Department of
Chemistry

UNIVERSITY OF BATH

CH30063: Chemistry Research

THE SYNTHESIS OF FOUR NOVEL POTENTIAL HIV INHIBITORS FROM *L*-ISOLEUCINE

Megan R. McInnes

Registration number: 1100748100

ABSTRACT

In this paper, four novel potential HIV inhibitors, Targets A-D, were synthesised from *L*-isoleucine for the use in combination therapy. The first step involved the conversion of the amino acid to its corresponding methyl ester hydrochloride, which was the starting reagent of each target. Targets A, a reverse transcriptase inhibitor, and D, a fusion inhibitor, were successfully synthesised as confirmed by ^{1}H NMR and IR data, in low yield (30 % and 34 % respectively). Synthesis of target B, an integrase inhibitor, was unsuccessful and the product isolated was the symmetrical biaryl side product 4,4'-dimethoxy-1,1'-biphenyl as determined by its characteristic ^{1}H NMR spectrum containing only three peaks. Target C, a protease inhibitor, was also unsuccessful as the product was too impure and low yielding, thought to be due to incompletion of the Robinson-Gabriel cyclisation. Alternative routes were proposed to increase the success and efficiency of the above syntheses, with the view that the activity of each target can be tested in future research.

Supervisor: Prof. Steven Bull

Plagiarism Declaration

Please complete the following declaration by adding your name and the date. If handing in a printed copy, please sign this page. Submission of an electronic copy of this report through e.g. Moodle indicates that you agree with the statement

I certify that I have read and understood the entry in the Student Handbook for the Department of Chemistry on Cheating and Plagiarism and that all material in this assignment is my own work, except where I have indicated with appropriate references.

I agree that, in line with Regulation 15.3(e), if requested I will submit an electronic copy of this work for submission to a Plagiarism Detection Service for quality assurance purposes.

Student Name ___Megan McInnes_____

Student Signature _____

Date: _____17/4/19_____

Introduction

HIV is a viral infection that is notoriously hard to treat, and thus research into developing new drugs is ongoing. HIV targets T-helper cells, leading to a decreased efficiency of the immune system causing the victim to be more susceptible to contracting other diseases. Combination therapy is the most effective treatment of HIV. This is because HIV is able to form resistance to all classes of HIV retroviral drugs quickly, therefore administering multiple drugs which act *via* different mechanisms increases the chance of inhibiting the virus. Highly active antiretroviral therapy (HAART) is a type of combination therapy consisting of a cocktail of three of more HIV drugs.[1] Combinations of an RT or protease inhibitor and an integrase inhibitor may be beneficial as integrase inhibitors are not metabolized by cytochrome P450 as RT and protease inhibitors are, reducing the potential for drug-drug interactions which may affect the pharmacological effects of the drug.[2] We have therefore synthesized four potential HIV drugs. (Figure 1).

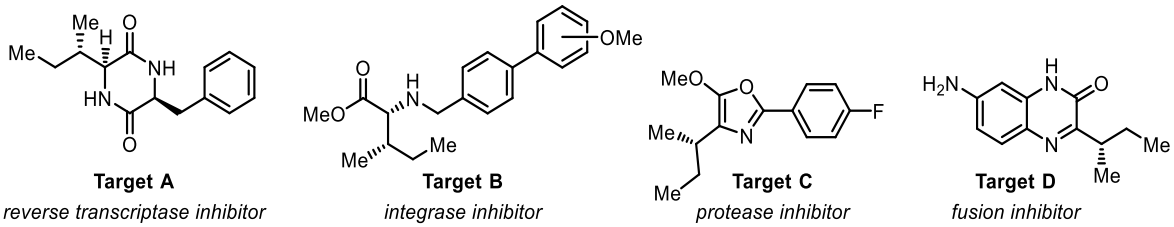
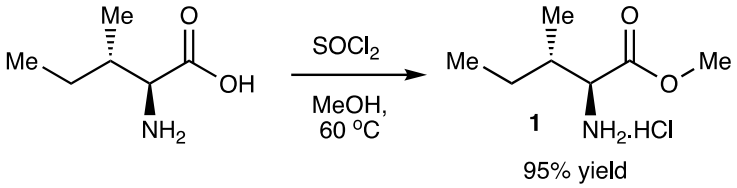


Figure 1. Targets A, B, C and D

Target A is a potential non-nucleoside reverse transcriptase inhibitor (NNRTI), a class of drug which targets an HIV enzyme called 'reverse transcriptase' (RT). NNRTIs prevent the enzyme from converting single-stranded HIV RNA to double stranded HIV DNA by binding to an allosteric pocket of HIV-1 RT, leading to a conformational change in the enzyme structure, dramatically decreasing the rate of reaction of the enzyme.[3] **Target B** is a potential integrase inhibitor, which inhibit the HIV enzyme integrase. Integrase is responsible for integrating viral DNA into the host DNA which occurs in three steps; [2] formation of the pre-integration complex, site specific cleavage of 3'ends of viral DNA, and insertion of viral DNA into host DNA. The active site of integrase contains a catalytic triad of amino acid residues which bind to two cationic magnesium ions. Therefore, an inhibitor should have the ability to bind to two divalent cations to inhibit integrase. A hydrophobic aromatic group is also required for binding to the pre-integrase complex.[2] **Target C** is a potential protease inhibitor which mimics polypeptide structures, by replacing the hydrolysable peptide bond with a non-cleavable isostere.[4] The HIV-1 enzyme is an aspartyl protease which is responsible for cleavage of peptide bonds in polyproteins into individual proteins. These then assemble to form new virus particles. Finally, **target D** is a possible fusion inhibitor. Fusion occurs when the HIV viral protein gp120 fuses with CD4 receptors on T-helper cells. This causes another HIV protein, gp41 to bind to a cofactor on the cell membrane, leading to the release of viral RNA into the host cell. The proteins gp120 and gp41 are found in the envelope glycoprotein complex, and fusion inhibitors bind to this, preventing the conformational change in gp41 necessary for fusion of viral DNA into the host cell.[5]

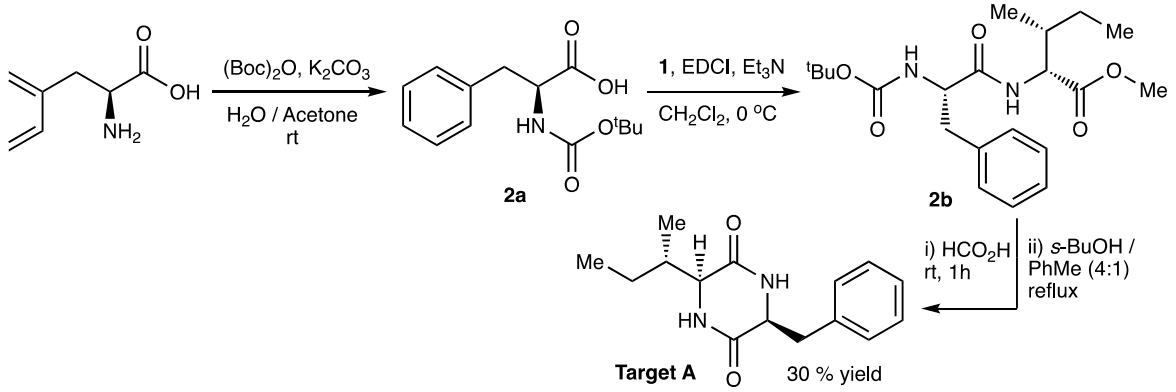
Results and Discussion

The first step was the conversion of L-isoleucine to its corresponding methyl ester hydrochloride (**1**) (scheme 1), which proceeded with a 95 % yield, and the structure was determined by the specific optical rotation of $[\alpha]_D^{20}$: + 18° (1 g / 100 mL, H_2O) characteristic of **1**. This compound was used as the starting material in the synthesis of **targets A-D**.



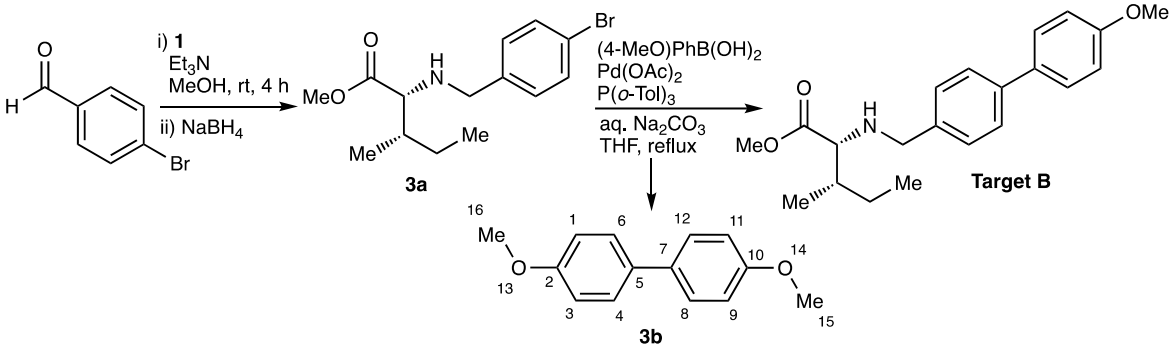
Scheme 1. The synthesis of *L*-isoleucine methyl ester hydrochloride from *L*-isoleucine

Target A was successfully synthesized (scheme 1) as confirmed by ^{1}H NMR and IR spectra. Comparison to literature was not possible due to this compound not being previously characterised. The reaction proceeded with low yield, 30 %, which may have been due to incomplete reaction in the final step, leaving uncyclised dipeptide. An alternative synthesis for **Target A** may be suggested as the three step synthesis reported by C. R. B. Rhoden *et al.*,[6] which entails an Ugi reaction using 1-(2,2-dimethoxyethyl)-2-isocyanobenzene as the isocyanide, an aldehyde with the L-phenyl alanine side chain, ammonia, and the N-Boc protected isoleucine methyl ester. The subsequent steps involve treatment with TFA to form a reactive intermediate, followed by addition of $NaHCO_3$ to force cyclisation respectively. This synthetic route may be beneficial as the process is carried out as a one pot reaction, and therefore it is easily automated, and may result in a better yield. (Appendix 1)



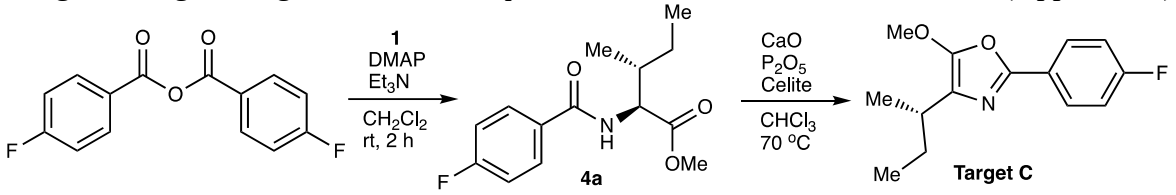
Scheme 2. The synthesis of target A from **1** and *N-tert*-butoxycarbonyl-*L*-phenylalanine

Target B was hypothesised to be a desirable structure for integrase inhibition due to the presence of the carbonyl and amine functional groups in a 1,2 position which may have been able to bind to the divalent Mg ions in the catalytic site of integrase. It additionally contains a hydrophobic benzene ring important for the stabilisation of the pre-integration complex, thus would indicate a potential inhibitor molecule. The isolated product of the reaction of scheme 2 however was not Target B as intended, but the side product **3b**, determined by its characteristic ^{1}H NMR spectrum. This is thought to have occurred due to the reaction being insufficiently degassed with N_2 providing the reaction with O_2. This would prevent the reduction of Pd(II) to Pd(0), inhibiting the desired aryl-aryl coupling cycle and promoting the oxidative homocoupling of the aryl-boronic acid reagent.[7] The reaction may have been further favoured due to the presence of the electron donating methoxy group on the benzene ring, making the boronic acid derivative more nucleophilic. To remedy this, the reaction must be degassed for longer to ensure this air sensitive reaction proceeds as intended. An alternative synthesis may substitute the first reductive amination step for a palladium catalysed Buchwald-Hartwig amination, although this would require both steps be carried out under an inert atmosphere and not be beneficial.



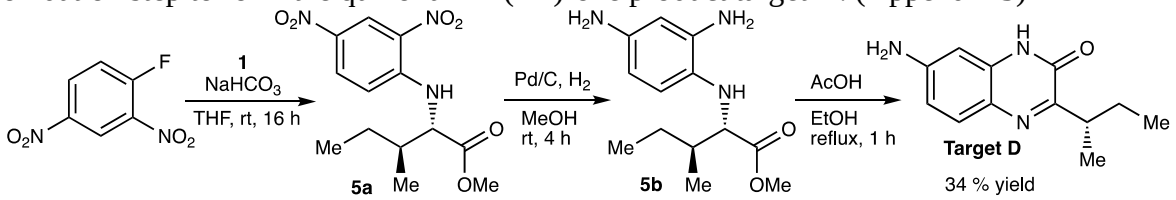
Scheme 3. The desired synthesis of target B from **1** and 4-bromobenzaldehyde and formation of side product

Synthesis of **Target C** (scheme 4), as a potential protease inhibitor was attempted due to the presence of the non-cleavable peptide isostere; the oxazole ring, however the synthesis was unsuccessful finding that the product was extremely impure, and therefore data was not reported. It may be suggested that this was due to incomplete reaction of step two (the Robison Gabriel cyclisation); although the reaction was monitored by TLC, this was not done so past four hours when the reaction was assumed to have gone to completion due to insufficient time, and therefore additional portions of P_2O_5 may have been required. An alternate route to Target C may have been to replace the second step with a cyclodehydration using the Burgess reagent, shown to require less time, and milder conditions.[8] (Appendix 2)



Scheme 4. The synthesis of target C from **1** hydrochloride and 4-fluorobenzoic anhydride

Target D was successfully synthesised (Scheme 5) determined by ^{1}H NMR spectra (Appendix 12) in low yield (34 %). This was unexpectedly low, possibly due to the compound being very soluble in many solvents, therefore a large amount of the product may have been lost in the trituration. A final reflux step was carried out to ensure that the cyclised product was achieved, providing a greater yield. An alternate two step synthesis has been reported by M. Imanishi *e. al.*,[9] in which 1-bromo-2,4-diaminobenzene and the methyl ester **1** may undergo a copper catalysed coupling to form the quinoxalin-2-one, followed by an oxidation step to form the quinoxalin-2(1H)-one product target D. (Appendix 3)



Scheme 5. The synthesis of target D from **1** and 1-fluoror-2,4dinitrobenzene

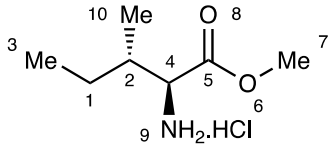
Conclusion

In conclusion the synthesis of **Targets A** and **D** were successful but require improvement due to the low yields. **Target B** was not isolated but may be successful upon repeating with more rigorous degassing in the Suzuki step. **Target C** was unsuccessful and may require a different cyclisation step as suggested above. Alternate syntheses for **targets A-D** have been proposed, and these should be carried out to determine the optimal route. Future work may

focus on testing **targets A-D** for activity on their respective biological targets by use of high-throughput screening. Structural optimisation of the compounds may be required, and therefore methods such as the Ugi reaction proposed for **target A** may be beneficial due to the possibility of forming many products in one pot by an automated process.

Experimental

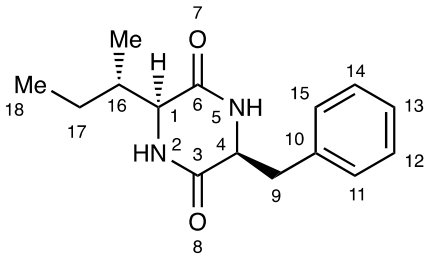
L-isoleucine methyl ester hydrochloride (1)



$SOCl_2$ (2.4 mL, 33 mmol, 1.1 equiv.) was added slowly to a suspension of L-isoleucine (3.94 g, 30 mmol, 1 equiv.) in methanol (10 mL) and refluxed for 2 h. Concentrating *in vacuo* yielded the methyl ester hydrochloride (5.1527 g, 95 %) as a white solid, with data in accordance to the literature.[10] mp 95-96°C {Lit.[10] 98 °C}; IR (film) ν_{max}: 2878 (C-H), 1739 (C=O), 1230 (C-O), ^{1}H NMR (300 MHz, Deuterium Oxide) δ_H: 0.85-1.04 (6H, m, H-3, 7), 1.19-1.55 (2H, m, H-1), 1.97-2.10 (1H, m, H-2), 3.81 (3H, s, H-7), 4.05 (1H, d, H-4); ^{13}C{^{1}H} NMR (75 MHz, $CDCl_3$) δ_C: 11.8, 15.0, 25.8, 37.0, 54.2, 58.2, 171.3, $[\alpha]_D^{20}$: + 18° (1 g / 100 mL, H_2O) {Lit.[11] + 18 ° (1 g / 100 mL, H_2O)}

***N-tert*-butoxycarbonyl-*L*-phenylalanine (2a).** L-Phenylalanine (2.64 g, 16 mmol, 1 equiv.), K_2CO_3 (8.84 g, 64 mmol, 4 equiv.), and $(Boc)_2O$ (5.5 mL, 24 mmol, 1.5 equiv.), were dissolved in a mixture of acetone : water (13 mL : 16 mL) and stirred overnight at rt. After acidifying to pH 3 with 2M HCl. The organic phase was extracted with ethyl acetate (3 x 30 mL), washed with brine (30 mL) and dried over $MgSO_4$. The solution was concentrated *in vacuo* to yield the product (3.7072 g, 87 %) as a clear oil. ***Isoleucine-phenyl*alanine dipeptide (2b)**. A solution of N-Boc phenylalanine, **2a** (3.7072 g, 14 mmol, 2.8 equiv.) in DCM (93 mL), was cooled to 0°C. The methyl ester, **1** (0.91 g, 5 mmol, 1 equiv.), and NEt_3 (0.70 mL, 5 mmol, 1 equiv.), were added to the solution (33.3 mL, 5 mmol, 1 equiv.), followed by slow addition of EDCl (0.9585 g, 5 mmol, 1 equiv.) and stirred for 16 h. After dilution with CH_2Cl_2 (50 mL), the organic portion was washed with NaOH (25 mL), H_2O (25 mL), HCl (25 mL), and brine (25 mL), dried over $MgSO_4$ and concentrated *in vacuo* to yield the dipeptide (1.1920 g, 61%) as a white solid.

Target A

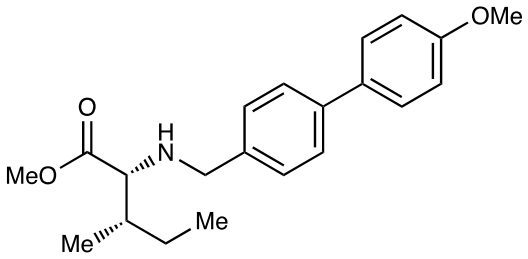


A solution of dipeptide, **2b** (1.1816g, 3 mmol, 1 equiv.) in formic acid (130 mL) was stirred at rt for 1 h. The solution was concentrated *in vacuo*, the solid dissolved in a 4:1 mix of sec-butanol : toluene and refluxed for 90 min. After cooling to rt, the solution was concentrated *in vacuo*, and triturated in EtOAc. **Target A** was collected by filtration as a white solid (0.2312 g, 30%). mp 260-261 °C; IR (film) ν_{max}: 3187 (C-H), 3044 (C-H), 2962 (C-H), 2930 (C-H), 2875 (C-H), 1659 (C=O), 1449 (C=O); ^{1}H NMR (300 MHz, DMSO-d_6) δ_H: 0.43-0.70 (8H, m, H-17, 18, 19), 1.29-1.48 (1H, m, H-16), 2.83 (1H, dd, J = 13.4, 5.0 Hz, H-9),

3.15 (1H, dd, J = 13.4, 4.0 Hz, H-9), 3.50-3.63 (1H, m, H-1); 4.14- 4.26 (1H, m, H-4), 7.11-7.29 (5H, m, H-10-15), 7.89 (1H, s, H-5), 8.13 (1H, s, H-2)

N-Aryl isoleucine methyl ester (3a). NEt_3 (0.70 mL) was added to a solution of methyl ester, **1** (0.91 g, 5 mmol, 1 equiv.) in MeOH (18.5 mL), and stirred for 5 min at rt. 4-bromobenzaldehyde (0.93 g, 5 mmol, 1 equiv.) was added and stirred for 4 h at rt. After cooling to 0 °C, $NaBH_4$ (0.3 g, 8 mmol, 1.6 equiv.) was added slowly and stirred for 30 min at rt. The reaction was quenched with 1M NaOH (20 mL) and extracted with EtOAc (2 x 30 mL). The organic portion was extracted with 1M HCl (2 x 30 mL). The aqueous was basified to pH 10 with 2M NaOH, extracted with EtOAc (3 x 30 mL) and washed with brine (30 mL). The organics were dried over $MgSO_4$ and concentrated *in vacuo* to yield a white solid (0.6336 g, 2 mmol, 40 %)

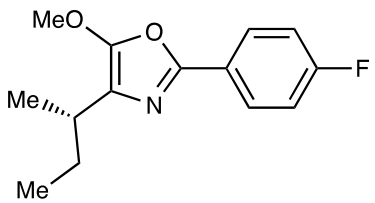
Target B



4-Methoxyphenylboronic acid (0.39 g, 2.6 mmol, 1.3 equiv.) and $P(o\text{-}Tol)_3$ (0.077 g, 0.25 mmol, 0.125 equiv.) were added to a solution of N-aryl ester **3a** (0.63 g, 2 mmol, 1 equiv.) in THF (4 mL) and degassed with N_2. $Pd(AcO)_2$ (0.023 g, 0.1 mmol, 0.05 equiv.) and degassed aq. Na_2CO_3 (3 mL, 1.5 equiv.) were added and refluxed for 2 h. Upon cooling, the solution was diluted with water (5 mL), the aqueous portion was extracted with EtOAc (3 x 15 mL), followed by washing of the organic portion with H_2O (2 x 15 mL) and brine (2 x 15 mL). The organic was dried over $MgSO_4$ and concentrated *in vacuo* to yield the brown solid which was purified by column chromatography (92 % Petroleum ether (40-60): 8 % ethyl acetate) to yield the white solid side product **3b** (0.0150g, 2 %). ^1H NMR (300 MHz, Chloroform-d) δ 3.85 (s, 3H, H-15, 16), 7.06 – 6.87 (m, 2H, H-1, 3, 9, 11), 7.61 – 7.41 (m, 2H, H-4, 6, 8, 12). $^{13}C\{^1H\}$ NMR (75 MHz, $CDCl_3$) δ_C: 55.5, 114.3, 127.9, 133.6. 158.8

***N-Acyl isoleucine methyl ester (4a)*.** The methyl ester, **1** (0.9083 g, 5 mmol, 1 equiv.) was dissolved in DCM (50 mL). NEt_3 (1.40 mL, 10 mmol, 2 equiv.), 4-fluorobenzoic anhydride (1.97 g, 7.5 mmol, 1.5 equiv.) and DMAP (0.061 g, 0.5 mmol, 0.1 equiv) were added. After stirring at rt for 2.5 h, MeOH (50 mL) wad added and stirred for 30 min. The solution was concentrated *in vacuo*, the oil was dissolved in EtOAc (50 mL) and washed with water (30 mL). The aqueous phase was extracted with EtOAc (2 x 30 mL) and the combined organics washed with aq. $CuSO_4$ (30 mL), 2M HCl (15 mL) and brine (30 mL). The organic portion was dried over $MgSO_4$, filtered and recrystallised in hexane. Concentration *in vacuo* yielded the white solid (0.4982g, 37 %).

Target C

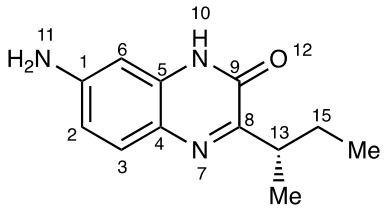


The N-acyl methyl ester **4a** (0.50g, 1.86 mmol, 1 equiv.) was dissolved in chloroform (6.64 mL), followed by addition of CaO (0.52 g, 9.3 mmol, 5 equiv.), celite (0.10g,

20%w/w), and P_2O_5 (1.06 g, 7.44 mmol, 4 equiv.). The reaction was refluxed for 2 h, monitored by TLC, (10 % EtOAc: 90 % Petroleum ether), and a further portion of P_2O_5 (1.06 g, 7.44 mmol, 4 equiv.) was added. After 2 h, the reaction was cooled to 0°C and diluted to half the concentration with 6M NaOH. The solution was filtered through celite, extracted with DCM (2 x 30 mL) and washed with brine (30 mL). The organic portion was dried over $MgSO_4$ and concentrated *in vacuo* to yield the crude brown solid (0.0198 g, 4 %). No data to report.

N-(2,4-dinitrophenyl)- isoleucine methyl ester (5a). Fluoro-2,4-dinitrobenzene (0.42 mL, 3.33 mmol, 1 equiv.), methyl ester, **1** (0.91 g, 5 mmol, 1.5 equiv.), and $NaHCO_3$ (0.56 g, 6.66 mmol, 2 equiv.) were dissolved in THF (15 mL) and stirred overnight at rt. EtOAc (30 mL) and H_2O (30 mL) were added and the aqueous phase was extracted with EtOAc (2 x 30 mL). The organics were washed with aq. $NaCO_3$ (30 mL) and brine (30 mL), dried over $MgSO_4$, and concentrated *in vacuo* to yield the crude product as a yellow solid. ***N-(2,4-diaminophenyl)- isoleucine methyl ester (5b).*** Pd/C (333 mg) was added to a solution of **5a** (1.04 g, 3.33 mmol, 1 equiv.) in methanol (33 mL). The reaction was stirred for 4 h at rt under H_2, filtered through celite and concentrated *in vacuo*.

Target D

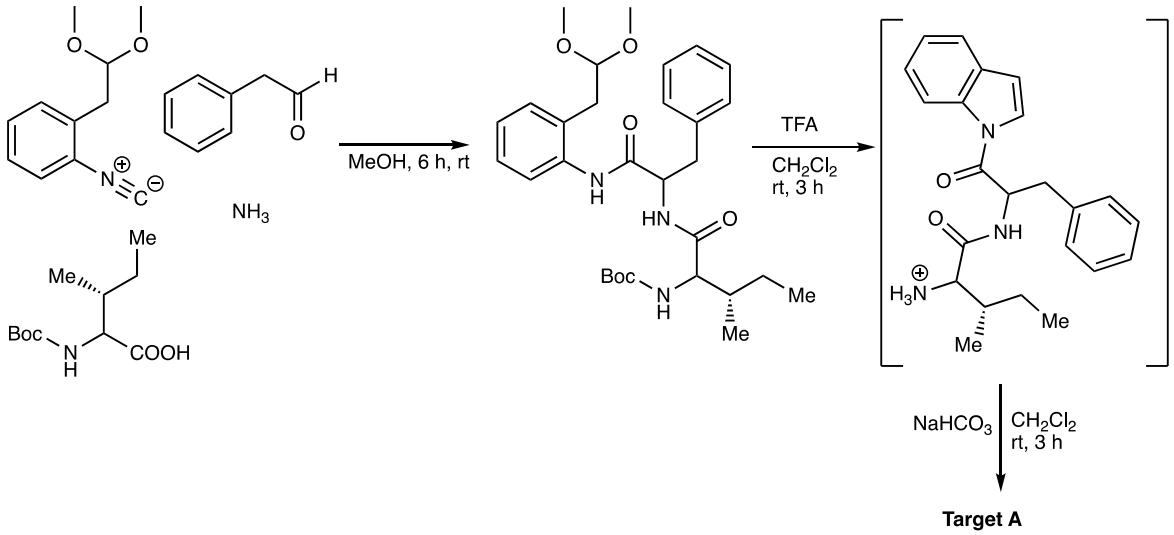


Acetic acid (1 mL) was added to a solution of the diaminophenyl ester **5b**, (0.836 g, 3.33 mmol, 1 equiv.) in ethanol (10 mL) and refluxed for 2 h. The black oil was dissolved in acetone and purified by trituration in ethyl acetate and petroleum ether (40-60) to yield the product as a brown solid (0.2460 g, 34 %). mp 173-184 °C; IR (film) ν_{max}: 3342 (N-H), 2961 (C-H), 2962 (C-H), 2930, 1651 (C=O), 1619 (C=O); ^{1}H NMR (300 MHz, DMSO-d_6) δ 0.83 (d, J = 7.4 Hz, 3H, H-16), 1.12 (d, J = 6.9 Hz, 3H, H-14), 1.83 – 1.37 (m, 2H, H-15), 3.16 (q, J = 6.9 Hz, 1H, H-13), 5.82 (s, 2H, H-11), 6.31 (d, J = 2.3 Hz, 1H, H-6), 6.50 (dd, J = 8.7, 2.4 Hz, 1H, H-2), 7.33 (d, J = 8.7 Hz, 1H, H-3), 11.88 (s, 1H, H-10).

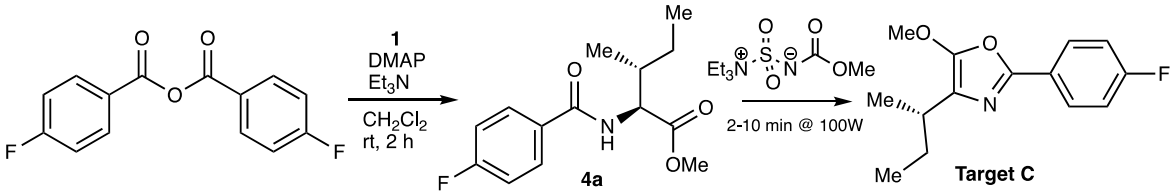
References

1. Aids info, https://aidsinfo.nih.gov/contentfiles/lvguidelines/adultandadolescentgl.pdf accessed April 2019
2. J-F. Mouscadet, L. Tchertanov, *Eur. J. Med. Res.*, 2009, **14**, 5-16
3. N. Sluis-Cremer, G. Tachedjian, *Virus Res.*, 2008, **134**, 147-156
4. Z. Lv, Y. Chu, Y. Wang, *HIV AIDS (Auckl)*, 2015, **7**, 95-104
5. D. Eggink, B. Berkhout, R.W. Sanders, *Curr. Pharm. Des.*, 2010, **16**, 3716-3728
6. C. R. B. Rhoden, D. G. Rivera, *et. al.*, *J. Comb. Chem.*, 2009, **11**, 1078–1082
7. C. Adamo, C. Amatore, I. Ciofini, A. Jutand, H. Lakmini, *J. Am. Chem. Soc.*, 2006, **128**, 6829-6836
8. C. T. Brain, J. M Paul, *Synlett* 1999, **10**, 1642
9. M. Imanishi, M. Sonoda, H. Miyazato, K. Sugimoto, M. Akagawa, S. Tanimori *ACS Omega*, 2017, **2**, 1875−1885
10. H. Masotti, G. Pfeiffer, *et. al.*, *Bulletin des Societes Chimiques Belges*, 1991, **100**, 63 – 77
11. A. Karim, A. Mortreux, *et. al.*, *J. Organomet. Chem.*, 1986, **317**, 93-104

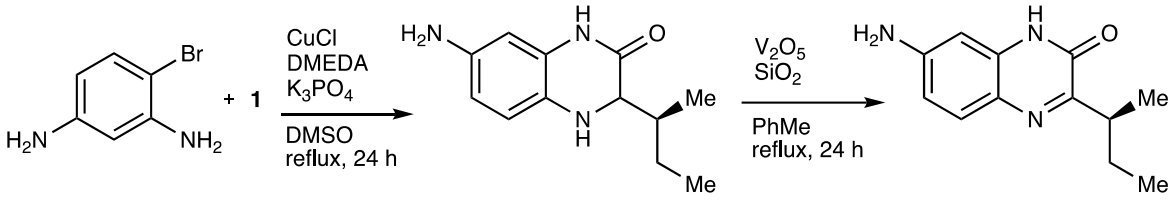
Appendix



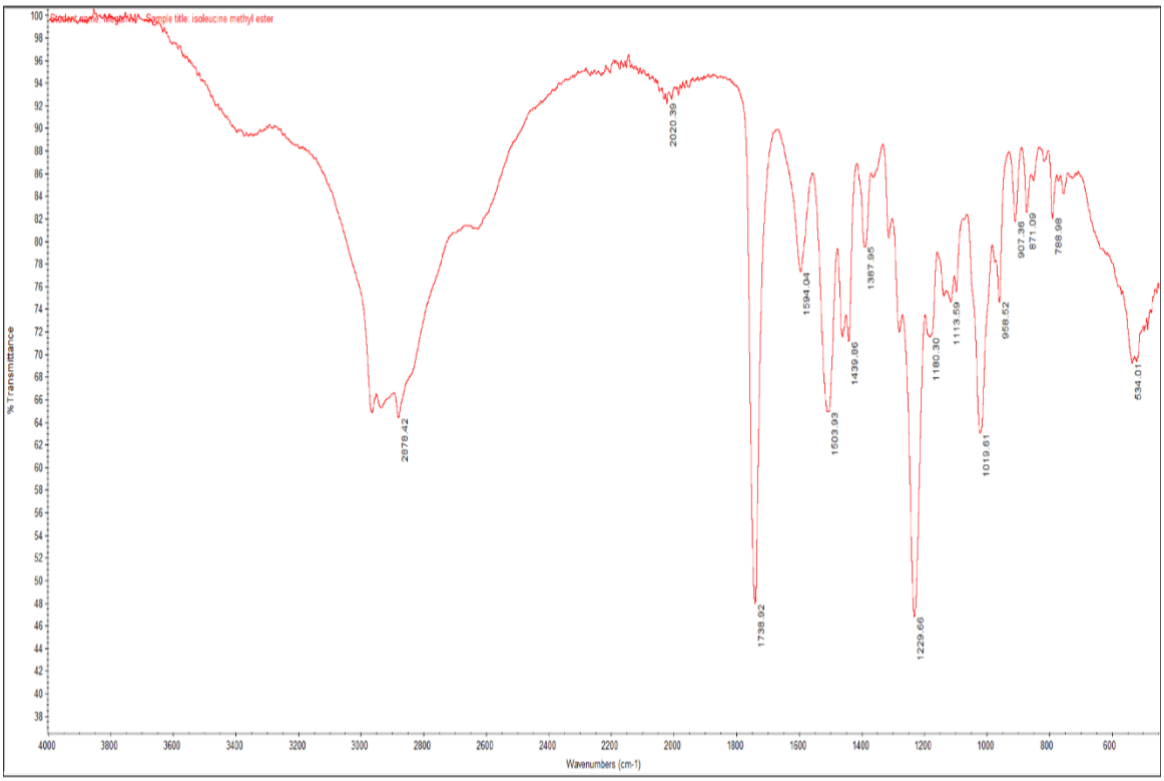
Appendix 1. An alternate synthesis of **Target A**[6]



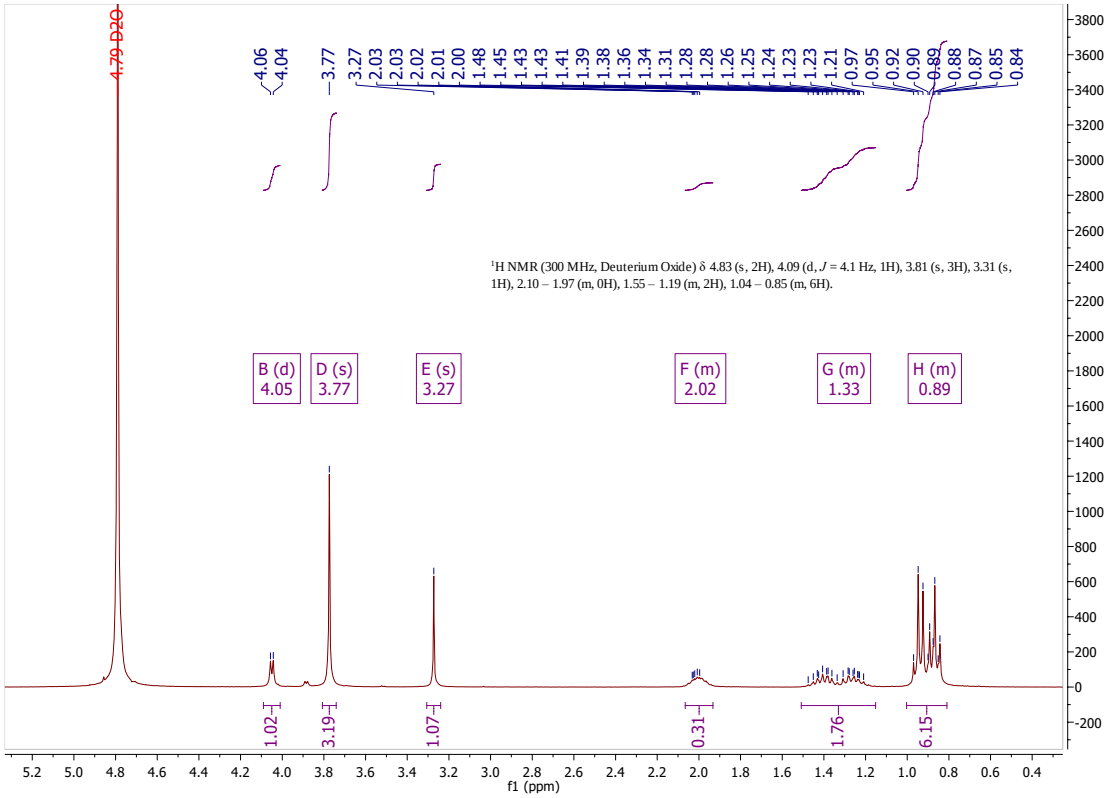
Appendix 2. An alternate synthesis of **Target C**[8]



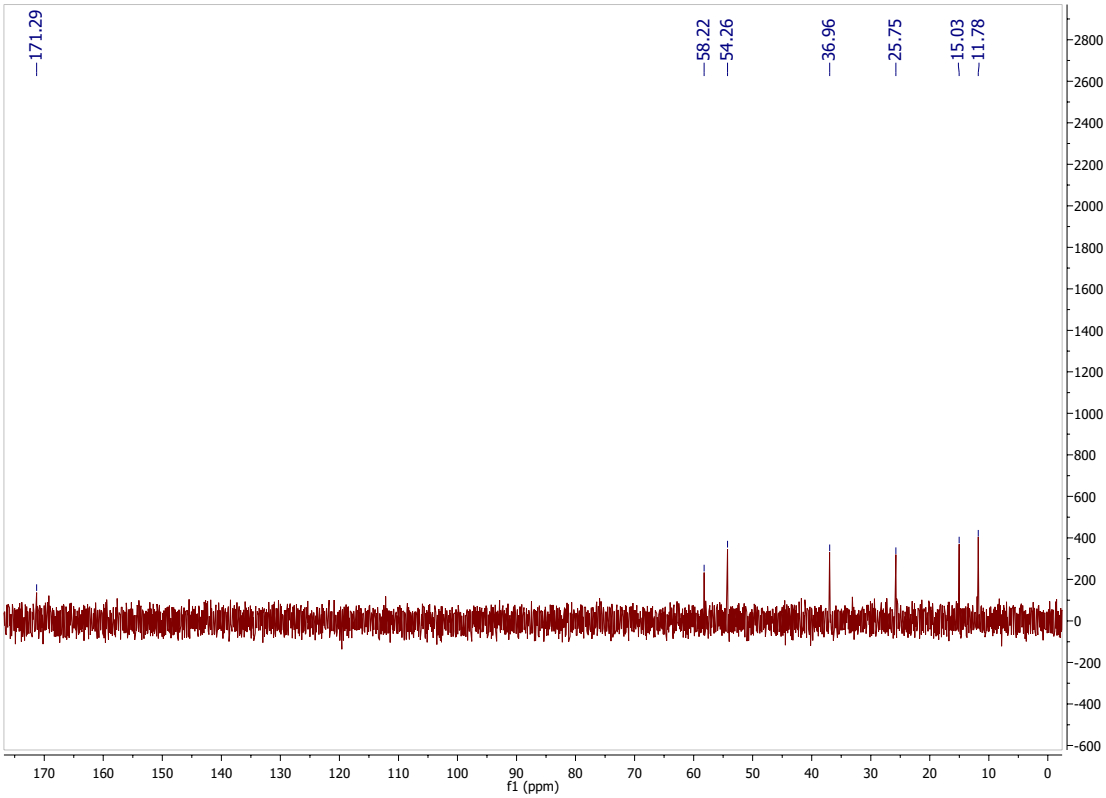
Appendix 3. An alternate synthesis of **Target D**[9]



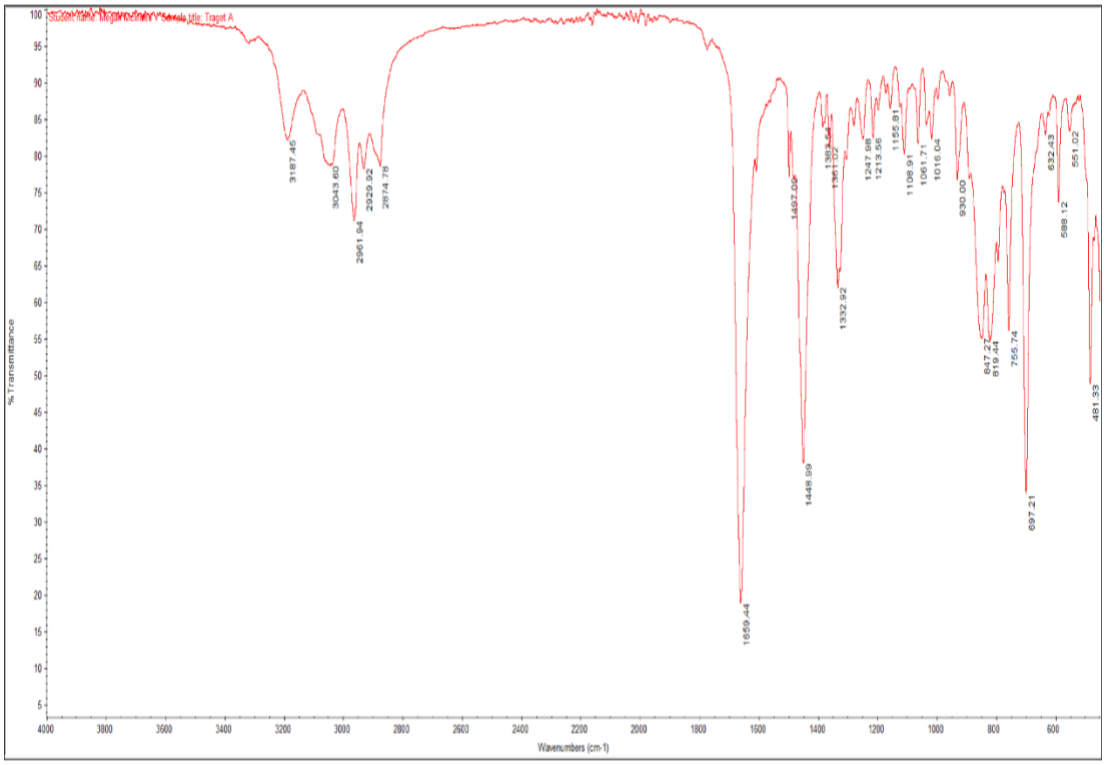
Appendix 4. The IR spectrum of isoleucine methyl ester hydrochloride **(1)**



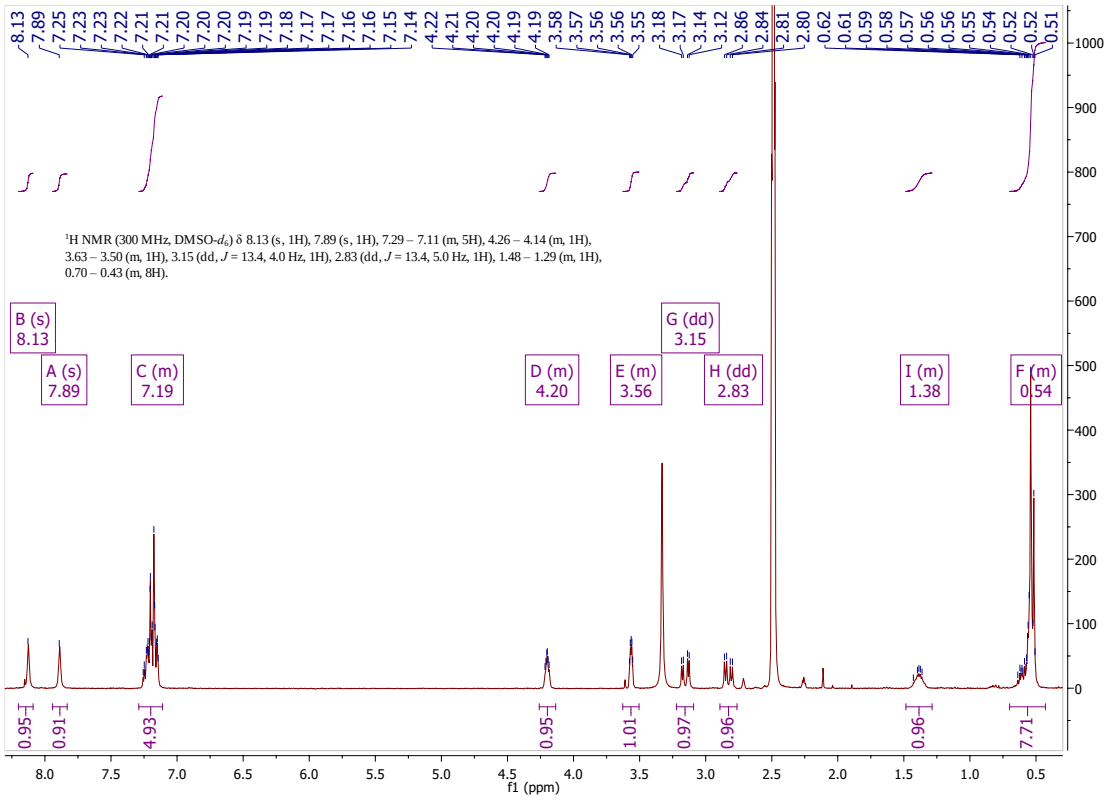
Appendix 5. The ^{1}H NMR spectrum of isoleucine methyl ester hydrochloride **(1)**



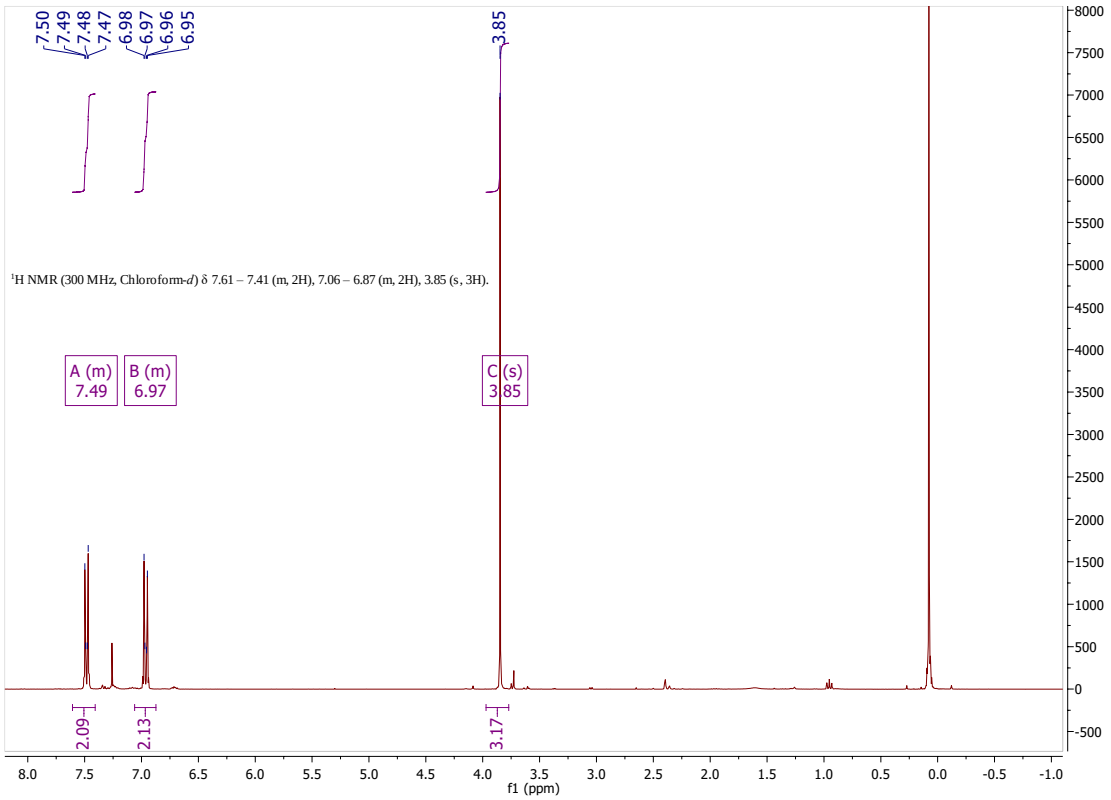
Appendix 6. The ^{13}C NMR spectrum of isoleucine methyl ester hydrochloride **(1)**



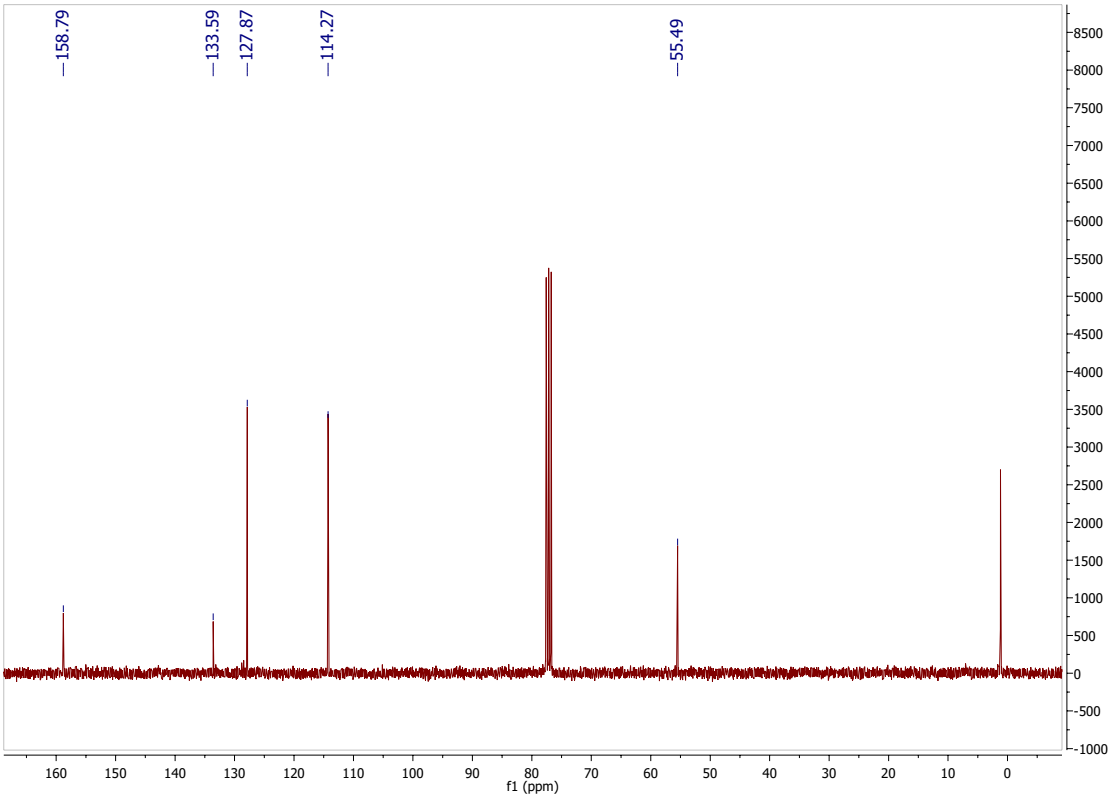
Appendix 7. The IR spectrum of **Target A**



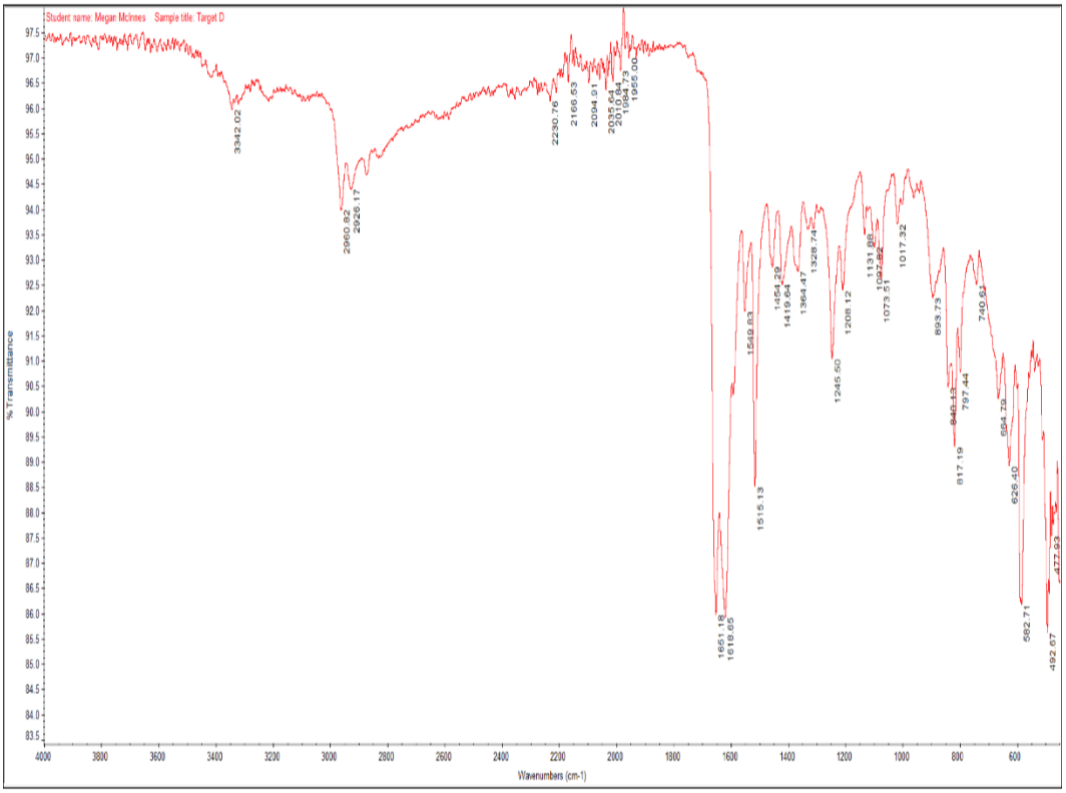
Appendix 8. The ^{1}H NMR spectrum of **Target A**



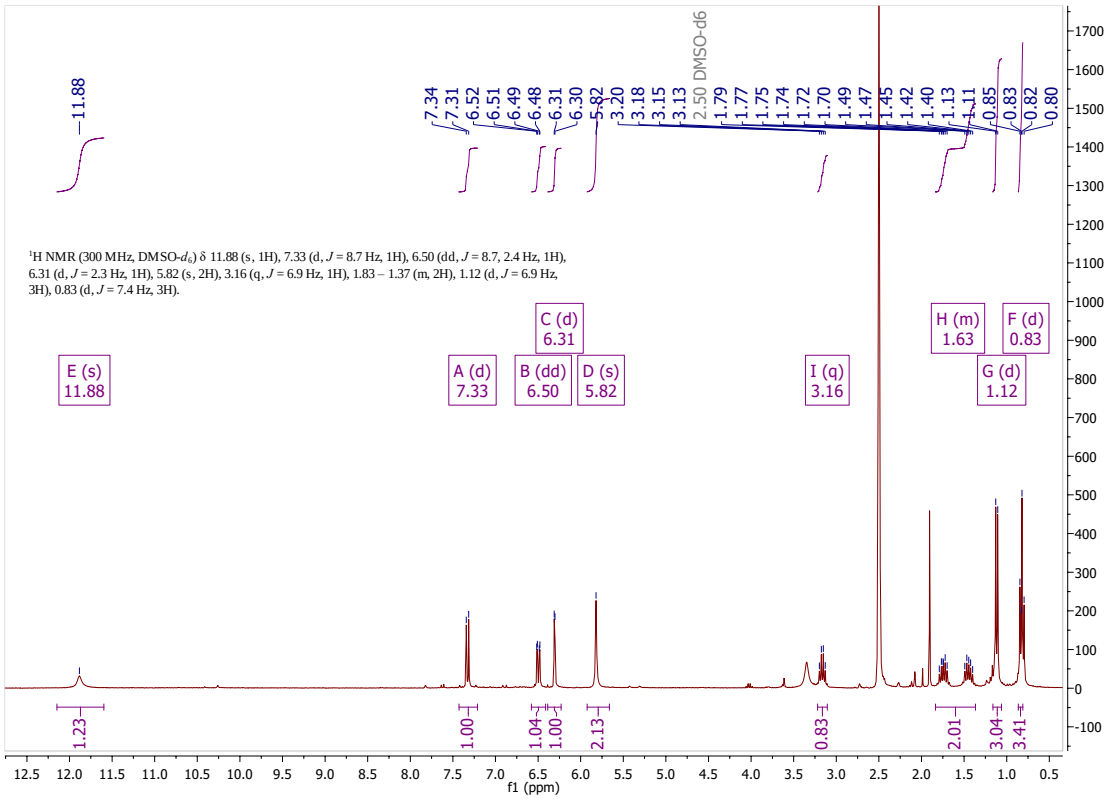
Appendix 9. The ^{1}H NMR spectrum of target B side product, 4,4'-dimethoxy-1,1'-biphenyl **3b**



Appendix 10. The ^{13}C NMR spectrum of target B side product, 4,4'-dimethoxy-1,1'-biphenyl **3b**



Appendix 11. The IR spectrum of **Target D**



Appendix 12. The ^{1}H NMR spectrum of **Target D**