–3).

Random conformers of the two alternative diastereomers of tumonolide (**1**), i.e. 2*S*,9*S*,23*R*,25*R*,27*R* (hereafter *RRR*-**1**) and 2*S*,9*S*,23*S*,25*S*,27*S* (hereafter *SSS*-**1**) were built and used as starting structures for conformational search. Conformational search was performed using molecular dynamics (MD) in the cvff force field. MD runs were performed setting the temperature at 600 K to observe possible slow conformational changes in the short duration of the simulation. Each MD run lasted 250 ns and coordinates were saved every 50 ps, yielding 5000 conformers per run. The conformers were optimized in the cvff and then in the MMFF force field, and duplicate conformers (RMSD $<$ 0.1 Å) were removed.

Because the conformations of the three amide bonds (*E*/*cis* for C8-N and C-21/N, *Z*/*trans* for C14-N) and the conformation of the C21-C28 segment ($gauche_{C22\text{-}C23}$/$trans_{C23\text{-}C24}$/$trans_{C24\text{-}C25}$/$trans_{C25\text{-}C26}$/$gauche_{C26\text{-}C27}$) were known from NMR studies, the relevant dihedral angles of all conformers were measured, and only the conformers that matched these conformations (i.e., those with $|\varphi_{C2\text{-}N\text{-}C8\text{-}C9}| < 30°$, $|\varphi_{C9\text{-}N\text{-}C14\text{-}C15}| > 150°$, $|\varphi_{C19\text{-}N\text{-}C21\text{-}C22}| < 30°$, $|\varphi_{C21\text{-}C22\text{-}C23\text{-}C24}| < 90°$, $|\varphi_{C22\text{-}C23\text{-}C24\text{-}C25}| > 150°$, $|\varphi_{C23\text{-}C24\text{-}C25\text{-}C26}| > 150°$, $|\varphi_{C24\text{-}C25\text{-}C26\text{-}C27}| > 150°$, and $|\varphi_{C25\text{-}C26\text{-}C27\text{-}C28}| < 90°$) were selected for subsequent analysis.

These "NMR-compliant" conformers were used as starting structure for density functional theory (DFT) calculations. The geometry of each conformer was optimized at the B3LYP/6-31G(d) level of theory; the energies calculated at the B3LYP/6-311G + (d,p) level of theory with the SMD model for the solvent, $CHCl_3$, were used to calculate the population of each optimized conformer with the Boltzmann distribution law at 298 K. The full protocol from the 250-nm MD run to DFT optimization was repeated until no new conformers with population $>$ 1 % were found, which required 4 runs both for *RRR*-**1** and *SSS*-**1**. Overall, the conformational searches identified 9 significantly populated conformers for *RRR*-**1** and 7 significantly populated conformers for *SSS*-**1**, which were used for NMR and UV/ECD calculations. The results of the conformational searches are summarized in Table S1, and the Cartesian coordinates of the optimized conformers of *RRR*-**1** and *SSS*-**1** are reported, respectively, in Tables S4 and S5.

NMR calculations were performed at two levels of theory, mPW1PW91/6-311 + G(d,p) using PCM solvent model for application of DP4+ and mPW1PW91/6-311 + G(2d,p) using SMD solvent model for evaluation of chemical shift deviations and calculation of ^{1}H and ^{13}C RMSD.[43] The isotropic shielding calculated for each nucleus were averaged over the conformers according to their respective populations, and average isotropic shielding σ were used directly for application of the DP4+ method, or converted into chemical shifts using the pre-calculated scaling factors proposed for the level of theory used, i.e., $\delta = (186.5242 - \sigma)/1.0533$ for ^{13}C and $\delta = (31.8018 - \sigma)/1.0936$ for ^{1}H.[44] ^{1}H chemical shift of methyl protons were calculated as the average of chemical shifts of the three protons.

When evaluating chemical shift deviations or applying the DP4+ method, it is important to assign the correct ^{1}H chemical shift value to the pro*R* and pro*S* protons of a diastereotopic pair. For diastereotopic protons at C-22, C-24, and C-26 stereospecific assignment was obtained experimentally as a consequence of the application of the Murata method. For diastereotopic methylene protons at C-3, C-15, C-16, C-17, for which experimental stereospecific assignment was not possible, the assignment was chosen that gave better agreement with the predicted chemical shifts (therefore, the assignment of experimental chemical shifts used for *RRR*-**1** was different from that used for *SSS*-**1**, see Table S4). The same best-fit approach was used to assign experimental chemical

shifts to the pro*R* and pro*S* geminal methyl pairs C-5/C-6, C-12/C-13, and C-29/C-30. Average isotropic shieldings, predicted chemical shifts, and stereospecific assignment of experimental chemical shifts for *RRR*-**1** and *SSS*-**1** are reported in Tables S2 and S3, and the detailed results of DP4+ calculations are reported in Table S4 and Figure S22.

The UV and ECD spectra of *RRR*-**1** and *SSS*-**1** were predicted using Time Dependent DFT (TDDFT) at the B3LYP/TZVP/PCM(MeOH) level of theory. The UV and ECD spectra of each conformer were generated using the program SpecDis setting the σ parameter (half the bandwidth at 1/*e* peak height) at 0.45 eV.[51] Average UV and ECD spectra of *RRR*-**1** and *SSS*-**1** were obtained from them using Boltzmann statistics (Figure S24). The predicted UV and ECD spectra were then shifted by −5 nm to match the predicted (237 nm) and experimental (232 nm) UV maxima and scaled with scaling factors 1.2 (UV) and 2.9 (ECD) to obtain the best match with experimental data (Figure 5).

Molecular Docking

The cryo-EM structure of TACR2 in complex with Neurokinin-A obtained at 2.70 Å resolution was retrieved from the Protein Data Bank (PDB ID 7XWO) and prepared for docking with Schrödinger Protein Preparation Workflow, considering a pH of 7.4.[50,52] The membrane position was obtained from the Orientation of Proteins in Membranes (OPM) database.[53] The ligand (**1**) was built with Maestro 2D Sketcher and prepared for docking with LigPrep. The ligand was docked into TACR2 with Schrödinger GlideSP with enhanced sampling and with and without Prime Macrocycle Sampling.[54–56] The poses obtained were rescored with MM-GBSA optimization using the OPLS4 force field and VSGB solvation model, where residues within 5 Å of any ligand atom were allowed to move.[57–59] The most likely binding pose was selected based on the MM-GBSA ΔG_{bind} energy.

To allow for an appropriate comparison of the tumonolide docking mode in the TACR2 pocket, a different procedure was applied to generate the structures of tumonolide docked to TACR1 and TACR3.

The coordinates of the targets were obtained from the Protein Data Bank with PDBIDs 6HLP (TACR1) and 8JBG (TACR3) and aligned with the structure for TACR2 obtaied after the MM-GBSA optimization as described in the main text. The solvent, lipids, and crystallization molecules were removed, and the protein was pre-prepared for docking using Maestro Protein Preparation Protocol at pH 7.4, in the presence of the co-crystalized ligands. The ligands were then removed, and the tumonolide ligand from the TACR2 structure copied to the TACR1 and TACR3 structures. The new complexes were refined with Prime Protein-Ligand refinement where all residues within 5 Å of the ligand were allowed to move. Next, a search grid for docking was created centered on tumonolide, and the ligand redocked into the grid with constraints to the refined structure. Finally, the complexes obtained were optimized with MM-GBSA, where all residues with atoms with-in 5 Å of tumonolide were allowed to move.

Molecular Dynamics

The final pose obtained from the MM-GBSA optimization of docking poses was used as initial inouut structures for the MD simulations with Desmond (v. 7.6.132), included in the Schrödinger Suite.[60,61] The system's initial configuration was embedded in a POPC bilayer with SPC water molecules and simulated for 10 ns with the OPLS4 force field in the NPγT ensemble at 300 K and 1 atm.[62] Before simulation, the system was allowed to relax using the standard Schrödinger protocol for membrane relaxation.

GPCR screening

GpcrMAX panel biosensor assays were used to characterize tumonolide (1, 20 μM) (agonist and antagonist mode). The experiments were carried out by DiscoveRx Corporation (Fremont, CA, USA), using PathHunter β-arrestin enzyme fragment complementation (EFC) technology.

Dose response curves were obtained by subjecting **1** to a secondary screening in antagonist mode using the three initial targets (TACR1, TACR2, TACR3, CNR2, NPSR1b). CHO–K1 cells expressing the TACR1, TACR2, TACR3 and CNR2 receptors, and U2OS cells expressing NPSR1b were incubated with **1** for 30 min. The assays were performed at 10-point concentrations using 3-fold serial dilutions in duplicate, where the highest concentration was 100 μM for TACR1 and TACR3 and 50 μM for the remaining.

GPCR Binding Assay

Cell membrane homogenates (8 μg protein) are incubated for 60 min at 22 °C with 0.1 nM [^{125}I]-NKA in the absence or presence of the test compound in a buffer containing 20 mM Hepes/NaOH (pH 7.4), 1 mM MnCl2 and 0.1% BSA. Nonspecific binding is determined in the presence of 0.3 μM [Nle^{10}]-NKA(4–10). Following incubation, the samples are filtered rapidly under vacuum through glass fiber filters (GF/B, Packard) presoaked with 0.3% PEI and rinsed several times with ice-cold 50 mM Tris-HCl using a 96-sample cell harvester (Unifilter, Packard). The filters are dried then counted for radioactivity in a scintillation counter (Topcount, Packard) using a scintillation cocktail (Microscint 0, Packard). The results are expressed as a percent inhibition of the control radioligand specific binding. The standard reference compound is [Nle10]-NKA(4–10), which is tested in each experiment at several concentrations to obtain a competition curve from which its IC50 is calculated.[63]

Cell Viability Assays

HCT116 (American Type Culture Collection, ATCC), RAW264.7 (ATCC) and HEK293-ARE (Signosis) cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum at 37 °C humidified air and 5% CO_2. Cells were seeded (9,000/well (HCT116), 20,000/well (HCT11610) and 10,000/well (HEK293)) in 96-well plates and left to attach overnight. The cells were then treated with compound or vehicle control for 24 h (HEK293 and RAW264.7) or 48 h (HCT116), followed by addition of MTT dye, according to the manufacturer's protocol (Promega, USA). IC_{50} values were calculated using GraphPad prism software.

RNA Extraction

HCT116 cells (ATCC) were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum at 37 °C humidified air and 5% CO_2. Cells were seeded (4×10 5/well) in 6-well plates and left to attach overnight. The cells were then treated with compound or vehicle control for 12 h. Total RNA was extracted using the RNAeasy Mini Kit l according to manufacturer's protocol (Qiagen, Germany).

RNA Extraction and Illumina Sequencing Library

Sample quality was assessed using the Agilent TapeStation 4200 (Agilent Technologies, Inc). 250 ng of total RNA was used for library construction using NEB Ultra II Directional RNA-seq library preparation kit for Illumina according to manufacturer's protocol. First, 250 ng of total RNA was used for mRNA isolated using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs, catalog # E7490). Then followed by RNA library construction with the NEBNext Ultra II Directional RNA Library Prep Kit (New England Biolabs, catalog #E7760) according to the manufacturer's user guide. Briefly, RNA was fragmented in NEBNext First Strand Synthesis Buffer via incubation at 94 °C for the desired time. This step was followed by first-strand cDNA synthesis using reverse transcriptase and random hexamer primer. Synthesis of ds-cDNA was performed using the 2nd strand master mix provided in the kit, followed by end-repair and adaptor ligation. At this point, Illumina adaptors were ligated to the sample. Finally, each library (uniquely barcoded) was enriched by 11 cycles of amplification, and purified with Agencourt AMPure beads (Beckman Coulter, catalog # A63881). 18 barcoded libraries were sized on the Bioanalyzer and quantified with the Qubit® 2.0 Fluorometer. Finally, these 18 individual libraries were pooled in equimolar concentration. RNASeq libraries were constructed at the UF ICBR Gene Expression Core (https://biotech.ufl.edu/gene-expression-genotyping/, RRID: SCR_019145). The Illumina NovaSeq 6000 was used to sequence the libraries for 2×150 cycles.

Illumina NovaSeq6000 Sequencing

Sequencing was performed at the ICBR NextGen Sequencing (https://biotech.ufl.edu/next-gen-dna/, RRID:SCR_019152). Normalized libraries were submitted to the "Free Adapter Blocking Reagent" protocol (FAB, Cat# 20024145) in order to minimize the presence of adaptor-dimers and index hopping rates. The library pool was diluted to 0.8 nM and sequenced on one S4 flow cell lane (2x150 cycles) of the Illumina NovaSeq6000. The instrument's computer utilized the NovaSeq Control Software v1.6. Cluster and SBS consumables were v1.5. The final loading concentration of the library was 120 pM with 1% PhiX spike-in control. One lane generated 2.5–3 billion paired-end reads (~950Gb) with an average Q30% > = 92.5% and Cluster PF = 85.4%. FastQ files were generated using the BCL2fastQ function in the Illumina BaseSpace portal.

Short reads were trimmed using trimmomatic (v 0.36), and QC on the original and trimmed reads was performed using FastQC (v 0.11.4) and MultiQC.[64–66] The reads were aligned to the transcriptome using STAR version 2.7.9a.[67] Transcript abundance was quantified using RSEM (RSEM v1.3.1).[68] Differential expression analysis was performed using DESeq2, with an FDR-corrected p-value threshold of 0.05.[69] The output files were further filtered to extract transcripts showing a 2.0-fold change in either direction. Results were reported for protein-coding genes only, and for all transcript types. The data were deposited with the GEO accession number GSE268527.

Ingenuity Pathway Analysis

The RNA-seq data was analyzed using Ingenuity Pathway Analysis (IPA) (Qiagen, Germany). For the RNA-seq analysis the samples were compared to the DMSO control. Using the IPA software, the analysis was done using a cutoff of 1.5-fold change in the gene expression and a p value < 0.05.

Supporting Information

The following supporting information can be downloaded at xxx Figure S1: Structures of selective TACR2 inhibitors; FigureS2–20: 1D/2D NMR data of tumonolide and tumonolide aldehyde in $CDCl_3$; Figure S21–22: Bioassay data of tumonolide; Figure S23–25: Computational data of tumonolide for the stereochemical evaluation; Figure S26: Interactions between TACR2 and NKA; Figure S27: RMSD data; Figures S28–29: Binding of tumonolide to TACR1 and TACR3; Table S1: Conformational search results; Table S2–3: Predicted NMR data; Table S4: Input data for DP4+ model; Table S5–6; Optimized Cartesian coordinated for *RRR*-1 and *SSS*-1.

Acknowledgements

This research was supported by the National Institutes of Health (NIH), grant RM1GM145426, and the Debbie and Sylvia DeSantis Chair Professorship (H.L.). We would like to thank Jim Rocca for assisting with the 2D NMR spectra, Jennifer Sneed and Jay Houk for assistance with sample collection and Jay Houk for assistance with the extraction of the 2016 collection. We thank the Guam Department of Agriculture Division of Aquatic and Wildlife Resources for collection permits. We thank Yanping Zhang from the Gene Expression and Genotyping Core and Alberto Riva from the Bioinformatics Core of the Interdisciplinary Center for Biotechnology Research for RNA sequencing and assistance with data analysis. A portion of this work was performed in the McKnight Brain Institute at the National High Magnetic Field Laboratory's Advanced Magnetic Resonance Imaging and Spectroscopy (AMRIS) Facility, which is supported by National Science Foundation Cooperative Agreement DMR-1644779 and the State of Florida.

Conflict of Interests

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: GPCR · TACR2 · marine natural products · macrocycle · configurational analysis

[1] S. B. Singh, *J. Nat. Prod.* **2022**, *85*, 666–687.
[2] A. R. Carroll, B. R. Copp, R. A. Davis, R. A. Keyzers, M. R. Prinsep, *Nat. Prod. Rep.* **2023**, *40*, 275–325.
[3] J. Demay, C. Bernard, A. Reinhardt, B. Marie, *Mar. Drugs* **2019**, *17*, 1–49.
[4] D. Klein, J. C. Braekman, D. Daloze, L. Hoffmann, G. Castillo, V. Demoulin, *J. Nat. Prod.* **1999**, *62*, 934–936.
[5] D. Klein, J. C. Braekman, D. Daloze, L. Hoffmann, V. Demoulin, *Tetrahedron Lett.* **1996**, *37*, 7519–7520.

[6] S. Matthew, L. A. Salvador, P. J. Schupp, V. J. Paul, H. Luesch, *J. Nat. Prod.* **2010**, *73*, 1544–1552.
[7] A. R. Pereira, Z. Cao, N. Engene, I. E. Soria-Mercado, T. F. Murray, W. H. Gerwick, *Org. Lett.* **2010**, *12*, 4490–4493.
[8] Y. Liu, B. K. Law, H. Luesch, *Mol. Pharmacol.* **2009**, *76*, 91–104.
[9] A. O. Paatero, J. Kellosalo, B. M. Dunyak, J. Almaliti, J. E. Gestwicki, W. H. Gerwick, J. Taunton, V. O. Paavilainen, *Cell Chem. Biol.* **2016**, *23*, 561–566.
[10] H. Luesch, W. Y. Yoshida, R. E. Moore, V. J. Paul, T. H. Corbett, *J. Am. Chem. Soc.* **2001**, *123*, 5418–5423.
[11] H. Luesch, W. Y. Yoshida, R. E. Moore, V. J. Paul, *Bioorg. Med. Chem.* **2002**, *10*, 1973–1978.
[12] P. Wu, W. Cai, Q. Y. Chen, S. Xu, R. Yin, Y. Li, W. Zhang, H. Luesch, *Org. Lett.* **2016**, *18*, 5400–5403.
[13] Y. Masuda, J. Suzuki, Y. Onda, Y. Fujino, M. Yoshida, T. Doi, *J. Org. Chem.* **2014**, *79*, 8000–8009.
[14] T. Doi, Y. Masuda, M. Yoshida, *Yuki Gosei Kagaku Kyokaishi* **2018**, *76*, 1170–1175.
[15] Q. Y. Chen, Y. Liu, H. Luesch, *ACS Med. Chem. Lett.* **2011**, *2*, 861–865.
[16] T. Doi, *Chem. Pharm. Bull.* **2014**, *62*, 735–743.
[17] T. Doi, Y. Numajiri, T. Takahashi, M. Takagi, K. Shin-Ya, *Chem. Asian J.* **2011**, *6*, 180–188.
[18] H. Luesch, W. Y. Yoshida, R. E. Moore, V. J. Paul, *Tetrahedron* **2002**, *58*, 7959–7966.
[19] S. C. Philkhana, S. Mehrotra, T. F. Murray, D. S. Reddy, *Org. Biomol. Chem.* **2016**, *14*, 8457–8473.
[20] D. M. Rosenbaum, S. G. F. Rasmussen, B. K. Kobilka, *Nature* **2009**, *459*, 356–363.
[21] K. Sriram, P. A. Insel, *Mol. Pharmacol.* **2018**, *93*, 251–258.
[22] P. Insel, A. Snead, F. Murray, L. Zhang, H. Yokouchi, T. Katakia, O. Kwon, D. Dimucci, A. Wilderman, *Br. J. Pharmacol.* **2012**, *165*, 1613–1616.
[23] E. Muratspahić, M. Freissmuth, C. W. Gruber, *Trends Pharmacol. Sci.* **2019**, *40*, 309–326.
[24] F. H. Al-Awadhi, B. Gao, M. A. Rezaei, J. C. Kwan, C. Li, T. Ye, V. J. Paul, H. Luesch, *J. Med. Chem.* **2018**, *61*, 6364–6378.
[25] V. Erspamer, *Trends Neurosci.* **1981**, *4*, 267–269.
[26] M. S. Steinhoff, B. von Mentzer, P. Geppetti, C. Pothoulakis, N. W. Bunnett, *Physiol. Rev.* **2014**, *94*, 265–301.
[27] G. A. Gaudreau, V. Plourde, *Neurosci. Lett.* **2003**, *351*, 59–62.
[28] J. N. Pennefather, A. Lecci, M. L. Candenas, E. Patak, F. M. Pinto, C. A. Maggi, *Life Sci.* **2004**, *74*, 1445–1463.
[29] A. Lecci, A. Capriati, M. Altamura, C. A. Maggi, *Auton. Neurosci.* **2006**, *126–127*, 232–249.
[30] W. Jianfeng, W. Yutao, B. Jianbin, *Cancer Cell Int.* **2021**, *21*, 1–12.
[31] L. Marson, R. Piatt, J. B. Cook, K. B. Thor, *Spinal Cord* **2021**, *59*, 1013–1017.
[32] Y. Shimizu, H. Matsuyama, T. Shiina, T. Takewaki, J. B. Furness, *Cell. Mol. Life Sci.* **2008**, *65*, 295–311.
[33] A. Lecci, A. Capriati, C. A. Maggi, *Br. J. Pharmacol.* **2004**, *141*, 1249–1263.
[34] M. Lördal, G. Navalesi, E. Theodorsson, C. A. Maggi, P. M. Hellström, *Br. J. Pharmacol.* **2001**, *134*, 215–223.
[35] C. Bassarello, G. Bifulco, A. Zampella, M. V. D'Auria, R. Riccio, L. Gomez-Paloma, *Eur. J. Org. Chem.* **2001**, 39–44.
[36] N. Matsumori, D. Kaneno, M. Murata, H. Nakamura, K. Tachibana, *J. Org. Chem.* **1999**, *64*, 866–876.
[37] A. Gutiérrez-Cepeda, A. Daranas, J. Fernández, M. Norte, M. Souto, *Mar. Drugs* **2014**, *12*, 4031–4044.
[38] L. Grauso, Y. Li, S. Scarpato, N. A. Cacciola, P. De Cicco, C. Zidorn, A. Mangoni, *J. Nat. Prod.* **2022**, *85*, 2468–2473.
[39] T. A. Halgren, *J. Comput. Chem.* **1999**, *20*, 730–748.
[40] L. Grauso, R. Teta, G. Esposito, M. Menna, A. Mangoni, *Nat. Prod. Rep.* **2019**, *36*, 1005–1030.
[41] Gaussian 16, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.
[42] A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B* **2009**, *113*, 6378–6396.
[43] N. Grimblat, M. M. Zanardi, A. M. Sarotti, *J. Org. Chem.* **2015**, *80*, 12526–12534.
[44] M. W. Lodewyk, M. R. Siebert, D. J. Tantillo, *Chem. Rev.* **2012**, *112*, 1839–1862.
[45] Y. Takeda, P. Blount, B. S. Sachais, A. D. Hershey, R. Raddatz, J. E. Krause, *J. Neurochem.* **1992**, *59*, 740–745.
[46] L. P. Nguyen, M. Cho, T. U. Nguyen, H. K. Park, H. T. Nguyen, K. Mykhailova, S. Hurh, H. R. Kim, J. Y. Seong, C. S. Lee, B. J. Ham, J. I. Hwang, *Cell Biosci.* **2023**, *13*, 212.
[47] J. L. Redburn, J. D. Leah, *Neuropeptides* **1999**, *33*, 115–119.
[48] S. Kiss, M. Yoshiyama, Y. Q. Cao, A. I. Basbaum, W. C. De Groat, A. Lecci, C. A. Maggi, L. A. Birder, *Neurosci. Lett.* **2001**, *313*, 57–60.
[49] M. Jocič, R. Schuligoi, E. Schöninkle, M. A. Pabst, P. Holzer, *Pain* **2001**, *89*, 147–157.
[50] W. Sun, Q. Yuan, H. Zhang, F. Yang, S. Ling, Y. Luo, P. Lv, H. Eric Xu, C. Tian, W. Yin, P. Shi, *Cell Discov.* **2022**, *8*, 72.
[51] T. Bruhn, A. Schaumlöffel, Y. Hemberger, G. Pescitelli, *SpecDis version 1.71* **2017**.
[52] H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J. D. Westbrook, C. Zardecki, *Acta Crystallogr. Sect. D* **2002**, *58*, 899–907.
[53] M. A. Lomize, I. D. Pogozheva, H. Joo, H. I. Mosberg, A. L. Lomize, *Nucleic Acids Res.* **2012**, *40*, D370–6.
[54] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, *47*, 1739–1749.
[55] T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J. Med. Chem.* **2004**, *47*, 1750–1759.
[56] D. Sindhikara, S. A. Spronk, T. Day, K. Borrelli, D. L. Cheney, S. L. Posy, *J. Chem. Inf. Model.* **2017**, *57*, 1881–1894.
[57] G. Weng, E. Wang, F. Chen, H. Sun, Z. Wang, T. Hou, *Phys. Chem. Chem. Phys.* **2019**, *21*, 10135–10145.
[58] C. Lu, C. Wu, D. Ghoreishi, W. Chen, L. Wang, W. Damm, G. A. Ross, M. K. Dahlgren, E. Russell, C. D. Von Bargen, R. Abel, R. A. Friesner, E. D. Harder, *J. Chem. Theory Comput.* **2021**, *17*, 4291–4300.
[59] J. Li, R. Abel, K. Zhu, Y. Cao, S. Zhao, R. A. Friesner, *Proteins Struct. Funct. Bioinf.* **2011**, *79*, 2794–2812.
[60] K. J. Bowers, D. E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan, D. E. Shaw, in *ACM/IEEE SC 2006 Conf.*, IEEE, **2006**, pp. 43–43.
[61] D. E. Shaw, "Schrödinger Release 2024–1: Desmond Molecular Dynamics System," **2024**.
[62] P. Mark, L. Nilsson, *J. Phys. Chem. A* **2001**, *105*, 9954–9960.
[63] D. Aharony, J. Little, S. Powell, B. Hopkins, K. R. Bundell, W. L. McPheat, R. D. Gordon, G. Hassall, R. Hockney, R. Griffin, *Mol. Pharmacol.* **1993**, *44*, 356–63.
[64] A. M. Bolger, M. Lohse, B. Usadel, *Bioinformatics* **2014**, *30*, 2114–20.
[65] S. Andrews, **2010**. Can be found under http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
[66] P. Ewels, M. Magnusson, S. Lundin, M. Käller, *Bioinformatics* **2016**, *32*, 3047–8.
[67] A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, *Bioinformatics* **2013**, *29*, 15–21.
[68] B. Li, C. N. Dewey, *BMC Bioinf.* **2011**, *12*, 323.
[69] M. I. Love, W. Huber, S. Anders, *Genome Biol.* **2014**, *15*, 550.

Version of record online: July 18, 2024