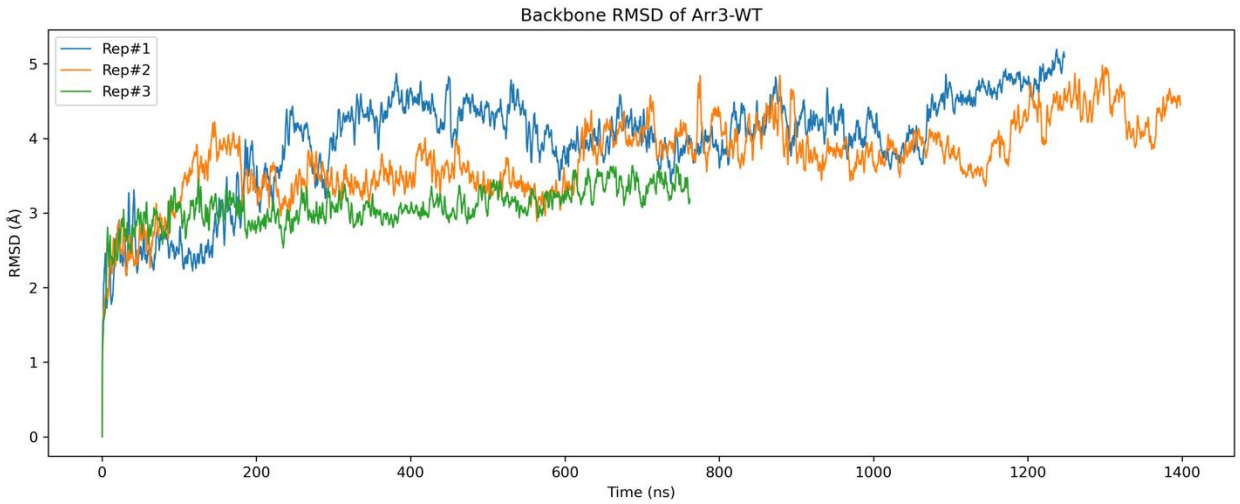
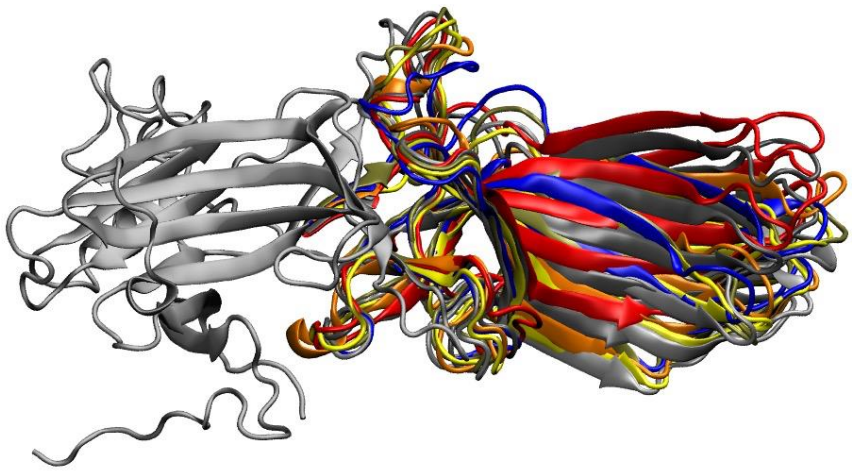
Supplementary Information

A small molecule enhances arrestin-3 binding to the β_2-adrenergic receptor

Han Kurt[Ŧ1,10], Ali Akyol[Ŧ3], Cagdas Devrim Son[3], Chen Zheng [7], Irene Gado[8], Massimiliano Meli[9], Erica Elisa Ferrandi[9], Ivan Bassanini[9], Francesca Vasile[8], Vsevolod V. Gurevich[7], Aylin Nebol[4,5,6] , Esra Cagavi[4,5,6], Giulia Morra*[9] and Ozge Sensoy*[2,4]



Supplementary Fig. 1. Backbone RMSD plots calculated from MD trajectories of basal arrestin-3 in water.



Supplementary Fig. 2. Selected conformations from arrestin-3 MD trajectories. The structures

are aligned with respect to the N domain of the protein shown in gray. The C domain in different conformations are shown in different colors.

Supplementary Table 1. Structures of arrestin subtypes in different states

PDB ID	Gate Loop Distance (Å)	Back Loop Distance (Å)	Status
7DF9	11.6	11.8	Receptor-bound (V2Rpp-1-Fab30)
2WTR	10.4	6.4	Inactive
1G4M	10.3	7.6	Inactive
1G4R	10.2	7.6	Inactive
3GC3	10.3	7.3	Inactive
5W0P	12.0	12.3	Receptor-bound (Rhodopsin)
3GD1	10.4	7.3	Inactive
1SUJ	10.2	7.6	Inactive
1CF1	10.3	7.7	Inactive
4ZRG	10.2	7.8	Pre-active (polar-core mutant, R175E)
3UGU	10.1	7.9	Pre-active (C-tail truncated)
1AYR	10.1	7.6	Inactive
4J2Q	12.0	12.9	Pre-activate (C-tail truncated)
3UGX	10.4	7.3	Inactive
3P2D	10.0	7.1	Inactive
1ZSH	10.2	7.6	Inactive
1JSY	10.2	7.6	Inactive
4ZWJ	11.9	12.7	Receptor-bound (Rhodopsin)
6U1N	12.1	8.1	Receptor-bound(M2 muscarinic receptor)
6TKO	13.3	12.2	Receptor-bound (beta 1 adrenoreceptor)
6K3F	16.4	12.1	Receptor-bound (CXCR7 phosphopeptide)
7JTB	10.4	7.6	Inactive
7F1W	10.3	7.7	Inactive
5TV1	14.0	9.4	Pre-active (inositol hexakiphosphate)

Supplementary Table 2. Structural parameters used to describe the conformational states of *the back loop* and *gate loop* as well as inter-domain rotation angle

Conformation #	Back loop (Å)	Gate Loop (GL) (Å)	Conformational States of SH and GL	Rotation angle (^{0})
1	4.2	10.5	Basal, pre-active	-0.5
2	6.9	10.5	Basal, pre-active	9.8
3	8.5	8.3	Receptor-bound, basal	9.6

4	5.9	11.2	Basal, pre-active	-10
5	10.3	9.9	Receptor-bound, basal	-6.7
6	11.5	10.5	Receptor-bound, pre-active	-6.3
7	10.9	10.6	Receptor-bound, pre-active	-6.2

Supplementary Table. 3. Properties of the pockets identified by SiteMap for selected conformations

1st conformation	S.Score	size	Dscore	volume	exposure	enclosure	contact	phobic	philic	balance	don/acc	
site_1	1.024	179	1.063	456.880	596	691	879	726	895	811	816	gate loop (291-296)
site_2	950	204	960	534.050	700	623	795	240	1.086	221	946	
site_3	1.019	134	1.009	534.390	666	726	960	313	1.125	279	747	
site_4	979	112	998	339.570	650	667	920	508	1.049	484	1.356	
site_5	943	92	868	193.110	598	657	924	52	1.313	40	678	
site_6	955	78	844	225.690	581	759	1.002	273	1.355	201	1.187	
site_7	885	71	875	206.490	608	693	836	810	1.022	792	492	
site_8	1.063	73	796	121.080	411	952	1.416	58	1.775	33	604	
site_9	887	76	923	217.800	740	592	714	621	770	806	904	
site_10	676	45	575	75.800	637	574	878	188	1.225	153	1.436	

2nd conformation	S.Score	size	Dscore	volume	exposure	enclosure	contact	phobic	philic	balance	don/acc	
site_1	1.009	334	945	737.790	528	711	888	169	1.293	131	920	helix (313-317)
site_2	1.027	240	995	656.840	527	738	998	515	1.190	433	592	
site_3	1.053	153	1.038	278.170	487	777	1.036	492	1.132	434	809	
site_4	855	73	819	150.920	594	641	806	90	1.132	79	815	
site_5	1.036	108	880	286.750	603	752	1.009	171	1.567	109	507	
site_6	810	58	792	134.800	570	657	828	601	977	615	759	
site_7	774	45	734	128.970	587	700	893	395	970	407	1.478	
site_8	789	51	651	150.580	619	697	819	133	1.344	99	212	
site_9	806	45	793	156.750	648	691	945	1.297	772	1.679	1.007	
site_10	599	26	511	49.730	634	636	906	283	1.073	264	1.366	

3rd conformation	S.Score	size	Dscore	volume	exposure	enclosure	contact	phobic	philic	balance	don/acc	
site_1	998	317	951	928.840	631	696	888	274	1.247	220	614	helix (313-317)
site_2	984	126	1.006	294.980	573	671	850	610	1.031	592	933	
site_3	1.014	106	1.019	381.420	619	718	930	437	1.080	405	882	
site_4	1.036	71	770	128.970	398	925	1.373	25	1.770	14	921	
site_5	738	54	717	142.690	667	585	771	446	997	448	1.134	
site_6	800	66	773	117.650	647	605	790	286	1.086	264	836	
site_7	766	48	723	88.150	442	684	926	355	1.038	342	772	
site_8	737	46	623	115.590	652	658	826	305	1.256	243	616	
site_9	639	31	615	72.370	726	561	714	519	752	690	1.161	
site_10	698	28	656	97.070	627	686	908	1.061	775	1.369	79	

4th conformation	S.Score	size	Dscore	volume	exposure	enclosure	contact	phobic	philic	balance	don/acc	
site_1	1.032	146	924	282.970	484	746	1.001	156	1.420	110	440	helix (313-317)
site_2	976	169	986	473.680	654	662	826	395	1.076	367	887	
site_3	1.004	139	920	449.670	621	704	885	200	1.356	148	824	
site_4	961	107	946	227.070	587	639	833	198	1.157	172	1.200	
site_5	898	84	899	263.770	729	637	824	332	1.056	314	775	gate loop (291-296)
site_6	858	63	850	150.920	611	682	907	573	938	611	1.791	
site_7	724	59	693	137.540	728	539	731	94	1.080	87	1.879	
site_8	712	38	624	71.690	591	688	922	351	1.130	310	1.314	
site_9	908	43	541	83.350	482	937	1.408	20	1.956	10	344	
site_10	656	41	528	99.130	721	578	772	0	1.290	0	1.393	

5th conformation	S.Score	size	Dscore	volume	exposure	enclosure	contact	phobic	philic	balance	don/acc	
site_1	994	220	997	471.620	540	689	916	364	1.095	333	1.000	
site_2	1.046	232	1.030	602.990	536	766	1.026	427	1.134	376	784	helix (313-317)
site_3	878	77	807	197.910	560	650	954	183	1.250	146	1.334	
site_4	887	68	781	199.630	616	721	948	282	1.308	215	807	
site_5	788	51	736	83.350	510	694	920	475	1.078	441	657	
site_6	837	58	720	188.990	618	715	853	66	1.305	51	1.034	
site_7	691	35	648	117.990	741	646	768	342	908	376	403	

site_8	849	28	340	45.280	391	988	1.522	0	2.303	0	454
site_9	725	39	511	106.330	705	699	1.005	18	1.521	12	860
site_10	659	22	614	112.850	758	675	758	702	711	987	659

6th conformation	S.Score	size	Dscore	volume	exposure	enclosure	contact	phobic	philic	balance	don/acc	
site_1	1.028	175	1.042	432.870	457	739	959	561	1.049	535	564	helix (313-317)
site_2	978	186	976	490.490	658	666	791	262	1.113	236	729	
site_3	1.020	96	924	272.340	522	750	981	310	1.372	226	556	gate loop (291-296)
site_4	975	81	863	197.570	460	770	1.096	492	1.364	361	552	
site_5	1.012	60	639	98.780	450	963	1.393	83	2.046	41	520	
site_6	756	53	697	122.790	616	631	904	155	1.124	138	1.001	
site_7	766	45	705	158.470	591	709	1.030	399	1.076	370	566	
site_8	687	30	636	112.500	651	680	833	343	880	389	428	
site_9	656	28	623	82.320	702	617	813	683	750	911	1.102	
site_10	753	36	757	122.450	692	633	836	1.088	574	1.894	938	

7th conformation	S.Score	size	Dscore	volume	exposure	enclosure	contact	phobic	philic	balance	don/acc	
site_1	1.026	213	1.050	508.670	532	722	935	506	993	509	1.091	
site_2	1.034	141	947	292.240	536	750	1.016	214	1.359	158	522	helix (313-317)
site_3	983	121	949	348.140	630	673	839	138	1.212	114	646	
site_4	976	82	968	217.800	457	766	1.040	758	1.051	722	632	
site_5	907	74	927	262.390	719	654	775	385	840	459	1.255	
site_6	717	60	691	114.220	725	522	669	82	1.075	76	1.926	
site_7	764	51	571	85.750	568	658	930	0	1.520	0	947	
site_8	752	40	640	68.260	420	731	1.142	541	1.207	448	408	
site_9	711	33	675	126.570	708	666	816	533	814	655	1.311	
site_10	627	32	573	117.990	761	597	779	335	973	345	1.621	

Supplementary Table 4. Druggability scores and ranking of selected pockets using *SiteMap*[1], *fpocket*[2], and *deepsite*[3]

Conformation#	Dscore	Ranking	Ranking GL/SH fpocket	Ranking GL/SH deepsite
1(gate loop)	1.06	1/10	37/26	2/1

2(Back Loop)	0.95	1/10	5/1	1/3
3(Back Loop)	0.95	1/10	32/5	2/3
4(Back Loop)	0.92	1/10	13/2	2/1
5(Back Loop)	1.03	2/10	21/31	2/1
6(Back Loop)	1.04	1/10	34/32	2/1
7(Back Loop)	0.95	2/10	30/4	3/1

The ranking of the top SiteMap pockets was also evaluated using two independent software tools developed for drug discovery purposes. As shown in Table S2, the SH (gate loop) pocket has a high ranking in most frames. Also, it almost always has greater druggability than the GL site.

Supplementary Table 5. Properties of pharmacophore groups for selected conformations

Rank	*Feature_label*	*Score*	*X*	*Y*	*Z*	*type*	*num*	*source*
1	A20	-2	-9.59	6.2959	3.5011	A	19	HBond+PhobEnHB
2	A414	-1.7	-7.9899	13.4701	0.3003	A	413	HBond+PhobEnHB
3	A552	-1.7	-7.214	6.8379	-0.4985	A	551	HBond+PhobEnPairHB
4	R2186	-1.66	-1.9276	7.2966	-1.6841	R	2185	RingChemscoreHphobe
5	N1588	-1.25	-9.8635	8.4155	3.2994	N	1587	HBond
6	R1941	-1.07	-4.6161	8.3876	1.1707	R	1940	RingChemscoreHphobe
7	D1037	-0.82	-5.958	12.0051	3.9704	D	1036	HBond+PhobEnHB

2nd conformation selected

Rank	*Feature_label*	*Score*	*X*	*Y*	*Z*	*type*	*num*	*source*
1	D931	-2.2	-8.2613	13.0688	0.9512	D	930	HBond+PhobEnHB
2	A96	-2	-11.8577	4.0212	1.1523	A	95	HBond+PhobEnHB
3	A540	-1.7	-8.6501	15.6195	2.0211	A	539	HBond+PhobEnHB
4	A1	-1.45	0.55	13.0384	1.08	A	0	HBond+PhobEn
5	R2088	-1.33	-5.3759	8.2438	-0.3826	R	2087	RingChemscoreHphobe
6	A544	-1.06	-6.9096	10.2815	0.1348	A	543	HBond+PhobEnPairHB
7	R1910	-1.05	-7.7721	17.5044	2.545	R	1909	RingChemscoreHphobe

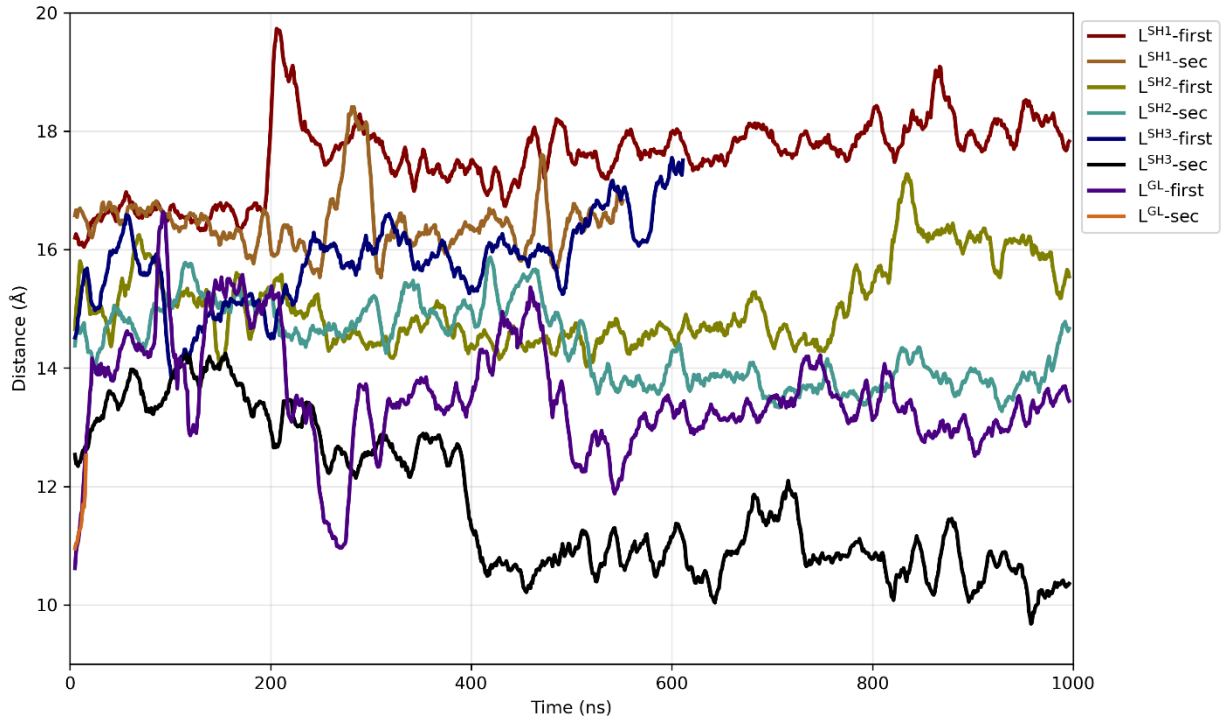
5th conformation selected

Rank	*Feature_label*	*Score*	*X*	*Y*	*Z*	*type*	*num*	*source*
1	A32	-2.2	-9.4997	14.9334	2.7846	A	31	HBond+PhobEnHB
2	D928	-2.2	-9.2382	13.17	1.0989	D	927	HBond+PhobEnHB
3	D1043	-2.17	-12.1161	4.4626	0.8786	D	1042	HBond+PhobEnHB
4	N1589	-2	-9.1744	7.8323	0.3601	N	1588	HBond+PhobEnPairHB
5	A108	-1.83	-7.1433	18.6963	2.5031	A	107	HBond+PhobEnHB
6	A551	-1.7	-7.6951	9.5702	-0.7328	A	550	HBond+PhobEnPairHB
7	D924	-1.45	-6.5192	13.0002	5.5923	D	923	HBond+PhobEnHB

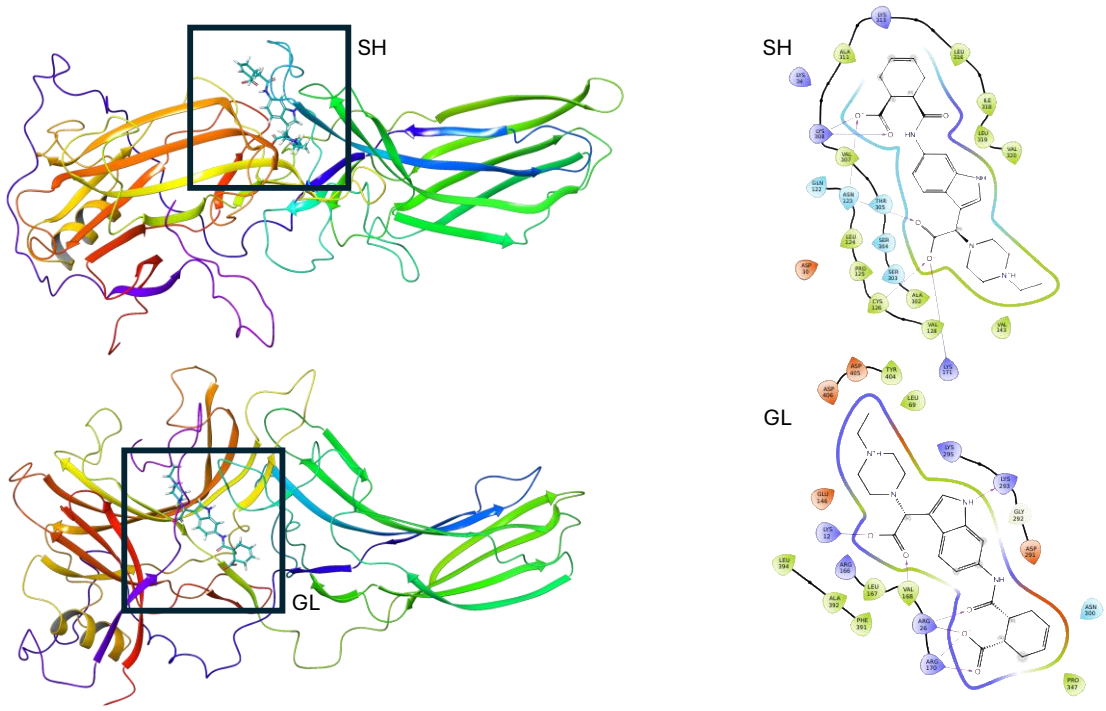
6th conformation selected

Supplementary Table 6. Glide docking scores of the compounds that target back loop and gate loop pockets and associated % rotated snapshots and average rotation angles over MD simulations

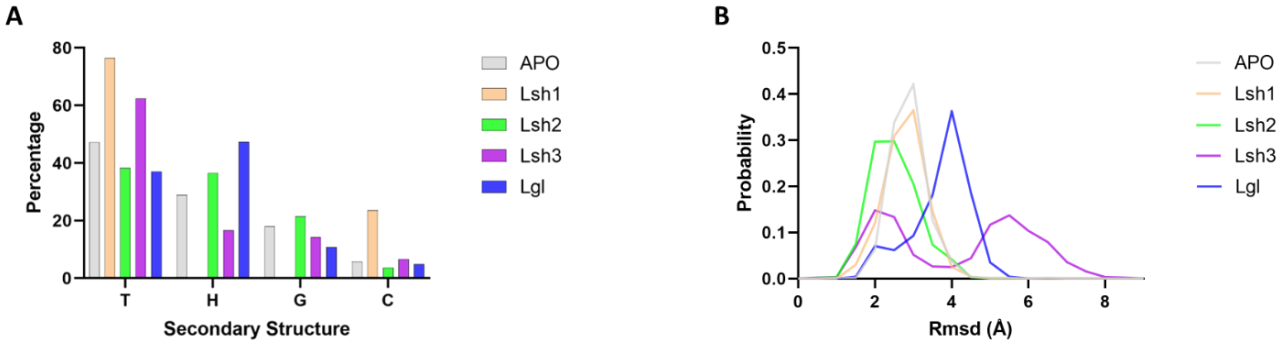
ZINC ID	Docking_Score (kcal/mol)	%rotated snapshots	<Rotation angle>pos
58950162	-8.9	37	3.1
11867617	-8.73	30	3.24
6812752	-8.22	27	2.38
67490967	-8.38	48	6.58
36075880	-7.97	28	3.21
2560972	-7.71	34	2.29
14971552	-7.20	46	3.25
31135800	-8.58	30	6.2
16925920	-9.84	13	2.89
36365956	-8.27	46	12.9
12915185 (gate loop)	-7.14	22	3.97



Supplementary Fig.3. Timeline plots showing the distance between the center of mass of compounds and the center of mass of the target sites (back loop and gate loop) over 1 μs MD simulation.



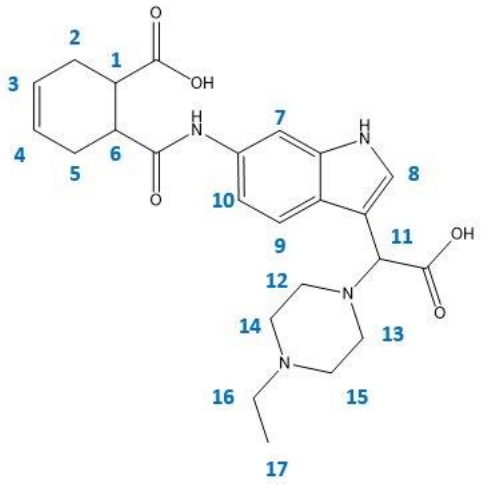
Supplementary Fig.4. Representative binding poses of the compound L^{SH-3} docked at the *back loop (SH)* and the *gate loop (GL)* binding pocket.



Supplementary Fig. 5. Percentage secondary structure content of the *back loop* sampled in *apo* and ligand-bound Arr3 trajectories calculated by `DSSP` algorithm implemented in VMD, T, H, G, and C correspond to turn, α-helix, 3_{10}-helix, and coil. **B.** The RMSD of the *gate loop* calculated using Cα atoms of residues 291-299**.**

NMR characterization

Supplementary Table 7. Structure, ^{1}H chemical shifts, multiplicity, J values and ^{13}C assignments of L^{SH-3} obtained in phosphate buffer at 298K.



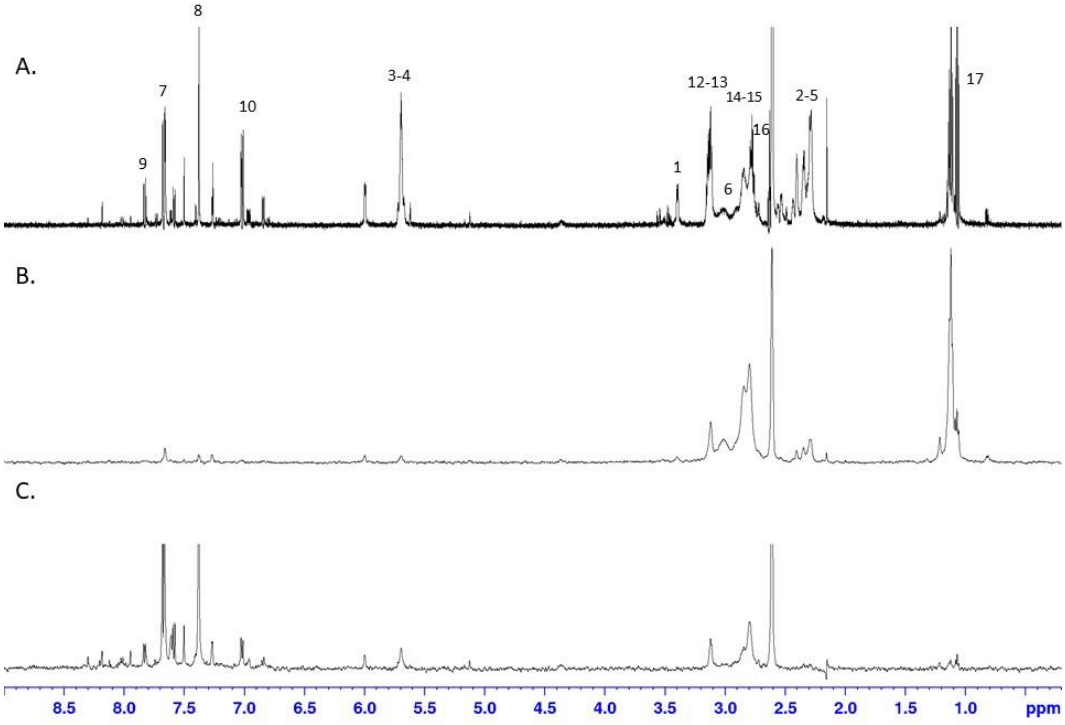
Atom	^{1}H ppm (main isomer)	M	J (Hz)	^{13}C ppm (main isomer)
1	3,396	m	/	39,9 (CH)
2 - 5	2,315 - 2,372	m	/	26,4 - 26,8 (CH2)
3 - 4	5,697	m	/	126,6 (CH)
6	3,105	q	/	41,6 (CH)
7	7,493	m	1,7	106,3 (CH)
8	7,282	d	/	126,2 (CH)
9	7,566	d	9	119,6 (CH)
10	6,930	m	1,7 / 9	116,0 (CH)
11	5,978	s	/	127,9 (CH)
12 - 13	3,204	m	/	42,6 (CH2)
14 - 15	2,944	m	/	42,3 - 43,3 (CH2)
16	2,754	q	7,6	52,1 (CH2)
17	1,094	t	7,6	9,9 (CH3)

STD-NMR

STD experiment maps protons close to the protein. STD spectrum is obtained by the difference of two spectra: *on resonance* and *off* resonance. In the *on resonance* experiment the protein receives a selective irradiation (at frequencies able to excite only the protons of the protein and not affecting

the atoms of the ligand) that saturates the macromolecule protons; then due to intramolecular nuclear Overhauser effect (NOE) the magnetization is transferred to the ligand, that is near the protein (4-5 Å), which in turn, by exchange, is moved into solution where it is detected. The *off-resonance* spectrum is acquired with a saturation frequency where no protein or ligand signals are detectable, usually around ~40 ppm. This off-resonance spectrum represents a normal 1H NMR. Subtraction of the on-resonance and off-resonance spectrum results in the final saturation transfer difference (STD) NMR 1D spectrum showing only signals from ligands that have binding affinity, while the peaks of the distant protons are absent.

The analysis of different STD spectra acquired by irradiating at different frequencies (corresponding to different residues of the protein) allows the identification of the type of protein residues contacting the ligand and allows to identify the protein binding pocket (see ref.[4])



Supplementary Fig. 6. A. ^{1}H-NMR spectrum of L^{SH-3} **B.** STD spectrum irradiating at 0 ppm **C.** STD spectrum irradiating at 8 ppm

Supplementary Table 8. STD Abs at 0 and 8 ppm; the total ligand saturation values are highlighted in the last row.

H	STDAbs at 0 ppm	STDAbs at 8 ppm
2 - 5	0,93	0,00
3 - 4	0,50	0,55
6	5,26	0,00
7	1,19	3,70
8	0,83	1,89
9	1,00	3,13
10	1,19	1,89
11	0,59	1,11
12 - 13	2,00	0,55
14 - 15	7,14	1,32
17	2,22	0,36
SUM	**22,86**	**14,50**

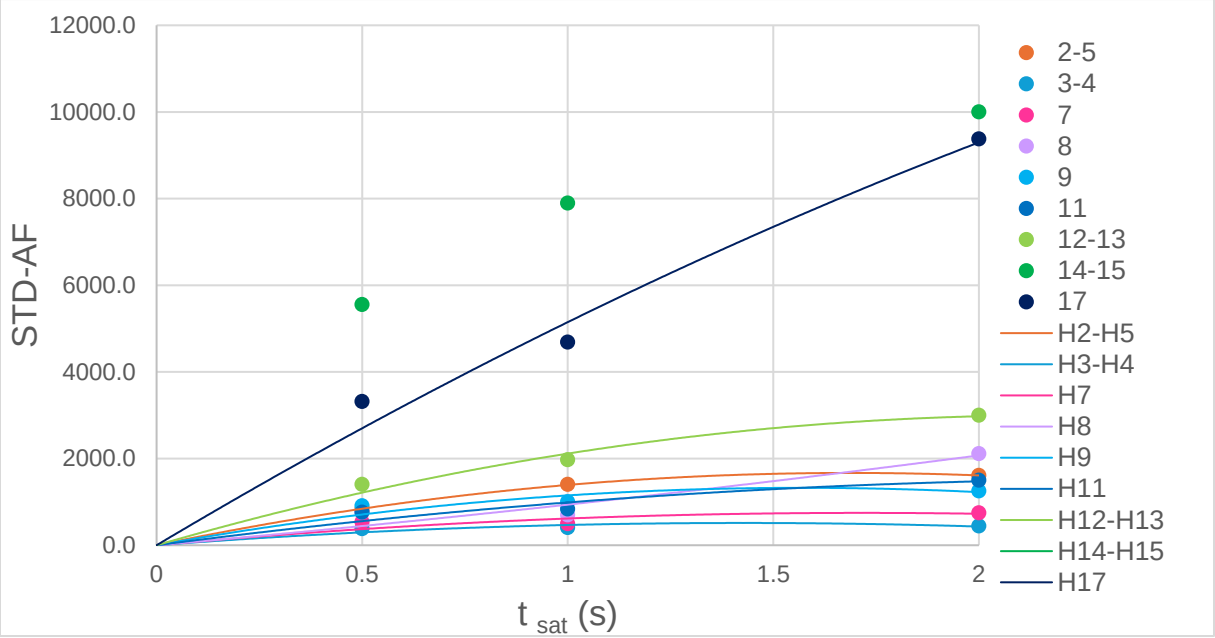
STD was also employed as a fast method for determination of K_D values for weakly binding ligands. To obtain protein-ligand association curves from STD NMR experiments, Mayer and Meyer introduced the conversion from observed experimental intensities to STD amplification factors (STD-AF) by multiplying the observed STD by the molar excess of ligand over protein[5].

Supplementary Table 9. STD (%) obtained with saturation time of 2s at irradiation frequency 0,0 ppm, 298 K

Atom	STD Abs	STD-AF	STD-AF$_0$
2-5	1,07	2145,92	305,87
3-4	0,29	588,24	191,17
7	0,50	995,02	180,52
8	1,41	2816,90	133,66
9	0,83	1666,67	299,77
11	1,00	2000,00	174,24
12-13	2,00	4000,00	413,80
14-15	6,67	13333,33	1982,58
17	6,25	12500,00	731,79

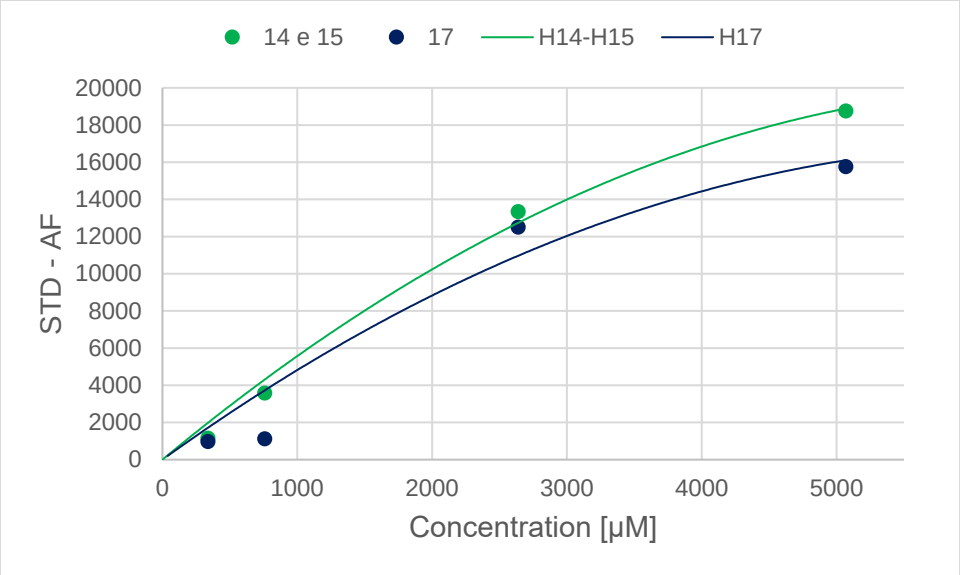
In the following figure, the STD-AF values of selected protons of L^{SH-3} are reported as a function of the saturation time. For each proton, STD-AF values were obtained at different saturation times

(0.5, 1 and 2 s) and fit by using the equation STD-AF(t_{sat})=STD-AF$_{max}$[1−exp(−$k_{sat}t_{sat}$)]. The initial slopes, STD-AF$_0$, are obtained from $STD - AF_0 = STDAF(t_{sat}) - (\alpha \times t_{sat})$, where α is the slope of the curve.



Supplementary Fig. 7. For each proton, STD-AF values were obtained and plotted at different saturation times (0.5, 1 and 2 s). The different coloured dots represent the experimental values and the lines are the mathematical fit to an exponential function.

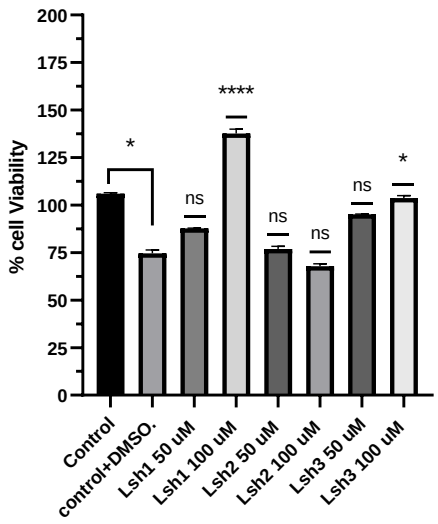
Therefore, a plot of STD-AF values at increasing ligand concentrations gives rise to the protein-ligand binding isotherms, from which the dissociation constant K_D can be derived (the K_D is the concentration of ligand necessary to reach 50% saturation of the protein).



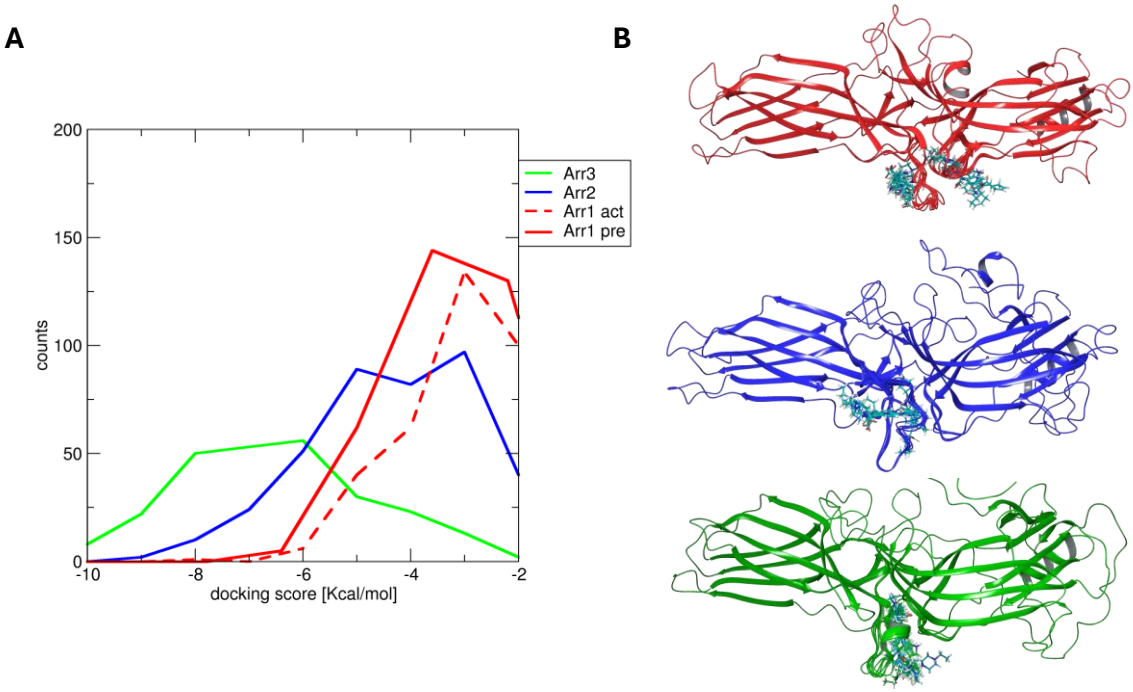
Supplementary Fig. 8. Binding curve of L^{SH-3} with Arr3. The dots represent the experimental values, the lines are the mathematical fit to an exponential function.

Cytotoxicity in N2a and HEK293 cells.

We assessed the toxicity of L^{SH-1}, L^{SH-2}, and L^{SH-3} by succinate dehydrogenase activity (MTT) assay[6] . Each compound was tested at 50 and 100 µM to determine the maximum non-toxic concentration that can be used (1% DMSO served as control). The toxicity of the compounds was determined using Neuro 2A (N2a) cell line, which is a mouse neural crest-derived cell line that has been extensively used to study neuronal differentiation, axonal growth, and GPCR signaling[7] . Tested compounds were found to be non-toxic. No toxicity was observed in HEK293 cells. Because some chemicals, such as flavonoids and phytochemicals, can reduce MTT reagent by non-enzymatic reactions[8,9] , we performed an MTT assay in the absence of N2a or HEK293 cells as a control. Chemicals were added to the sample and compared against a solution that contained 1% DMSO. We did not observe interference.



Supplementary Fig. 9. Toxicity test of the compounds in N2a cells. The final DMSO concentration is 1% in all groups except the control group. *p<0.05, ****p<0.0001 statistically significant differences. ns: no significant differences. Data are shown as means ± S.E.M



Supplementary Fig. 10. Induced Fit Docking of compound L^{SH3} to arrestin subtypes. Comparative Induced Fit Docking of compound L^{SH3} was carried out on representative structures of the most populated clusters obtained from 3μs MD simulations of arrestin-3 and arrestin-2, as well as on crystal structures of pre-activated (pre) and rhodopsin-bound arrestin-1 (act), using Schrodinger Glide[10]. **A**. Histogram showing distribution of docking scores of all identified poses on each receptor. **B.** Superposition of best docking score complexes, color coded as in A: red arrestin-1, blue arrestin-2, green arrestin-3.

References

1. Halgren, T. New method for fast and accurate binding-site identification and analysis. *Chem Biol Drug Des* 69, 146–148 (2007).
2. Le Guilloux, V., Schmidtke, P. & Tuffery, P. Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics* 10, (2009).
3. Jiménez, J., Doerr, S., Martínez-Rosell, G., Rose, A. S. & De Fabritiis, G. DeepSite: Protein-binding site predictor using 3D-convolutional neural networks. *Bioinformatics* 33, 3036–3042 (2017).
4. Monaco, S., Tailford, L. E., Juge, N. & Angulo, J. Differential Epitope Mapping by STD NMR Spectroscopy To Reveal the Nature of Protein–Ligand Contacts. *Angewandte Chemie* 129, 15491–15495 (2017).
5. Meyer, B. & Peters, T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angewandte Chemie - International Edition* 42, 864–890 (2003).
6. Morgan, D. M. L. Tetrazolium (MTT) Assay for Cellular Viability and Activity. in *Polyamine Protocols* (ed. Morgan, D. M. L.) 179–184 (Humana Press, Totowa, NJ, 1998). doi:10.1385/0-89603-448-8:179.
7. Fichter, K. M., Flajolet, M., Greengard, P. & Vu, T. Q. Kinetics of G-protein-coupled receptor endosomal trafficking pathways revealed by single quantum dots. *Proc Natl Acad Sci U S A* 107, 18658–18663 (2010).
8. Peng, L., Wang, B. & Ren, P. Reduction of MTT by flavonoids in the absence of cells. *Colloids Surf B Biointerfaces* 45, 108–111 (2005).
9. Hsu, S. *et al.* Green Tea Polyphenols Induce Differentiation and Proliferation in Epidermal Keratinocytes. *Journal of Pharmacology and Experimental Therapeutics* 306, 29 (2003).
10. Friesner, R. A. *et al.* Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem* 49, 6177–6196 (2006).