ELECTRONIC SUPPLEMENTARY INFORMATION

Camphor-based CCR5 blockers lead compounds – a computational and experimental approach

Gonçalo C. Justino[a]*, Pedro F. Pinheiro[a], Alexandra P.S. Roseiro[a], Ana S.O. Knittel, João Gonçalves[b,c], Marta C. Justino[e] and M. Fernanda N.N. Carvalho[a]

a- Centro de Química Estrutural, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1049-001 Lisboa, Portugal
b- URIA-Centro de Patogénese Molecular, Faculdade de Farmácia, Universidade de Lisboa, Lisboa, Portugal
c- Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal.

CONTENTS

Computational details of the homology modelling procedures

A library of camphor derivatives (Figure S1) was created and their ability to bind CCR5 evaluated by docking to a CCR5 model (a homology modelling approach was used to obtain a structure of the protein), since no experimental CCR5 structure exists.

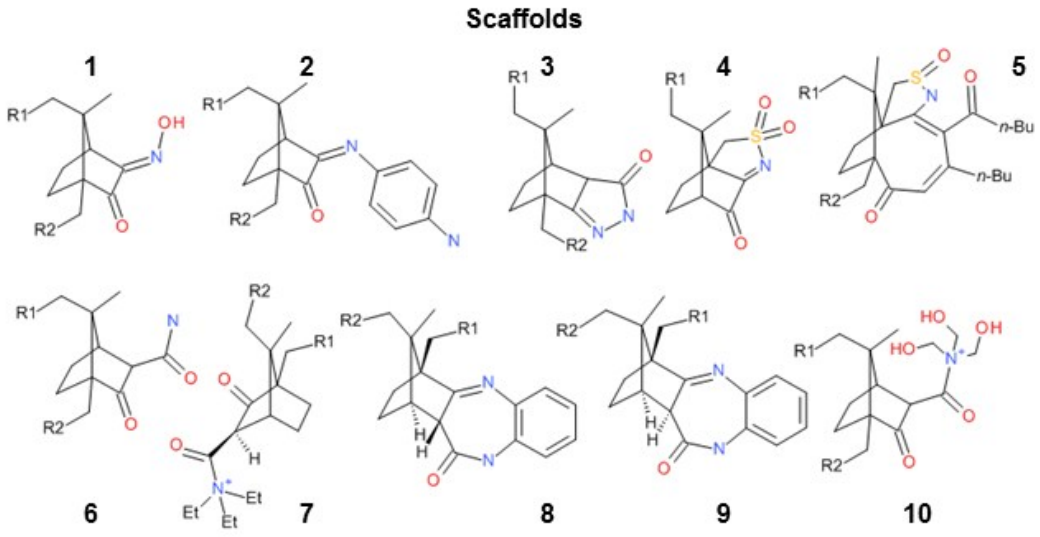

Figure S1. Structure of the scaffolds used to build the virtual library.

Modelling the CCR5 receptor

Homology modelling

The sequence of the CCR5 protein was taken from the NCBI Protein database (accession code AAB57793.1). A protein-protein BLAST search (blastp)[1] was performed to find other proteins with regions of similarity to this sequence, using the database of non-redundant protein sequences of human origin. In parallel, a template search was

performed within the Swiss-Model workspace using the default parameters.[2–5] The MUlti-Source ThreadER (MUSTER) of Wu and Zhang[6] was also used to identify the adequate templates. All approaches indicated that the sequence codes for a member of the Rhodopsin-like GPCR superfamily.
For all identified templates, aligned sequences with an E value below 10^{-10} in Swiss-model (for a minimum of 10^{-15}, from the gapped blast approach) or below 10^{-75} in NCBI's BLAST, the available tri-dimensional structures were retrieved from the PDB database[7] and used as templates for homology modelling using the Swiss-Model server's Automated Mode; all ten MUSTER predicted templates were also used for modelling using Swiss-Model. All obtained models were analysed in terms of their QMean[8] scoring and local model quality estimation by Anolea,[9] Qmean and Gromos.[10]

Stereochemical quality of the model

The stereochemical quality was checked using the PROCHECK[11] package as implemented both in Swiss-Model and in PSVS[11]. The structural motifs in the protein were analysed using PROMOTIF,[12] as implemented in Swiss-Model.
The stereochemical analysis of the 3oe6A-based model indicates that the few residues raising concerns are, except two, located in the intracellular and extracellular domains of the protein. The Ramachandran analysis did not find residues in the disallowed regions, and only two are found in generously allowed regions, Asn268^{EC3} and Arg232^{IC3}, where EC and IC stand for extracellular and intracellular loops, respectively, numbered from the N terminal.
Of all the non-glycine, non-proline, non-start and non-end 256 protein residues, 232 (90.6 %) are located in the most favoured regions, and 8.6 % in the additionally allowed regions. All C-α ζ values are within 2.0 standard deviations of the ideal values. ω torsion deviations are all within accepted values except for those pertaining to residues Ala91^{EC1}, Ala92^{EC1}, Val134^{IC2}, Ile165$^{TM/EC2}$, Lys171^{EC2}, His231^{IC3}, and Asn267^{EC3}, where TM denotes the transmembrane location of these residues. Ile165, a transmembrane residue, is located at the extracellular top of an helix formed by residues 144 to 165 or 166, and is not close to the described binding residues. Deviation from mean χ values is also largely within acceptable values, except for residues Met64^{IC1}, Asn98^{EC1}, Ile295 and Lys303^{EC}, of which none is located within the membrane; Ile295 belongs to the intracellular terminal C-segment of the protein. G factors are acceptable for all residues except Met64^{IC1} and Trp190^{EC2}. No residues within 15 Å of the center of the described binding sites show deviations from ideal values for main-chain bond lengths and angles, and planarity. The GROMOS energy environment analysis of the model indicates that the vast majority of the amino acid residues are located within a favourable environment, with energies ranging from 0 to -300 kcal/mol, while there are two residues that are in highly unfavourable environments, with energies at aprox. 200 kcal/mol – Ser63^{IC1} and Ser180^{EC2}.

Upon loop refinement, no residues were found in the generously allowed and forbidden areas of the Ramachandran plot, one main-chain bond deviation of 0.1 Å from the ideal value was found (bond Met64^{TM}-Thr65^{TM}, for an acceptable difference of 0.05 Å), two residues had main-chain bond angles deviation of 14.3 (N-C_A-C_B angle in Met 64^{TM}) and 15.0 degrees (N-C_A-C_B angle in Trp 190^{EC2}), for an acceptable deviation

of 10.0 degrees, and residues Tyr78^{TM} and Phe189^{EC2} presented a sidechain deviation from planarity of 0.118 and 0.104 Å, for an acceptable value of 0.03 Å. Residues Met64^{TM}, Asn98^{EC1}, Ile295^{Ct} and Lys303^{Ct} showed higher than allowed χ_1 deviations from the main value, and residue Ile164^{TM} has a higher than allowed ω torsion deviation from mean value; none of these TM residues are close to the binding site.

Ab initio derived models M1 and M2, using the MUSTER approach, show a better Qmean6 score and also present lower binding scores towards the five controls used. However, in both models the side chains of some critical residues, in particular of Trp86 and Trp248, are oriented away from the protein core and towards the membrane, and these residues have been described as being fundamental to establish protein-ligand interactions, and so these models were not selected to be used for the docking studies. This model, obtained using the 3oe6 template, also features a disulphide bridge between Cys101 and Cys178; this bridge, essential for proper folding of the extracellular loop, is essential for HIV fusion to the target cell. In our model, the putative binding site (Figure S2) is located closer to the extracellular side of the CCR5 receptor, and the obtained binding pocket is in agreement with results from the literature for other models of this protein and for other chemokine receptors.[11-13]

Docking - binding affinities of CCR5 positive controls

The CCR5 models obtained from the various templates were analysed for their quality using the Qmean6 and the Z scores as implemented in Swiss-Model, and these quantitative results are presented in Table S1.
PROMOTIF and PROCHECK analyses were performed to confirm that the models were structurally valid. In addition, as the obtained models will be used to dock CCR5 ligands, control ligands were docked to all the models; the predicted binding affinities are displayed in Table S1.

On the basis of all the quality parameters assayed and of all the predicted binding affinities, the CCR5 model obtained using as template the A chain of the 3oe6[13] protein (the CXCR4 chemokine receptor in complex with a small molecule antagonist) was chosen. The obtained structure was analysed for multiple stereochemical and conformational parameters, such as ω angles, C-α chirality and χ angles, and compared to the values from the statistical analysis of various high-level X-ray structures.[30] Separation of the protein into transmembrane, intracellular, and extracellular domains was performed with TMHMM[14,15] and the MEMSAT/PSIPRED[16–20] approaches, as well as by direct analysis of the modelled structure. The main results of this analysis are also included in Figure S2.

Table S1. Results from docking of selected CCR5 models and clinically relevant ligands.

	CCR5 Model			AutoDock Vina screening				
				Predicted Binding Affinity (kcal/mol)				
		Qmean6 score	Z score	Maraviroc	TAK779	Aplaviroc	Vicriviroc	Ancriviroc
Homology Modelling	1f88A [14]	0.365	-4.400	-9.0	-8.8	-8.7	-7.2	-7.4
	1l9hA [15]	0.309	-5.001	-9.3	-9	-9.3	-8.9	-7.7
	2ks9A [16]	0.332	-5.014	-8.1	-9.1	-7.7	-7.0	-7.3
	2rh1A [17]	0.348	-4.565	-9.8	-12	-10.1	-8.3	-8.2
	2vt4A [18]	0.371	-4.318	-9.6	-7.8	-8.1	-6.9	-8.9
	2vt4B [18]	0.382	-4.185	-11.1	-13.5	-8.9	-9.2	-9.6
	2x72A [19]	0.363	-4.388	-9.0	-10.1	-8.7	-8.6	-8.4
	2ydvA [20]	0.387	-4.132	-8.4	-8.4	-7.4	-8.3	-7.1
	2ziyA [21]	0.371	-4.610	-10.8	-11.1	-9.5	-8.4	-9.1
	3c9lA [22]	0.248	-5.980	-9.7	-8.8	-9.0	-8.0	-8.7
	3d4sA [23]	0.403	-3.977	-8.9	-8.4	-7.4	-8.0	-6.8
	3emlA [24]	0.393	-4.070	-7.3	-8.6	-7.3	-6.7	-7.0
	3kj6A [25]	0.267	-5.429	-9.2	-9.8	-8.9	-7.7	-8.6
	3oduA [26]	0.431	-3.734	-10.8	-8.9	-9.7	-9.5	-9.4
	3oe0A [26]	0.450	-3.450	-10.5	-11.5	-9.1	-9.0	-9.2
	3oe8A [26]	0.431	-3.665	-11.0	-10.2	-9.6	-9.3	-9.7
	3oe9A [26]	0.440	-3.540	-10.6	-10.5	-10.1	-9.7	-9.9
	3pblA [27]	0.369	-4.334	-9.9	-9.0	-9.0	-8.8	-9.3
	3rzeA [28]	0.331	-4.744	-11.3	-9.4	-9.2	-7.8	-9.4
	3uonA [29]	0.404	-3.960	-11.2	-13.2	-10.2	-9.9	-10.6
	3oe6A [26]	0.459	-3.363	-10.9	-11.6	-10.3	-9.7	-9.8
	3oe6A-FG	0.412	-3.866	-9.0	-10.2	-9.4	-8.8	-9.1
	3oe6A-MR	0.420	-3.782	-10.4	-9.0	-9.1	-8.8	-8.2
	3oe6A-L6	0.367	-4.347	-9.7	-10.8	-10.0	-10.8	-10.2
	3oe6A-L6-FG	0.325	-4.801	-10.7	-11.8	-11.0	-9.1	-10.8
	3oe6A-L6-MR	0.393	-4.069	-10.6	-9.4	-9.2	-9.6	-9.3
GPCR DB	Q5QIP1	0.368	-4.714	-6.7	-7.8	-6.3	-6.6	-5.8
	O14694	0.331	-5.086	-9.1	-10.7	-9.6	-8.8	-8.7
	Q5QIN9	0.346	-4.973	-7.2	-6.8	-6.1	-5.5	-5.8
	Q5QIP0	0.380	-4.580	-6.8	-7.7	-6.3	-5.6	-5.5
Ab-initio Modelling	T1	0.395	-4.402	-8.7	-11.8	-9.6	-10.1	-10.1
	T2	0.392	-4.432	-10.1	-11.1	-10.6	-10.0	-9.0
	T3	0.382	-4.552	-8.7	-10.5	-8.7	-8.8	-8.7
	T4	0.408	-4.248	-7.8	-5.3	-6.4	-6.6	-6.8
	T5	0.380	-4.578	-9.6	-12.1	-10.7	-9.3	-10.3
	M1	0.326	-5.200	-15.5	-16.8	-15.2	-12.0	-13.4
	M2	0.481	-3.401	-16.3	-18.2	-17.4	-14.1	-15
	M3	0.448	-3.787	-9.4	-9.9	-8.9	-8.5	-9.3
	M4	0.341	-5.032	-8.3	-7.9	-6.7	-7.5	-6.5
	M5	0.479	-3.429	-11.0	-10.0	-10.1	-9.0	-9.1
	M6	0.423	-4.073	-11.0	-11.6	-11.2	-9.7	-7.9
	M7	0.466	-3.578	-10.0	-11.3	-9.7	-7.9	-8.5
	M8	0.430	-3.991	-9.8	-10.4	-8.6	-9.9	-9.3
	M9	0.429	-4.002	-6.6	-6.5	-5.9	-6.5	-7.7
	M10	0.427	-4.035	-10.3	-11.5	-10.2	-10.0	-8.6

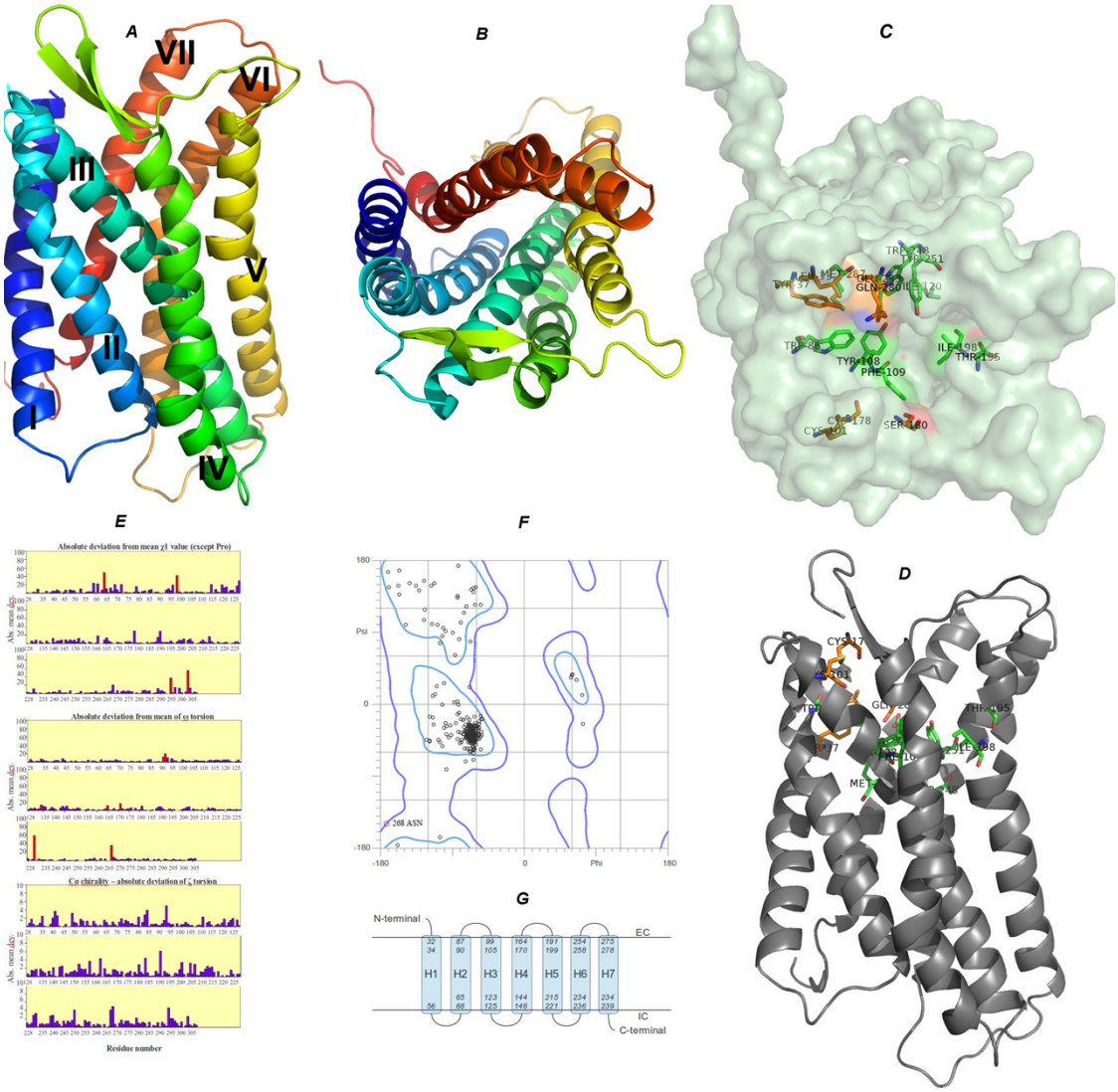

Figure S2 - Structure of the CCR5 model obtained by homology modelling using as template chain A of protein 3oe6 from the PDB database. A) side view of the model, with helix identification; B) top view of the model; C) top view of the protein in a surface representation, with the relevant residues presented in sticks, D) same as C, in a side view; E) stereochemical analysis of the model; F) Ramachandran plot of the model; G) helix content prediction from the protein sequence; for each helix region, the starting and ending residues, from different prediction tools, are shown in italics.

Ab-initio modelling

In parallel with the homology modelling procedures, an *ab-initio* modelling approach was also conducted using the MUSTER identified templates via the I-TASSER[21–23] server. Also, the GPCR database[2425] was browsed, and four model structures of the native human CCR5 protein were retrieved (accession codes O14694, Q5QIN9, Q5QIP0 and Q5QIP1).

All models were refined using three different approaches: ModRefiner,[26] which uses a two-step atomic-level energy minimization to improve the hydrogen bonding network, backbone topology and side-chain positioning of theoretical models; FG-MD[27]

(Fragment-Guided MD Simulation), which uses analogous fragments from available protein structures via structural alignment to derive spatial restraints that will be used to re-shape the funnel of the MD energy landscape and guide the MD conformational sampling; and ModLoop,[28,29] that relies on the loop modeling routine in MODELLER to predict the loop conformations by satisfaction of spatial restraints. As all obtained models contain six loops, three in the extracellular side and three in the intracellular side, ModLoop was used as follows: the models were subject to an EC2 refinement, as this is a loop involved in ligand access from the extracellular environment to the binding site; this refined model was then used to refine, one by one, in order, loops EC1, EC3, IC1, IC2 and IC3. All refined models were analysed using Swiss-Model tools and PSVS, as described before. Various attempts were made to use the three refinements for the various models, in various combinations, but all the final structures were severely distorted in terms of the orientation of the side-chains that are relevant to the binding processes.

References

(1) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215* (3), 403–410.

(2) Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. The SWISS-MODEL Workspace: A Web-Based Environment for Protein Structure Homology Modelling. *Bioinformatics* **2006**, *22* (2), 195–201.

(3) Bordoli, L.; Kiefer, F.; Arnold, K.; Benkert, P.; Battey, J.; Schwede, T. Protein Structure Homology Modeling Using SWISS-MODEL Workspace. *Nat Protoc* **2009**, *4* (1), 1–13.

(4) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. SWISS-MODEL: An Automated Protein Homology-Modeling Server. *Nucleic Acids Res* **2003**, *31* (13), 3381–3385.

(5) Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Cassarino, T. G.; Bertoni, M.; Bordoli, L.; Schwede, T. SWISS-MODEL: Modelling Protein Tertiary and Quaternary Structure Using Evolutionary Information. *Nucleic Acids Res.* **2014**, *42* (Web Server issue), W252–W258.

(6) Wu, S.; Zhang, Y. MUSTER: Improving Protein Sequence Profile-Profile Alignments by Using Multiple Sources of Structure Information. *Proteins* **2008**, *72* (2), 547–556.

(7) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242.

(8) Benkert, P.; Biasini, M.; Schwede, T. Toward the Estimation of the Absolute Quality of Individual Protein Structure Models. *Bioinformatics* **2011**, *27* (3), 343–350.

(9) Melo, F.; Feytmans, E. Assessing Protein Structures with a Non-Local Atomic Interaction Energy. *J Mol Biol* **1998**, *277* (5), 1141–1152.

(10) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A* **1999**, *103* (19), 3596–3607.

(11) Bhattacharya, A.; Tejero, R.; Montelione, G. T. Evaluating Protein Structures Determined by Structural Genomics Consortia. *Proteins Struct. Funct.*

Bioinforma. **2007**, *66* (4), 778–795.
(12) Hutchinson, E. G.; Thornton, J. M. PROMOTIF--a Program to Identify and Analyze Structural Motifs in Proteins. *Protein Sci* **1996**, *5* (2), 212–220.
(13) Wu, B.; Chien, E. Y.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. *Science (80-.).* **2010**, *330* (6007), 1066–1071.
(14) Sonnhammer, E. L.; von Heijne, G.; Krogh, a. A Hidden Markov Model for Predicting Transmembrane Helices in Protein Sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **1998**, *6*, 175–182.
(15) Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L. Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes. *J Mol Biol* **2001**, *305* (3), 567–580.
(16) Jones, D. T. Improving the Accuracy of Transmembrane Protein Topology Prediction Using Evolutionary Information. *Bioinformatics* **2007**, *23* (5), 538–544.
(17) Jones, D. T. Protein Secondary Structure Prediction Based on Position-Specific Scoring Matrices. *J Mol Biol* **1999**, *292* (2), 195–202.
(18) Jones, D. T.; Taylor, W. R.; Thornton, J. M. A Model Recognition Approach to the Prediction of All-Helical Membrane Protein Structure and Topology. *Biochemistry* **1994**, *33* (10), 3038–3049.
(19) Buchan, D. W.; Ward, S. M.; Lobley, A. E.; Nugent, T. C.; Bryson, K.; Jones, D. T. Protein Annotation and Modelling Servers at University College London. *Nucleic Acids Res* **2010**, *38* (Web Server issue), W563–W568.
(20) Nugent, T.; Jones, D. T. Transmembrane Protein Topology Prediction Using Support Vector Machines. *BMC Bioinformatics* **2009**, *10* (1), 159.
(21) Roy, A.; Kucukural, A.; Zhang, Y. I-TASSER: A Unified Platform for Automated Protein Structure and Function Prediction. *Nat Protoc* **2010**, *5* (4), 725–738.
(22) Yang, J.; Yan, R.; Roy, A.; Xu, D.; J, P.; Zhang, Y. The I-TASSER Suite: Protein Structure and Function Prediction. *Nat. Methods* **2015**, *12* (1), 7–8.
(23) Zhang, Y. I-TASSER Server for Protein 3D Structure Prediction. *BMC Bioinformatics* **2008**, *9* (1), 40.
(24) Horn, F.; Bettler, E.; Oliveira, L.; Campagne, F.; Cohen, F. E.; Vriend, G. GPCRDB Information System for G Protein-Coupled Receptors. *Nucleic Acids Res* **2003**, *31* (1), 294–297.
(25) Vroling, B.; Isberg, V.; Vroling, B.; van der Kant, R.; Li, K.; Vriend, G.; Gloriam, D. GPCRDB: An Information System for G Protein-Coupled Receptors. *Nucleic Acids Res.* **2014**, *42* (December 2013), 422–425.
(26) Xu, D.; Zhang, Y. Improving the Physical Realism and Structural Accuracy of Protein Models by a Two-Step Atomic-Level Energy Minimization. *Biophys. J.* **2011**, *101* (10), 2525–2534.
(27) Zhang, J.; Liang, Y.; Zhang, Y. Atomic-Level Protein Structure Refinement Using Fragment-Guided Molecular Dynamics Conformation Sampling. *Structure* **2011**, *19* (12), 1784–1795.
(28) Fiser, A.; Do, R. K.; Sali, A. Modeling of Loops in Protein Structures. *Protein Sci.* **2000**, *9* (9), 1753–1773.
(29) Fiser, A.; Sali, A. ModLoop: Automated Modeling of Loops in Protein Structures. *Bioinformatics* **2003**, *19* (18), 2500–2501.

Compound ranking

Total number of docked compounds: 2505

Table S2. Ranking of selected compounds in the docking results.

Ligand	Protein structure	
	Homology model	X-ray structure
CCA-H	1916th	1806th
CC-NH	1916th	1329th
CC-NLi	1863rd	1525th
Cl-OH	2122nd	2043rd
CSI	1916th	1630th
CS-Cl	1796th	1806th
C-phenaz	176th	399th

ESI-MS Spectra

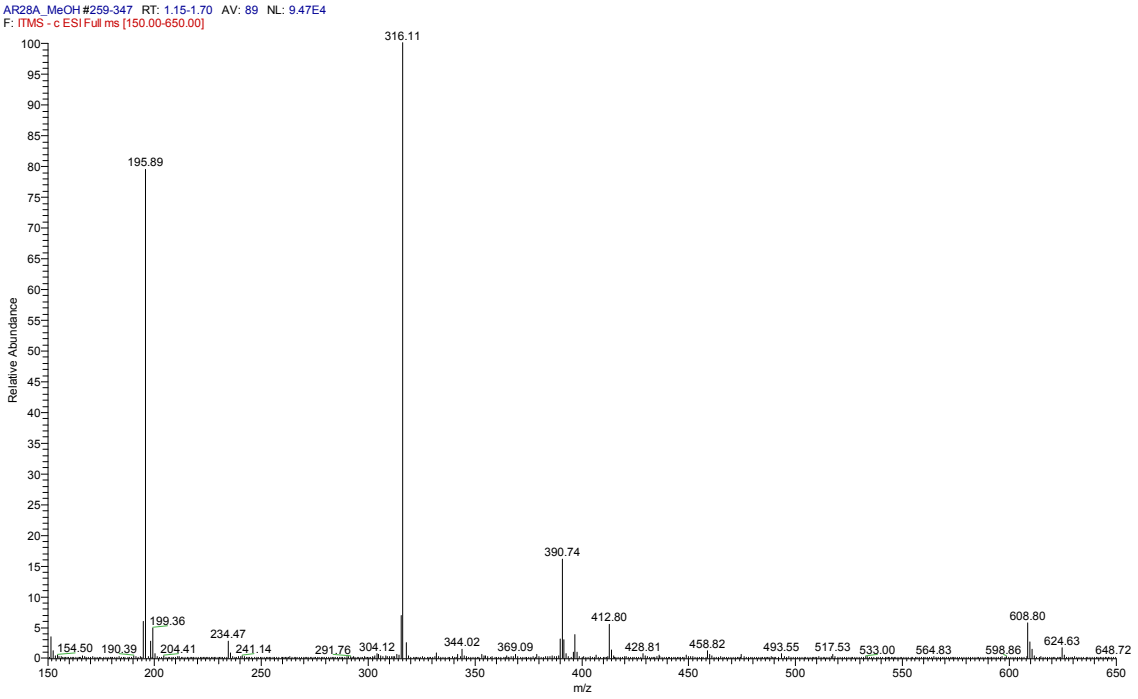

Figure S3. ESI-MS full scan (negative mode) of CC-NH sample.

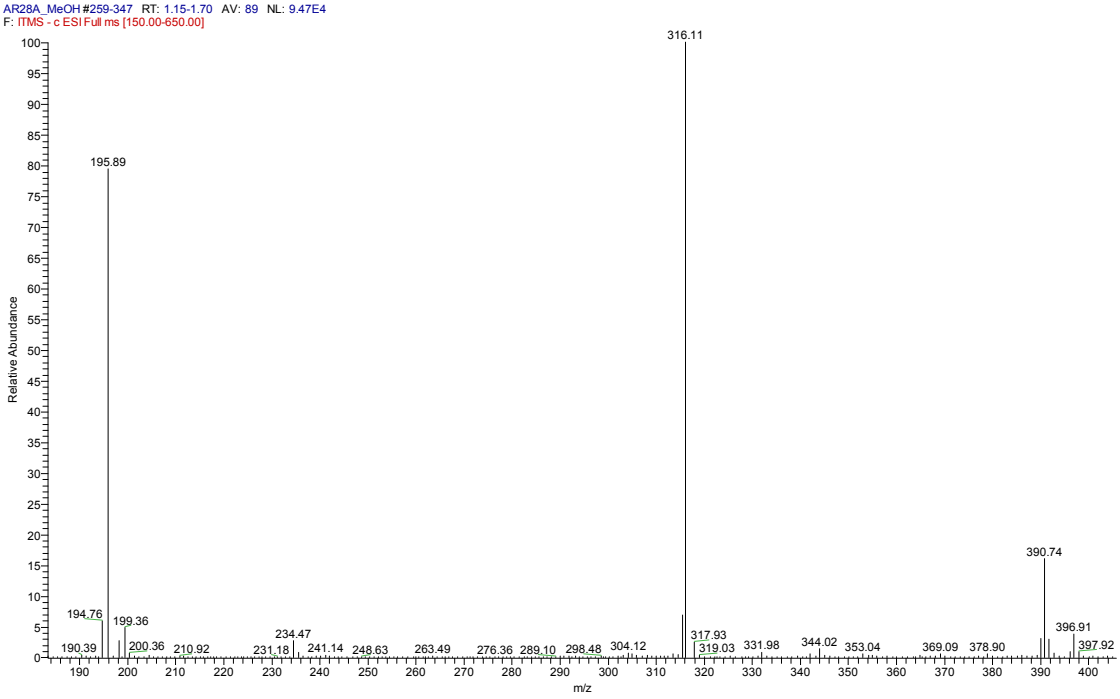

Figure S4. ESI-MS full scan (negative mode) of CC-NH sample – zoom-in on the 190-400 range.

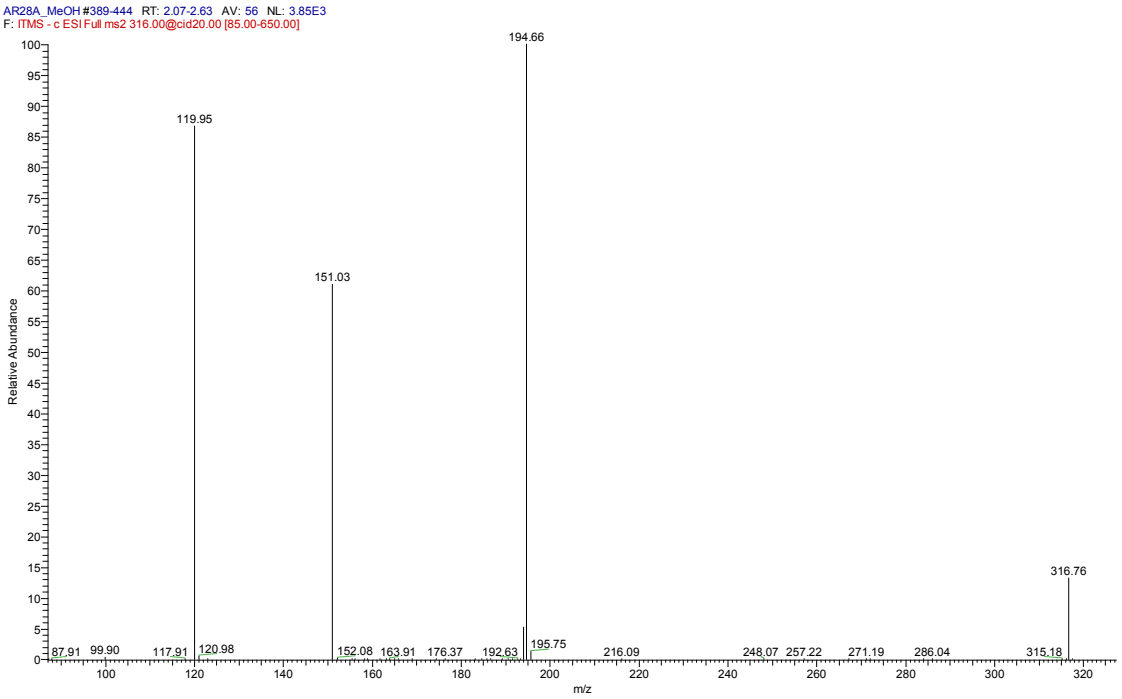

Figure S5. ESI-MS/MS scan (negative mode) of precursor ion at $m/z = 316$ of CC-NH.

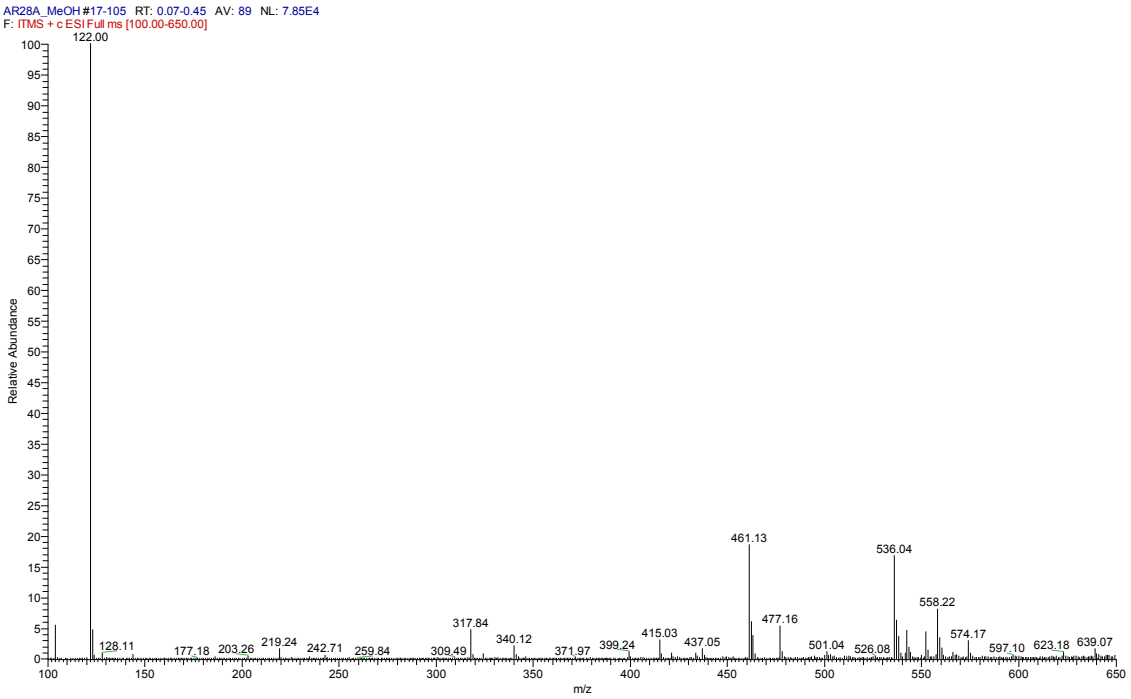

Figure S6. ESI-MS full scan (positive mode) of CC-NH sample.

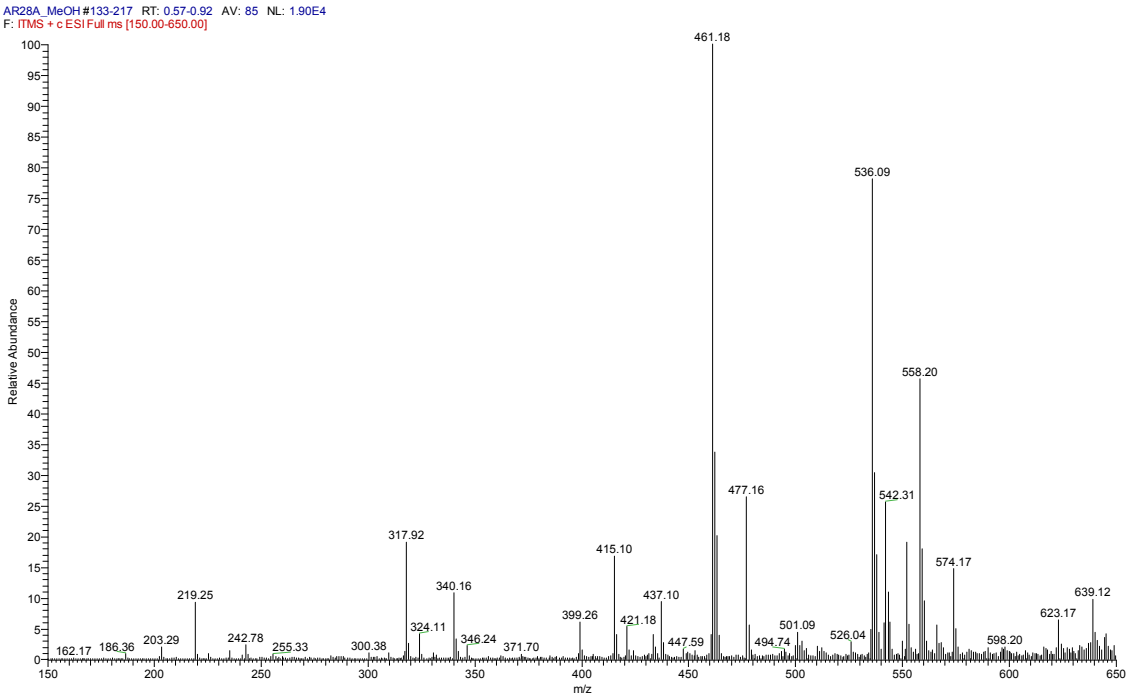

Figure S7. ESI-MS full scan (positive mode) of CC-NH sample – zoom-in on the 160-650 range.

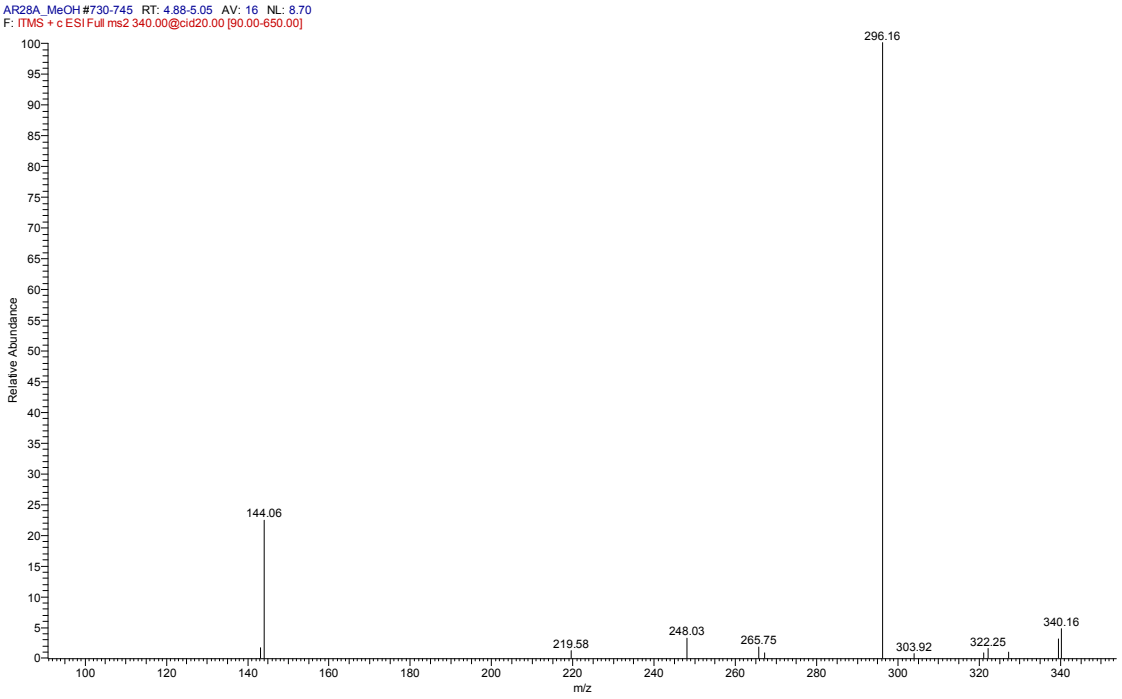

Figure S8. ESI-MS/MS scan (positive mode) of precursor ion at $m/z = 340$ of CC-NH.

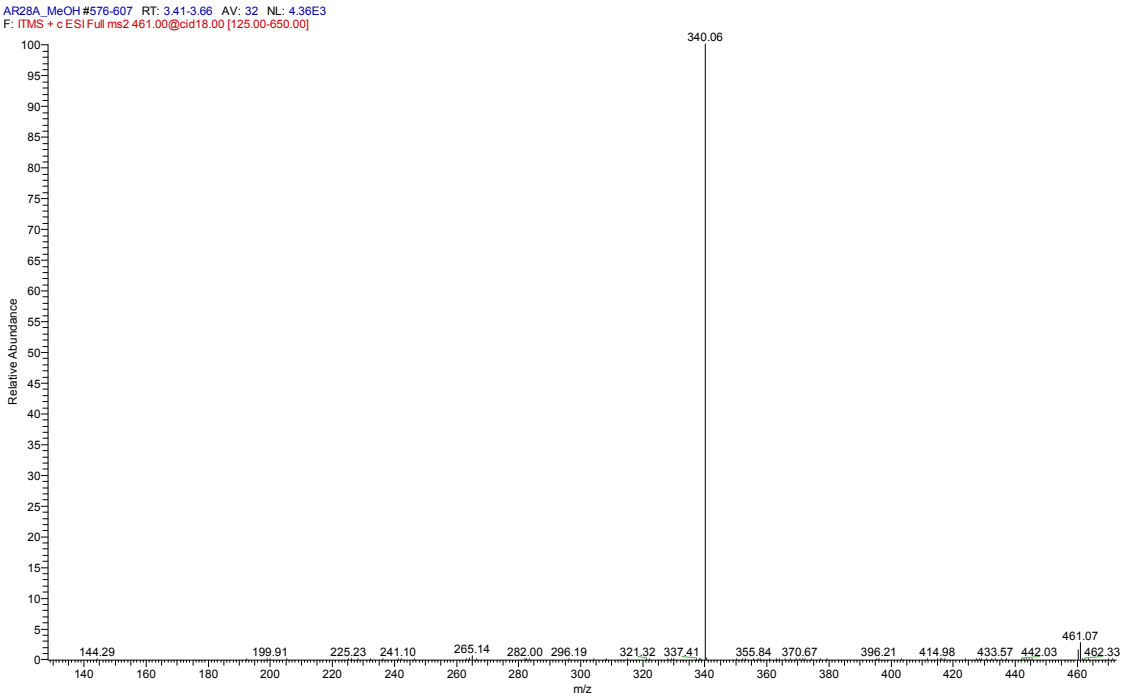

Figure S9. ESI-MS/MS scan (positive mode) of precursor ion at $m/z = 461$ of CC-NH.

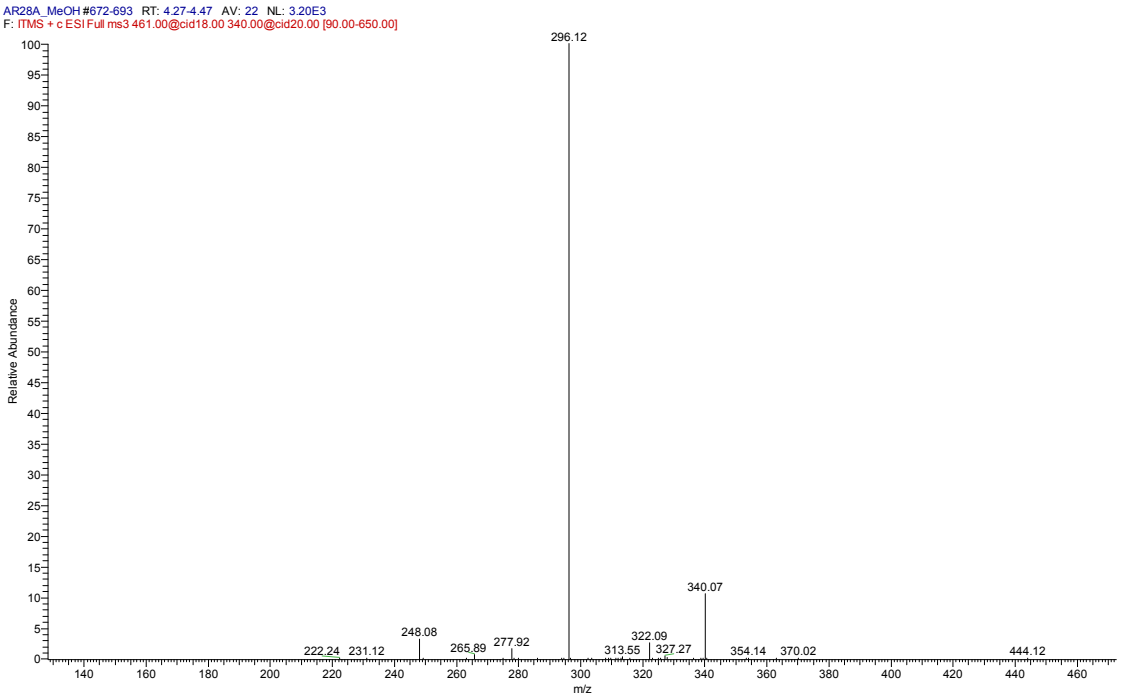

Figure S10. ESI-MS^3 scan (positive mode) of ion $m/z = 461 > 340$ of CC-NH.

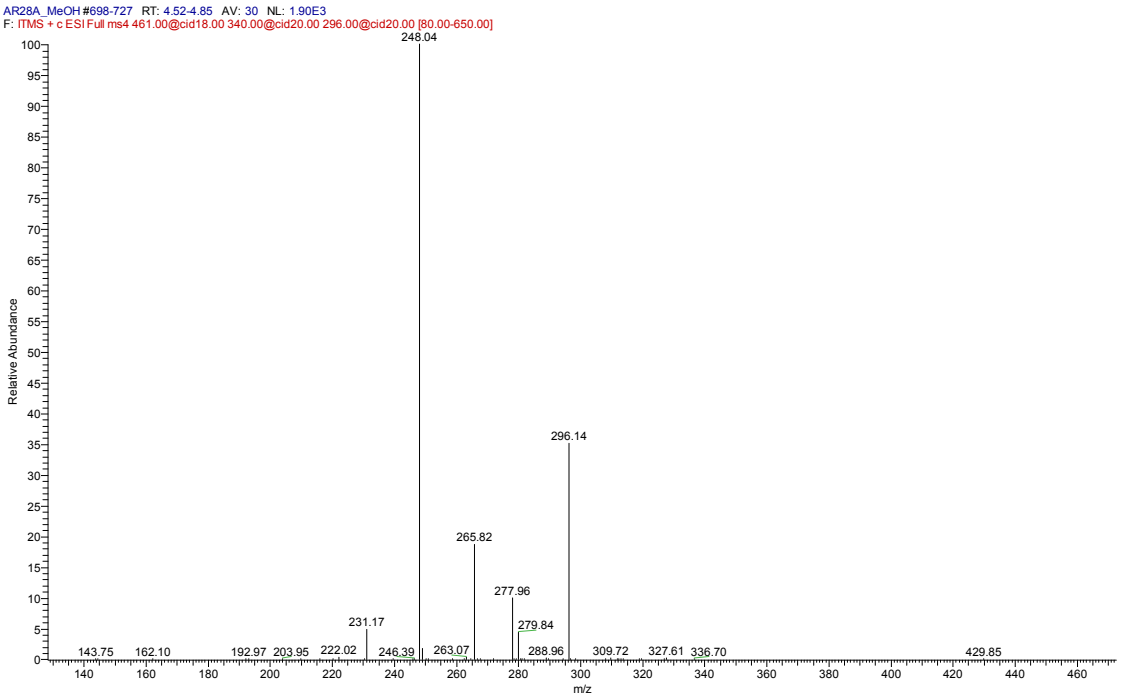

Figure S11. ESI-MS^4 scan (positive mode) of ion $m/z = 461 > 340 > 296$ of CC-NH.

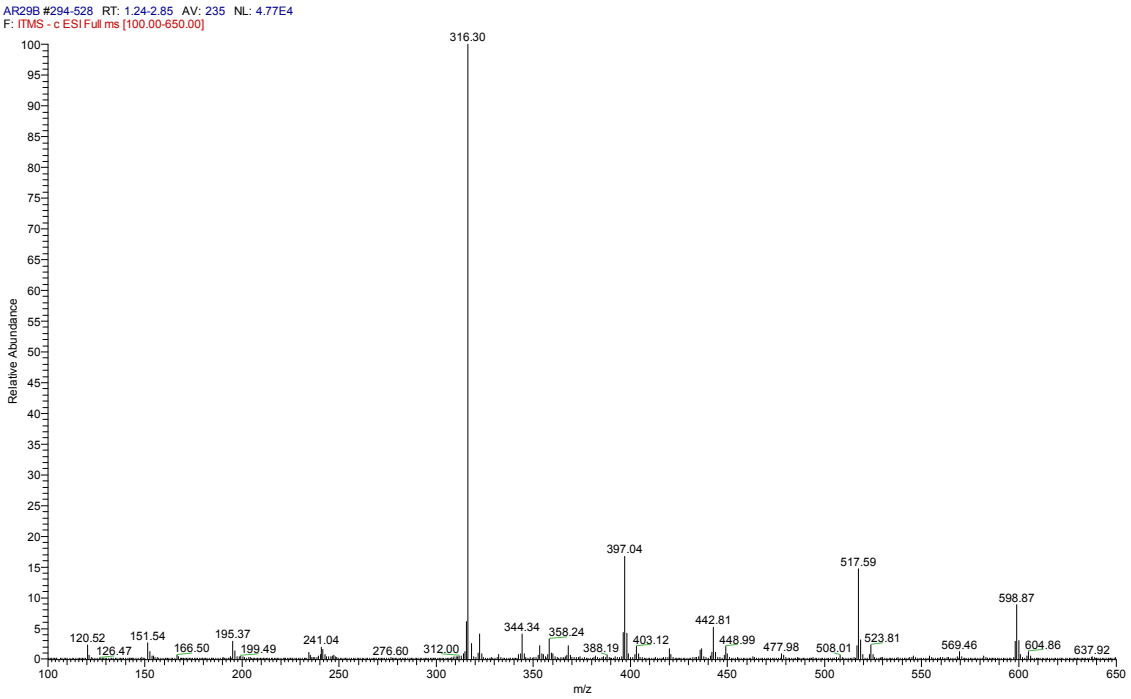

Figure S12. ESI-MS full scan (negative mode) of CC-NLi sample.

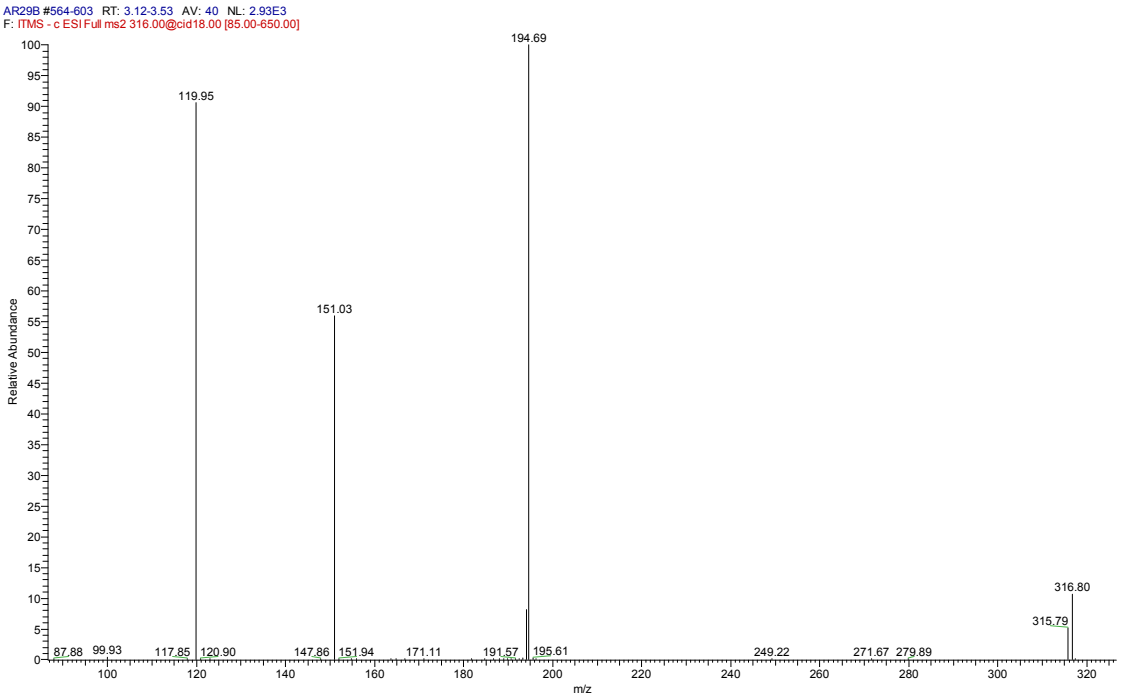

Figure S13. ESI-MS/MS scan (negative mode) of precursor ion at $m/z = 316$ of CC-NLi.

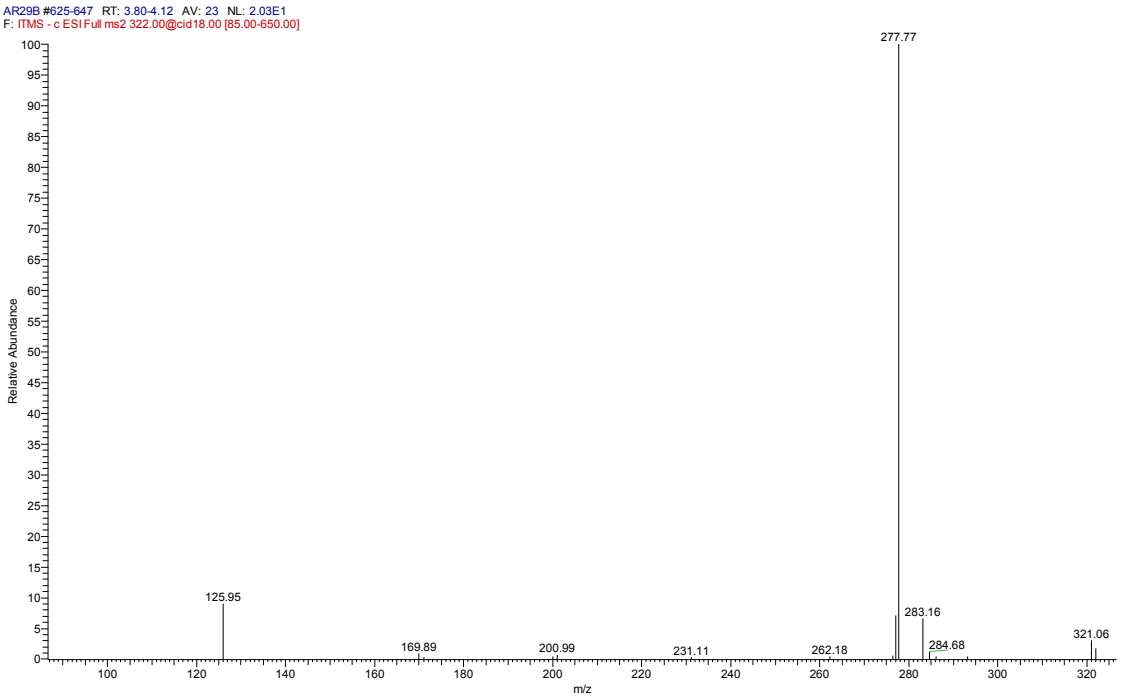

Figure S14. ESI-MS/MS scan (negative mode) of precursor ion at $m/z = 322$ of CC-NLi.

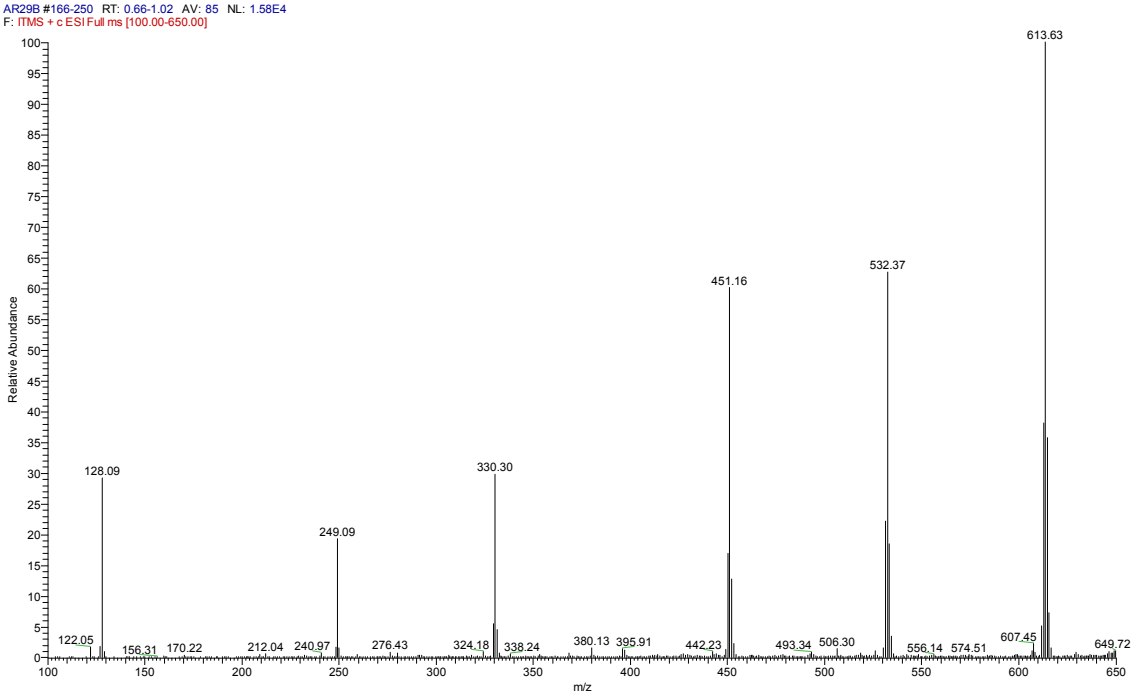

Figure S15. ESI-MS full scan (positive mode) of CC-NLi sample.

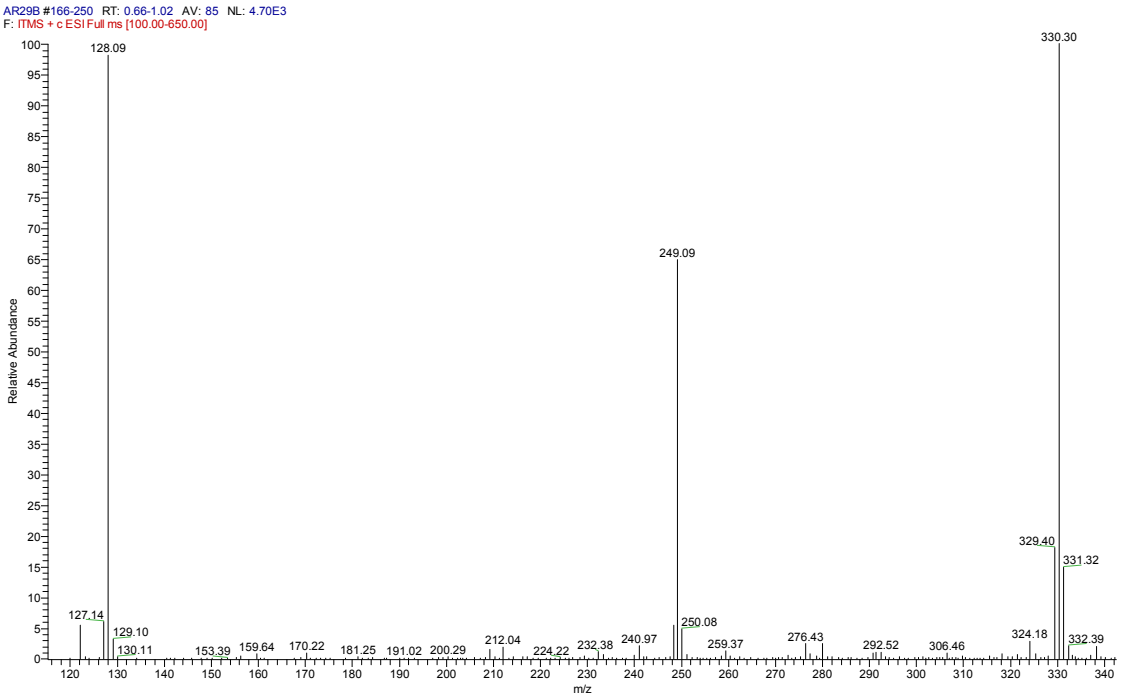

Figure S16. ESI-MS full scan (positive mode) of CC-NLi sample – zoom-in on the 120-340 range.

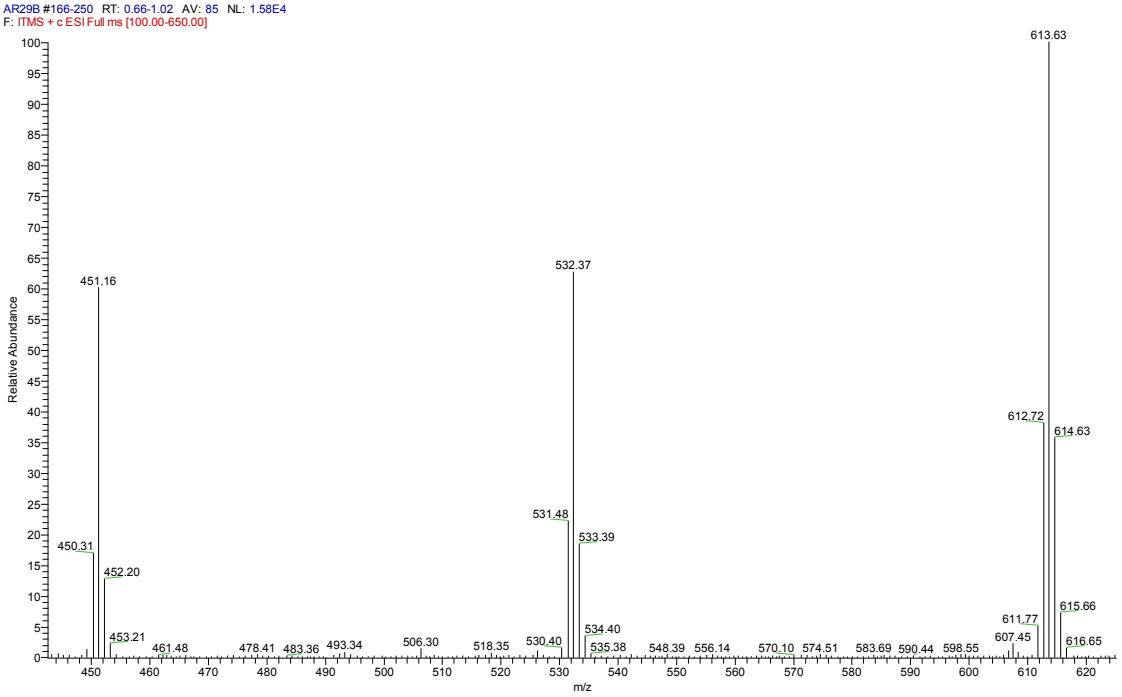

Figure S17. ESI-MS full scan (positive mode) of CC-NLi sample – zoom-in on the 450-620 range.

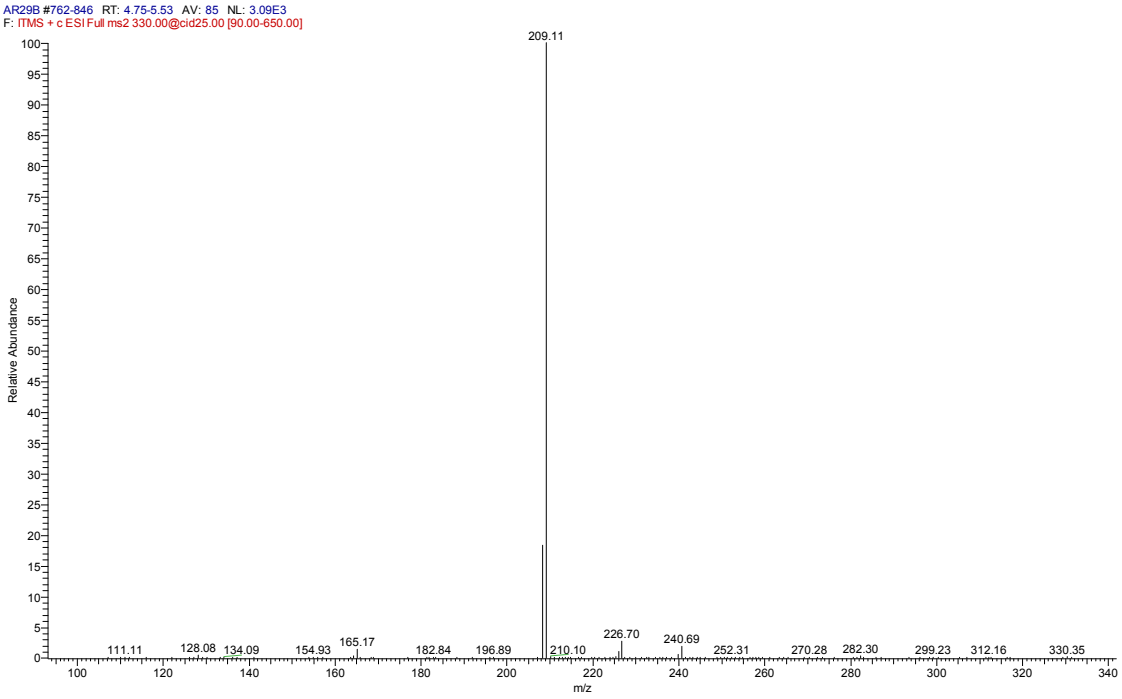

Figure S18. ESI-MS/MS scan (positive mode) of precursor ion at $m/z = 330$ of CC-NLi.

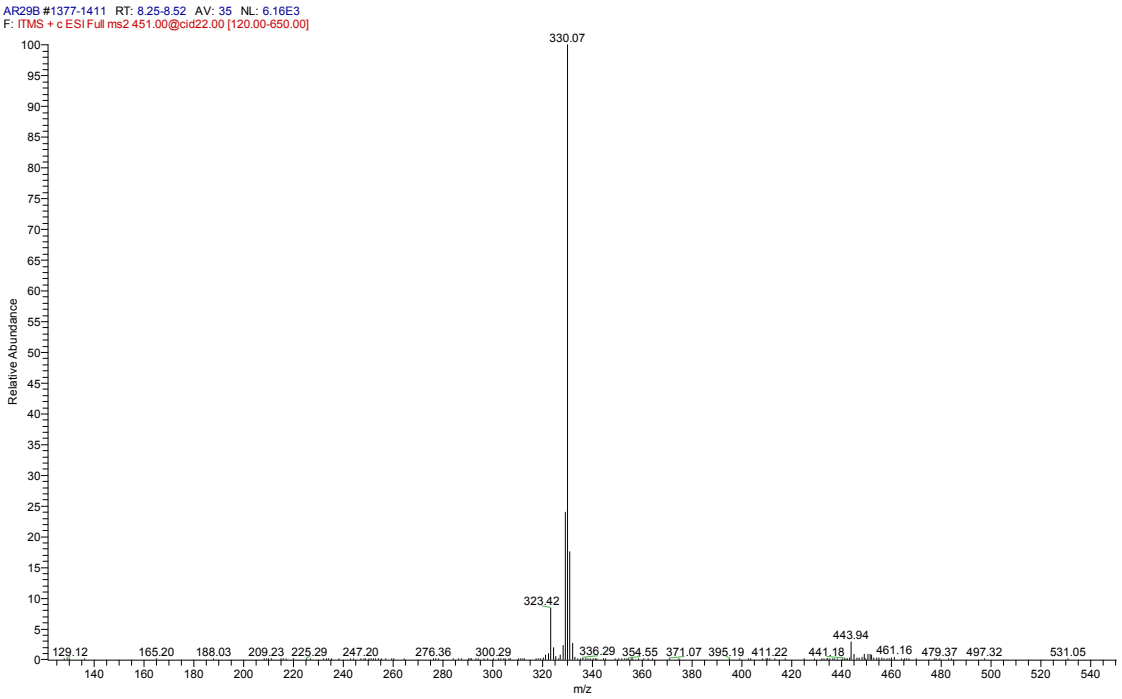

Figure S19. ESI-MS/MS scan (positive mode) of precursor ion at $m/z = 532$ of CC-NLi.

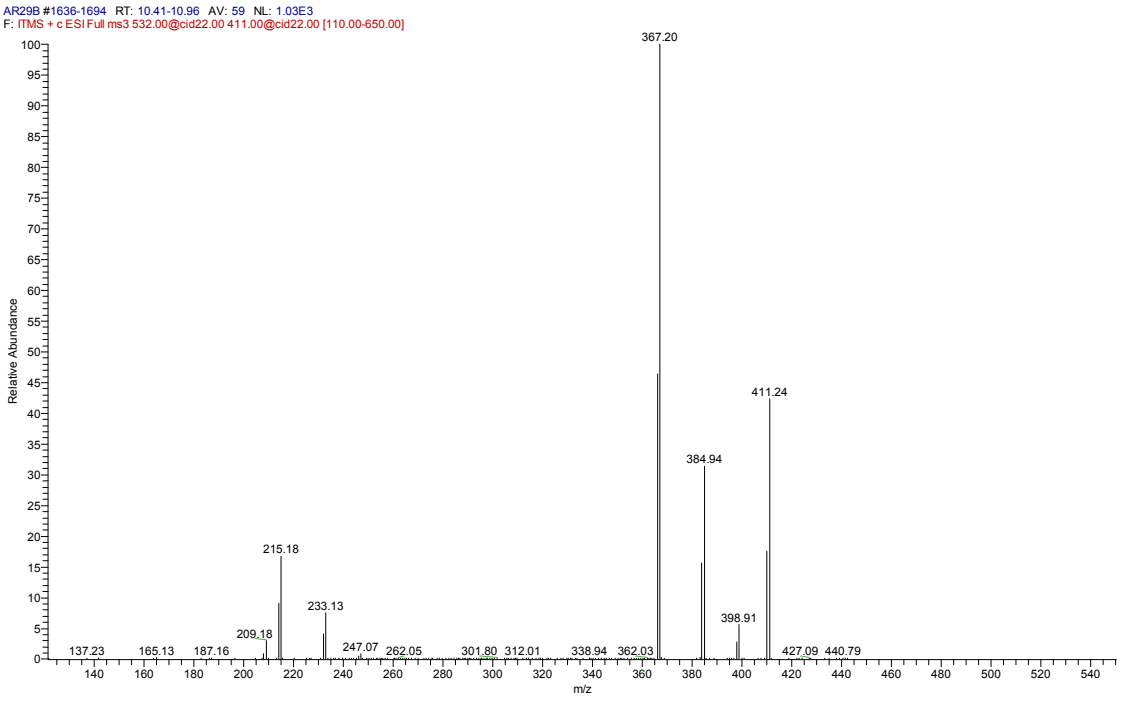

Figure S20. ESI-MS^3 scan (positive mode) of ion $m/z = 523 > 330$ of CC-NLi.

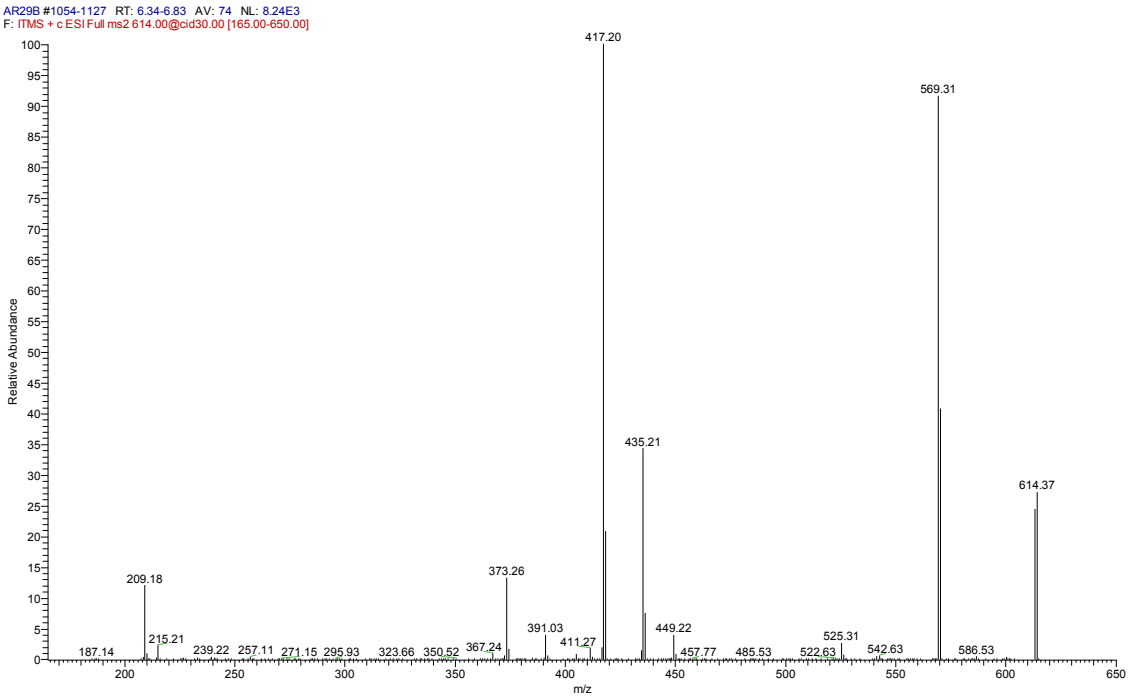

Figure S21. ESI-MS/MS scan (positive mode) of precursor ion at $m/z = 614$ of CC-NLi.

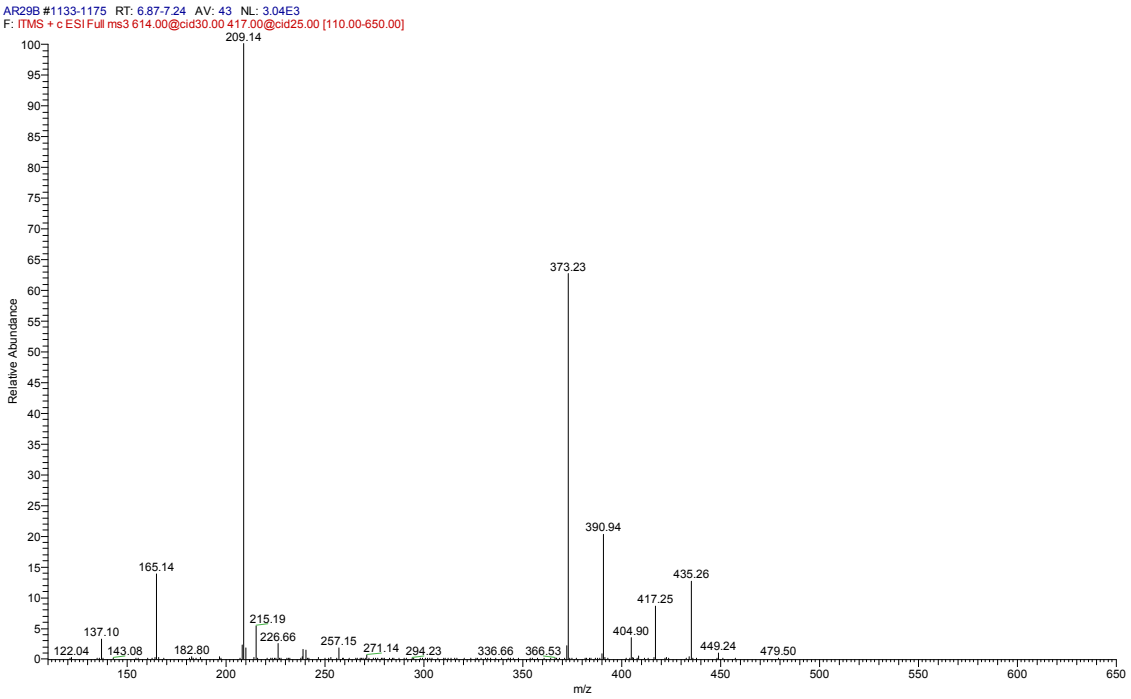

Figure S22. ESI-MS^3 scan (positive mode) of ion $m/z = 614 > 417$ of CC-NLi.

NMR spectra

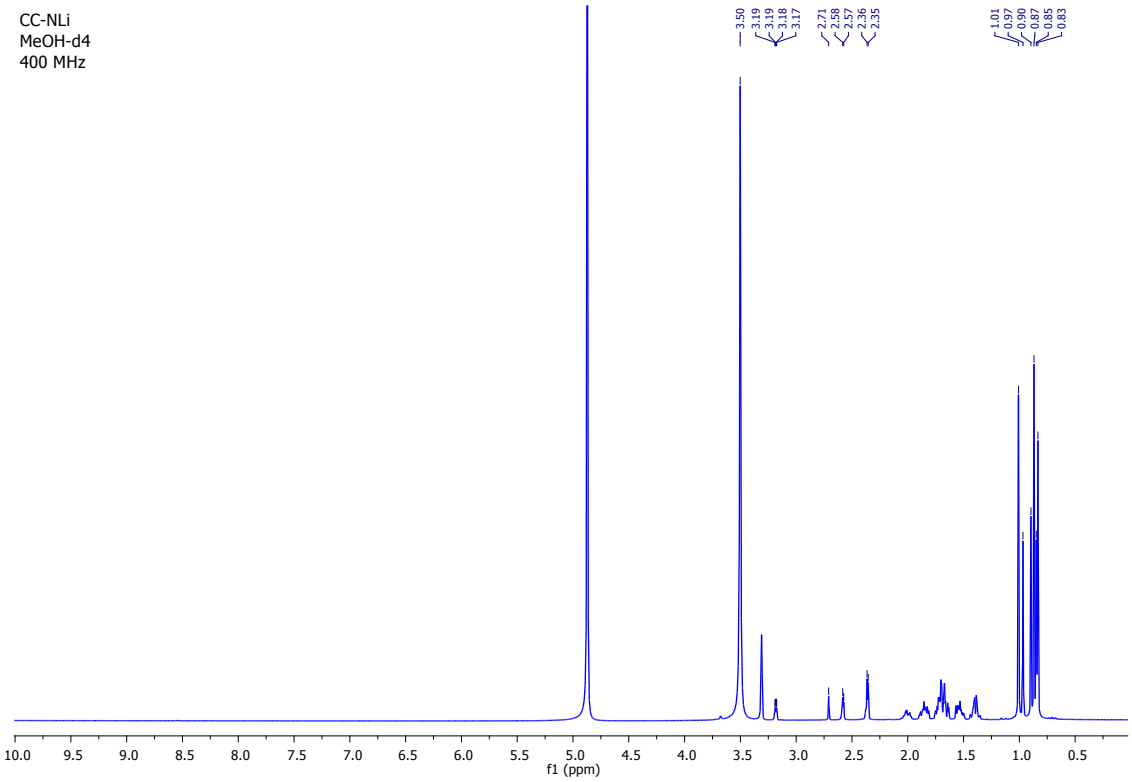

Figure S23. ^{1}H NMR spectra of CC-NLi (fresh solution)

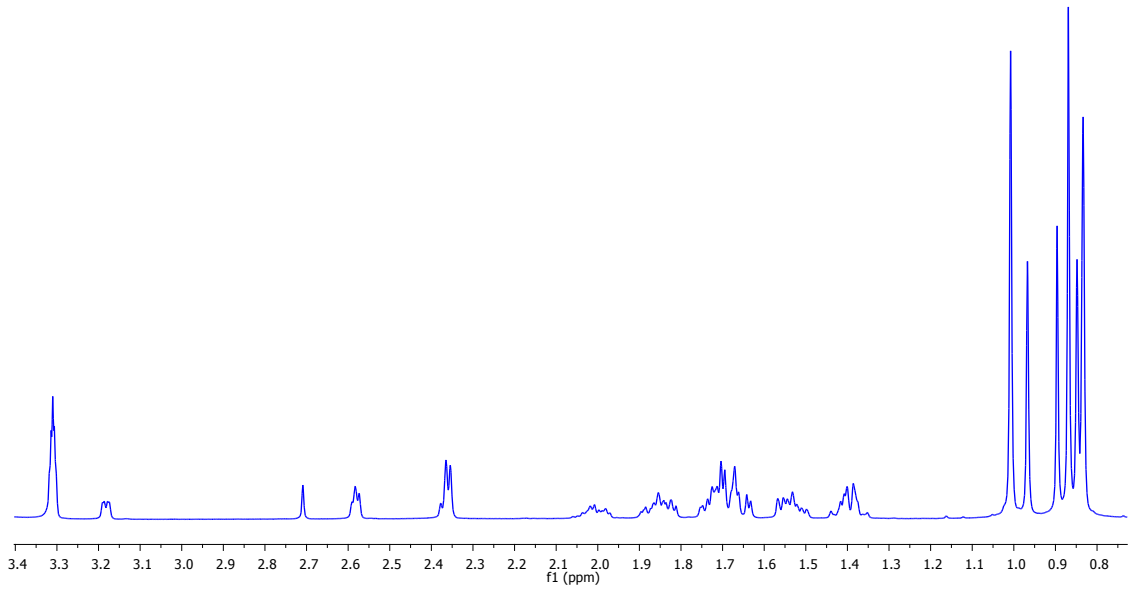

Figure S24. Detailed ^{1}H NMR spectra of CC-NLi (fresh solution)

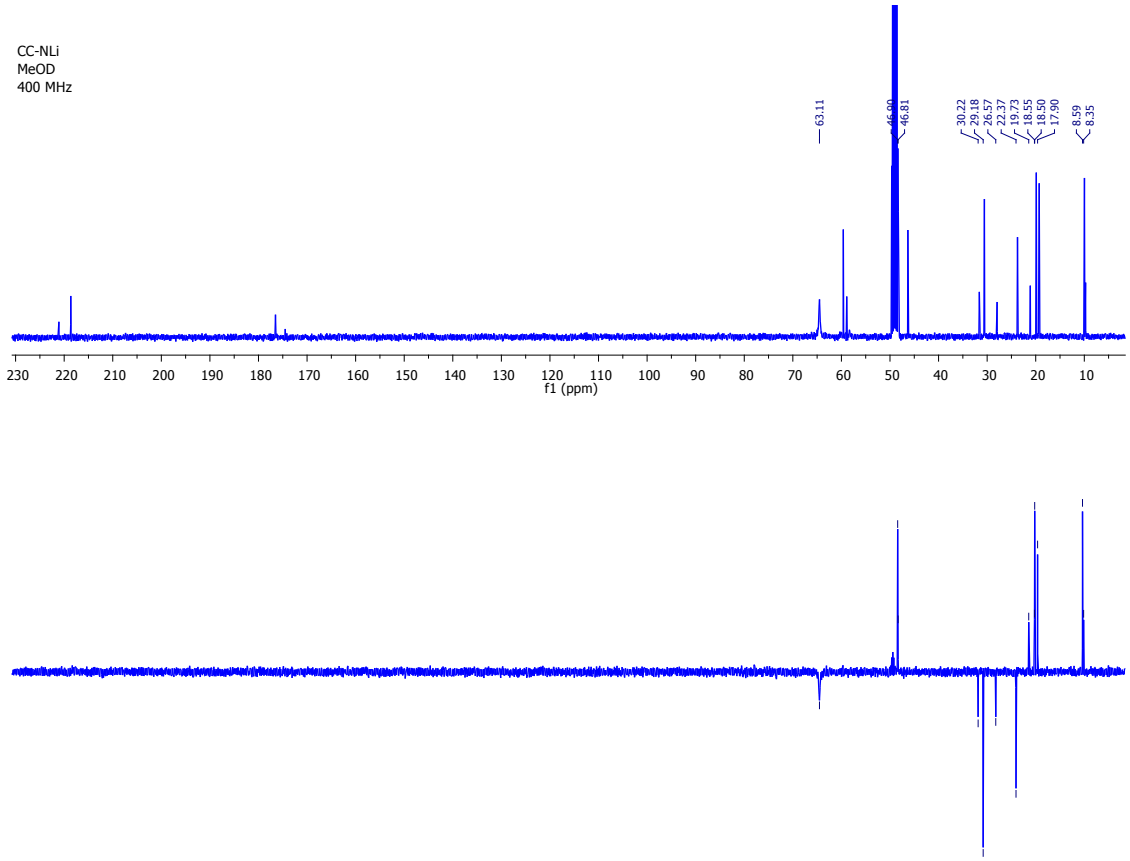

Figure S25. ^{13}C NMR and DEPT spectra of CC-NLi. Two isomers are visible.

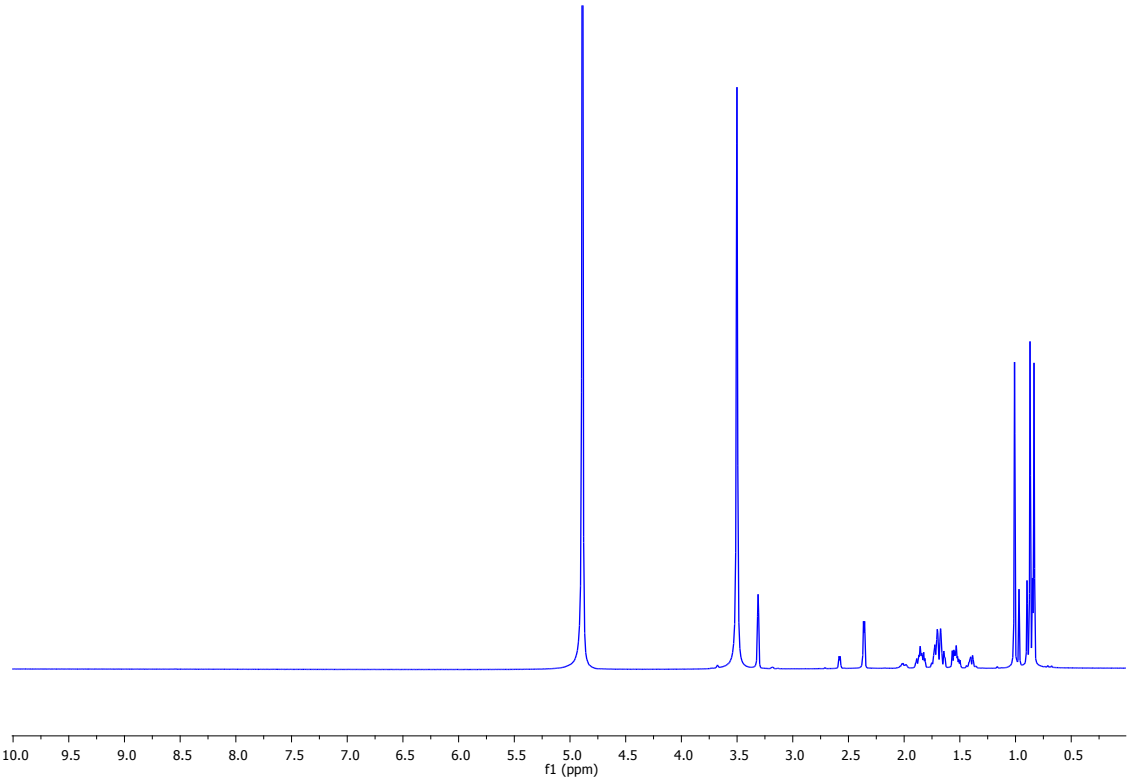

Figure S26. ^{1}H NMR of CC-NLi (after 5 hours in solution). Deuteration of the *exo* isomer occurs (Li^+ is replaced by D^+) and the spectra becomes similar to that of CC-NH.

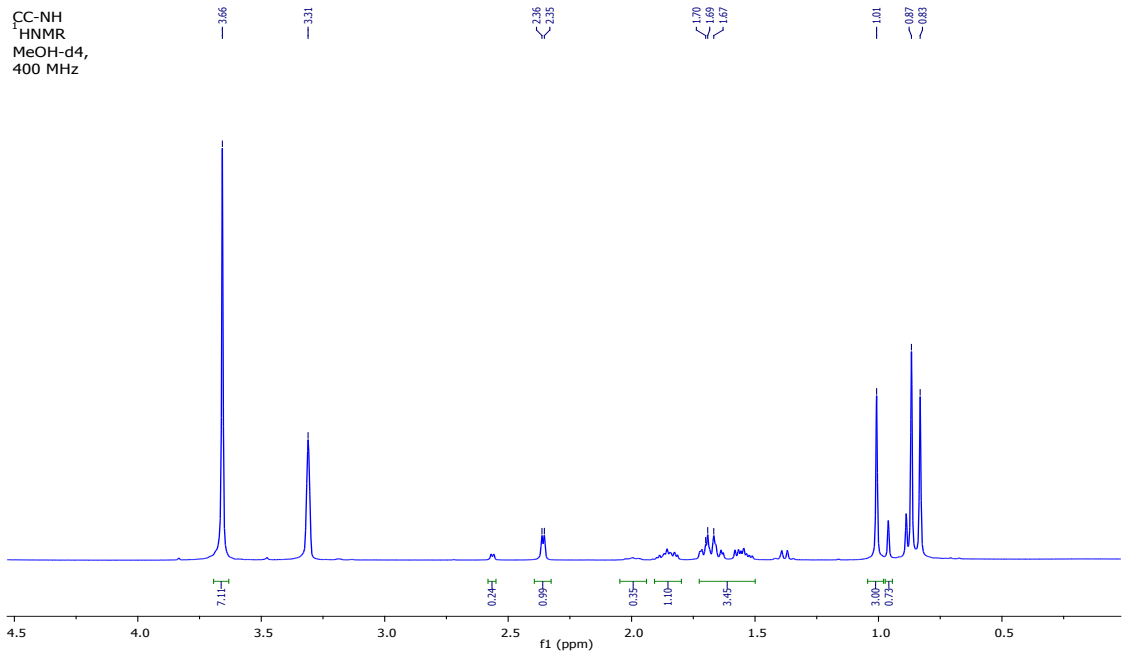

Figure S27. ^{1}H NMR spectrum of CC-NH.

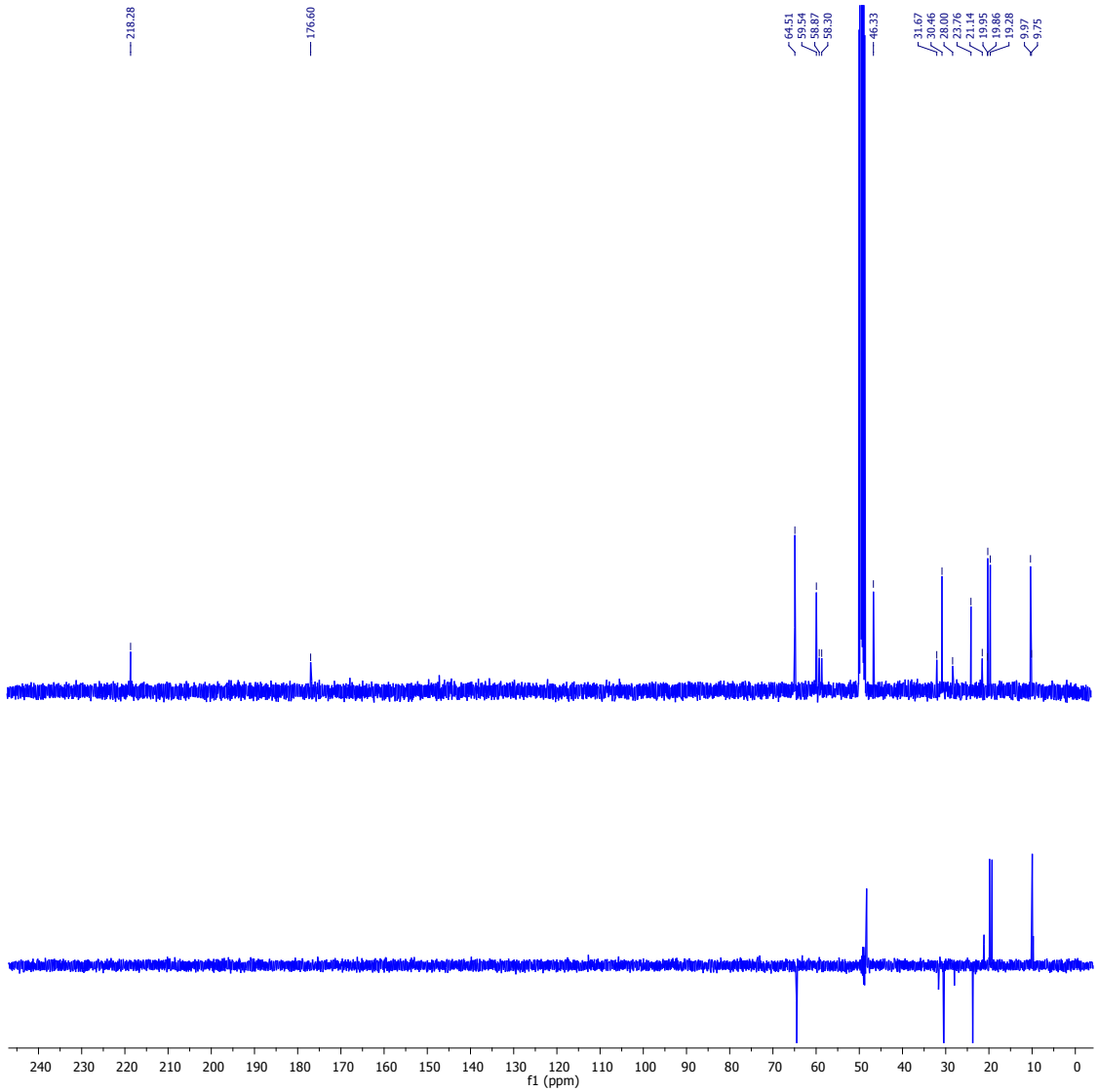

Figure S28. ^{13}C NMR and DEPT of CC-NH.