Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

A molecular receptor for zwitterionic phenylalanine

Francisco Gómez Herrero,[a]; Omayra H. Rubio,[a]; Laura M. Monleón,[a]; Ángel L. Fuentes de Arriba,[a] Luís Simón Rubio,[b] and Joaquín R. Morán[a]

[a] Organic Chemistry Department, Plaza de los Caídos 1-5, University of Salamanca, 37008 Salamanca, Spain

[b] Engineering Chemistry Department, Plaza de los Caídos 1-5, University of Salamanca, 37008 Salamanca, Spain

*Corresponding author: romoran@usal.es

Table of Contents

Figure S1. ^{1}H NMR spectrum of compound 3 (200 MHz, $CDCl_3$).

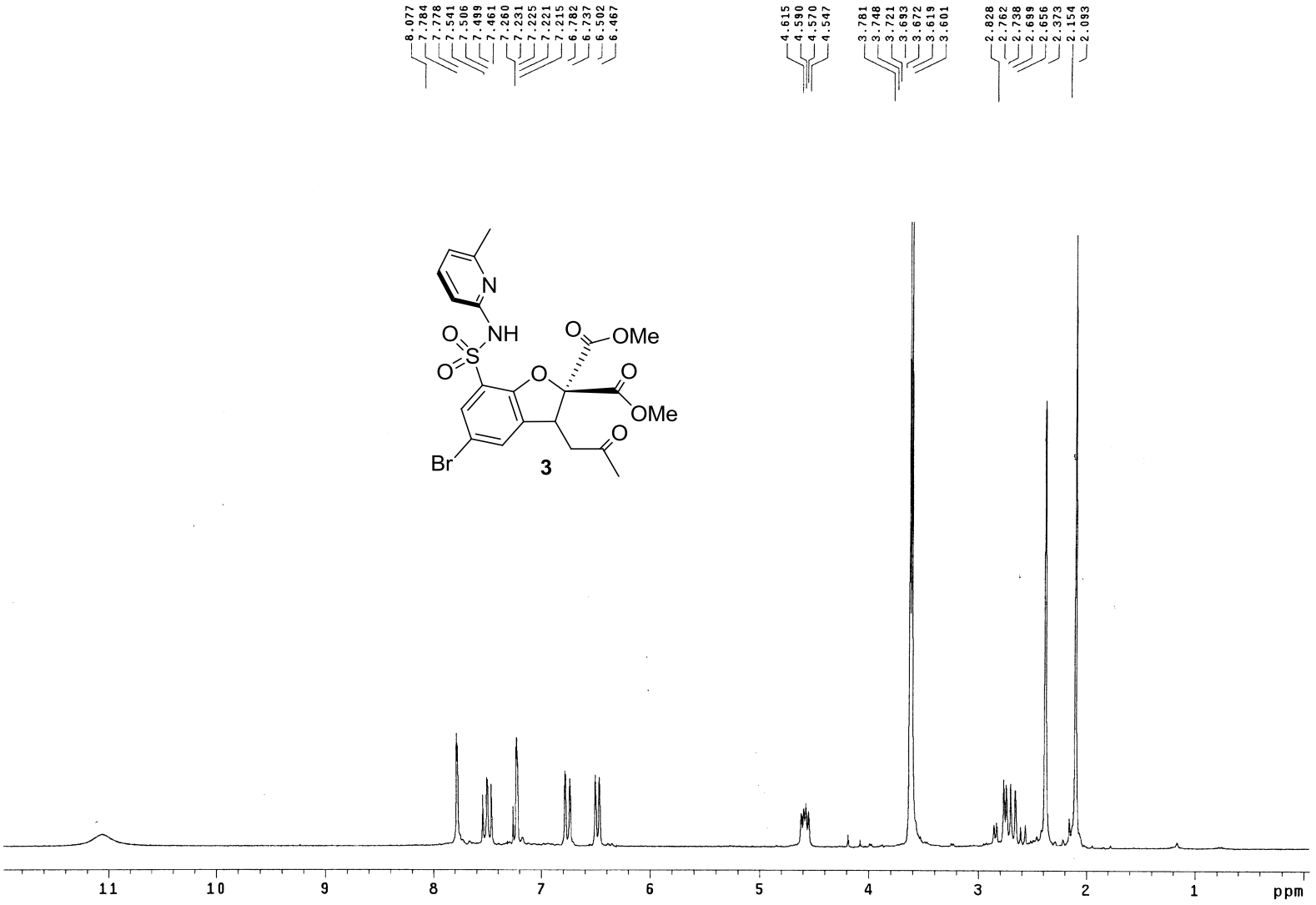


Figure S2. ^{13}C NMR spectrum of compound 3 (100 MHz, $CDCl_3$).

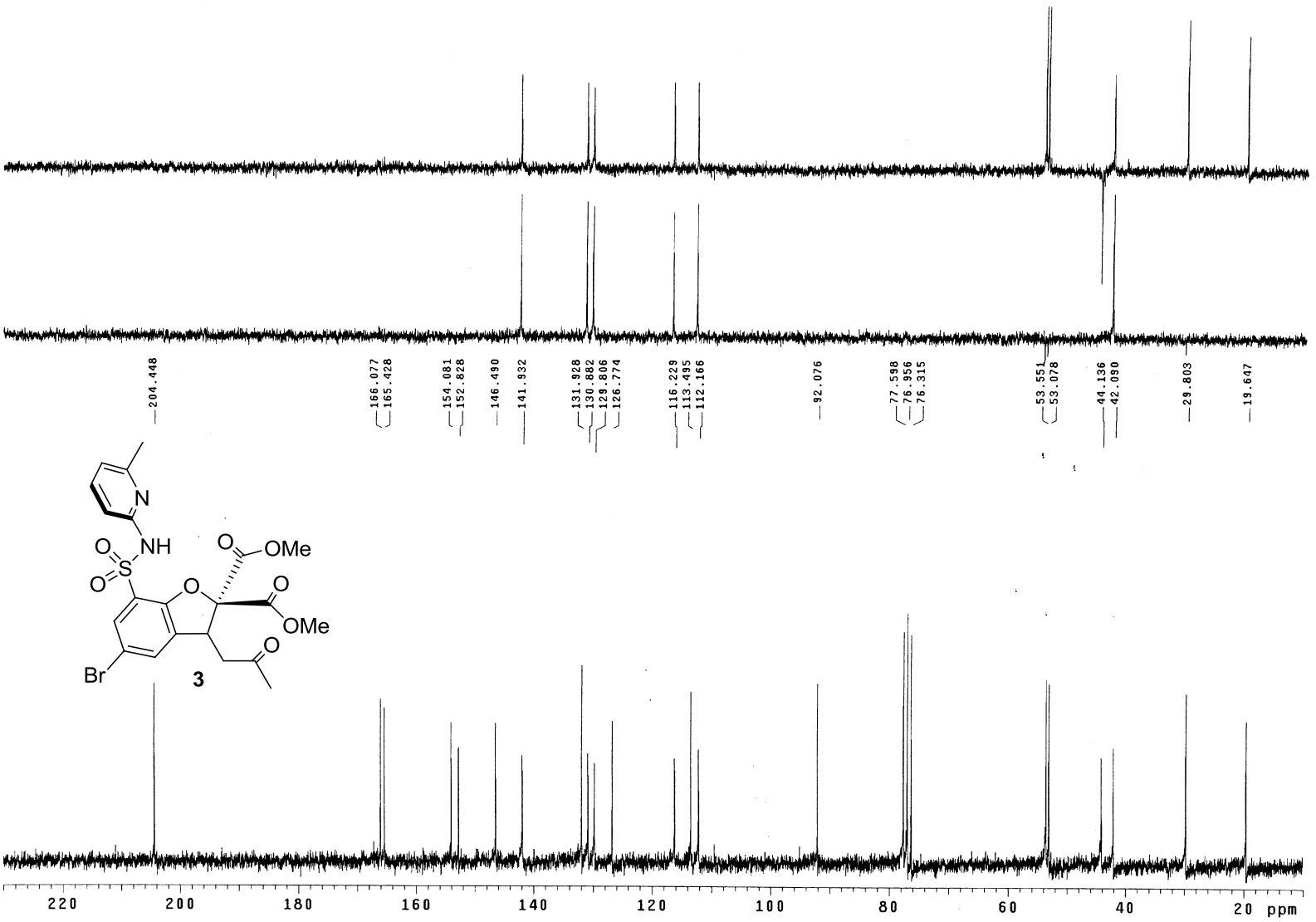


Figure S3. IR spectrum of compound 3.

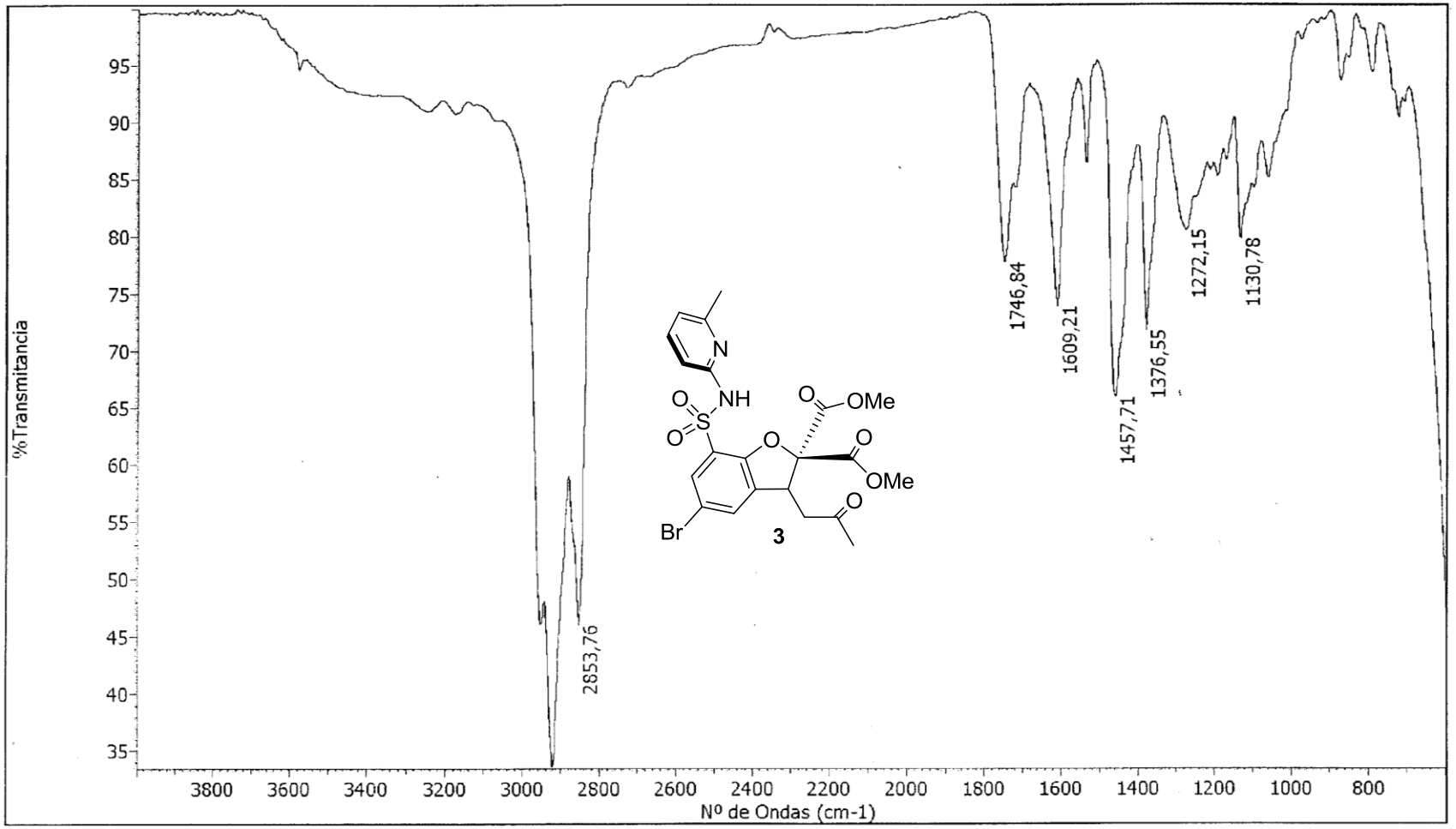


Figure S4. HRMS spectrum of compound 3 (ESI-QTOF).

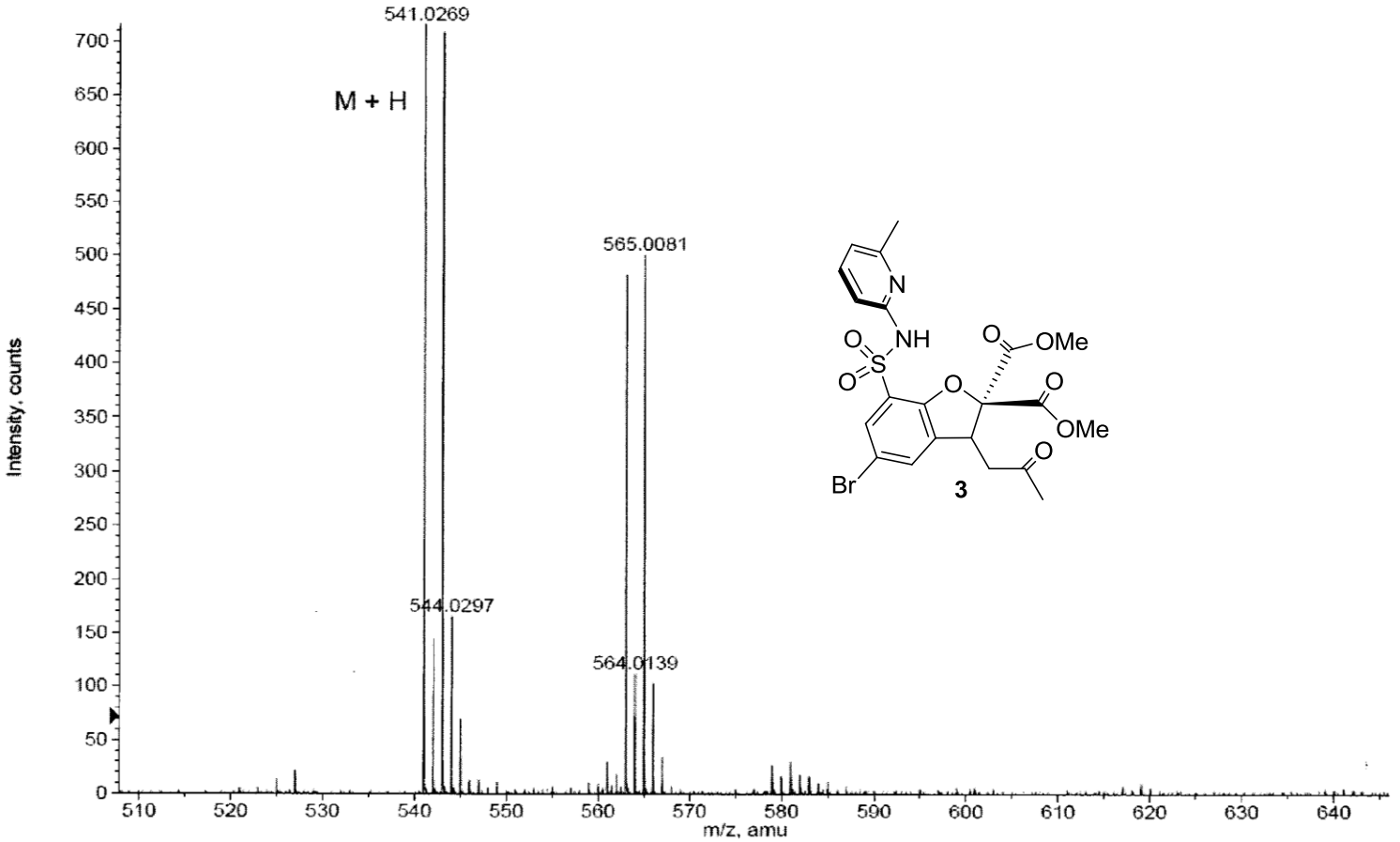


Figure S5. ^{1}H NMR spectrum of receptor 1 (400 MHz, $CDCl_3$).

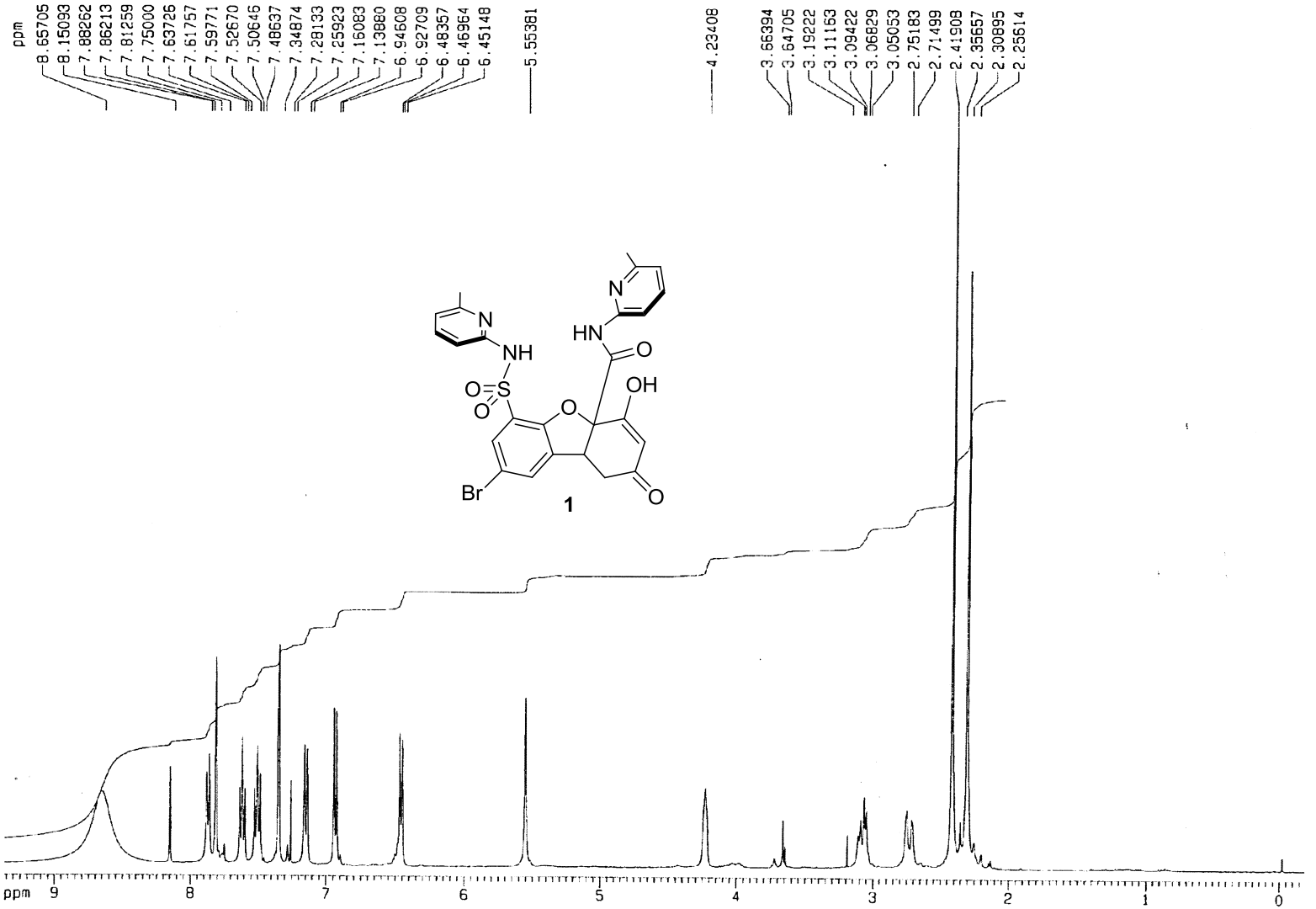


Figure S6. ^{13}C NMR spectrum of receptor 1 (100 MHz, $CDCl_3$).

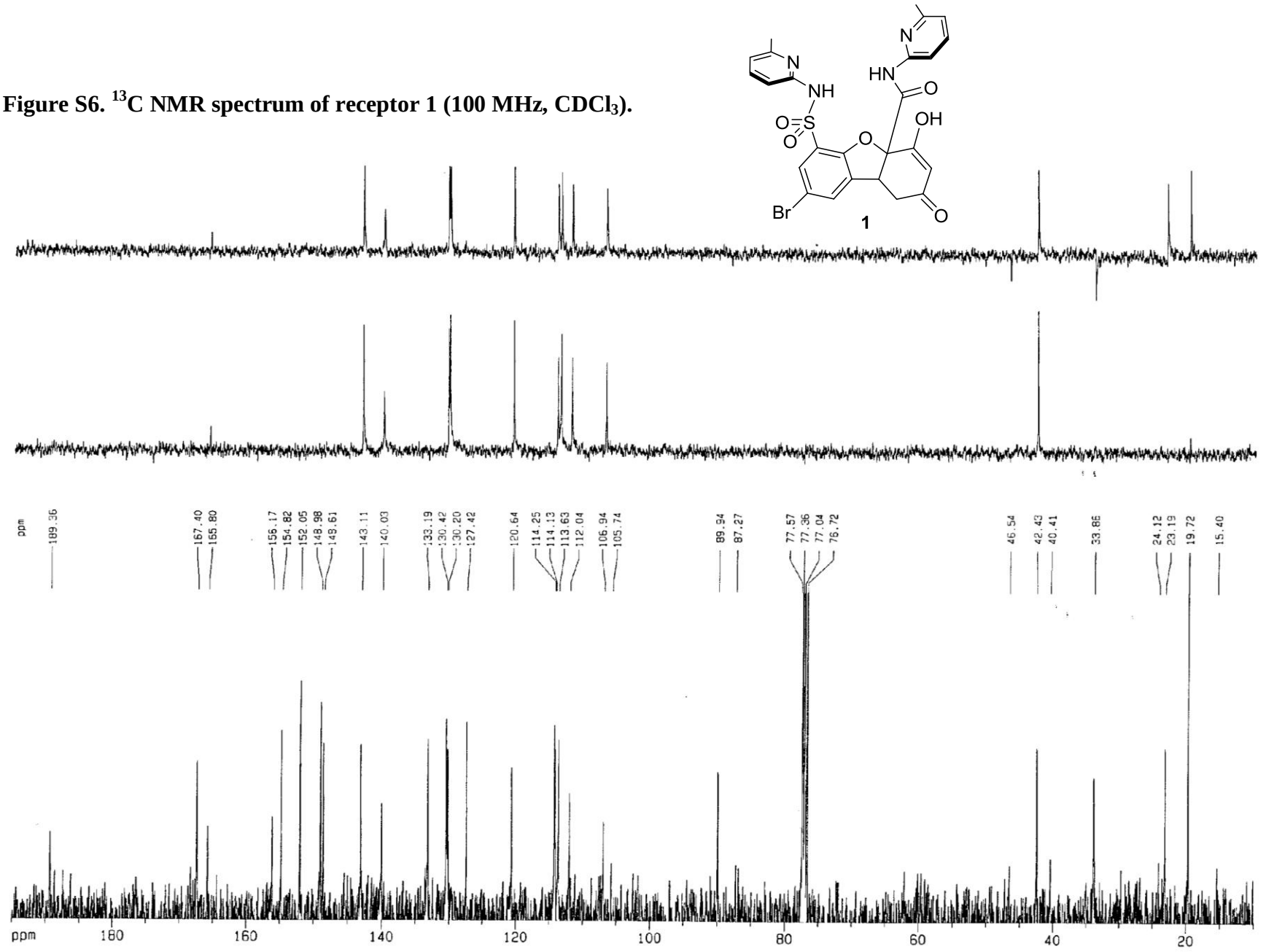


Figure S7. COSY spectrum of receptor 1 ($CDCl_3$).

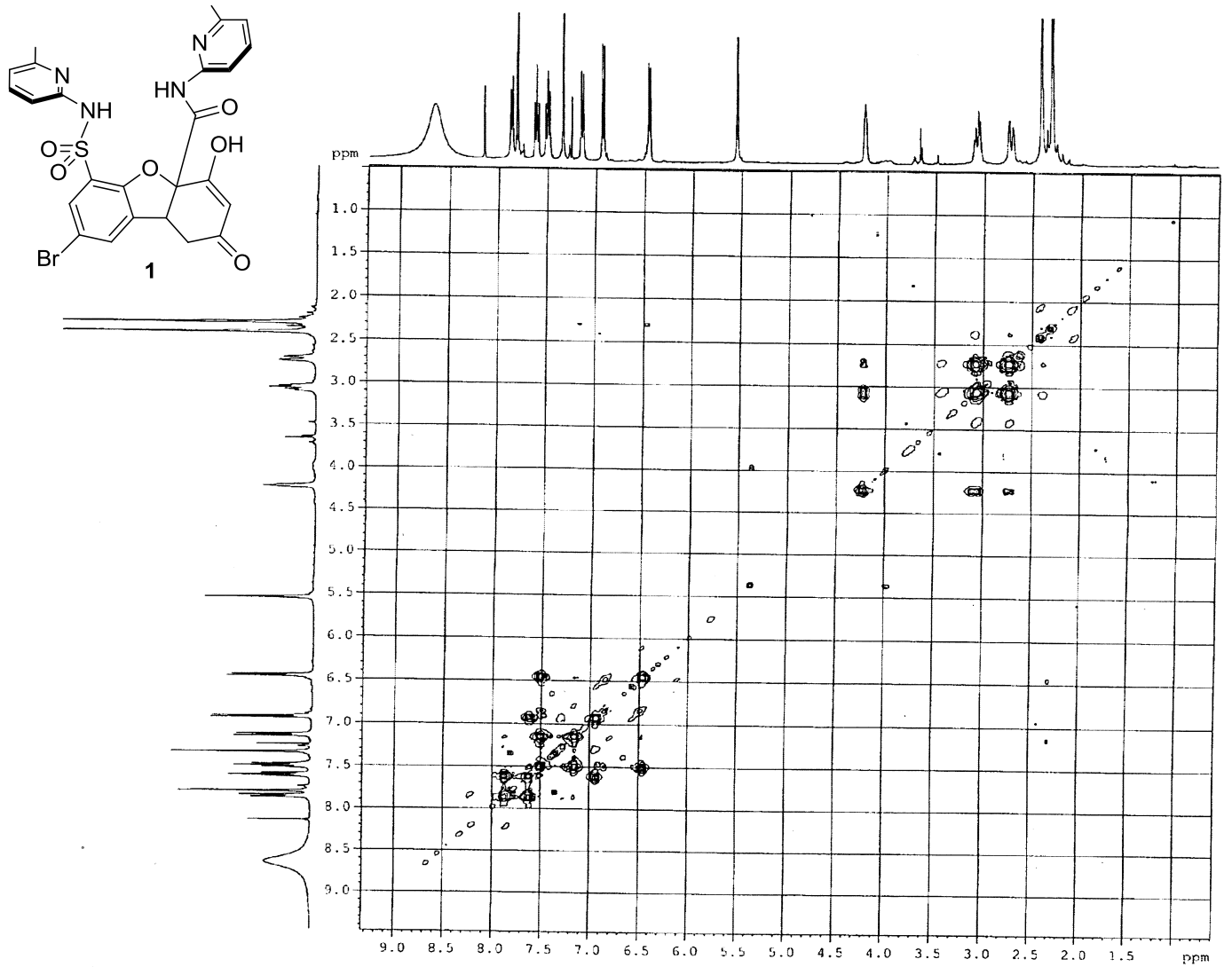


Figure S8. ROESY spectrum of receptor 1 ($CDCl_3$).

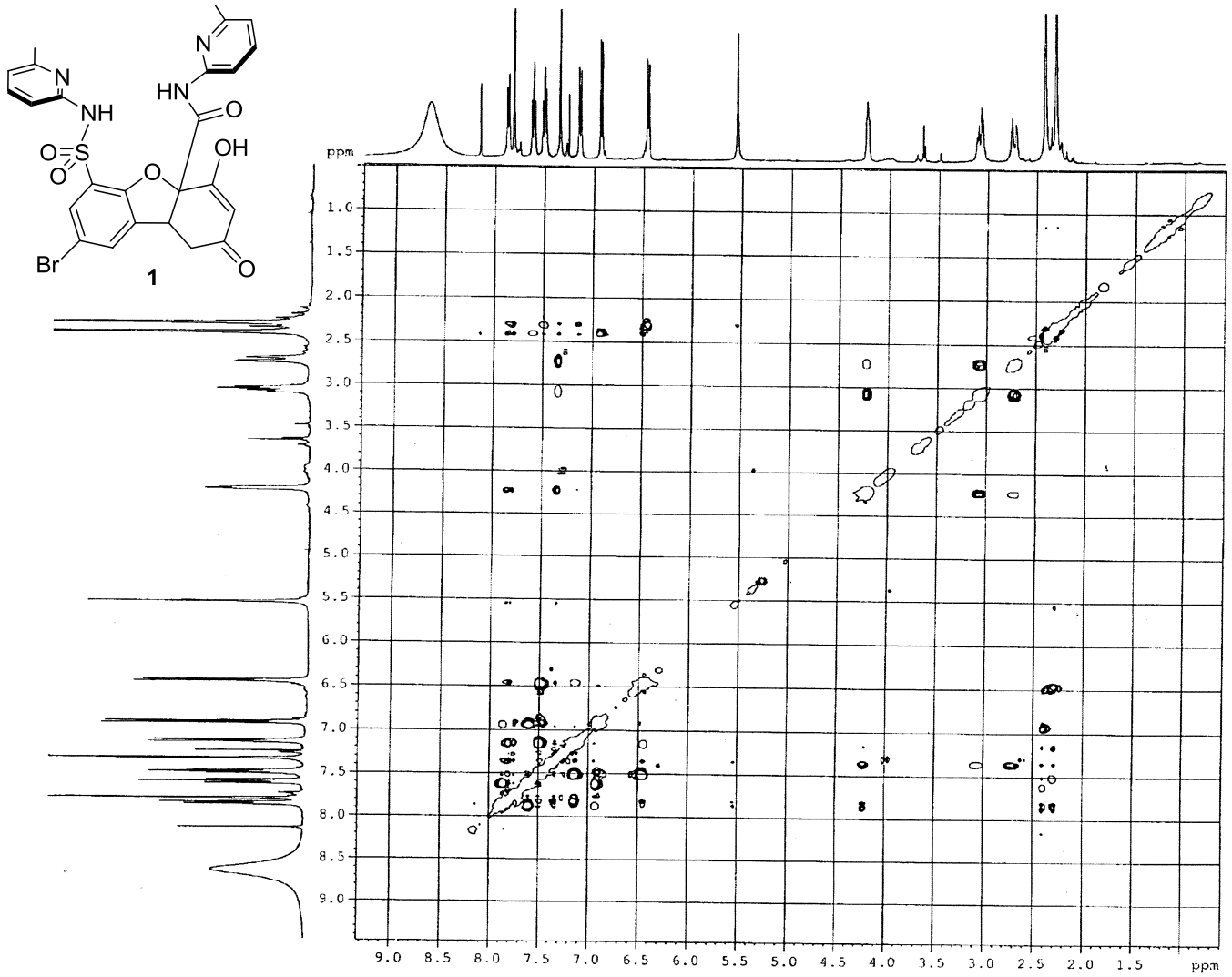


Figure S9. HMQC spectrum of receptor 1 ($CDCl_3$).

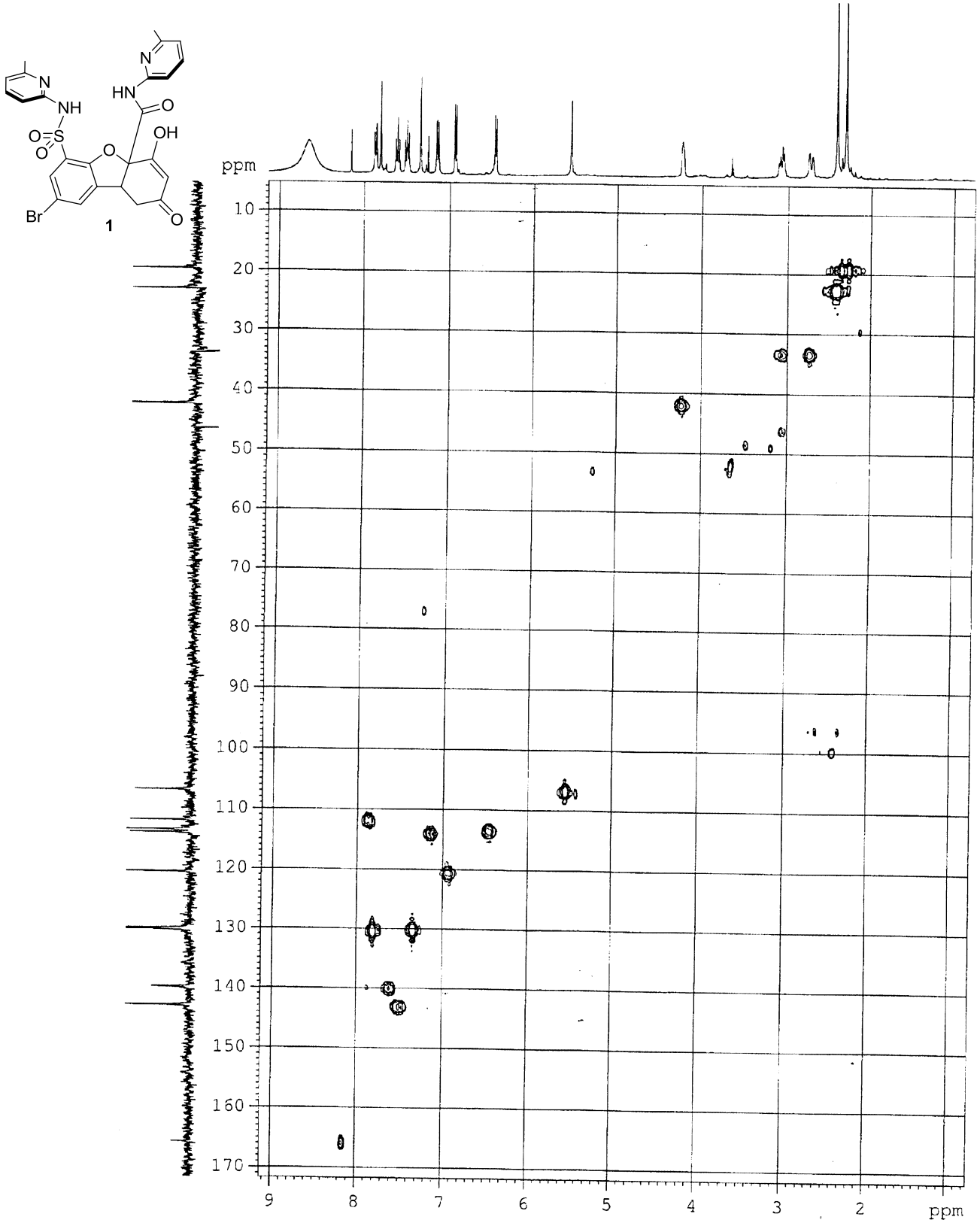


Figure S10. HMBC spectrum of receptor 1 ($CDCl_3$).

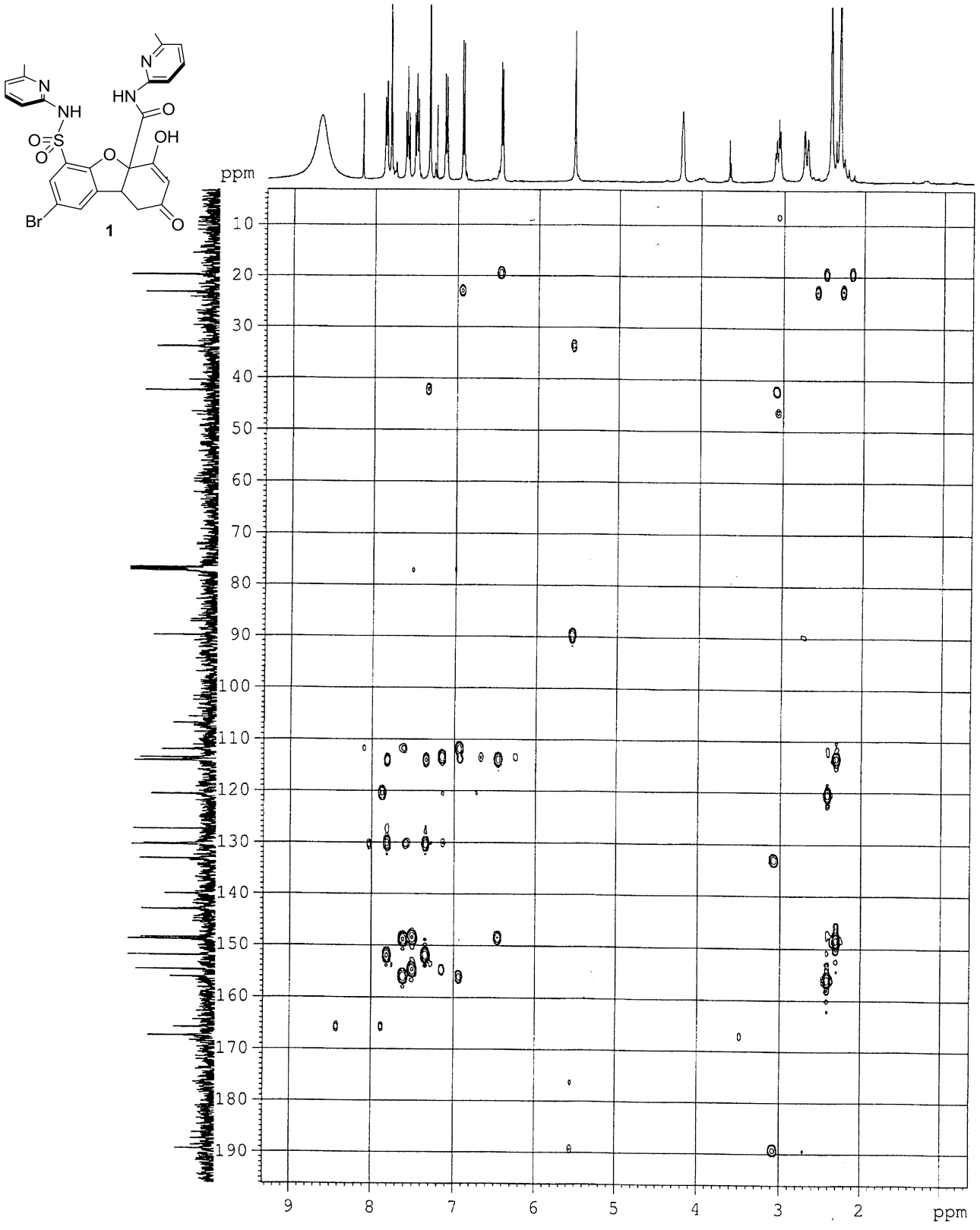


Figure S11. IR spectrum of receptor 1.

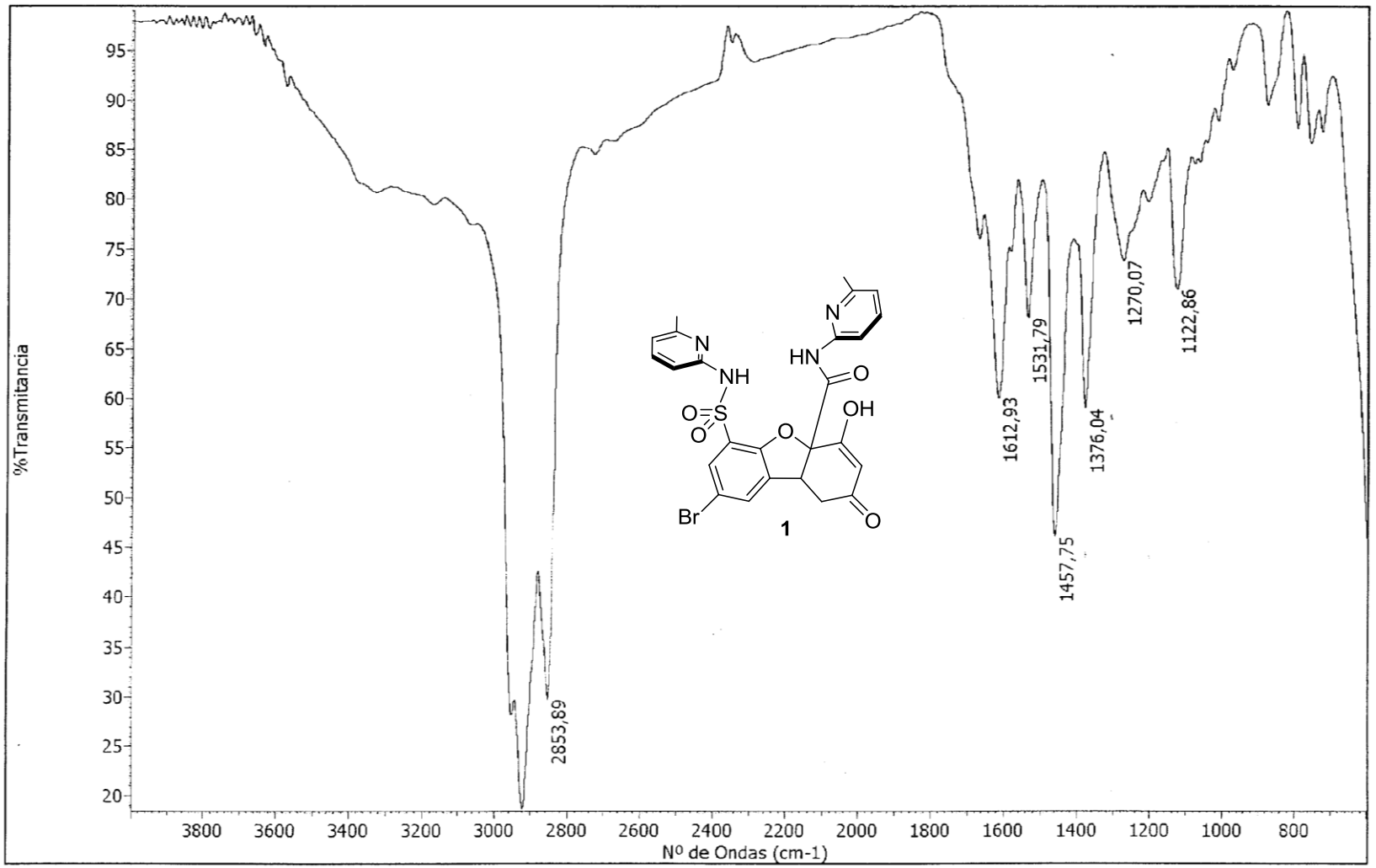


Figure S12. HRMS spectrum of receptor 1 (ESI-QTOF).

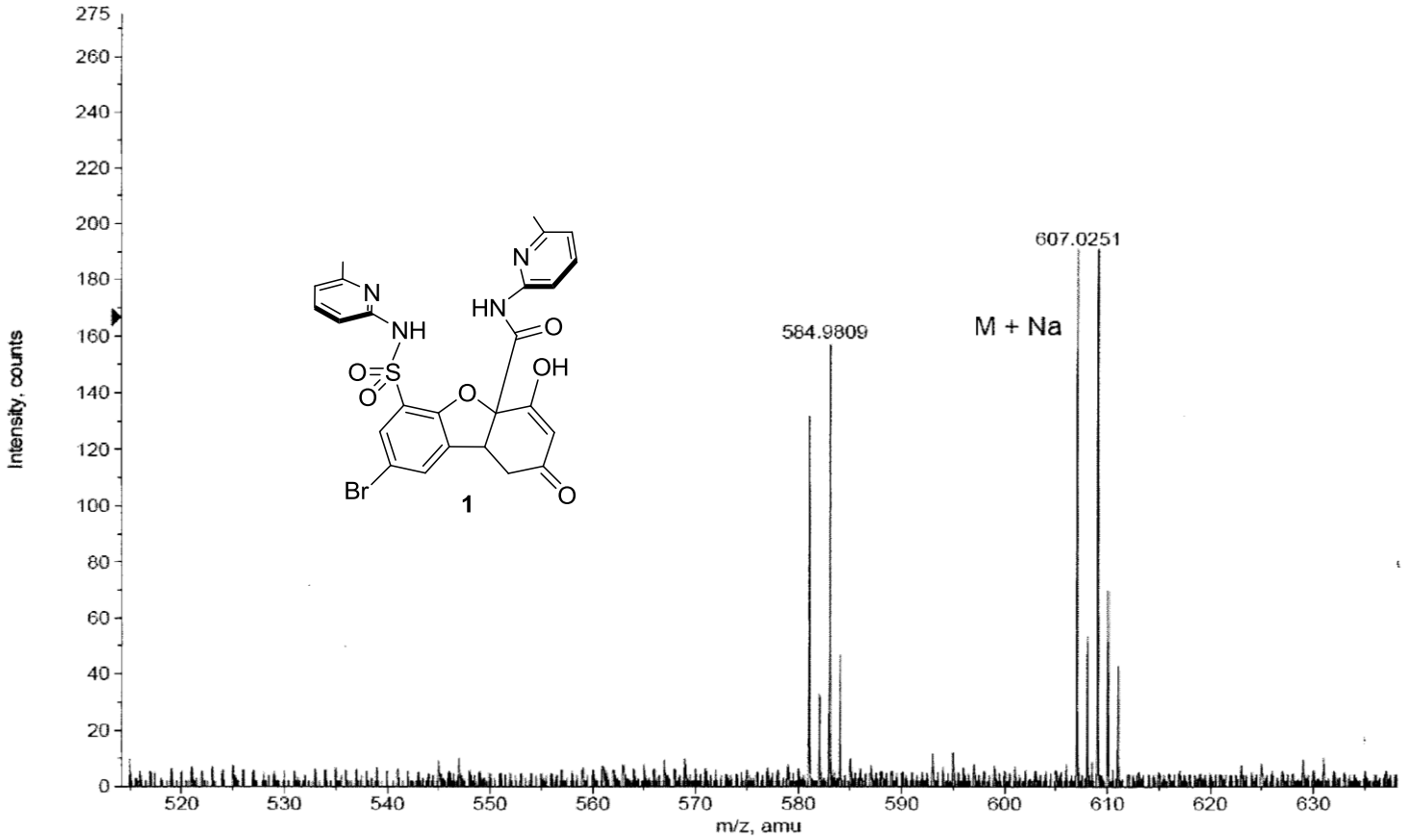


Figure S13. ^{1}H NMR spectrum of receptor 1 with *rac*-phenylalanine (200 MHz, $CDCl_3$).

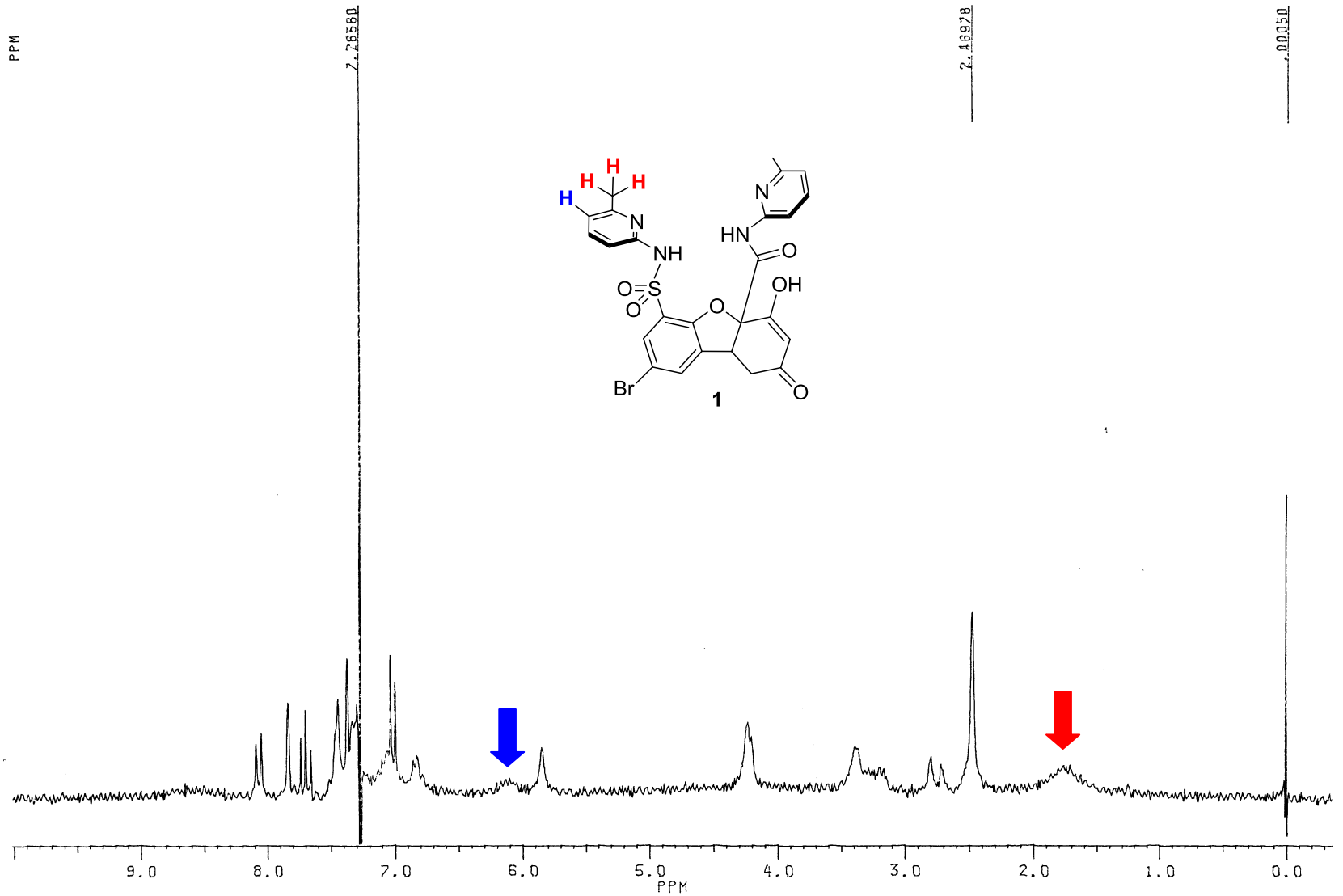


Figure S14. ^{1}H NMR spectrum of receptor 1 with *rac*-phenylalanine (DMSO-d_6).

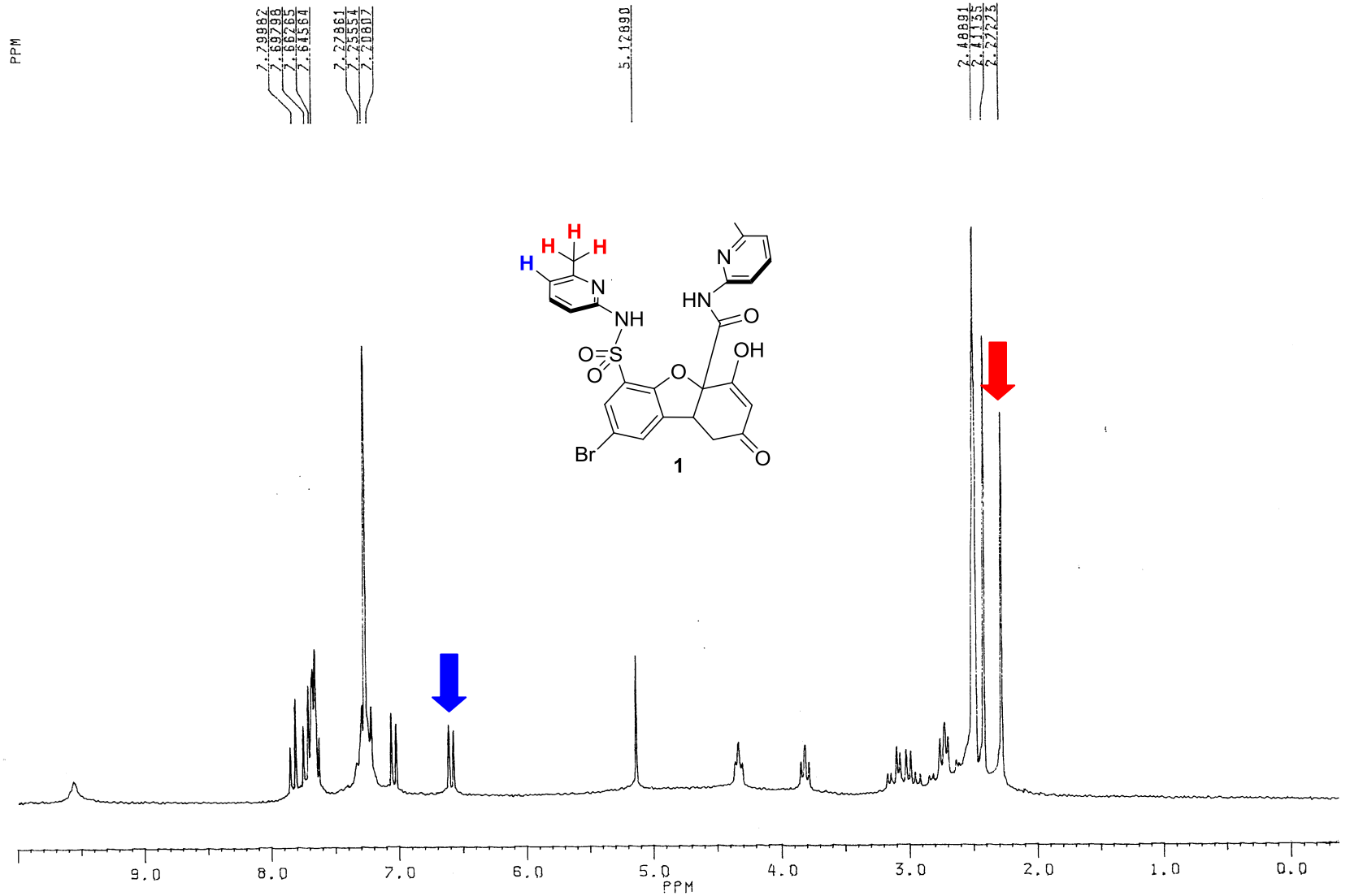


Figure S15. ^{1}H NMR spectrum of receptor 1 and the associates with *L*-phenylalanine (200 MHz, $CDCl_3$).

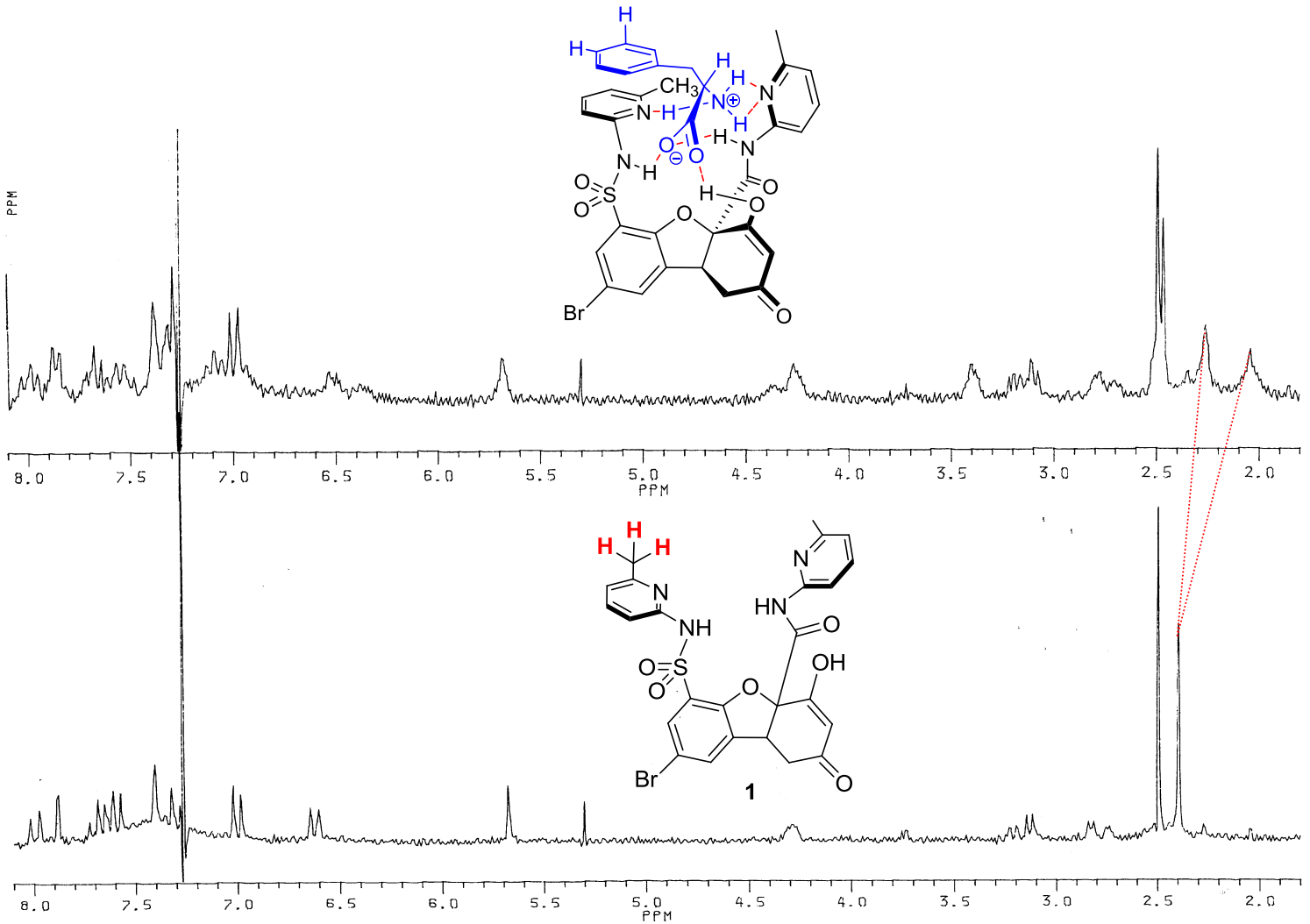
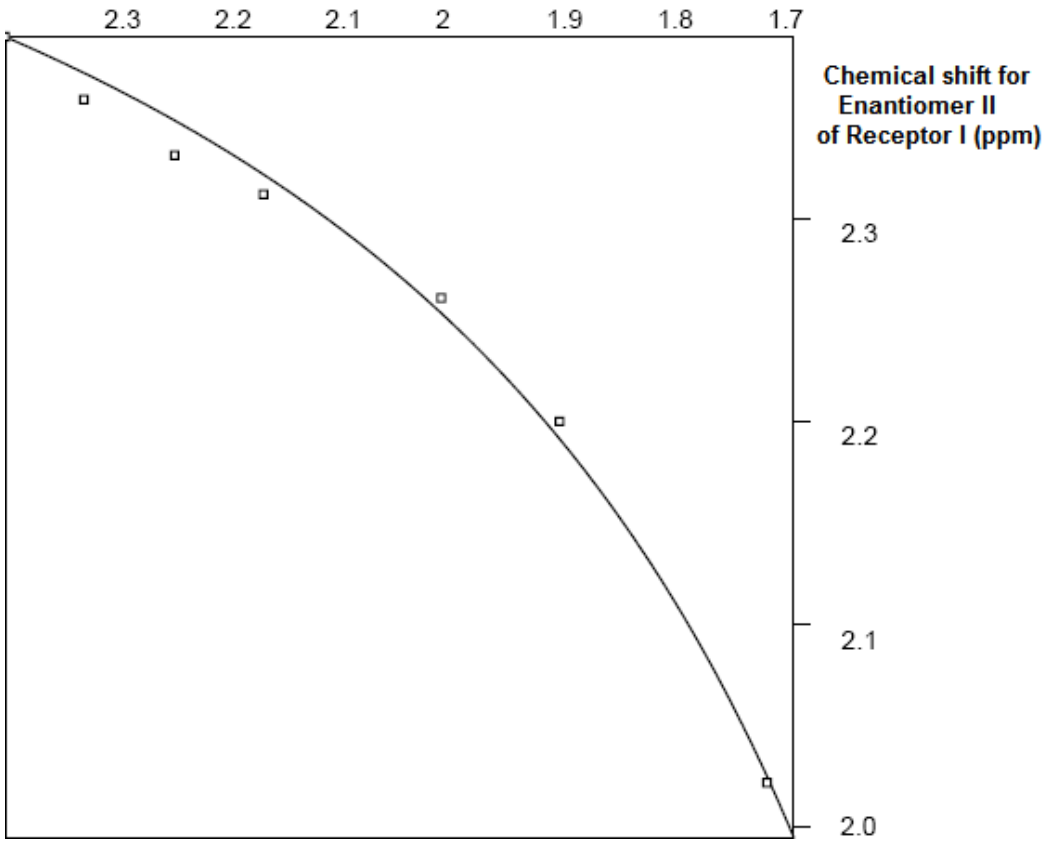


Figure S16. ^{1}H NMR titration of receptor 1 with *L*-phenylalanine ($CDCl_3$, 20°C).

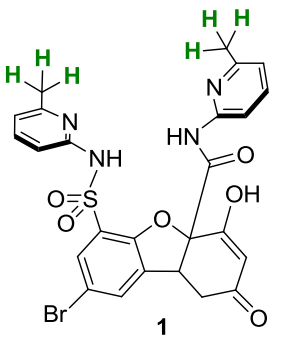
Competitive titration
Constant: 2.3 1/M
Max. shift 1 (ppm): 1.6759
Max. shift 2 (ppm): 1.9951

Host 1: Enantiomer I of Receptor 1
Host 2: Enantiomer II of Receptor 1

Guest: L-Phenylalanine



Shift Enantiomer I	Shift Enantiomer II
2.3896	2.3896
2.3184	2.3586
2.2362	2.3312
2.1559	2.3119
1.9951	2.2608
1.888	2.2001
1.7001	2.0221



1

Figure S17. ^{1}H NMR spectrum of solid residue obtained after preparative TLC of receptor 1 with *L*-phenylalanine impregnated plates (a) and the receptor 1 (b) (DMSO-d_6).

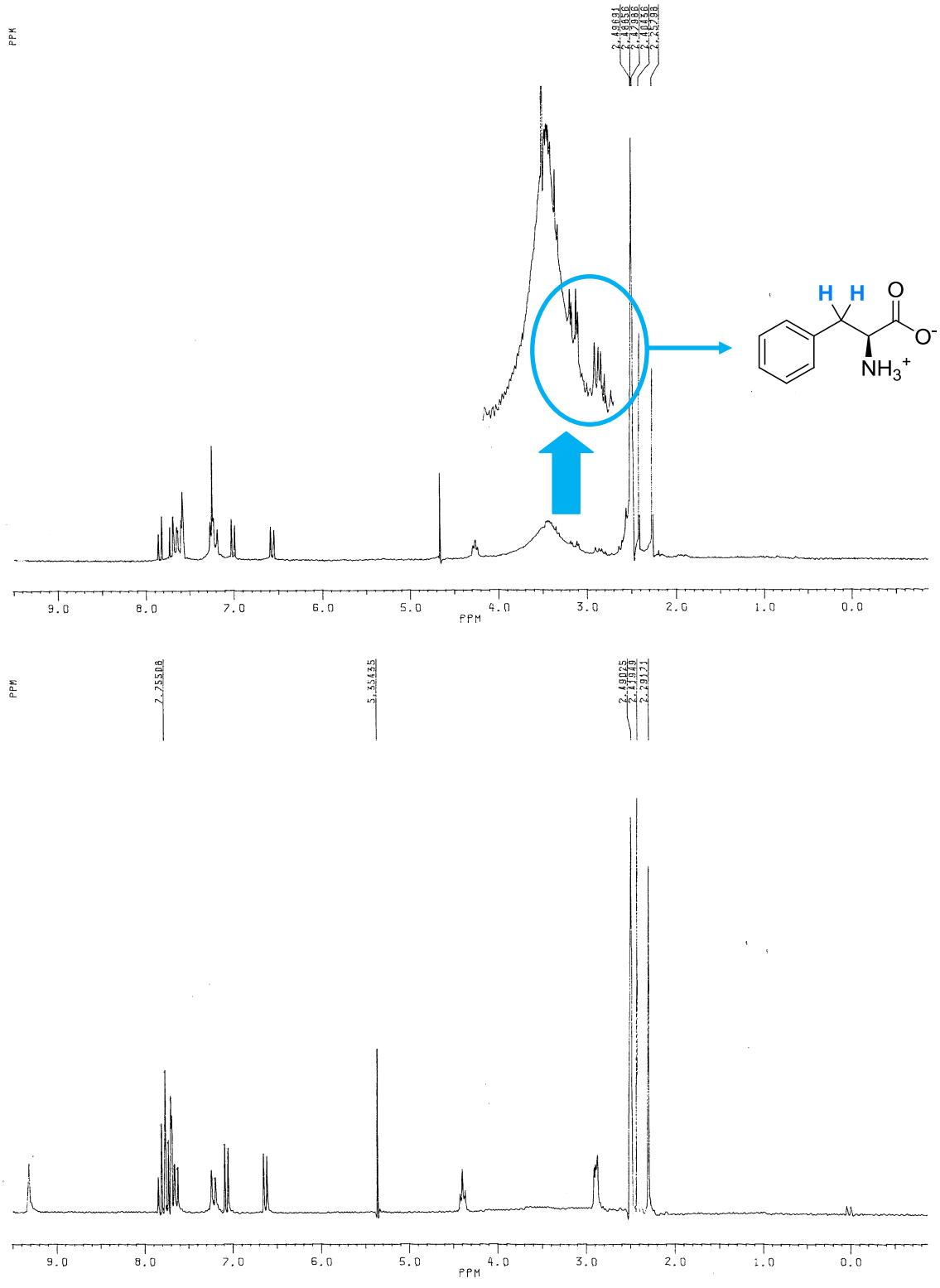


Figure S18. ^{1}H NMR spectrum of weak complex of receptor 1 with *L*-phenylalanine ($CDCl_3$).

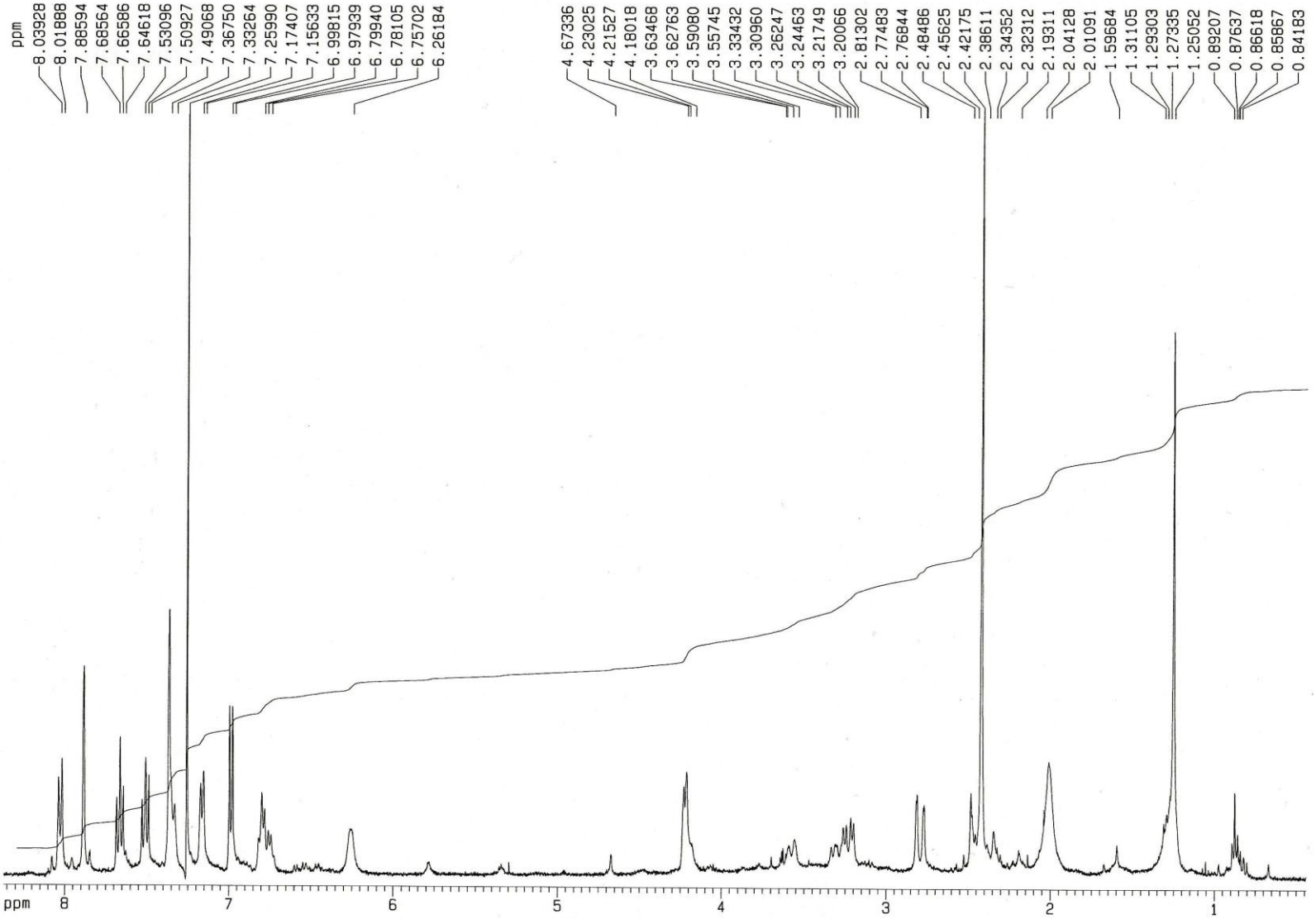


Figure S19. ROESY spectrum of weak complex of receptor 1 with *L*-phenylalanine ($CDCl_3$).

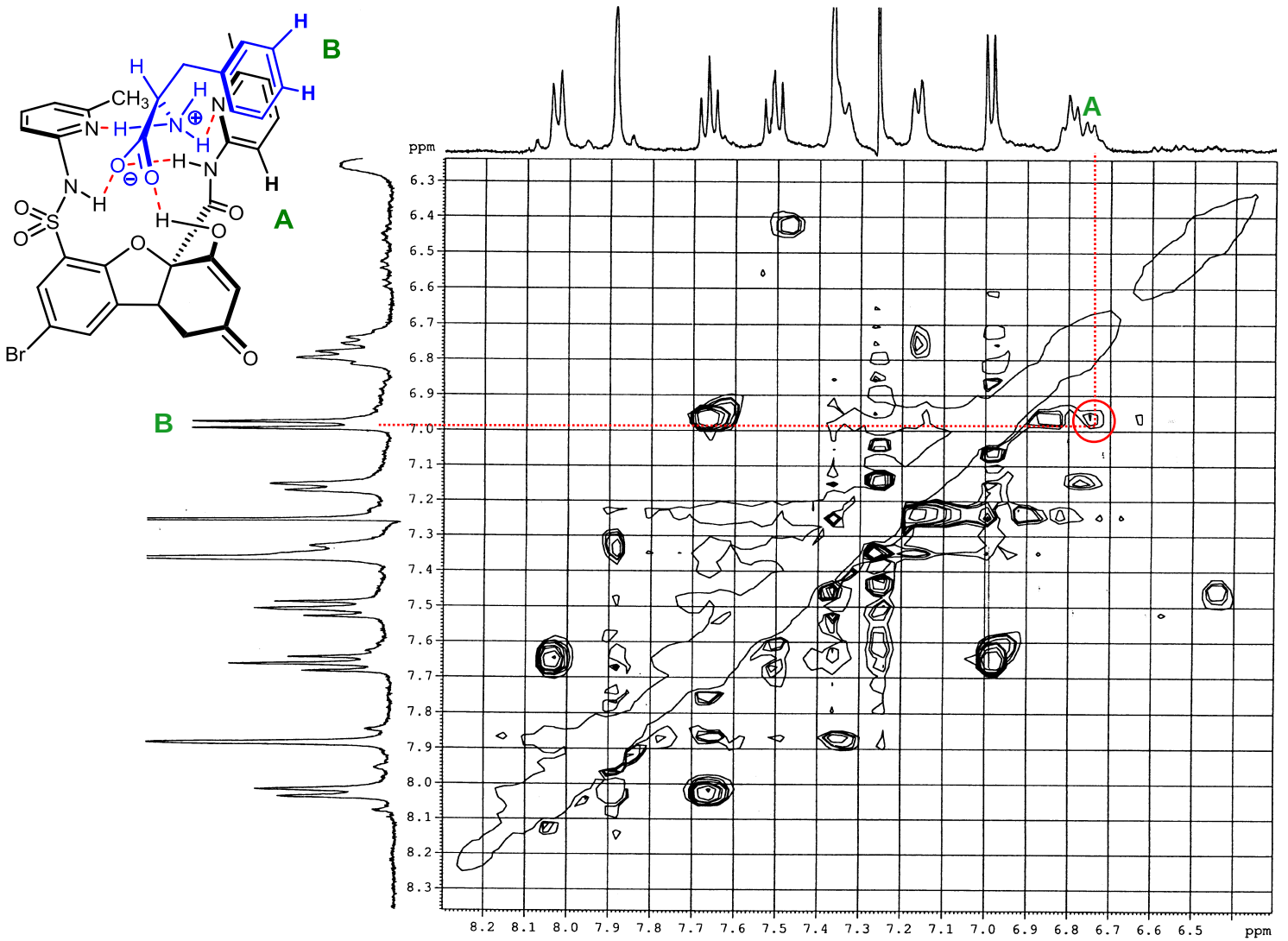


Figure S20. **^{1}H NMR spectrum of strong complex of receptor 1 with *L*-phenylalanine ($CDCl_3$)**.

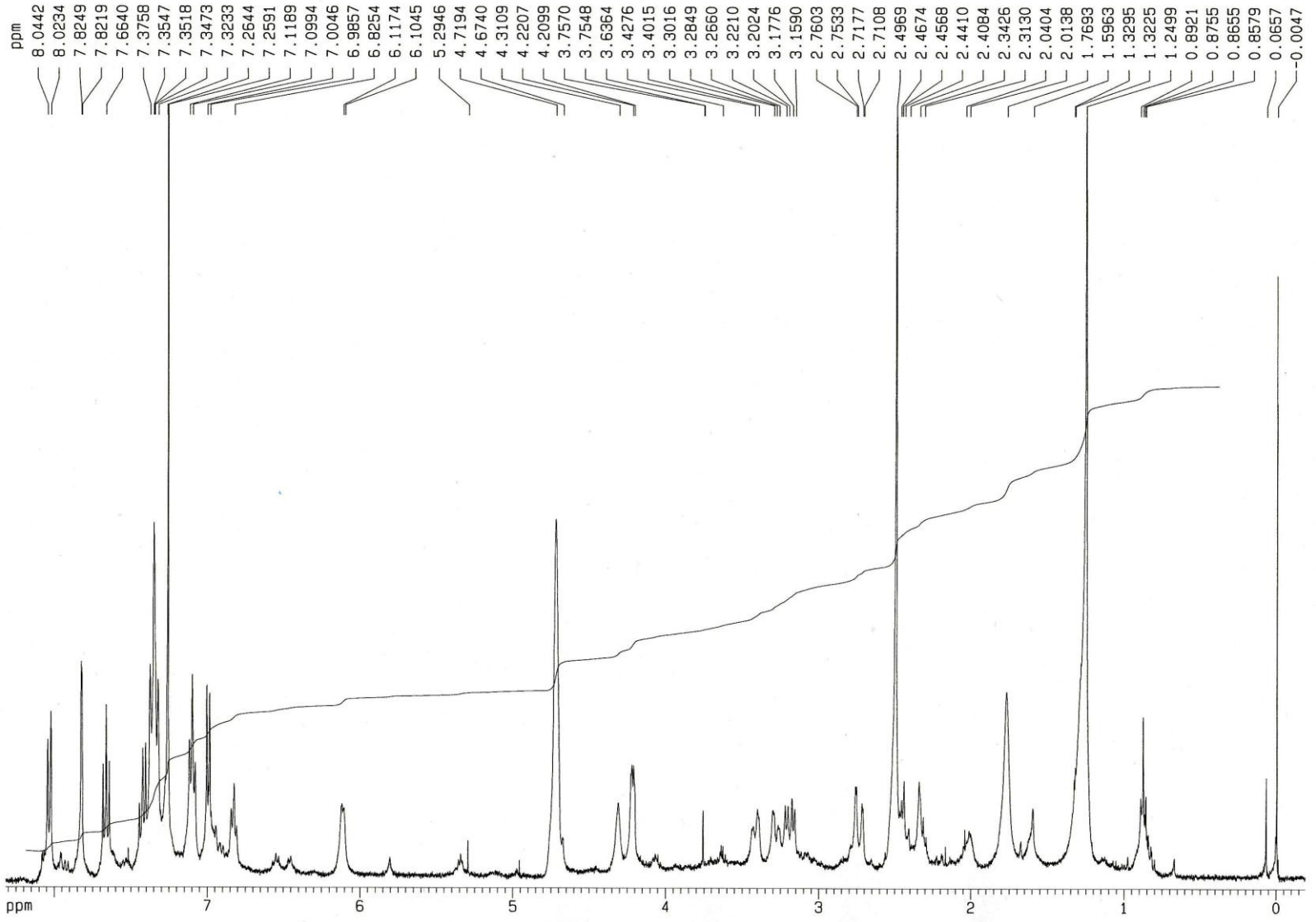


Figure S21. ROESY spectrum of strong complex of receptor 1 with *L*-phenylalanine ($CDCl_3$).

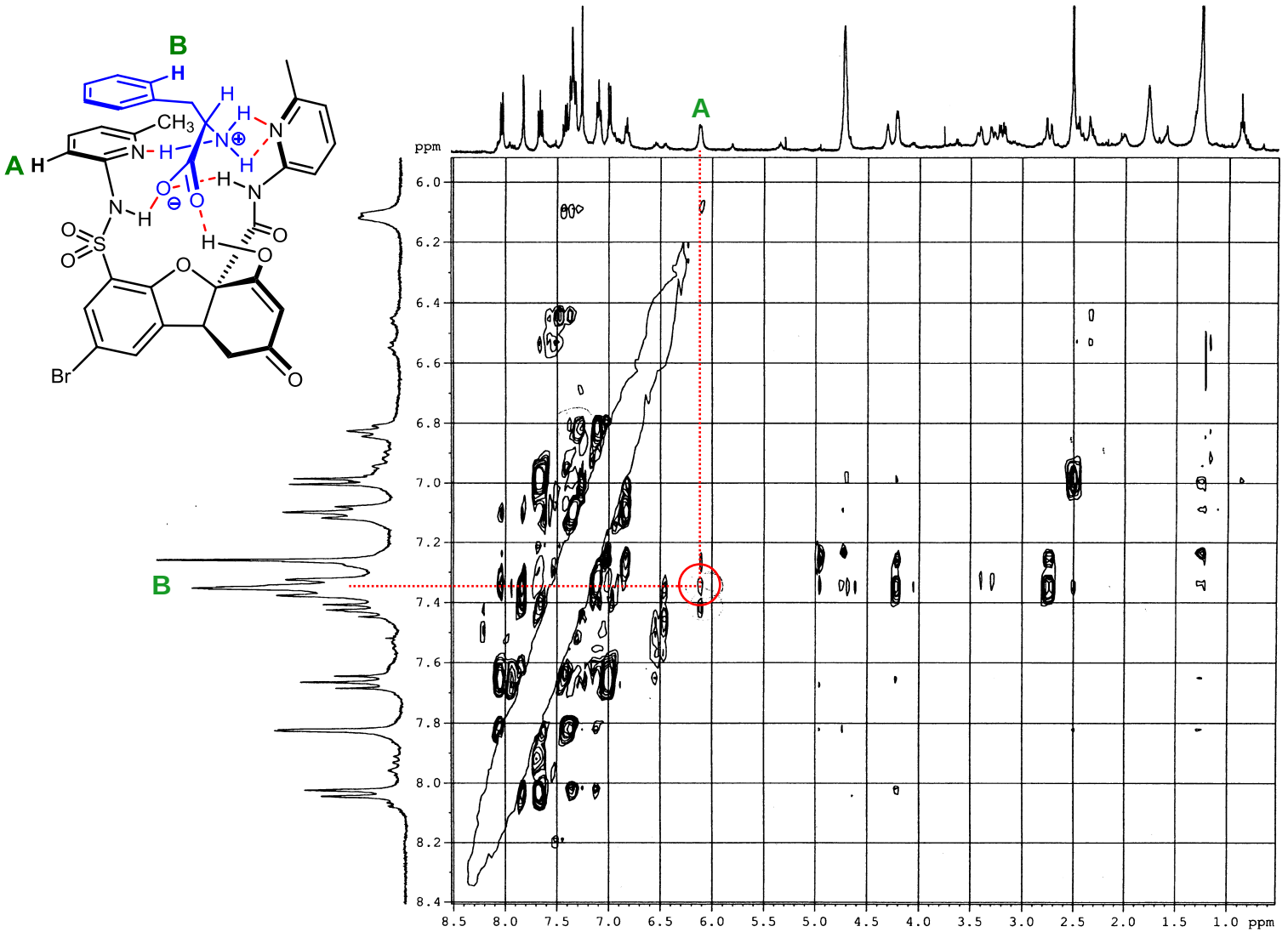


Figure S22. HPLC of extraction of racemic phenylalanine with the receptor 1.

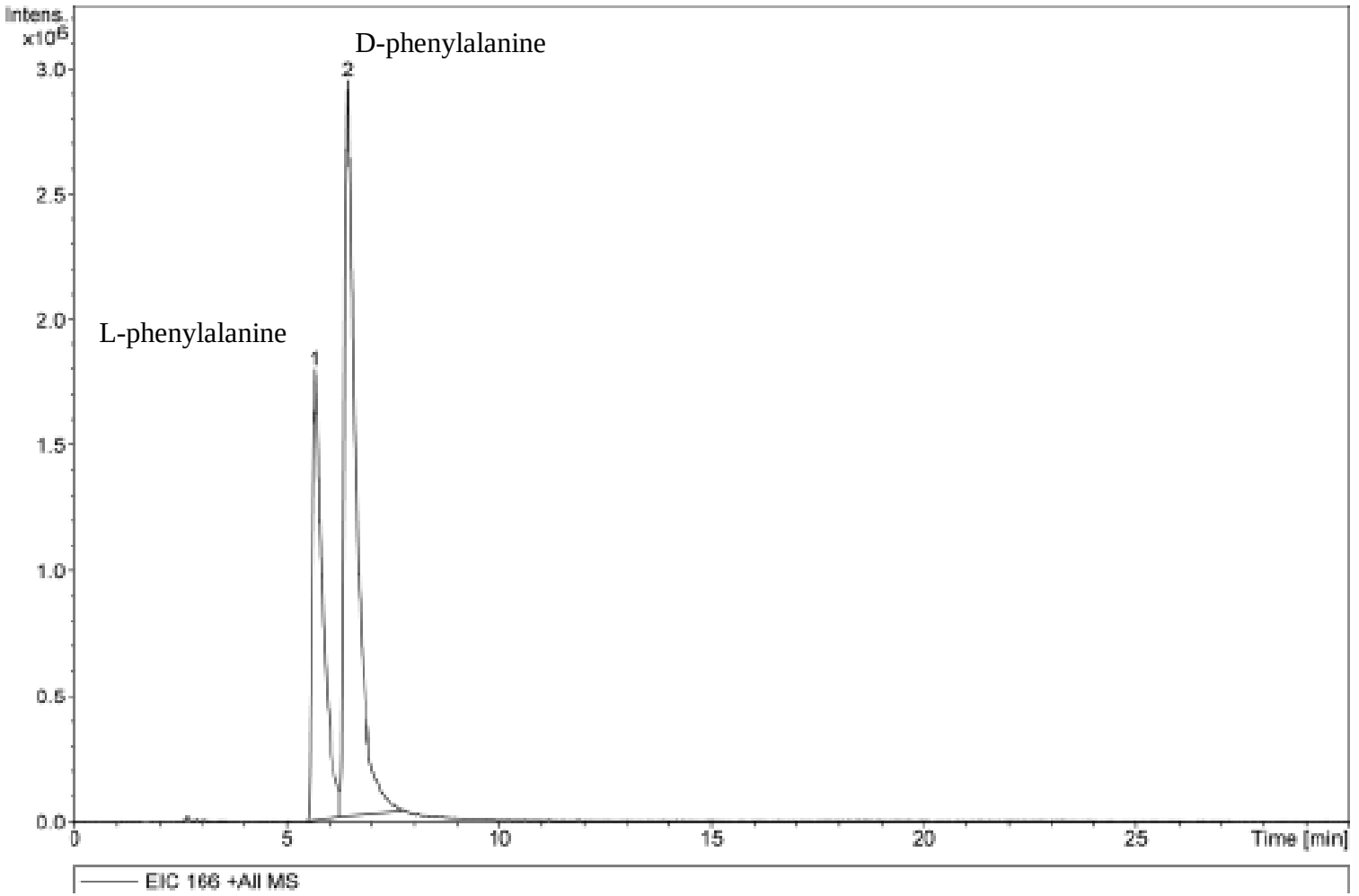
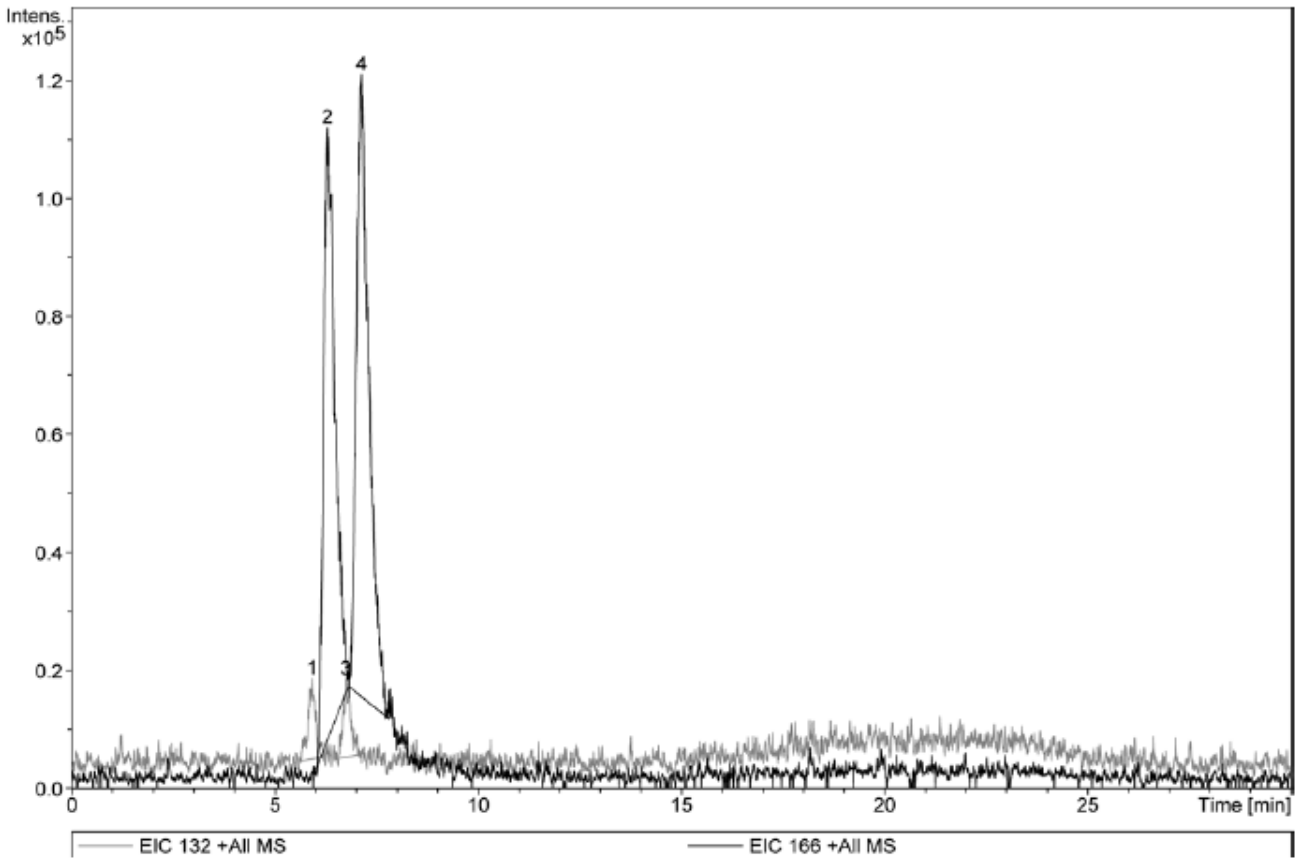


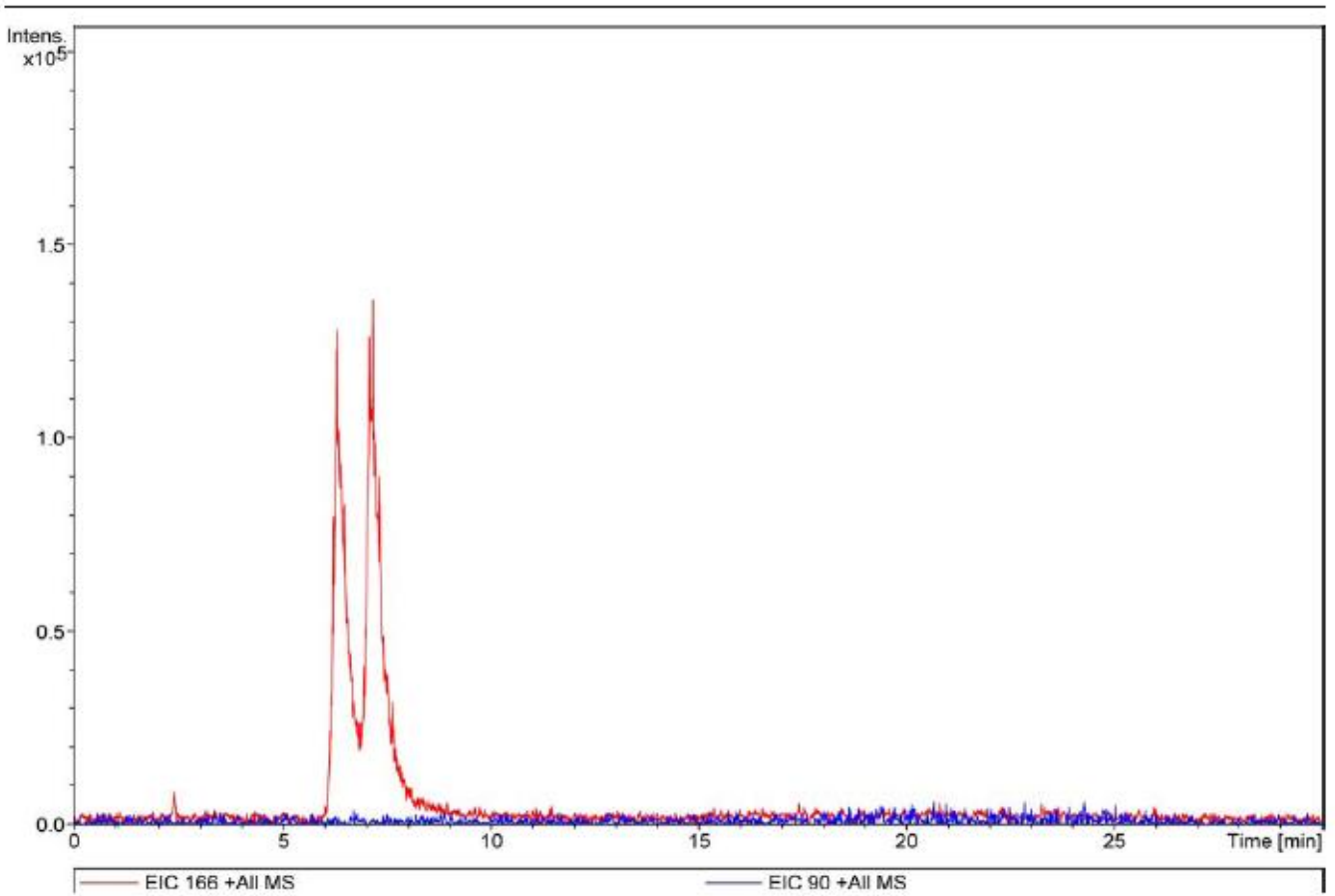
Figure S23. HPLC of extraction of *DL*-phenylalanine and *DL*-leucine with the receptor 1.



Compound List:

#	RT [min]	Range [min]	Height	Area	Area Frac %
1	5.9	5.6 - 6.2	13569	171601	3.7
2	6.3	6.0 - 6.8	103746	2060066	44.1
3	6.7	6.5 - 7.0	13231	159493	3.4
4	7.1	6.8 - 7.8	105429	2275637	48.8

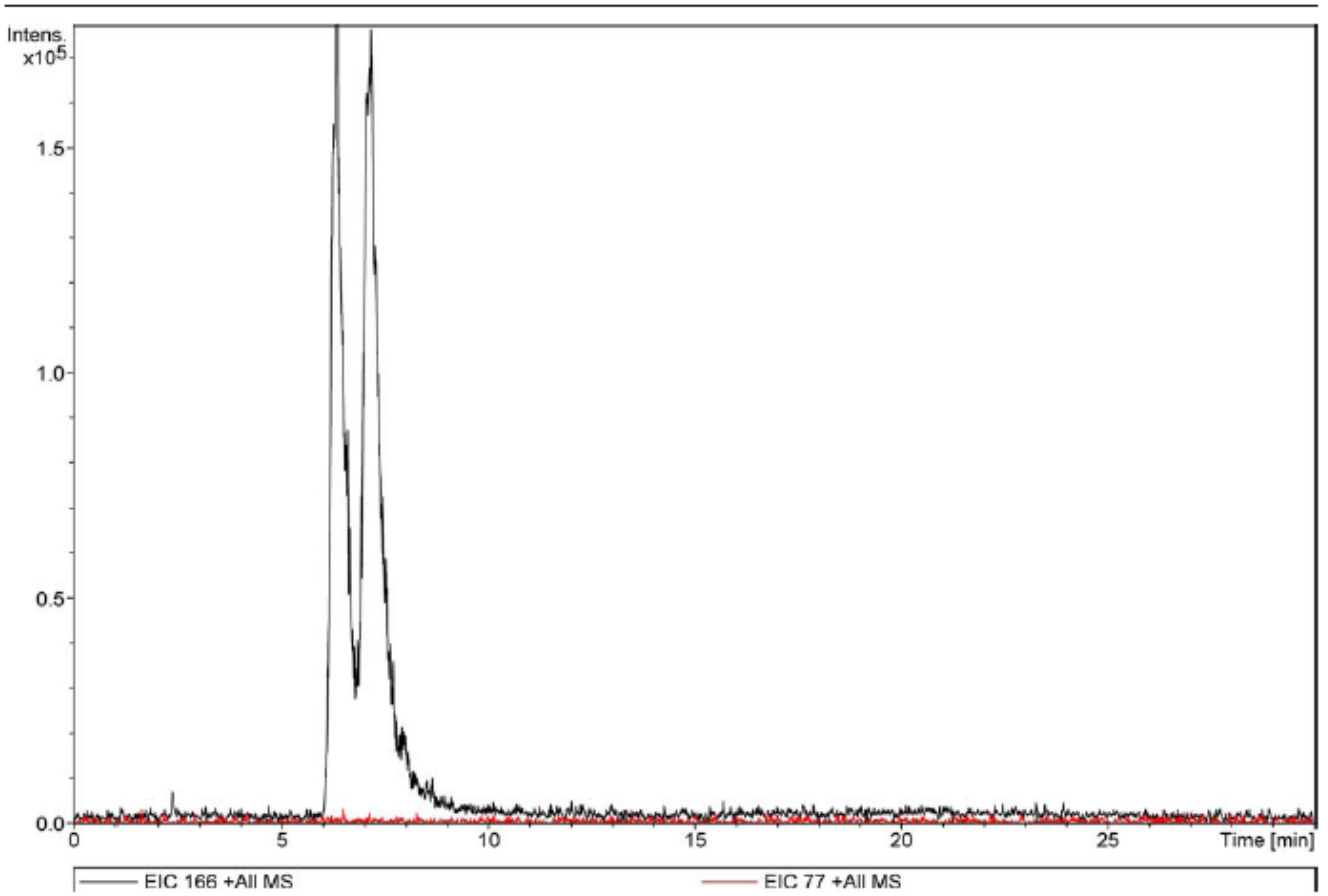
Figure S24. HPLC of extraction of *DL*-phenylalanine and *DL*-alanine with the receptor 1.



Compound List:

#	RT [min]	Range [min]	Height	Area	Area Frac %

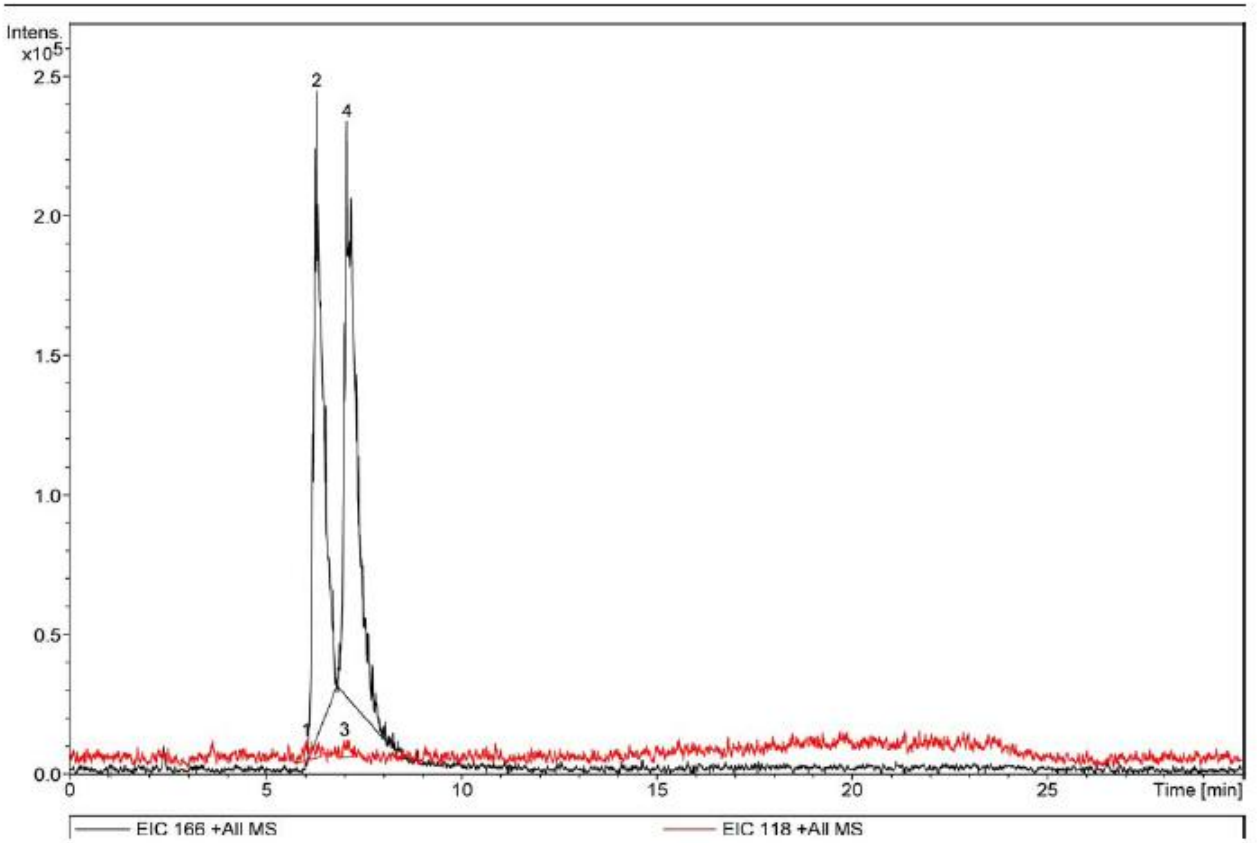
Figure S25. HPLC of extraction of *DL*-phenylalanine and glycine with the receptor 1.



Compound List:

#	RT [min]	Range [min]	Height	Area	Area Frac %

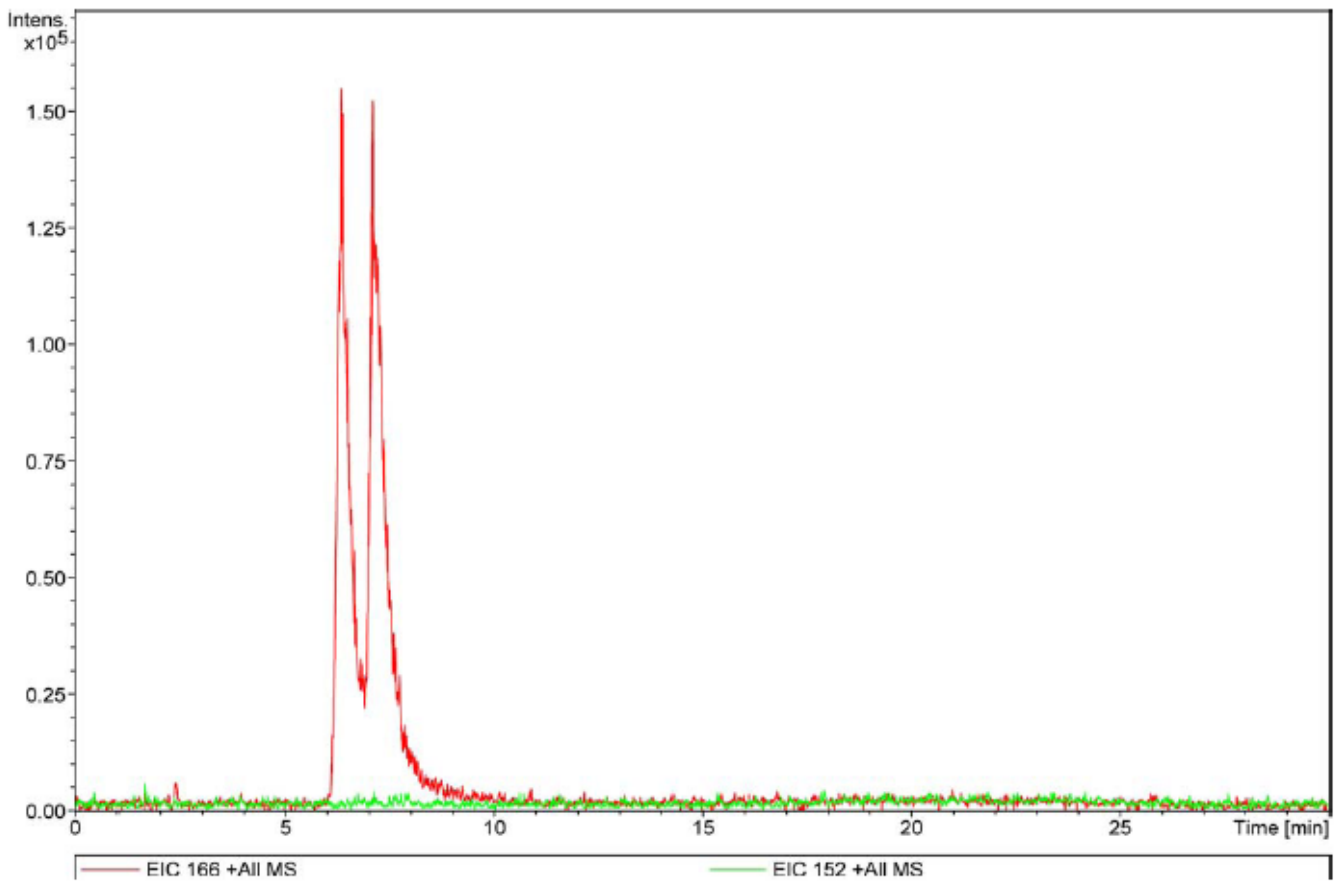
Figure S26. HPLC of extraction of *DL*-phenylalanine and *DL*-valine with the receptor 1.



Compound List:

#	RT [min]	Range [min]	Height	Area	Area Frac %
1	6.0	5.8 - 6.4	7057	109641	1.3
2	6.3	6.0 - 6.8	231830	3782135	46.3
3	7.0	6.8 - 7.3	6259	78253	1.0
4	7.1	6.8 - 8.1	206129	4196467	51.4

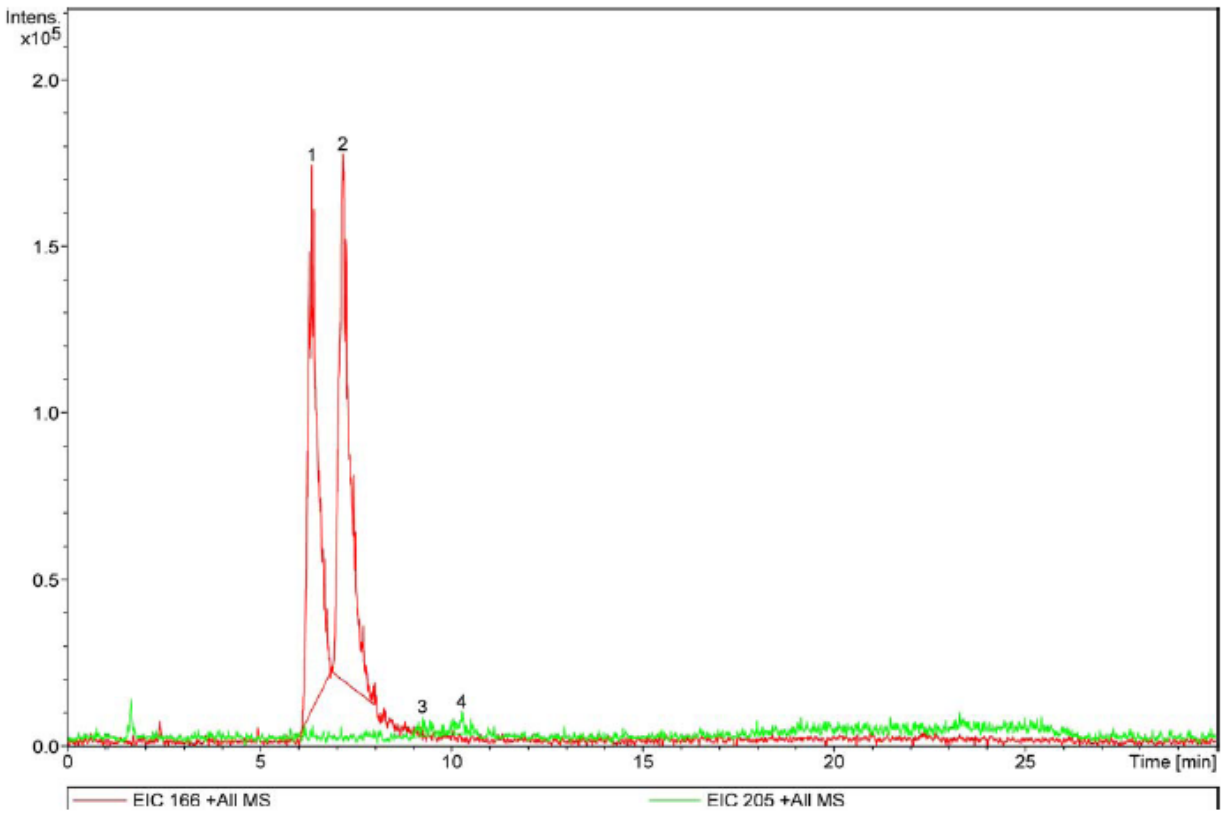
Figure S27. HPLC of extraction of *DL*-phenylalanine and *DL*-phenylglycine with the receptor 1.



Compound List:

#	RT [min]	Range [min]	Height	Area	Area Frac %

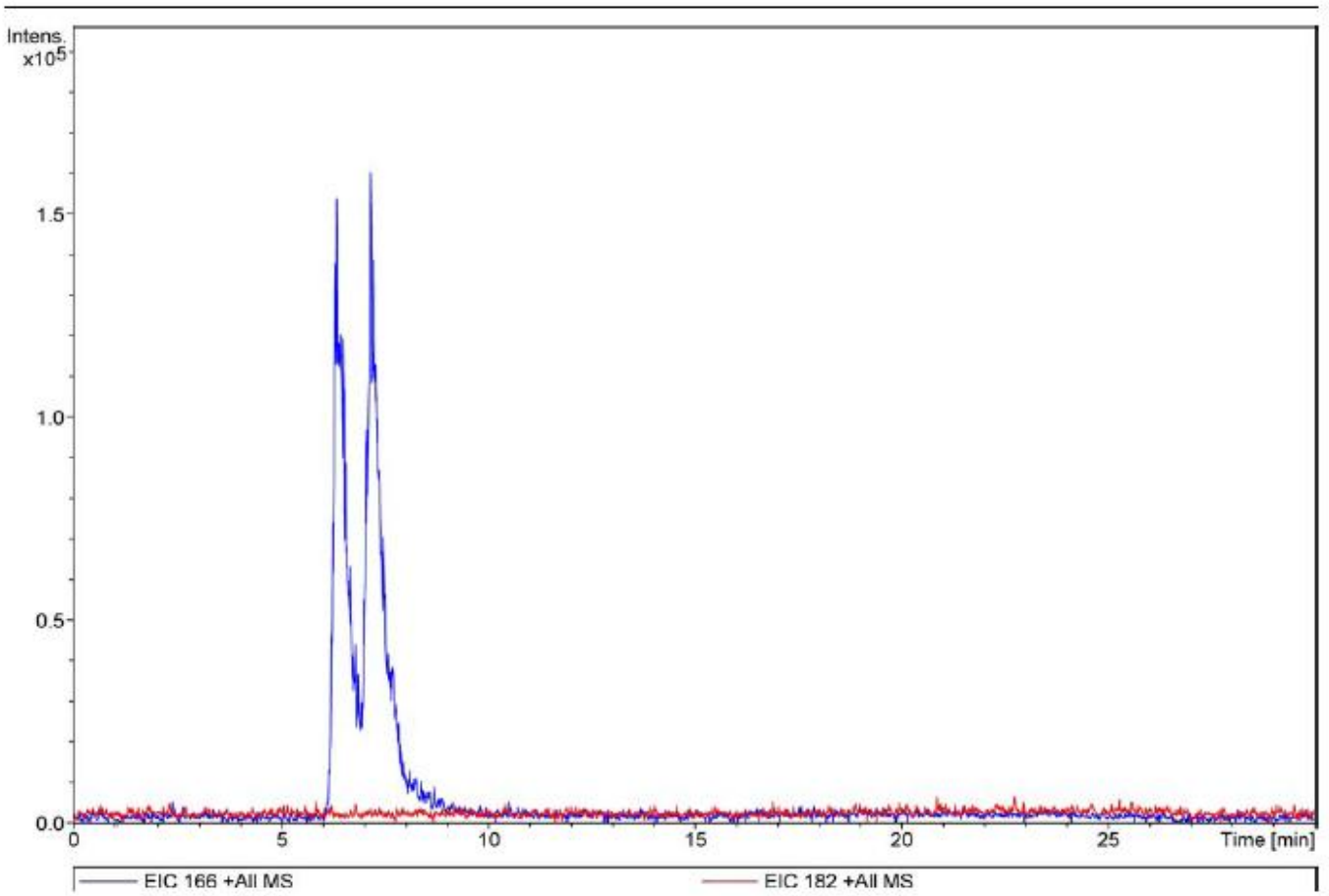
Figure S28. HPLC of extraction of *DL*-phenylalanine and *DL*-tryptophan with the receptor 1.



Compound List:

#	RT [min]	Range [min]	Height	Area	Area Frac %
1	6.3	6.0 - 6.8	163380	2614351	44.7
2	7.2	6.9 - 8.0	157848	3063395	52.4
3	9.2	8.9 - 9.6	5738	77297	1.3
4	10.3	9.8 - 10.6	6538	94775	1.6

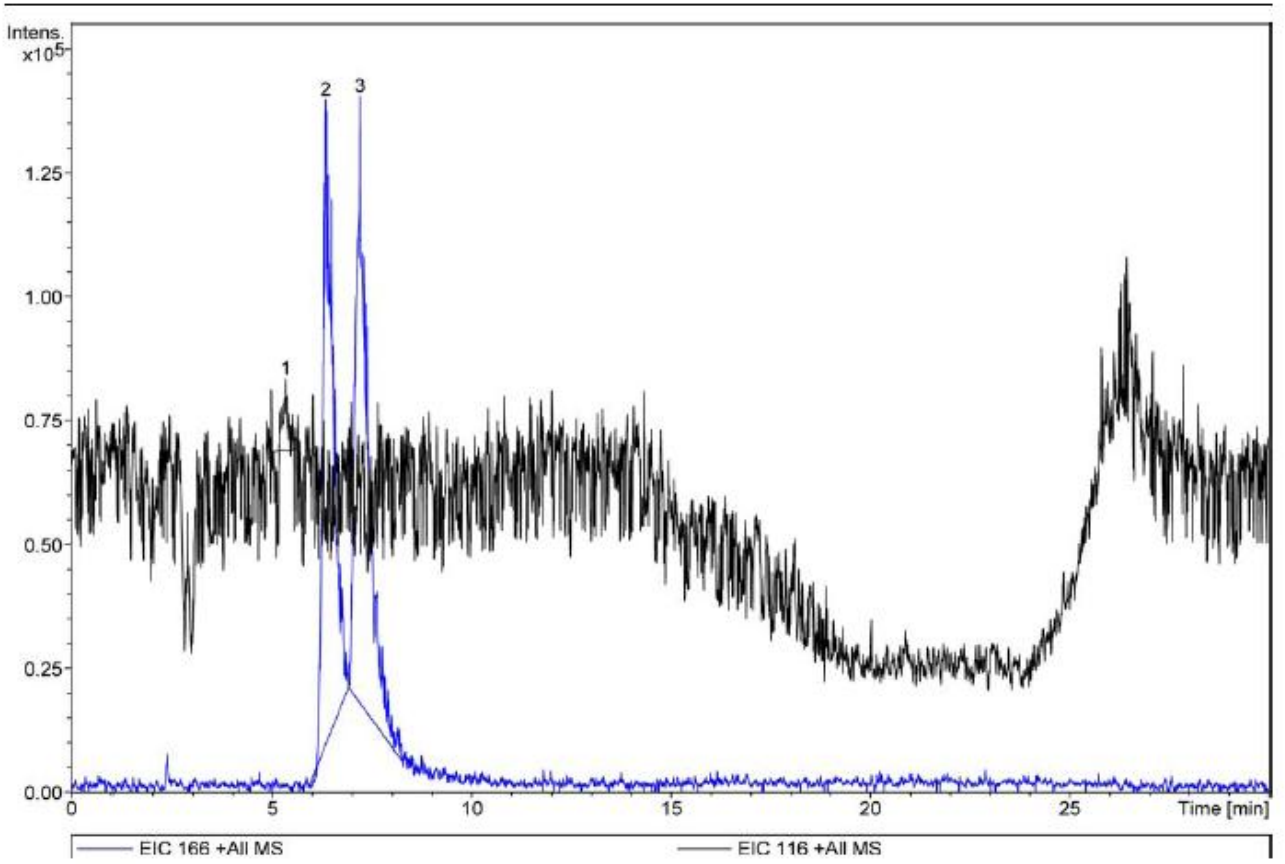
Figure S29. HPLC of extraction of *DL*-phenylalanine and *L*-tyrosine with the receptor 1.



Compound List:

#	RT [min]	Range [min]	Height	Area	Area Frac %

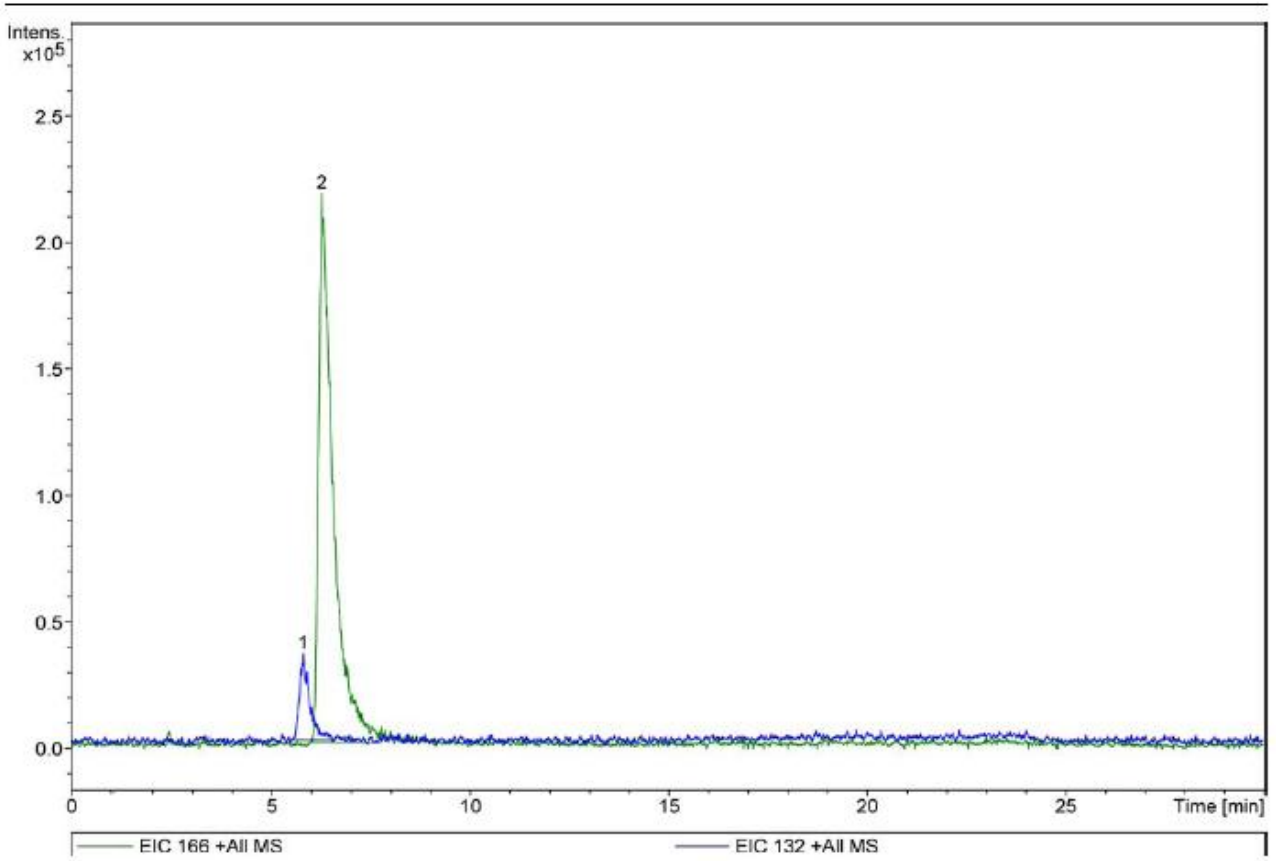
Figure S30. HPLC of extraction of *DL*-phenylalanine and *L*-proline with the receptor 1.



Compound List:

#	RT [min]	Range [min]	Height	Area	Area Frac %
1	5.3	5.0 - 5.6	14418	139387	2.8
2	6.3	5.9 - 6.9	129996	2278896	46.4
3	7.2	6.9 - 8.2	122380	2494640	50.8

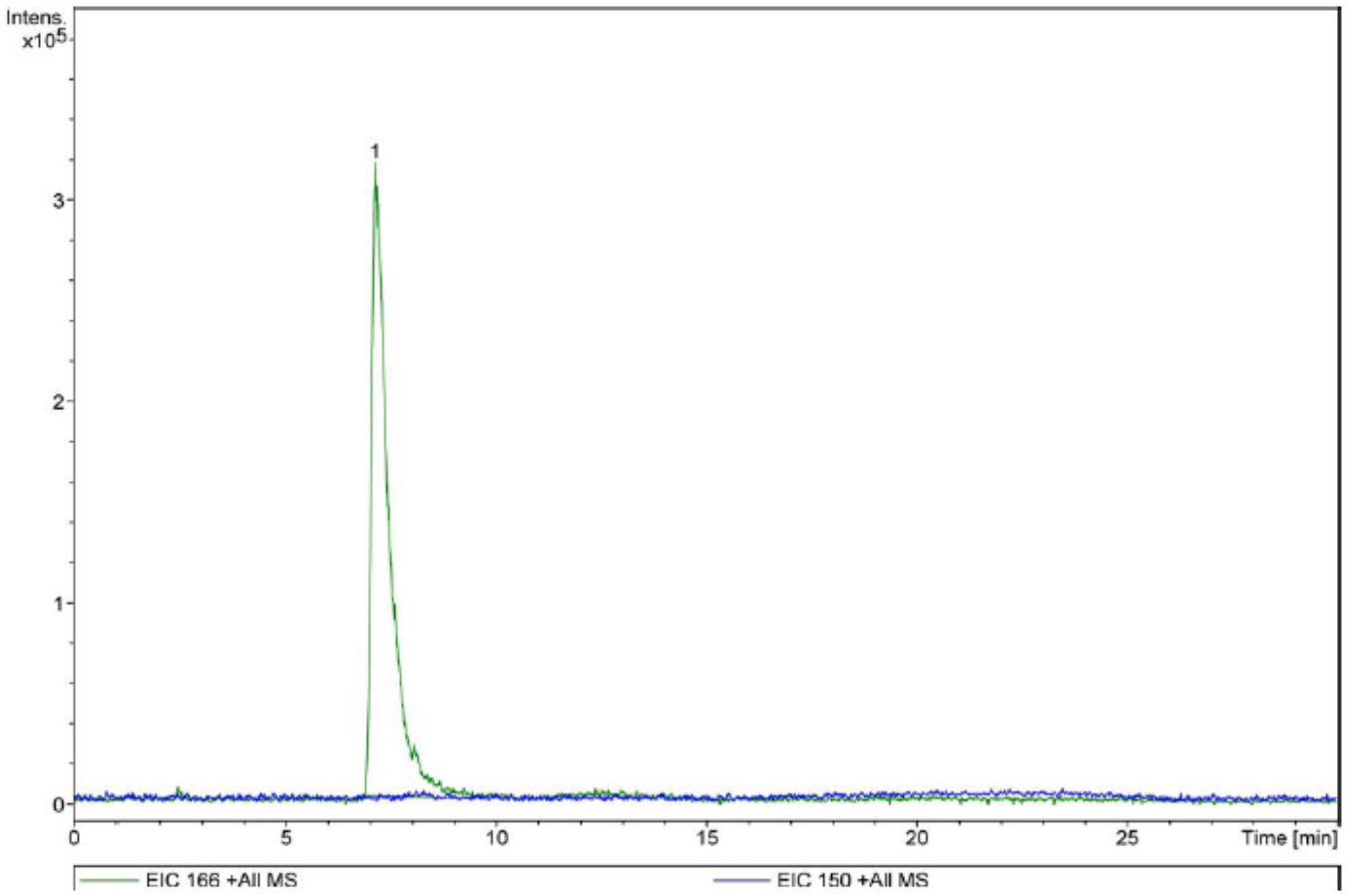
Figure S31. HPLC of extraction of *L*-phenylalanine and *L*-isoleucine with the receptor 1.



Compound List:

#	RT [min]	Range [min]	Height	Area	Area Frac %
1	5.8	5.5 - 6.5	34334	594900	9.7
2	6.3	6.0 - 7.7	217015	5558542	90.3

Figure S32. HPLC of extraction of *L*-phenylalanine and *L*-methionine with the receptor 1.



Compound List:

#	RT [min]	Range [min]	Height	Area	Area Frac %
1	7.1	6.7 - 9.4	314477	9118080	100.0

Figure S33. Circular dichroism spectrum for (2*S*, 3*R*)-receptor 1 (dioxane).

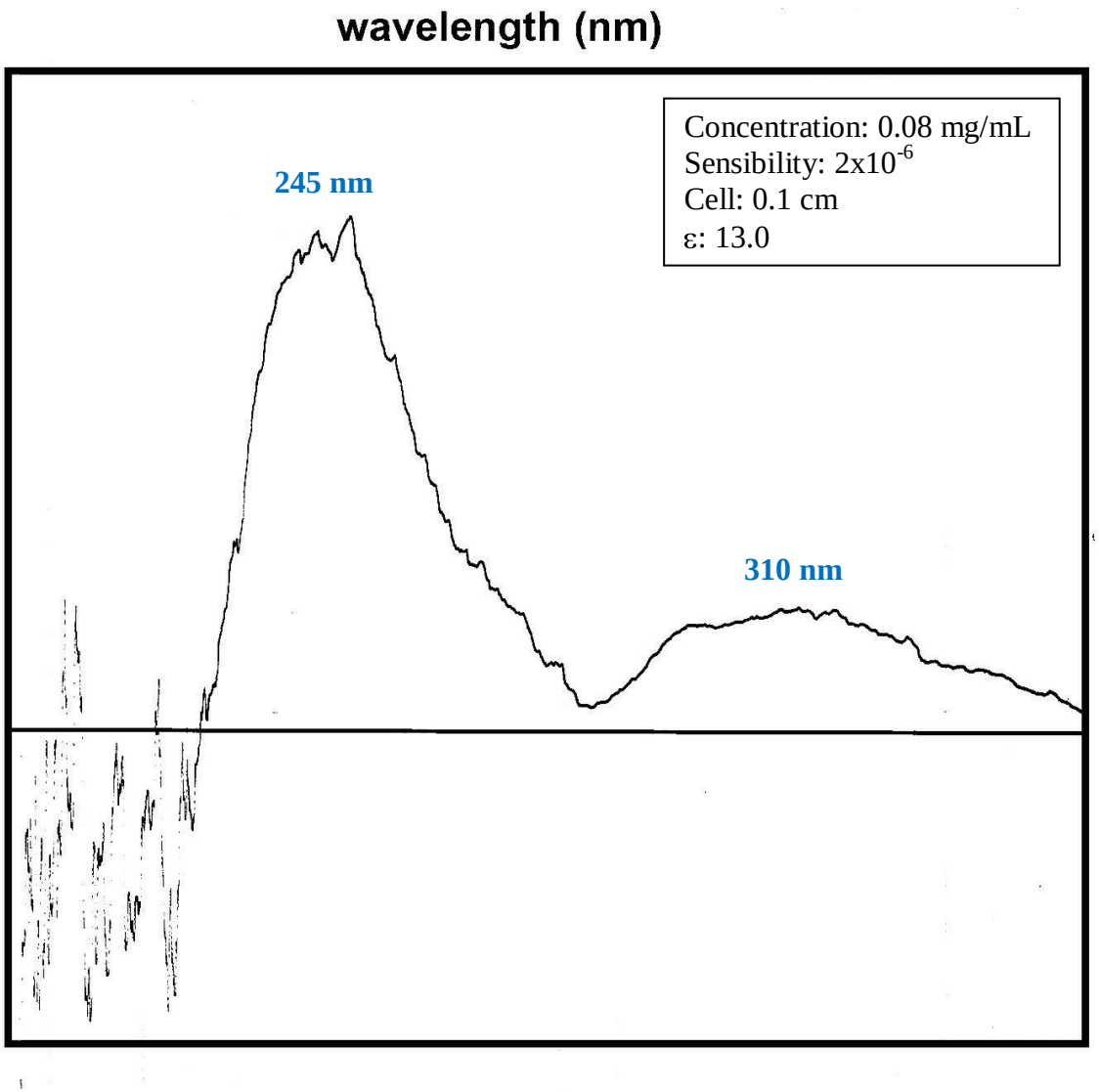


Figure S34. Circular dichroism spectrum for the complex of (2*S*, 3*R*)-receptor 1 and *L*-phenylalanine (CH_2Cl_2).

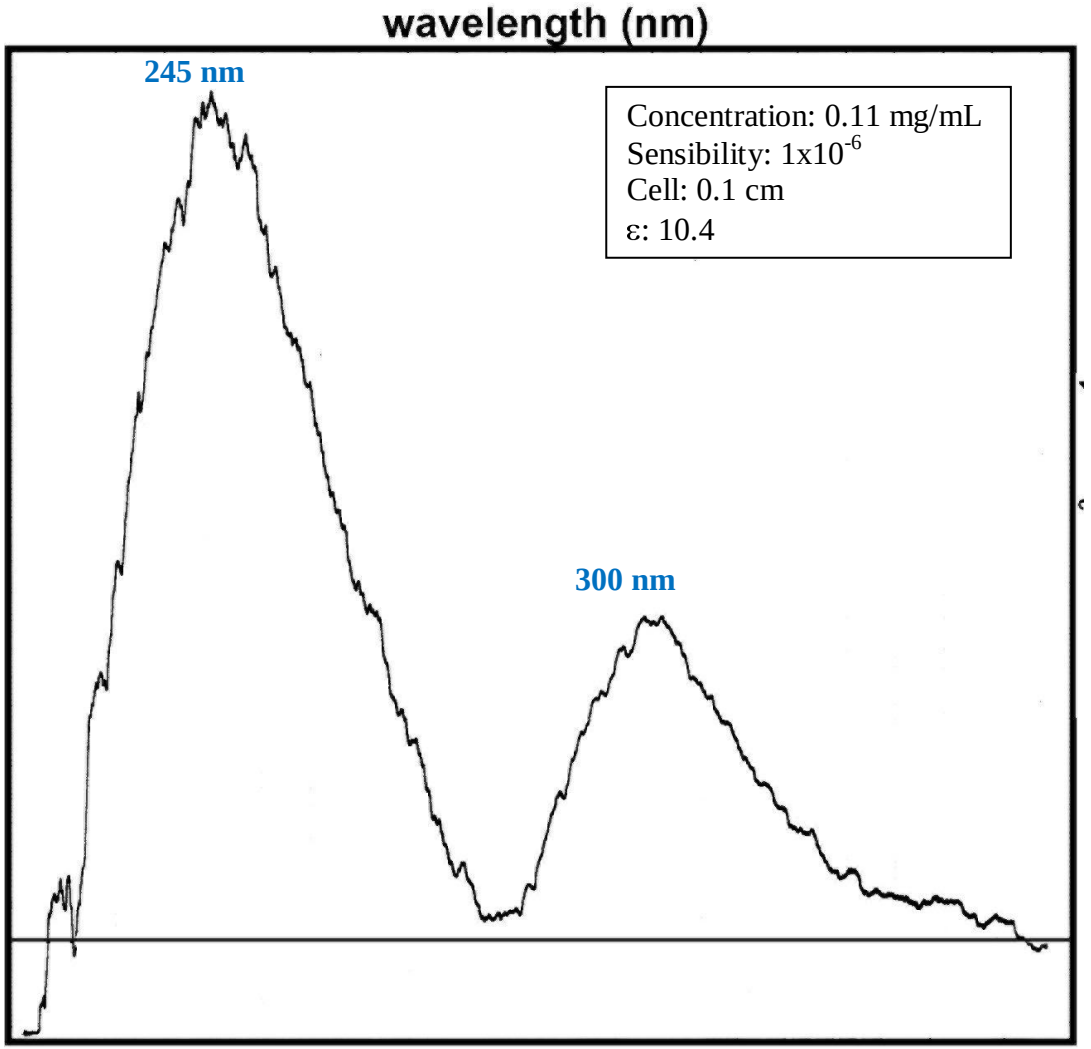


Figure S35. Circular dichroism spectrum for (2*R*, 3*S*)-receptor 1 (dioxane).

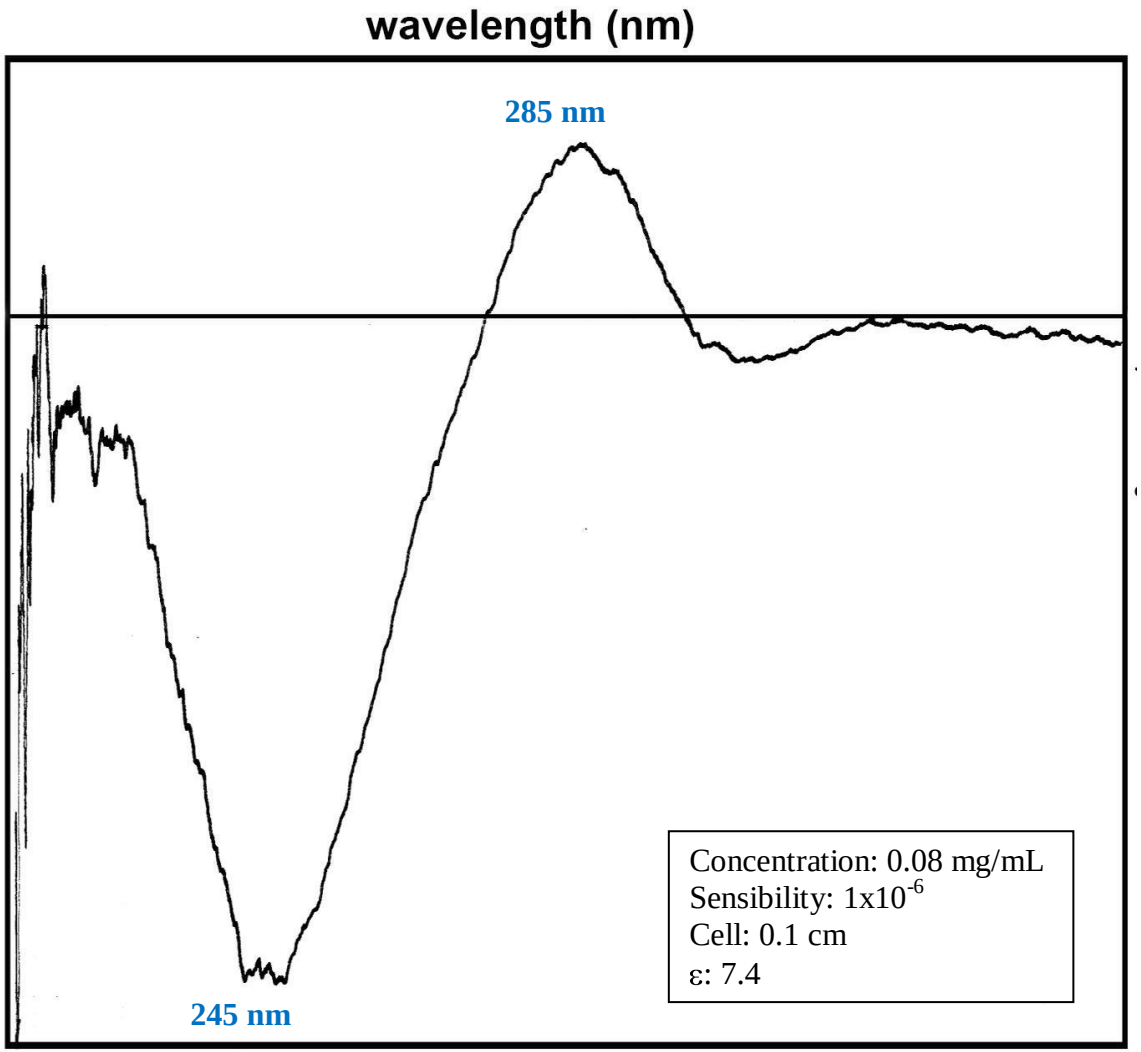


Figure S36. Circular dichroism spectrum for the complex of (2*R*,3*S*)-receptor 1 and *L*-phenylalanine (CH_2Cl_2).

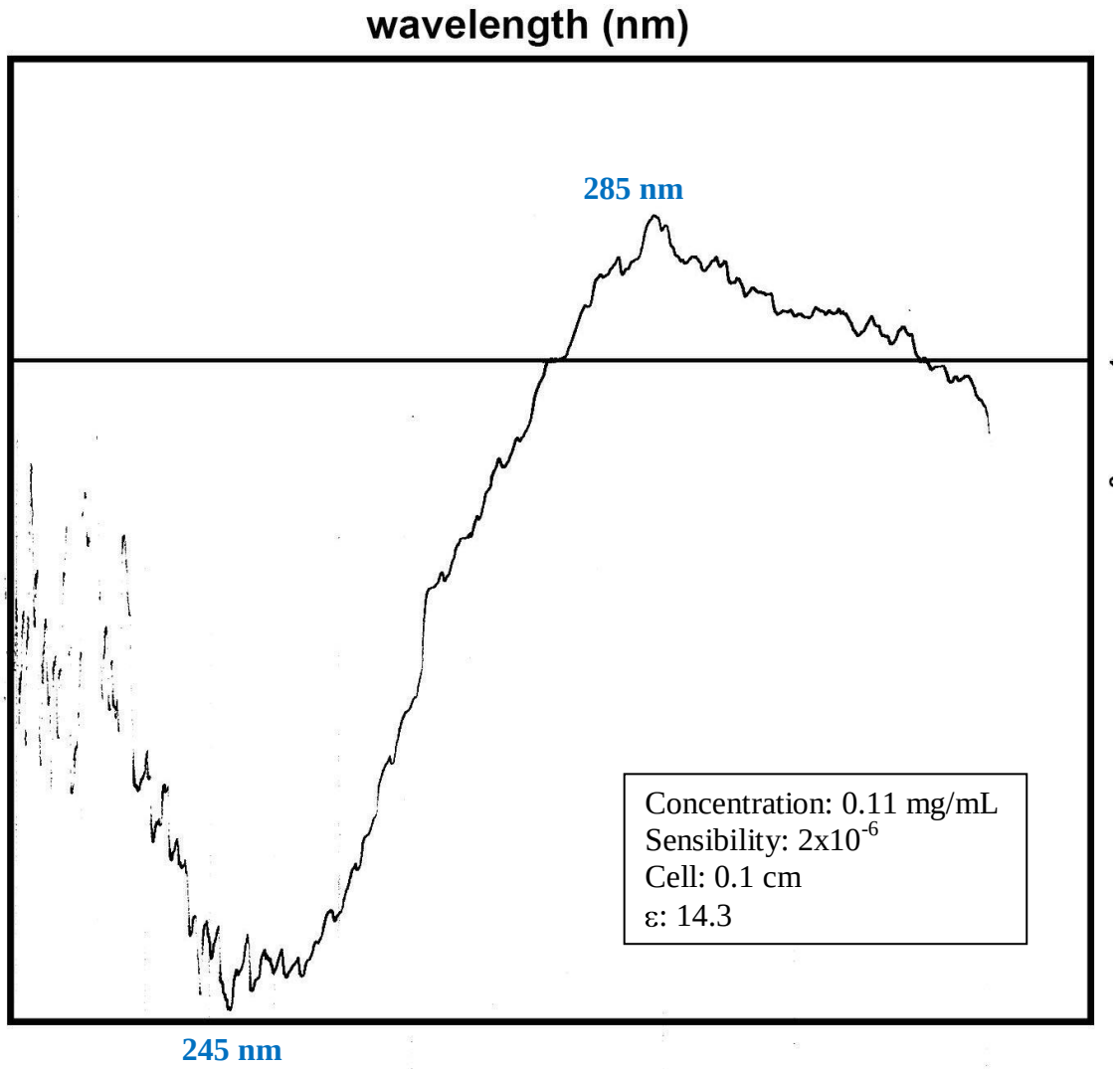
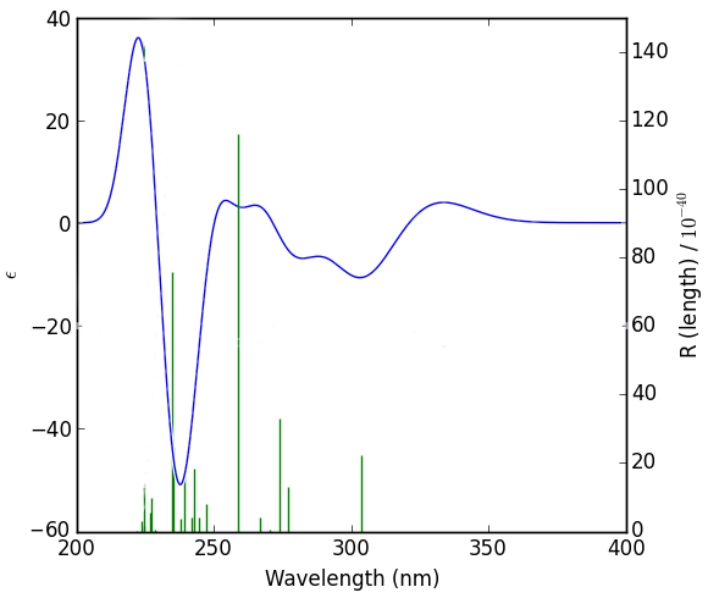


Figure S37. ECD spectra simulation for associates between receptor 1 and *L*-phenylalanine.

All calculations were performed using Gaussian09 software[1]. The structure of the complexes was optimized using M06-2X DFT functional [2] and 6-31G** basis set.[3] A time-dependent DFT calculation, including the 50 more stable singlet excited states, was performed using B98 functional [6] and 6-31G** basis set. This functional was chosen after considering its performance in the Truhlar benchmark database VES21 for electronic excitation energies involving valence excited states.[2] For both structures, the ECD spectra was simulated using a value of σ=0.4 using Gaussum software[7].

ECD spectra for the complex between *L*-phenylalanine and (2*R*,3*S*) receptor **1**. Comparison with experimental ECD spectra reveal that this absolute configuration corresponds to (-) receptor **1**.



ECD spectra for the complex between *L*-phenylalanine and (2*S*,3*R*) receptor **1**. Comparison with experimental ECD spectra reveal that this absolute configuration corresponds to (+) receptor **1**.

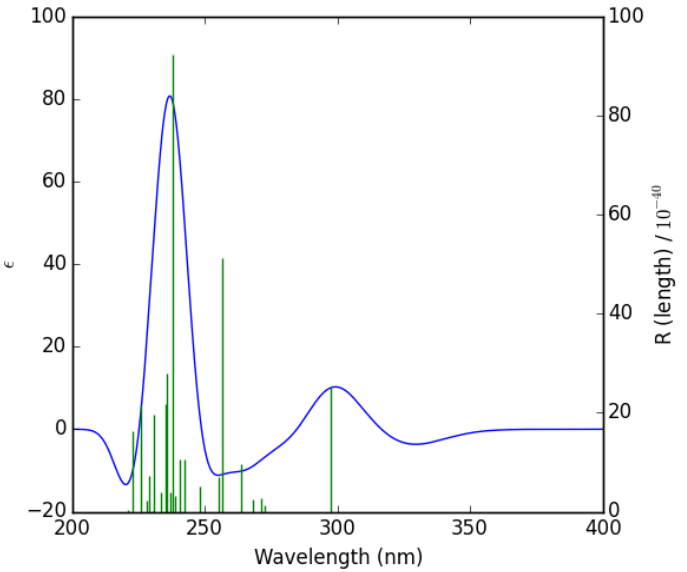


Figure S38. Modelling studies of associates between receptor 1 and *L*-phenylalanine.

All calculations were performed using Gaussian09 software.[1] Since it does not participate in the complex formation and in order to keep the calculations within reasonable limits, the bromine atom in the catalyst structure was substituted by a hydrogen atom. Geometry optimizations were performed using M06-2X DFT functional[2] and 6-31G** basis set[3] in the gas phase. Gibbs free energy correction at 298.15 K was calculated with this method and corrected according to the so-called "quasi-harmonic approach" using a free-rotor approximation for vibrational modes below 100 cm^{-1} and a rigid rotor approximation above this cutoff.[4] Single point energy was calculated on the optimized structures using M06-2X DFT functional[2] and 6-311+G** basis set. Solvent (chloroform) was modeled in this single-point calculation using SMD solvation model.[5] The Gibbs free energy of each structure is calculated adding to this single point energy the ΔG correction term calculated as described above.

Strong complex
ΔG: 0.00 kcal/mol

O	-1.77077	-0.49642	-0.29187
O	-3.55350	2.40108	-1.15114
N	-1.38858	2.21298	-0.40399
C	-2.08709	-4.13404	-0.74738
H	-1.50873	-5.04531	-0.63632
C	-3.40151	-4.16236	-1.19801
C	-4.14313	-2.97966	-1.31650
H	-5.17549	-3.01156	-1.65174
C	-3.54636	-1.77392	-0.99371
C	-4.02227	-0.33919	-1.04922
H	-3.88749	0.03798	-2.07049
C	-5.45519	-0.05798	-0.61405
H	-6.16122	-0.71905	-1.12358
C	-5.68998	-0.18485	0.88242
C	-4.56134	0.09763	1.76673
H	-4.74438	0.07320	2.83501
C	-3.30693	0.32018	1.32448
C	-2.95144	0.33056	-0.15938
C	-2.22636	-1.75447	-0.55513
C	-1.48765	-2.91668	-0.41208
C	-2.65726	1.76140	-0.63039
H	-0.72588	1.59458	0.08622
H	-5.70069	0.97181	-0.89460
C	-0.91171	3.49615	-0.66683
C	-1.71598	4.53864	-1.15587
C	0.95415	4.86818	-0.59852
C	-1.12062	5.77103	-1.35536

H	-2.76098	4.36353	-1.35792
C	0.23124	5.95262	-1.07462
H	-1.71728	6.59801	-1.72707
H	0.71537	6.91135	-1.21890
C	2.42225	4.98993	-0.28135
H	3.01936	4.37367	-0.96299
H	2.76093	6.02156	-0.38395
H	2.62460	4.66886	0.74594
N	0.39291	3.66362	-0.40372
O	-2.31259	0.55944	2.16858
O	-6.79228	-0.47690	1.30796
H	-3.86039	-5.11266	-1.44584
C	2.24195	1.71155	1.96815
H	2.40241	0.89925	-0.01533
N	2.40916	1.83052	0.48986
S	0.11026	-2.89043	0.35112
O	-0.05237	-2.55453	1.75178
O	0.80276	-4.10769	-0.04668
N	0.85274	-1.57361	-0.33538
H	0.69084	-0.71706	0.22620
H	1.61136	2.40390	0.10461
C	1.88227	-1.55097	-1.26294
C	2.17265	-2.61688	-2.13160
C	3.21546	-2.45307	-3.02100
H	1.60097	-3.53241	-2.08536
C	3.56538	-0.25044	-2.17082
C	3.93264	-1.25638	-3.05111
H	3.46734	-3.25854	-3.70372
H	4.74969	-1.10365	-3.74636
N	2.56129	-0.39907	-1.29055
C	4.28442	1.07106	-2.14036
H	3.57055	1.89680	-2.23191
H	4.82791	1.17476	-1.19359
H	5.00536	1.15095	-2.95515
C	0.93661	0.93754	2.26976
O	0.70427	0.68460	3.43601
O	0.22814	0.65801	1.23018
H	-1.40174	0.40485	1.80407
H	3.29097	2.29092	0.25903
C	3.46113	1.06436	2.62789
H	3.16886	0.88450	3.66618
H	4.30212	1.76806	2.63059
C	3.87079	-0.21753	1.94115
C	5.01315	-0.24546	1.13635
C	3.09138	-1.37410	2.05378
C	5.36746	-1.39713	0.43707
H	5.64282	0.64123	1.07380
C	3.45377	-2.53046	1.36703
H	2.19977	-1.37445	2.67747
C	4.58175	-2.54091	0.54787
H	6.25492	-1.40078	-0.18892
H	2.83979	-3.42171	1.45820
H	4.84689	-3.44054	0.00169
H	2.10354	2.72699	2.35612

Weak Complex

ΔG: 1.2 kcal/mol

O	-1.64820	0.39295	-0.17957
O	0.21210	2.88874	-1.87826
N	0.93145	0.91092	-0.96050
C	-5.12989	-0.80976	-0.10972
H	-5.72738	-1.62366	0.28776

C	-5.70539	0.20137	-0.86882
C	-4.92655	1.26109	-1.34810
H	-5.38289	2.06346	-1.92018
C	-3.57023	1.27557	-1.07765
C	-2.46463	2.22185	-1.48030
H	-2.14280	1.97360	-2.49928
C	-2.77835	3.71313	-1.42410
H	-3.71169	3.93946	-1.94623
C	-2.87584	4.27839	-0.01736
C	-2.03747	3.66237	1.01083
H	-2.01673	4.13094	1.98837
C	-1.33394	2.53075	0.81317
C	-1.34657	1.77116	-0.50719
C	-2.99818	0.24922	-0.33260
C	-3.76057	-0.78933	0.17851
C	0.02219	1.90560	-1.18428
H	0.70164	0.20273	-0.25589
H	-1.97072	4.25088	-1.93107
C	2.20151	0.82860	-1.53901
C	2.66749	1.73351	-2.50672
C	4.16166	-0.40229	-1.64167
C	3.92527	1.51600	-3.04015
H	2.05514	2.56665	-2.81077
C	4.69192	0.43758	-2.61279
H	4.30967	2.19798	-3.79205
H	5.68169	0.25074	-3.01339
C	4.94018	-1.58952	-1.14176
H	4.54141	-2.52362	-1.55087
H	5.98925	-1.51694	-1.43407
H	4.89961	-1.63860	-0.05067
N	2.93857	-0.20880	-1.11946
O	-0.55651	2.01895	1.76124
O	-3.58850	5.23700	0.21431
H	-6.76944	0.17691	-1.07412
C	2.18857	-1.59189	2.46585
H	1.41095	-2.35843	0.63073
N	2.31616	-1.99710	1.03125
S	-3.12267	-1.93946	1.37382
O	-2.92856	-1.24272	2.62628
O	-3.95084	-3.13074	1.28567
N	-1.58638	-2.30917	0.86594
H	-0.89042	-1.60657	1.18729
H	2.55397	-1.18664	0.39742
C	-1.25569	-2.97653	-0.30644
C	-2.20175	-3.59207	-1.14279
C	-1.73813	-4.22243	-2.28131
H	-3.25223	-3.58181	-0.88470
C	0.49046	-3.61377	-1.67933
C	-0.37359	-4.23627	-2.56811
H	-2.44266	-4.70513	-2.95124
H	0.01311	-4.71965	-3.45744
N	0.05835	-2.99867	-0.56578
C	1.97566	-3.58316	-1.91129
H	2.32693	-2.54679	-1.89068
H	2.49841	-4.14452	-1.12762
H	2.23945	-4.02604	-2.87310
C	0.94049	-0.69768	2.62173
O	0.32894	-0.44866	1.51382
O	0.69507	-0.27877	3.73528
H	-0.38385	1.03659	1.66978
H	3.03752	-2.70792	0.90351
C	3.42965	-0.84253	2.97059
H	4.29108	-1.52010	3.00761
H	3.18397	-0.56148	3.99932

C	3.78065	0.37890	2.14399
C	2.87481	1.43617	1.97300
C	5.02653	0.46165	1.51878
C	3.18887	2.50387	1.13889
H	1.92869	1.44951	2.50388
C	5.34694	1.53744	0.69264
H	5.75981	-0.32522	1.68428
C	4.41797	2.55182	0.48340
H	2.46135	3.29885	1.00377
H	6.31864	1.57771	0.20930
H	4.65488	3.38326	-0.17278
H	2.03724	-2.50221	3.05232

1. R. B. Gaussian 09, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. A., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Journal* **2009**.
2. H. Zhao and D. G. Truhlar, *Theor Chem Acta* **2007**, *120*, 215-241.
3. R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.* **1980**, *72*, 650-654.
4. (a) S. Grimme, *Chem. Eur. J*. **2012**, *18*, 9955-9964; (b) R. Paton, https://github.com/bobbypaton/compchem.
5. A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B* **2009**, *113*, 6378-6396.
6. H. L. Schmider and A. D. Becke, *J. Chem. Phys.*, 1998, **109**, 8188-8199.
7. N. M. O'Boyle, A. L. Tenderholt and K. M. Langner, *J. Comput. Chem.*, 2008, **29**, 839-845.