The role of solid-state reactivity in the prebiotic selection of nucleobases and formation of peptides

Stolar, Tomislav

University of Zagreb
FACULTY OF SCIENCE

Tomislav Stolar

THE ROLE OF SOLID-STATE REACTIVITY IN THE PREBIOTIC SELECTION OF NUCLEOBASES AND THE FORMATION OF PEPTIDES

DOCTORAL DISSERTATION

Supervisor: Dr. Krunoslav Užarević

Zagreb, 2022.

Sveučilište u Zagrebu
PRIRODOSLOVNO-MATEMATIČKI FAKULTET

Tomislav Stolar

ULOGA REAKTIVNOSTI U ČVRSTOM STANJU U PREBIOTIČKOM ODABIRU NUKLEOBAZA I STVARANJU PEPTIDA

DOKTORSKI RAD

Mentor: Dr. sc. Krunoslav Užarević

Zagreb, 2022.

Acknowledgements

I am thankful to my wife Jasna Alić Stolar, my mum Snježana, and my brothers Dado, Josip, and Angelo for their support. Also, to Rea for being the best dog ever and doing what only dogs can do. I love you with all my heart.

Thanks to Krunoslav Užarević for being the best possible mentor for me. For giving me space to grow and being alongside me through ups and downs.

Thanks to Ernest Meštrović for supporting me throughout my professional career: for advice, motivation, and access to experts and equipment to work with.

Thanks to José G. Hernández for being a great mentor and friend. José is one of the most talented scientists that I know, and he will do great things for the Colombian and Latin American scientific community.

Thanks to Saša Grubešić for helping and teaching me about analytical chemistry. Saša will soon obtain his PhD degree which he undoubtedly deserves given his expertise.

Thanks to Luca Grisanti for believing in my crazy ideas and working with me. I am looking forward to publishing the results that we obtained.

Thanks to Martin Etter for mentoring me about crystal structure solution from powder X-ray diffraction data and providing me with the software to work with. You helped me tremendously.

Thanks to my scientific brother Stipe Lukin for joining me in many scientific adventures. I enjoyed the time, discussions, and learned a lot from you.

Thanks to my collaborators: Dr. Nikola Cindro (University of Zagreb, Chemistry Department, Croatia), Dr. Ivica Đilović (University of Zagreb, Chemistry Department, Croatia), Dr. Nenad Tomašić (University of Zagreb, Geology Department, Croatia), Dr. Mario Cindrić (Ruđer Bošković Institute, Croatia), Mrs. Lucija Dončević (Ruđer Bošković Institute, Croatia), Dr. Krešimir Molčanov (Ruđer Bošković Institute, Croatia), Dr. Ivor Lončarić (Ruđer Bošković Institute, Croatia), Dr. Ana Palčić (Ruđer Bošković Institute, Croatia), Dr. Gregor Mali (National Institute of Chemistry, Slovenia), Dr. Andraž Krajnc (National Institute of Chemistry, Slovenia), Dr. Ben K. D. Pearce (Johns Hopkins University, USA). I hope we will soon publish our results if we already didn't do that.

Thanks to: Dr. N. V. Hud (Georgia Tech, USA), Dr. Jason Dworkin (NASA), Dr. Michael Callahan (Boise State University, USA), Dr. Ramanarayanan Krishnamurthy (The Scripps Research Institute, USA), Prebiotic Chemistry & Early Earth Environments community supported by NASA Astrobiology programme, and participants of Molecular Origins of Life conference organized by Emergence of Life Initiative for the discussion.

Thanks to my PhD committee members: Prof. Mirta Rubčić, Prof. Sanda Rončević, and Prof. Oliver Trapp. I enjoyed the discussion.

Thanks to my best man Mario Špadina for being supportive whenever needed. Thanks to all other friends who helped me in my life.

Contents

ABSTRACT ..... **IX**
SAŽETAK ..... **XIII**
PROŠIRENI SAŽETAK ..... **XIII**
§ 1. INTRODUCTION **1**
§ 2. LITERATURE REVIEW **5**
2.1. Prebiotic selection of nucleobases **5**
2.2. Prebiotic peptide synthesis **8**
2.3. Mechanochemical reactions in prebiotic chemistry **11**
§ 3. EXPERIMENTAL **15**
3.1. Molecular recognition of nucleobases **15**
3.1.1. General conditions *15*
3.1.2. Variable-temperature synchrotron powder X-ray diffraction monitoring *16*
3.1.3. Crystal structure solution from powder X-ray diffraction data *16*
3.1.4. Other solid-state analytical techniques *18*
3.2. Peptide synthesis **18**
3.2.1. General conditions *18*
3.2.2. Ion-pairing high-performance liquid chromatography *18*
3.2.3. Mass spectrometry *19*
§ 4. RESULTS AND DISCUSSION **23**
4.1. Prebiotic emergence of molecular recognition on a nucleobase level **23**
4.1.1. 2,6-Diaminopurine and uracil *23*
4.1.2. 2,6-Diaminopurine and thymine *27*
4.1.3. Adenine and hypoxanthine *33*
4.1.4. Selectivity of base-pairing and other combinations *43*
4.2. Mechanochemical prebiotic peptide synthesis **46**
4.2.1. Amino acid oligomerization *46*
4.2.2. Amino acid and α-hydroxy acid oligomerization *60*
§ 5. CONCLUSION **64**
§ 6. REFERENCES **66**
§ 7. APPENDIX ..... **XXIII**
§ 8. CURRICULUM VITAE ..... **LXXXV**

University of Zagreb
Faculty of Science
Department of Chemistry

Doctoral Thesis

ABSTRACT

THE ROLE OF SOLID-STATE REACTIVITY IN THE PREBIOTIC SELECTION OF NUCLEOBASES AND THE FORMATION OF PEPTIDES

Tomislav Stolar
Ruđer Bošković Institute, Bijenička c. 54, 10000 Zagreb, Croatia

Within this thesis, solid-state reaction medium and mechanochemical methodologies were explored to unveil weather nucleobases can self-assemble and amino acids can oligomerize under these conditions. X-ray diffraction and differential scanning calorimetry analysis showed that 2,6-diaminopurine (D) selectively self-assembles with uracil (U) and thymine (T), while adenine (A) self-assembles with hypoxanthine (H). Crystal structures of D-U cocrystal hydrate, D-T cocrystal hydrate, D-T cocrystal anhydrate, A-H cocrystal hydrate, and A-H cocrystal anhydrate were determined from powder X-ray diffraction and single-crystal X-ray diffraction data. Ion-pairing high performance liquid chromatography and mass spectrometry analysis showed that glycine and alanine undergo oligomerization by mechanical activation. Based on the results, it can be concluded that minerals as additives, time, mechanical loading, and temperature are all important parameters for the mechanochemical prebiotic peptide synthesis. In summary, solid-state reactivity of nucleobases and amino acids might have been an important factor in their prebiotic chemical evolution.

(160 pages, 119 figures, 14 tables, 161 references, original in english)

Thesis deposited in Central Chemical Library, Horvatovac 102A, Zagreb, Croatia and National and University Library, Hrvatske bratske zajednice 4, Zagreb, Croatia.

Keywords: amino acids/ mechanochemistry/ molecular recognition/ nucleobases/ peptides/ prebiotic chemistry/ solid-state reactivity

Supervisor: Dr. Krunoslav Užarević, Senior Research Associate

Thesis accepted: 29.06.2022.

Reviewers:
Prof. dr. sc. Mirta Rubčić
Prof. dr. sc. Sanda Rončević
Prof. dr. Oliver Trapp

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Kemijski odsjek

Doktorska disertacija

SAŽETAK

ULOGA REAKTIVNOSTI U ČVRSTOM STANJU U PREBIOTIČKOM ODABIRU NUKLEOBAZA I STVARANJU PEPTIDA

Tomislav Stolar
Institut Ruđer Bošković, Bijenička c. 54, 10000 Zagreb, Hrvatska

U ovoj doktorskoj disertaciji koristili smo čvrsto stanje kao reakcijski medij i mehanokemijske metodologije da bismo istražili dolazi li do prepoznavanja nukleobaza i oligomerizacije aminokiselina u tim uvjetima. Analiza difrakcijom rentgenskog zračenja i diferencijalnom pretražnom kalorimetrijom pokazala je selektivno molekulsko prepoznavanje 2,6-diaminopurina (D) s uracilom (U) i timinom (T), kao i adenina (A) s hipoksantinom (H). Riješene su kristalne strukture D-U kokristalnog hidrata, D-T kokristalnog hidrata, D-T kokristalnog anhidrata, A-H kokristalnog hidrata te A-H kokristalnog anhidrata iz podataka difrakcije rentgenskog zračenja. Analiza tekućinskom kromatografijom visoke djelotvornosti i spektrometrijom mase pokazala je da glicin i alanin oligomeriziraju u peptide uslijed mehanokemijske aktivacije. Rezultati su pokazali da su minerali kao aditivi, vrijeme, prenošenje mehaničke energije, i temperatura ključni parametri za mehanokemijsku prebiotičku sintezu peptida. Zaključeno je da je reaktivnost nukleobaza i aminokiselina u čvrstom stanju mogla biti važan faktor u njihovoj prebiotičkoj kemijskoj evoluciji.

(160 stranica, 119 slika, 14 tablica, 161 literaturnih navoda, jezik izvornika: engleski)

Rad je pohranjen u Središnjoj kemijskoj knjižnici, Horvatovac 102a, Zagreb i Nacionalnoj i sveučilišnoj knjižnici, Hrvatske bratske zajednice 4, Zagreb.

Ključne riječi: aminokiseline/ čvrsto stanje/ mehanokemija/ molekulsko prepoznavanje/ nukleobaze/ peptidi/ prebiotička kemija/

Mentor: dr. sc. Krunoslav Užarević., viši zn. sur.

Rad prihvaćen: 29.06.2022.

Ocjenitelji:
1. prof. dr. sc. Mirta Rubčić
2. prof. dr. sc. Sanda Rončević
3. prof. dr. Oliver Trapp

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Kemijski odsjek

Doktorska disertacija

PROŠIRENI SAŽETAK

1. UVOD

Prebiotička kemija je kemija koja prethodi nastanku života [L. E. Orgel, *Crit. Rev. Biochem. Mol. Biol.* **39** (2004) 99–123.]. Kroz kemijsku evoluciju raste kompleksnost na molekulskoj razini te ona prethodi biološkoj evoluciji. S obzirom da je površina Zemlje danas znatno izmijenjena uslijed različitih geoloških procesa tijekom vremena, nema povijesnog dokaza koji bi upućivao na način na koji je život nastao. Kao najizglednije lokacije na kojoj je život mogao nastati smatraju se kopno ili plitki kopneni bazeni [B. Damer, D. Deamer, *Astrobiology* **20** (2020) 429–452.]. Jedna od nepoznanica je razlog iz kojeg su kanonske nukleobaze odabrane i ugrađene u nukleinske kiseline s obzirom da se međusobno ne povezuju Watson-Crick načinom prepoznavanja u vodenim otopinama [D. M. Fialho, T. P. Roche, N. V. Hud, *Chem. Rev.* **120** (2020) 4806–4830.]. Naša hipoteza je da su druge kombinacije nukleobaza preferirane u odnosu na kanonske kombinacije. S druge strane, povezivanje slobodnih aminokiselina u oligomerne lance zahtjeva eliminaciju molekula vode. Iz tog razloga je oligomerizacija aminokiselina do peptida termodinamički nepovoljan proces u vodenom mediju [G. Danger, R. Plasson, R. Pascal, *Chem. Soc. Rev.* **41** (2012) 5416–5429.]. Naša hipoteza je da mehanokemijska aktivacija slobodnih aminokiselina na površinama minerala može omogućiti nastanak peptida.

2. LITERATURNI PREGLED

2.1. Prebiotički odabir nukleobaza

Nukleobaze (dušične baze) djeluju kao prepoznavajuće jedinice unutar nukleinskih kiselina i njihova komplementarnost je najpoznatiji primjer molekulskog prepoznavanja u prirodi. Pretpostavlja se da je na ranoj Zemlji bio dostupan veliki izbor dušikovih heterocikličkih spojeva koju su mogli poslužiti kao nukleobaze [D. M. Fialho, T. P. Roche, N. V. Hud, *Chem. Rev.* **120** (2020) 4806–4830.]. Argumenti za takvu hipotezu su brojni sintetski eksperimenti koji rezultiraju velikim brojem različitih purinskih i pirimidinskih baza [C. Menor-Salván, M. R. Marín-Yaseli, *Chem. Soc. Rev.* **41** (2012) 5404–5415.]. Također, pronalazak različitih nukleobaza u meteoritima [M. P. Callahan, K. E. Smith, H. J. Cleaves II, J. Ruzicka, J. C. Stern, D. P. Glavin, C. H. House, J. P. Dworkin, *Proc. Natl. Acad. Sci. U.S.A.* **108** (2011) 13995–13998.], kao i njihova sinteza u interstelarnim uvjetima pri niskim temperaturama [Y. Oba, Y. Takano, H. Naraoka, N. Watanabe, A. Kouchi, *Nat. Commun.* **10** (2019) 4413.], snažno upućuju u širok bazen nukleobaza koje su vjerojatno bile prisutne prije u Sunčevom sustavu prije nastanka života. Prema tome, velika je nepoznanica na koji način je došlo do prebiotičkog odabira nukleobaza koje su ugrađene u nukleinske kiseline s obzirom da se kanonske nukleobaze međusobno ne prepoznaju u vodenom mediju.

2.2 Prebiotička sinteza peptida

Smatra se da su peptidi zbog svojih katalitičkih svojstava i simbioze s nukleinskim kiselinama igrali važnu ulogu na ranoj Zemlji [M. Frenkel-Pinter, M. Samanta, G. Ashkenasy, L. J. Leman, *Chem. Rev.* **120** (2020) 4707–4765.]. S obzirom da je u vodenom mediju nastajanje peptidne veze termodinamički nepovoljan proces, prebiotička sinteza peptida u vodi je izazovna. Postoje brojni međusobno oprečni modeli koji objašnjavaju prebiotičku sintezu peptida, međutim oko njih ne postoji konsenzus u znanstvenoj zajednici [T. D. Campbell, R. Febrian, H. E. Kleinschmidt, K. A. Smith, P. J. Bracher, *ACS Omega* **4** (2019) 12745–12752.]. Brojni dokazi upućuju da je aminokiselina glicin bila prisutna na ranoj Zemlji, u vanzemaljskim tijelima [J. R. Cronin, C. B. Moore, *Science* **172** (1971) 1327–1329.], i interstelarnom mediju [S. Ioppolo, G. Fedoseev, K. J. Chuang, H. M. Cuppen, A. R. Clements, M. Jin, R. T. Garrod, D. Qasim, V. Kofman, E. F. van Dishoeck, H. Linnartz, *Nat Astron* **5** (2021)197–205.] prije nastanka života.

2.3. Mehanokemijske reakcije u prebiotičkoj kemiji

Mehanokemija je održiva sintetska metoda koja potiče kemijsku reaktivnost putem mehaničke aktivacije [K. J. Ardila-Fierro, J. G. Hernández, *ChemSusChem* **14** (2021) 2145–2162.]. Neki od izvora mehaničke energije na ranoj Zemlji mogli su uključivati udare meteorite i kometa, tektonske poremećaje, eroziju tla, potrese, i vremenske utjecaje poput vjetra ili kiše. Nedavno je objavljeno nekoliko radova koji pokazuju važnost mehanokemijske reaktivnosti za prebiotičku kemiju. Oni uključuju sintezu prekursora aminokiselina [C. Bolm, R. Mocci, C. Schumacher, M. Turberg, F. Puccetti, J. G. Hernández, *Angew. Chem. Int. Ed.* **57** (2018) 2423–2426.], modifikaciju aminokiselina [J. E. Elsila, N. M. Johnson, D. P. Glavin, J. C. Aponte, J. P. Dworkin, *Meteorit. Planet. Sci.* **56** (2021) 586–600.], i sintezu šećera [S. Lamour, S. Pallman, M. Hass, O. Trapp, *Life* **9** (2019) 52.; M. Hass, S. Lamour, S. B. Christ, O. Trapp, *Commun. Chem.* **3** (2020) 140.].

3. EKPERIMENTALNI DIO

3.1. Molekulsko prepoznavanje nukleobaza

3.1.1. Opći dio

Nukleobaze koje su korištene u ovome radu kupljene su od proizvođača Carbosynth. Mehanokemijske reakcije provedene su koristeći InSolido Technologies kuglični vibracijski mlin i posudice od stakloplastike, zapremnine 14 mL. Dvije čelične kuglice mase 1,4 g i promjera 7 mm korištene su za mljevenje. Reakcije su provedene pri frekvenciji od 30 Hz u trajanju od 2 sata. Za provođenje tekućinski potpomognutih reakcija mljevenja, 50 μL H_2O je ukupno dodano u 200 mg mase reakcijske smjese. Mehanokemijska reakcija u kriogenim uvjetima provedena je u istim uvjetima. Posudice s reakcijskom smjesom su uronjene u tekući dušik prije početka mljevenja te je tijekom reakcije osiguran kontinuirani dotok tekućeg dušika. Reakcije u vodi provedene su miješajući ekvimolarne količine reaktanata (ukupna masa reakcijske smjese iznosila je 100 mg) u 3 mL vode pri sobnoj temperaturi. Miješanje uz pomoć magnetske miješalice provedeno je u okruglim tikvicama volumena 10 mL, a trajalo je nekoliko dana. Jedinični kristali spojeva **2** i **10** dobiveni su hlapljenjem metanolnih otopina pripravljenih spojeva **1** i **4** (10 mg u 5 mL metanola). Hidrotermalna sinteza **9** provedena je zagrijavajući

ekvimolarne količine adenina (A) i hipoksantina (H) na 120 °C preko noći u kvarcnoj epruveti koja podnosi visoki tlak. Za analizu nuklearnom magnetskom rezonancijom u čvrstom stanju, uzorak je pripremljen suhim mljevenjem ekvimolarne količine A i H u čeličnoj posudici koristeći dvije čelične kuglice mase 1,4 g te naknadnim grijanjem tako priređene smjese do 300 °C u eksperimentu diferencijalne pretražne kalorimetrije (posudica od platine, 5 °C/min brzina zagrijavanja, atmosfera dušika). Difraktogrami praškastih uzoraka, koji su korišteni za rješavanje strukture odgovarajućih uzoraka, prikupljeni su u rotirajućoj staklenoj kapilari na laboratorijskom difraktometru (Stoe Stadi-P, λ = 1,54059 Å). Podaci difrakcije na jediničnom kristalu prikupljeni su na Oxford Diffraction Xcalibur Nova R instrumentu (λ = 1,54719 Å) opremljenom s Oxford CryoJet instrument koji koristi hlađenje tekućim dušikom. Programski paket CrysAlis PRO12 [CrysAlisPro, Rigaku Oxford Diffraction, 2018, version: 1.171.39.46, Rigaku Corporation, Oxford, UK] je korišten za redukciju podataka i numeričku absorpcijsku korekciju. Strukture iz podataka na monokristalu su određene i utočnjene pomoću SHELXL [G. M. Sheldrick, *Acta Cryst.* **A71** (2015) 3–8.]. U **2**, kristalizacijska molekula vode (O3) je u neredu kroz dvije pozicije (označene s A i B) te okupancijama od 63 % i 37 %.

3.1.2. Difrakcija sinkrotronskog rentgenskog zračenja na polikristalnom uzorku pri promijenjivim temperaturama

Mjerenja su provedena na P02.1 beamline na PETRA III, DESY pri čemu su korištene staklene kapilare promjera 0,5 mm. Uzorci su grijani pomoću Oxford Cryostream sustava u temperaturnom području 300-500 K. Udaljenost detektora i uzorka je iznosila 2010 mm. Korišteno je rentgensko zračenje od 0.20741 Å i difrakcijska slika je prikupljena na vrtećoj kapilari. Svaki uzorak je termalno uravnotežen na 300 K te je grijan brzinom od 5 K/min. Difrakcijske slike su prikupljene u razmaku od 10 K. Nakon što je uzorak postigao željenu temperaturu, termalno je uravnotežen, te je potom prikupljena difrakcijska slika.

3.1.3. Rješavanje kristalnih struktura iz difrakcijskih podataka na polikristalnom uzorku

Nakon indeksiranja, ćelija koja je bila najbolja kandidat je korištena za Pawley fit [G. S. Pawley, *J. Appl. Crystallogr.* **14** (1981) 357.]. Struktura je rješena metodom kaljenja krutog tijela u direktnom prostoru. Konačno utočnjavanje po Rietveldu [H. M. Rietveld, *J. Appl. Crystallogr.* **2** (1969) 65–71.] je napravljeno koristeći geometrijska ograničenja za duljine veza

i iznose kuteva, te planarnih ograničenja po potrebi. Svi računi su napravljeni u programu Topas [TOPAS, version 4.2; Bruker-AXS: Karlsruhe, Germany, 2009.].

3.1.4. Ostale analitičke tehnike u čvrstom stanju

ATR-FTIR spektri su snimljeni pomoću spektrometra PerkinElmer SpectrumTwo opremljenog s dijamantnom ćelijom u rasponu 4000–400 cm^{-1}. Difraktogrami praha su prikupljeni na PanAlytical Aeris difraktometru (CuKα zračenje, Ni filter) u Bragg-Bretano geometriji, pri čemu je uzorak bio nanesen na pločicu od silicija. Eksperimenti diferencijalne pretražne kalorimetrije provedeni su na Simultaneous Thermal Analyzer (STA) 6000 (PerkinElmer, Inc.) u platinskim posudicama te brzinom zagrijavanja od 5 °C/min u atmosferi dušika protoka 20 mL/min.

3.2. Sinteza peptida

3.2.1. Opći dio

Glicin, 2,5-diketopiperazin, TiO_2, SiO_2, i Montmorilonit K10 kupljeni su od Sigma-Aldrich. Gly_2–Gly_4 i L-Ala-L-Ala kupljeni su od Bachem. L-Ala kupljena je od Carbosynth. Mica (MICA-V-4) kupljena je od SPI. Svi mehanokemijski eksperimenti provedeni su koristeći InSolido Technologies vibracijski kuglični mlin. Za eksperimente s TiO_2, reakcije su provedene pri frekvenciji od 30 Hz u trajanju od 18 h. Ukupna masa reakcijske smjese iznosila je oko 400 mg (0,8 mmol Gly i 4 mmol TiO_2). Korištene su 14 mL posudice od čelika specifično dizajnirane za provođenje mehanokemijskih reakcija pri povišenoj temperaturi uz tri čelične kuglice mase 1 g i promjera 6 mm za mljevenje. 25 µL H_2O dodano je kao tekućinski aditiv u tekućinski potpomognute reakcije mljevenja. Za eksperimente sa SiO_2, reakcije su provedene pri različitim frekvencijama i vremenima mljevenja. Ukupna masa reakcijske smjese iznosila je 500 mg (1,33 mmol Gly i 6,66 mmol SiO_2). Korištene su čelične posudice volumena 28 mL specifično dizajnirane za provođenje mehanokemijskih reakcija pri povišenoj temperaturi te dvije čelične kuglice mase 1,4 g i promjera 7 mm za mljevenje.

3.2.2. Tekućinska kromatografija visoke djelotvornosti

Oligopeptidi su analizirani tekućinskom kromatografijom visoke djelotvornosti na Agilent 1290 Infinity instrumentu koji je bio opremljen Phenomenex Luna C18 kolonom (250 mm dužine, 4,6 mm promjer, i 3 µm veličina čestica). Mobilna faza je bila vodena otopina 50 mM

KH_2PO_4 i 7,2 mM natrijev heksansulfonat. pH je podešen na 2,5 dodatkom 85 % H_3PO_4 (HPLC čistoća). Mobilna faza je korištena izokratno s protokom od 1 mL/min. Uzorak je injektiran u volumenu od 1 µL, a vrijeme metode je bilo 20 min. Kolona je uravnotežena na 30 °C. Analiza se temeljila na UV-VIS detektoru pri absorbanciji od 195 nm. 2,5 mg Gly, 2,5-diketopiperazina, Gly_2, Gly_3, i Gly_4 prebačeno je u 1,5 mL HPLC vialu i napunjeno s mobilnom fazom. Viale su protresene na IKA Vortex u trajanju od 1 min kako bi se pospješilo otapanje. Za analizu uzoraka nakon provedenih reakcija, oko 40 mg svakog uzorka prebačeno je u Falcon epruvetu i napunjeno s 1,5 mL mobilne faze. Svaka epruveta je protresena na IKA Vortex u trajanju od 1 min kako bi se pospješilo otapanje oligopeptida. Nakon toga, epruvete su centrifugirane pri 15000 rpm u trajanju od 5 min. Naposljetku, 0,8 mL supernatanta prebačeno je u staklenu vialu te je analizirano pomoću HPLC.

3.2.3. Spektrometrija mase

Oligopeptidi su analizirani sa spregnutim sustavom tekućinske kromatografije koristeći Agilent 1290 Infinity instrument spojen na Q-TOF MS (Model 6545) instrument opremljen s Waters Acquity BEH HILIC amidnom kolonom (duljine 150 mm, promjera 2,1 mm i veličine čestica 1.7 µm). Mobilna faza (MF) je bila otapalo A: ACN i B: 5 mM $HCOONH_4$ pri pH 3,0 podešeno s HCOOH. Korišteno je izokratno eluiranje kroz 2 min (10 % MF B) nakone čega je uslijedilo gradijentno eluiranje (2-20 min, 10 % MF B do 50 % MF B). Protok je bio 0,4 mL/min. Volumen injektiranog uzorka iznosio je 3 µL, ukupno vrijeme metode bilo je 25 min te je kolona ekvilibrirana na 30 °C. Korišten je dvostruki AJS ESI izvor iona u pozitivnom (+) ESI modalitetu. Parametri temperature plina (300 °C), plina za sušenje (6 L/min), nebulizatora (60 psi), kapilarne struje (3000 V), fragmentora (100 V) i skimmera (65 V) optimizirani su koristeći standard glicina. Q-TOF MS je korišten u spektralnom masenom području 40-1700 m/z te je korišteno akvizicijsko vrijeme od 1,4 spektra/s (akvizicijsko vrijeme = 714 ms/spektru). Napravljeno je unutarnje utočnjavanje mase i kalibracija instrumenta Q-TOF prije analiziranja uzoraka. Instrument je kalibriran u visokom rezolucijskom modalitetu (4 GHz, 0-1700 m/z). Uzorci koji su korišteni za analizu tekućinskom kromatografijom visoke djelotvornosti su razrijeđeni u acetonitrilu (100 µL uzorka i 900 µL acetonitrila).

4. REZULTATI I RASPRAVA

4.1. Molekulsko prepoznavanje nukleobaza

4.1.1. 2,6-Diaminopurin i uracil

Provedeni su eksperimenti prepoznavanja nukleobaza u različitim uvjetima prebiotičke zastupljenosti vode. Za opisivanje uvjeta velikog suviška vode korištene su reakcije miješanja krute reakcijske smjese u zasićenoj vodenoj otopini. Za opisivanje uvjeta katalitičke prisutnosti vode korištene su tekućinski potpomognute reakcije mljevenja. Za opisivanje uvjeta odsustva vode korištene su reakcije grijanja u čvrstome stanju. Ovisno o reakcijskim uvjetima, 2,6-diaminopurin i uracil tvore kokristalni hidrat ili bezvodnu formu kokristala. Tehnikom difrakcije rentgenskog zračenja na jediničnom kristalu određena je kristalna struktura hidratnog kokristala 2,6-diaminopurina i uracila.

4.1.2. 2,6-Diaminopurin i timin

Ovisno o reakcijskim uvjetima, 2,6-diaminopurin i timin tvore kokristalni hidrat ili bezvodnu formu kokristala. Tehnikom difrakcije rentgenskog zračenja na jediničnom kristalu određena je kristalna struktura hidratnog kokristala 2,6-diaminopurina i timina. Tehnikom difrakcije rentgenskog zračenja na polikristalnom uzorku određena je kristalna struktura bezvodne forme kokristala 2,6-diaminopurina i timina.

4.1.3. Adenin i hipoksantin

Ovisno u reakcijskim uvjetima, adenin i hipoksantin tvore kokristalni hidrat ili bezvodnu formu kokristala. Tehnikom difrakcije rentgenskog zračenja na polikristalnom uzorku određene su kristalne strukture kokristalnog hidrata i bezvodne forme kokristala adenina i hipoksantina. Kao komplementara metoda difrakciji na polikristalnom uzorku korištena je nuklearna magnetska rezonancija u čvrstom stanju uz pomoć koje je utvrđeno u kojoj je tautomernoj formi prisutan hipoksantin.

4.1.4. Selektivnost molekulskog prepoznavanja

Rezultati su pokazali da u svim navedenim reakcijskim uvjetima dolazi do selektivnog prepoznavanja tri kombinacije nukleobaza: 2,6-diaminopurina i uracila, 2,6-diaminopurina i timina, te adenina i hipoksantina. Selektivnost u prepoznavanju je utvrđena u prisustvu kompetirajuće nukleobaze.

4.2. Prebiotička sinteza peptida

4.2.1 Oligomerizacija aminokiselina

Za prebiotičku sintezu peptida korištene su mehanokemijske reakcije u čvrstom stanju. Kao reaktanti su korišteni aminokiselina glicin i alanin. Reakcije su provedene uz prisustvo mineralnih aditiva. Reakcije su provedene u uvjetima različitog vremena trajanja reakcije, različite frekvencije mljevenja te različite temperature mljevenja. Rezultati su pokazali da povećanje temperature mljevenja uvelike pospješuje nastanje dužih oligomera glicina uz povećanje udjela diketopiperazina kao primarnog produkta reakcijske smjese. Tako je primjerice najveći udio ukupnih linearnih oligomera iznosio 10,2 % u uvjetima suhog mljevenja glicina uz TiO_2, frekvenciju od 30 Hz, reakcijsko vrijeme od 18 sati te temperaturu mljevenja od 100 °C. Pokazano je da povećanje duljine mljevenja povoljno utječe na uspješnost oligomerizacije glicina. Isto tako, ključnim se pokazala frekvencija mljevenja s obzirom da se veći unos mehaničke energije u reakcijski sustav očituje većim udjelom nastanka lineranih oligomera glicina. Međutim, čak i pri niskim frekvencijama mljevenja utvrđena je prisutnost nastalih peptida.

4.2.2. Oligomerizacija aminokiselina i α-hidroksilnih kiselina

Na temelju eksperimentalnih podataka utvrđeno je da je najjednostavniji oblik α-hidroksilne kiseline, glikolna kiselina, oligomerizira do poliesterskog oligomera duljine lanca četiri glikolne podjedinice. Također, zajedničkim mljevenjem glikolne kiseline i glicina nastaju njihovi heterooligomeri koji se sastoje od peptidnih i esterskih veza u oligomernom lancu.

5. ZAKLJUČAK

Rezultati istraživanja provedenih u okviru ove disertacije ukazuju da se neki od obrazaca molekulskog prepoznavanja koji su danas prisutni u biokemiji također očituju i na razini nukleobaza u prebiotičkim uvjetima. Molekulsko prepoznavanje 2,6-diaminopurina i uracila, 2,6-diaminopurina i timina, te adenina i hipoksantina je selektivno i odvija se pri različitim prebiotičkim reakcijskim uvjetima. S obzirom da mnogi prebiotički sintetski putevi rezultiraju smjesom nukleobaza, za očekivati je da su različite nukleobaze bile prisutne u smjesi na ranoj Zemlji prije nastanka života. Tijekom ranih faza kemijske evolucije, nukleobaze s

komplementarnim supramolekulskim interakcijama mogle su biti odabrane zbog svojih termodinamičkih i kinetičkih svojstava. Te karakteristike mogle su utjecati na njihovu kovalentnu reaktivnost i ugradnju u nukleinske kiseline.

Što se tiče prebiotičke sinteze peptida, pokazano je da mehanokemija vodi ka njihovoj sintezi. Mehanokemijska aktivacija aminokiselina potaknula je njihovu oligomerizaciju na mineralnim površinama. Pokazano je da je diketopiperazin produktivan intermedijer za sintezu peptida te da mehanokemijska sinteza peptidne veze tolerira prisustvo vlage. Reakcije s Gly_2, Gly_3, i diketopiperazinom su pokazale su kako je mehanokemijska sinteza peptidne veze dinamičan proces sa usklađenim nastajanjem i pucanjem peptidne veze. Zajedničko mljevenje glicina i alanine rezultiralo je nastankom njihovih hetero oligopeptida te je pokazalo da se ostale aminokiseline mogu ugraditi u rastući peptidni lanac. Rezultati ovog istraživanja pokazuju da su reakcijsko vrijeme i unos mehaničke energije važni parametri za prebiotičku mehanokemijsku sintezu peptida koja je u skladu s različitim geološkim i okolišnim uvjetima.

University of Zagreb
Faculty of Science
Department of Chemistry

Doctoral Thesis

§ 1. INTRODUCTION

Prebiotic chemistry is the chemistry that precedes the origin of life (Figure 1).[1] The hypothesis behind its underlying research is that chemical evolution, i.e., organization of simple matter into a more complex one, could have served as a transition to biological evolution, e.g., reproduction, variation, and selection of living systems over time.[2] Due to the alteration of Earth's surface by geological processes, origin of life "cannot be discovered, it has to be re-invented".[3] Among the proposed sites where life on Earth could have originated, hydrothermal vents[4] and land-based pools[5] are the most popular.

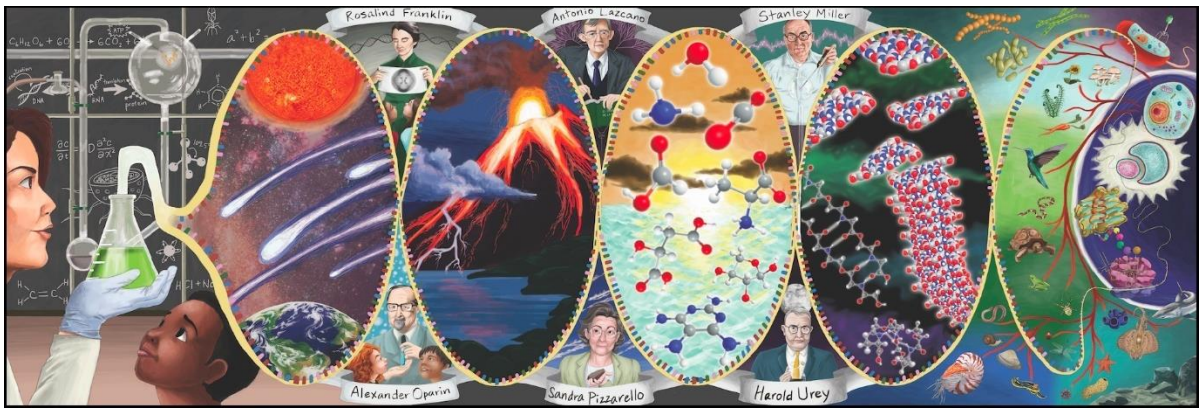

Figure 1. Artistic depiction of origin of life research that includes prebiotic chemistry. Downloaded from: https://centerforchemicalevolution.com/mural-origins-life-0.

Forming important chemical connections and linking monomers of organic building blocks into corresponding biological polymers requires the formation of covalent bonds through condensation reactions where water molecules are eliminated. Hence, their formation is thermodynamically unfavorable in water media.[6] This has led to the growing realization that, although water is essential for life, it is not necessarily desirable for the synthesis of life's

building blocks.[7] To connect the conditions required for their synthesis with the plausible early Earth environment, wet/dry cycles have emerged as the most used experimental approach in the field of prebiotic chemistry (Figure 2).[8] In the dry phase, chemical synthesis and the formation of covalent bonds is enabled whereas, in the wet phase, combinatorial selection through compartmentalization is operational. The idea behind using wet/dry cycles is that they closely resemble hydration (due to precipitation) and dehydration cycles (evaporation due to heating) in land-based pools.

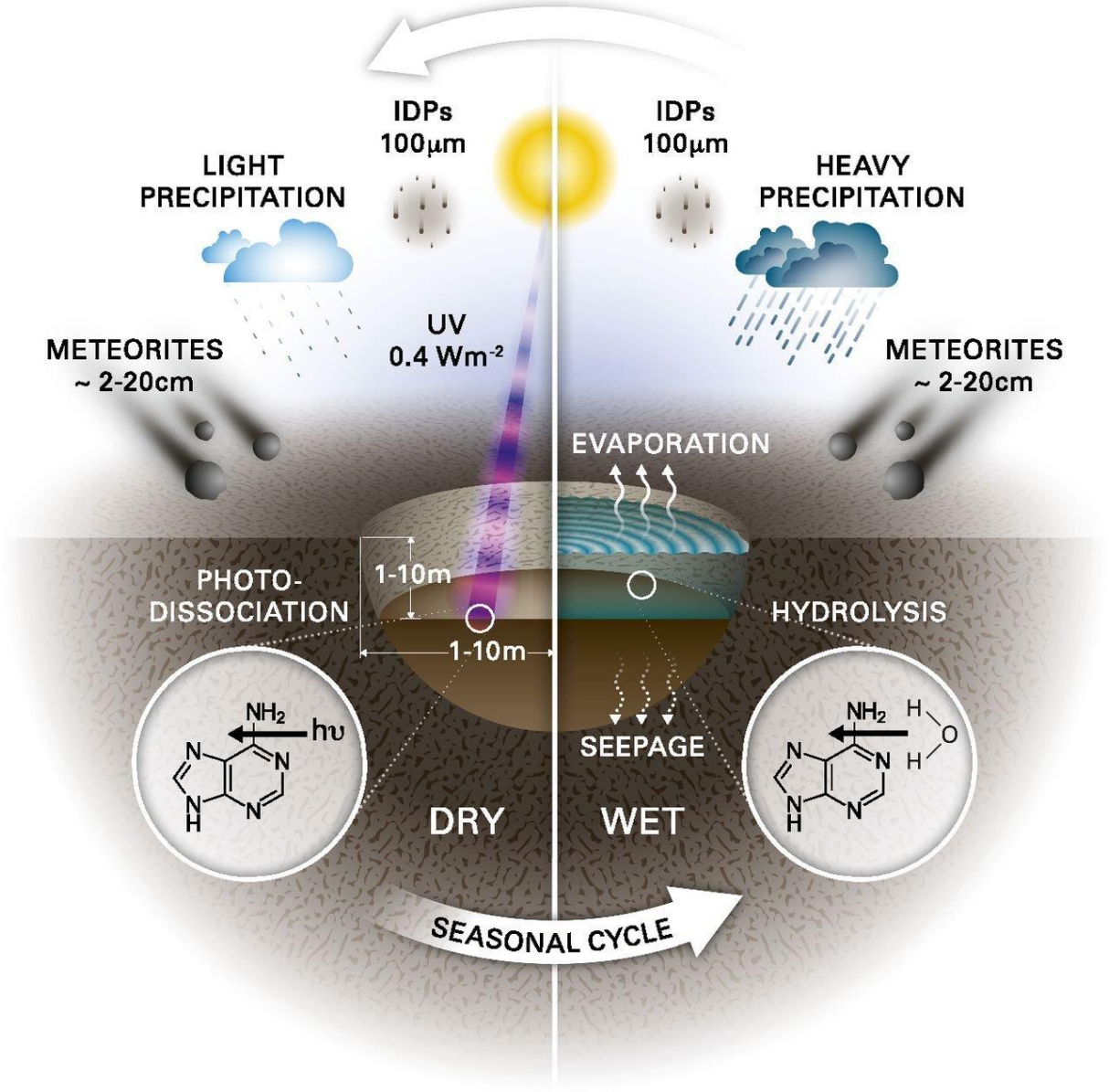

Figure 2. Wet/dry cycling and its central role in the land-based hypothesis for the origin of life on Earth. Reprinted from the reference[9] under the terms of Creative Commons public use license.

The availability of dry land on an early Earth is often discussed. Although it is estimated to have been only a fraction of today's landmass, it is considered that there was some land available on the emerging continents and near volcanoes.[10] Among the reasons why dry land is thought to have played an important role in prebiotic chemistry is its capability to concentrate organic building blocks and induce their chemical transformations.[11]

On the other hand, prebiotic chemistry is also a part of astrobiology, which includes interdisciplinary studies on the origin of life in the universe. Research efforts in prebiotic chemistry have been greatly influenced by the fall of Murchison meteorite which was proven to contain extraterrestrial amino acids,[12] nucleobases,[13] and sugars,[14] among other organic matter.[15] These studies have offered us a glimpse into the organic inventory that was carried to Earth during the time of the late heavy bombardment.[16] Meteorites, comets, and asteroids are solid planetary bodies whose composition is influenced not only by accretion, but also post-accretion events such as thermal metamorphism,[17] aqueous alteration,[18] and hydrous metamorphism.[19] Taken together with millions of years of space weathering,[20] these processes present opportunities for the evolution of organic matter in extraterrestrial solid-state reaction environments. Therefore, recent Hayabusa2[21] and OSIRIS-REx[22] asteroid sample return missions have the potential to offer us unprecedented insight into the origin and evolution of organic matter in our solar system.

In this doctoral dissertation, solid-state reaction medium was used with the aim of solving some of the open question in prebiotic chemistry. In this regard, the inability of canonical nucleobases to self-assemble via Watson-Crick hydrogen-bonding in water is a long-standing bottleneck. It makes it unclear as to how did nature know which nucleobases to incorporate into nucleic acids if they were previously unable to base-pair and select between themselves. We hypothesized that canonical nucleobase pairs are not favored, and instead, other nucleobase pairs might form selectively. We worked with a pool of commercially available and plausibly prebiotic nucleobases. To characterize the reaction products, we primarily relied on powder X-ray diffraction (PXRD) and differential scanning calorimetry (DSC). When possible, crystal structures of novel crystalline phases were solved from single-crystal X-ray diffraction (SCXRD) and PXRD data using available software.

Furthermore, linking free amino acids into the corresponding oligomers requires the formation of peptide bonds along with the elimination of water molecules. Hence, oligomerization is thermodynamically unfavorable in water media. We hypothesized that

mechanochemical activation of free amino acids might enable their oligomerization on mineral surfaces. For studies on oligomerization of free amino acids on solid mineral surfaces, we mostly used glycine (Gly). The reason to choose Gly is its presumed availability on early Earth, its presence in meteorites, and the availability of analytical methods to characterize reaction products. We used diverse class of minerals that were available to us. For analysis, we primarily used the ion-pairing high-performance liquid chromatography (IP-HPLC) with ultraviolet–visible spectroscopy (UV-VIS) detection. Furthermore, the mass spectrometry (MS) was used to unambiguously determine the presence of different oligomers. Mechanochemical ball milling reactions were used to investigate the oligomerization of amino acids on mineral surfaces and some of these were performed as thermally controlled neat grinding reactions.

Some of the results presented in this thesis have been published while others are still in the publishing process.[23]

§ 2. LITERATURE REVIEW

2.1. Prebiotic selection of nucleobases

All extant life utilizes DNA for storage and inheritance of genetic information. Besides contributing to the stability of the DNA double helix, specific hydrogen-bonding interactions between nucleobases act as a basis for molecular recognition. From the perspective of the origin of life, there are many puzzling questions regarding the abiotic origin of nucleic acids.[24–28] Among them is the assumption that, on the prebiotic Earth, there was likely a pool of structurally similar nitrogen-bearing heterocyclic compounds that could have acted as nucleobases (Figure 3).[29,30] Such a hypothesis is corroborated by many synthetic experiments that produce other purines and pyrimidines besides the canonical nucleobases.[31–35] Furthermore, the discovery of the presence of extra-terrestrial nucleobases in carbonaceous meteorites,[13,36] as well as nucleobase synthesis in interstellar ice analogues,[37,38] strongly points to a pool of nucleobases that may have been present throughout the solar system. Interestingly, canonical nucleobases are unable to base-pair and select between themselves.[39,40] Canonical nucleobases compete for hydrogen-bonding with water molecules, and vertical stacking is energetically favoured compared to horizontal hydrogen-bonding.[41,42] This has led to the idea that other base-pairing combinations might have preceded canonical base-pairing that we observe today.[39,43]

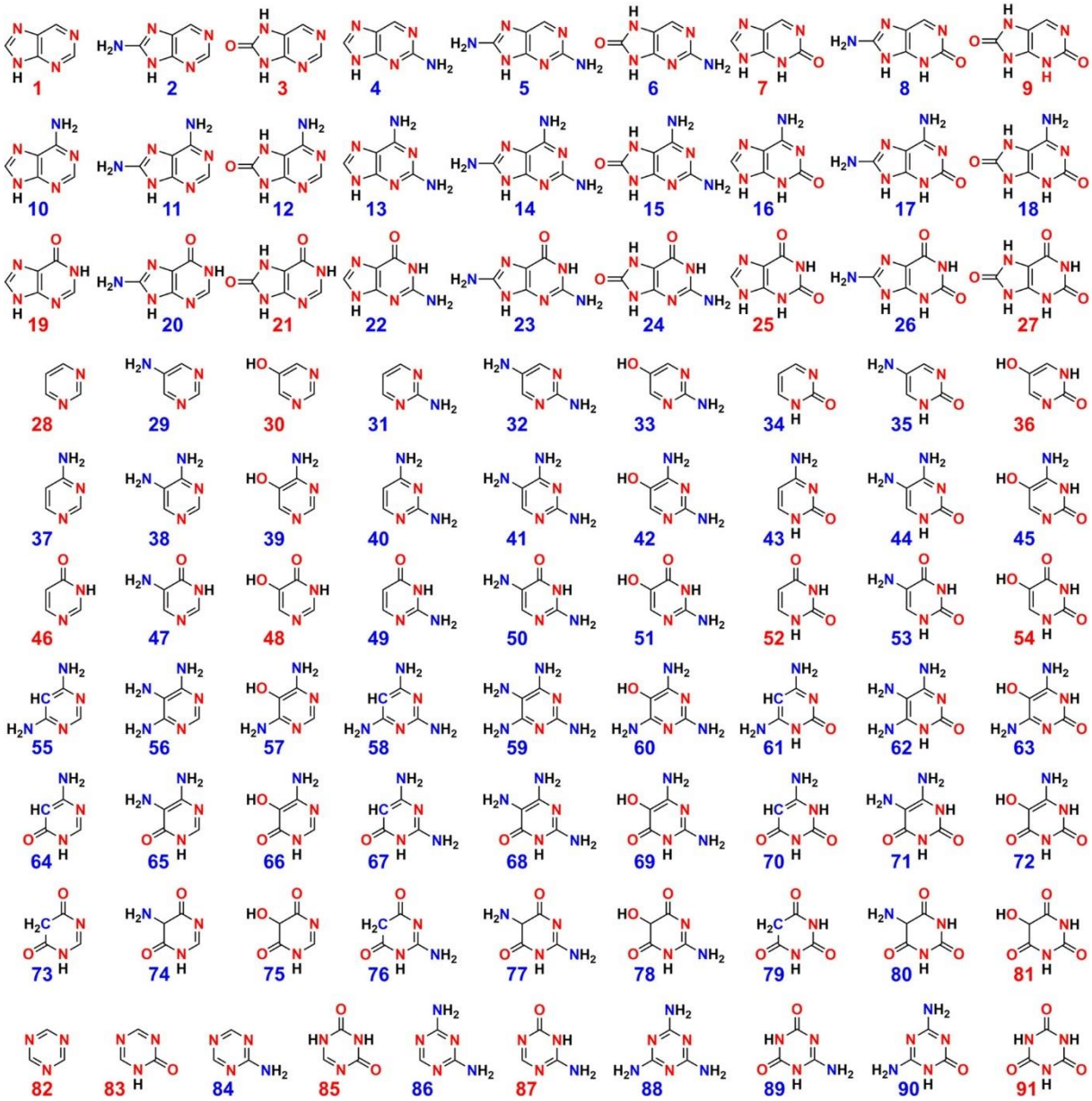

Figure 3. Chemical space of molecules that could have acted as nucleobases on an early Earth. Reprinted with permission from the reference.[30] Copyright 2020 American Chemical Society.

As an example, 2,6-diaminopurine (D) could have been a competitor to adenine (A) because of its ability to form three Watson-Crick hydrogen bonds with both uracil (U) and thymine (T), while adenine (A) can only form two (Figure 4a and 4b). In fact, there are examples from contemporary biology showing that certain viruses use D instead of A as its DNA nucleobase.[44,45] Recently, more examples and detailed mechanistic understanding have been found of how D replaces A in viral genomes.[46–48] Additionally, a ribozyme that catalyses template-directed ligation of two RNA strands with 3',5'-phosphodiester linkages and is

composed of only D and U nucleotides, can be obtained in vitro.[49] Moreover, F. H. C. Crick and L. Orgel have argued that an earlier version of the genetic code may have consisted of only A and hypoxanthine (H) as its bases.[50] Their arguments were based on the relative ease of the prebiotic synthesis of A whilst H could have been produced from A by deamination. Importantly, such purine-only nucleic acid would not violate the principle of continuity in its plausible evolution to the extant nucleic acids.[51] Their hypothesis has been brought up to the light by reports showing that synthetic DNA duplexes with only purine-purine base pairs,[52,53] including A and H paired via Watson-Crick hydrogen-bonding,[54] have comparable stability with the extant DNA. Nevertheless, H is biologically relevant as its nucleoside inosine is commonly found in transfer RNAs (tRNAs) where it is involved in codon-anticodon wobble base-pairing (Figure 4c).[55,56] Collectively, there is a large body of evidence for the presence of diverse nucleobases in prebiotic times. However, what kind of selection pressures operated between them remain a mystery.

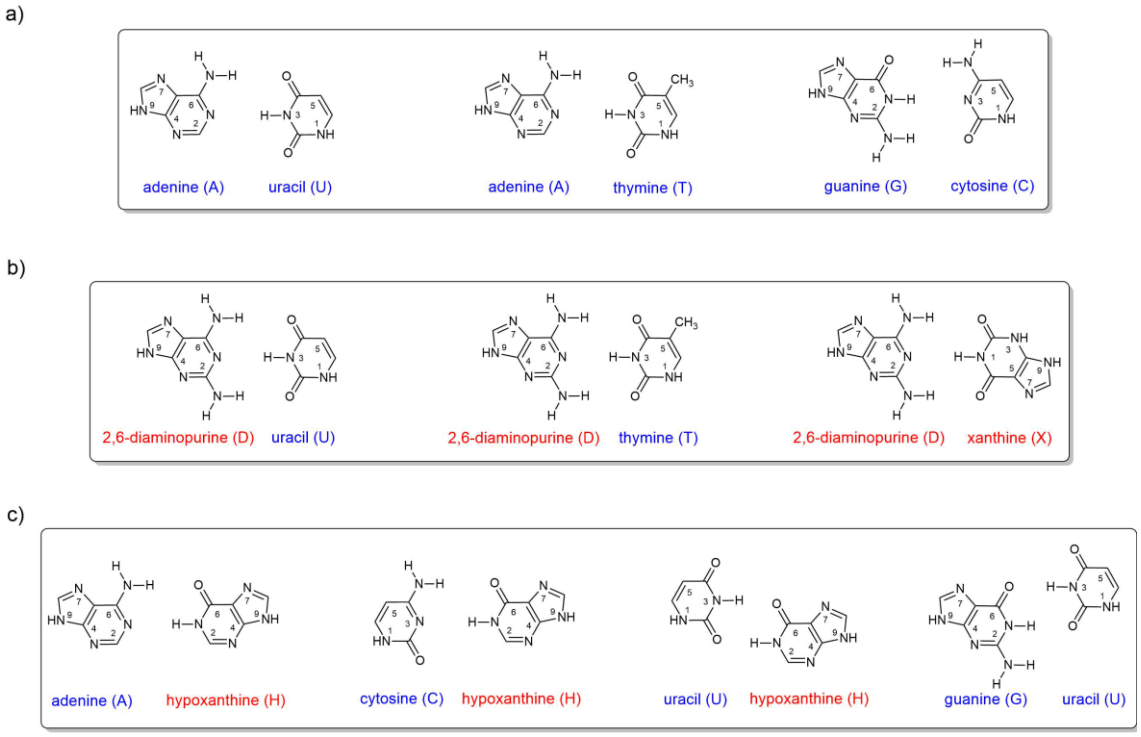

Figure 4. Canonical nucleobases are labeled with blue color, whereas three nucleobases that have been detected in meteorites and are present in contemporary biology, are labeled with red color: a) canonical Watson-Crick base-pairing, b) nucleobases capable of forming three Watson-Crick hydrogen bonds, and c) nucleobases involved in codon anticodon wobble base-pairing.

2.2. Prebiotic peptide synthesis

Among many open questions in the field of prebiotic chemistry, the formation of the peptide bond is one of the most important ones. In an aqueous environment, difficulties arise when condensing free amino acids due to thermodynamic and kinetic unfavourability of the peptide bond formation.[57] On one hand, multiple evidence shows that amino acids such as glycine were present on Earth,[58] in extraterrestrial bodies,[59,60] and in interstellar medium before life emerged.[61] On the other hand, peptides are an important class of biological polymers with catalytic properties and are hypothesized to be included in early molecular symbiosis with nucleic acids (Figure 5).58

Notably, reports are suggesting the presence of extraterrestrial peptides in meteorites.[62] Prebiotic peptide bond synthesis strategies include thermal condensation,[63–65] different wet/dry cycling approaches,[66–68] salt-induced peptide bond formation,[69] condensation in liquid SO_2,[70] catalysis by carbonyl sulfide,[71] α-aminonitrile ligation in water,[72] mineral catalysis,[73] hypervelocity impact studies,[74,75] and high-pressure shearing (Figure 6).[76]

In the search for energy needed to drive prebiotic formation of peptides,[77] the use of mechanical energy appears as a promising solid-state alternative that avoids the limitations related to solvation, solubility, and diffusion.[6] In relation, it is known that the primordial Earth underwent a heavy meteoritic bombardment,16 which likely delivered diverse pool of extraterrestrial amino acids in ppm concentrations.[78,79] Moreover, extraterrestrial impactors or terrestrial lithospheric activity could have provided mechanical and thermal energy to drive formation of peptides. As an example, studies simulating impact events in aqueous solutions of amino acids,[74] and those mimicking natural comet impacts under cryogenic conditions have revealed that glycine and alanine oligomerize due to the shock impact.[80] However, mostly linear dipeptides and tripeptides are formed in the process. In addition, simulations of glycine subjected to shear stress have shown that formation of linear diglycine could occur under 10 GPa compressive loads at RT, thus indicating the possibility for impact-based and tectonic shearing prebiotic scenarios.76 On the other hand, extreme mechanical stress can also cause peptide cleavage as demonstrated in hypervelocity impacts (mechanical loading of 2.4–5.8 km/s).[81] This might explain the relatively short peptides that are formed under harsh mechanical conditions.

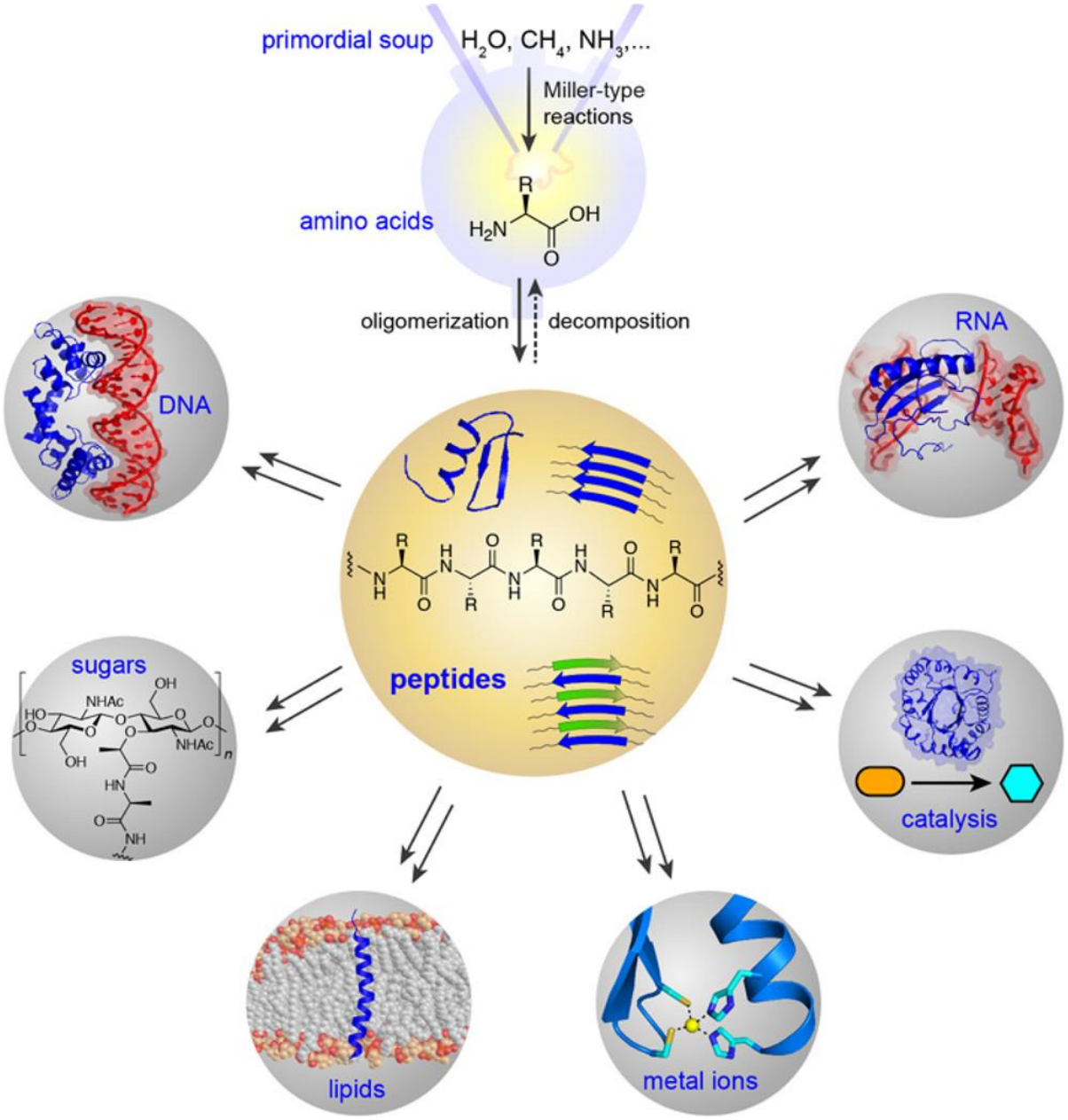

Figure 5. Peptides are hypothesized to have played an important role in the origin of life. Reprinted with permission from the reference.[58] Copyright 2020 American Chemical Society.

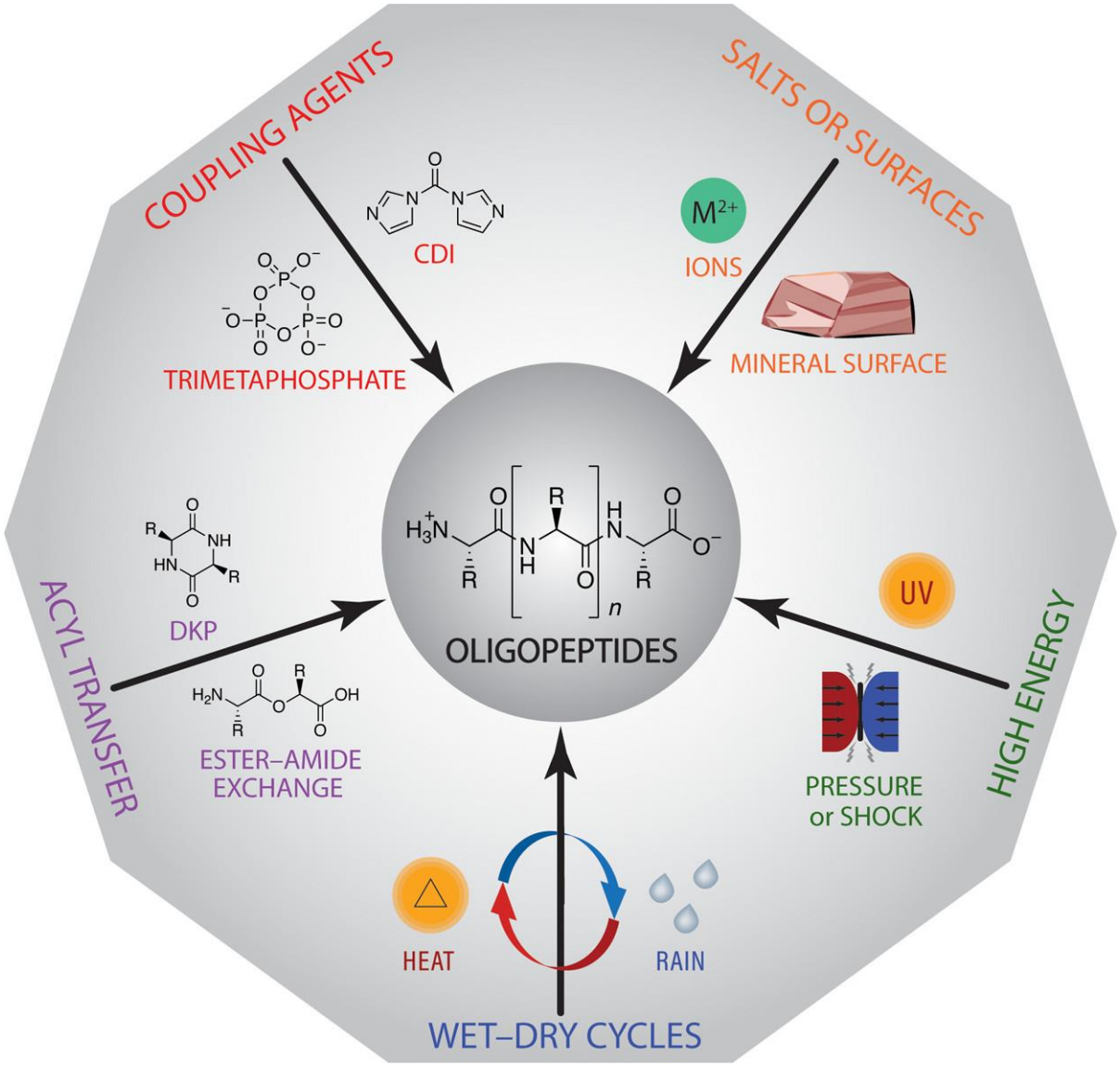

Figure 6. Some of the synthetic strategies towards prebiotic peptides. Reprinted with permission from the reference[82] under the terms of Creative Commons public use license.

2.3. Mechanochemical reactions in prebiotic chemistry

The use of mechanical action to run chemical reactions has been used since antiquity[83] and modern synthesis technique is called mechanochemistry.[84–87] It is highly efficient in inducing chemical reactivity in the absence of bulk solvents and kinetically accelerates solid-state reactions by overcoming slow solid-state diffusion.[88] Vibratory ball milling, planetary milling, and twin-screw extrusion are among the most used types of milling instruments (Figure 7). The field of mechanochemistry has experienced a significant growth in recent years due to its compliance with the principles of green chemistry and sustainability.[89] It has since been used in virtually all areas of chemical science due to different reactivity and selectivity when compared to the solution-based synthetic approaches.[90–113]

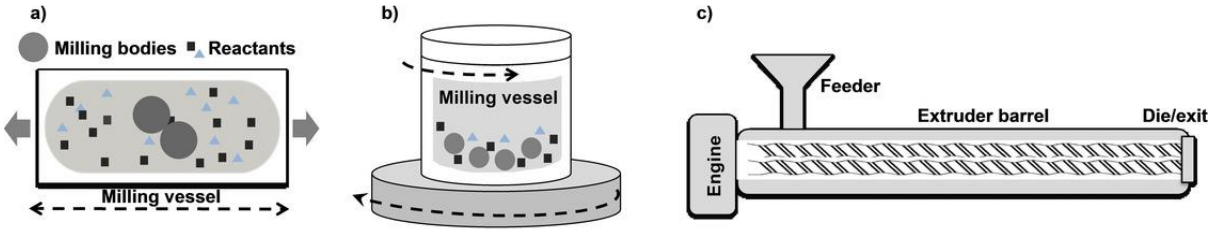

Figure 7. Three main types of mechanochemical instruments: a) vibratory ball milling, b) planetary ball milling, c) twin-screw extrusion. Adapted with permission from the reference.[89] Copyright 2021 John Wiley & Sons, Inc.

More than two decades of intense research resulted in many innovations within the field. In liquid-assisted grinding (LAG) methodology, small amounts of liquids are added to solid reaction mixtures.[114,115] It can have a strong reaction accelerating effect or can direct reaction courses into specific pathways. Importantly, several examples of temperature monitoring showed that bulk temperature of the reaction mixture rises during the milling process.[116–118] This is most likely due to collisions and friction that occur in milling reactions. Notably, temperature oscillations can also be due to endergonic or exergonic nature of specific reactions, albeit to a lesser extent.[119] Using sub-ambient temperature ball milling, it was established that temperature differences of 5 °C affect both yields and enantioselectivity on several reactions such as asymmetric aldol condensation catalyzed by (S)-proline, enantioselective multicomponent reactions between an aldehyde, an amine, and an alkyne, as well as Diels Alder reactions.[120]

Furthermore, experimental setup was devised that used modified jars, thermocouple, and a ring heater element managed through a proportional−integral−derivative (PID) controller device (Figure 8).[121] It was shown that temperature during milling can be precisely controlled up to 250 °C and with a deviation of ±0.2 to ±5 °C, depending on the milling reactor material. The utility and generality of such an experimental setup was showed for the synthesis of C-C, C-N bonds, and porous metal-organic frameworks. Notably, except for faster conversion rates to products, milling temperature had a profound effect on the selectivity as demonstrated on obtaining four different derivatives containing amido and urea groups in the mechanochemical reaction between 4-nitrobenzoyl azide and benzene-1,4-diamine just by varying milling temperature parameter. Furthermore, other experimental setups for high-temperature ball milling have been reported, for example using a heating gun that blows hot-air to the milling jar. First examples of such setup were described for the mechanochemical palladium-catalyzed Suzuki-Miyaura cross-coupling of insoluble aryl halides[122] as well as for the synthesis of magnesium-based carbon nucleophiles, i.e., Grignard reagents, in air.[123] Overall, literature reports show that high temperature ball milling allows access to products that are characterized by high activation energies.

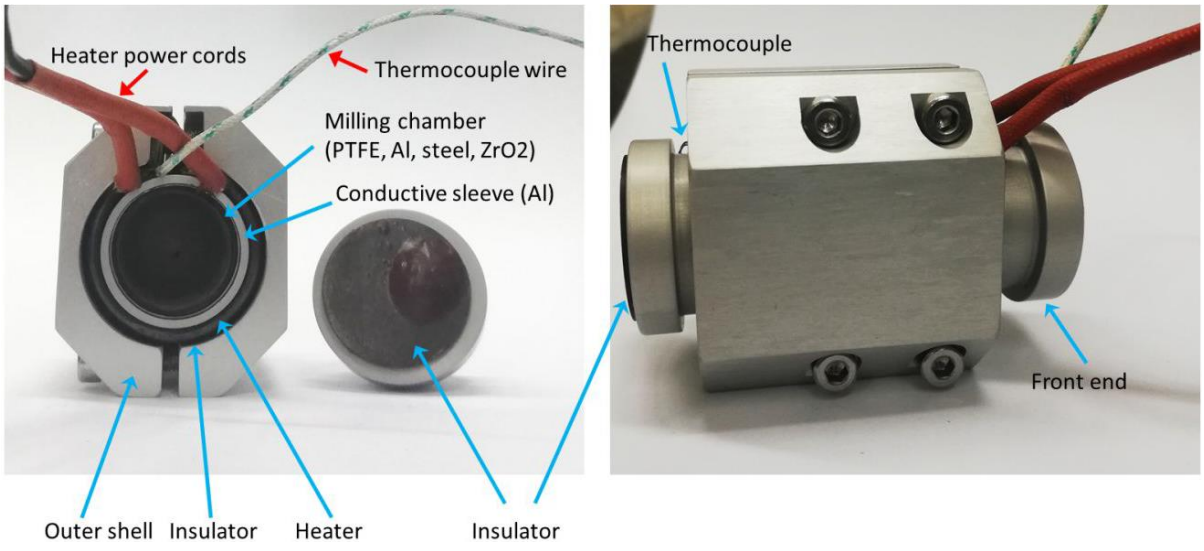

Figure 8. Thermally controlled ball milling setup that uses modified jars, thermocouple, and a ring heater element managed through a PID controller device. Reprinted with permission from the reference.[121] Copyright 2019 American Chemical Society.

Additionally, ball milling photocatalytic setups have been developed which merge light and mechanical activation.[124,125] Importantly, a growing number of mechanochemical reactions

include gaseous reactants and show great promise for the discovery of new reactivity of solid-gas interfaces.[126,127] Furthermore, *in situ* techniques for monitoring mechanochemical reaction courses have been established and are based mostly on powder X-ray diffraction[128–130] and Raman spectroscopy.[131–133] These offer mechanistic insights and means of kinetic analysis. Recently, mechanoredox reactions have been discovered and offer new avenues for redox chemistry.[134,135] Finally, in accelerated ageing methodology, solids are initially mixed and are left to react on a scale longer than typical laboratory experiments.[136] The latter might be compatible with long-time prebiotic reactions.

Some of the primordial sources of energy that are coupled with prebiotic solid-state syntheses of various classes of molecules are mechanical and thermal energy. Mechanical energy on the prebiotic Earth could have been provided by impacts, tectonics, erosion, earthquakes, and weathering (Figure 9), whereas geothermal activity could have provided thermal energy. Recently, there have been several literature examples of rich solid-state reactivity relevant to prebiotic chemistry. The first reported use of mechanical activation in a prebiotic chemistry context has enabled the activation of highly stable iron cyanide complexes.[137] Under ball milling conditions with SiO_2, the authors observed mechanochemical activation of potassium ferricyanide(III) *via in situ* production of HCN that was subsequently incorporated into the corresponding α-aminonitriles. Furthermore, hydrolysis of α-aminonitriles into amino amides successfully showed that an impact scenario on an early Earth might offer a route towards amino acids. These conclusions were further supported by a different study where grinding processing of iron meteorites resulted in the formation of amino acids.[138] This confirmed that mechanical action might have offered a mechanism for the delivery of amino acids by iron meteorites.

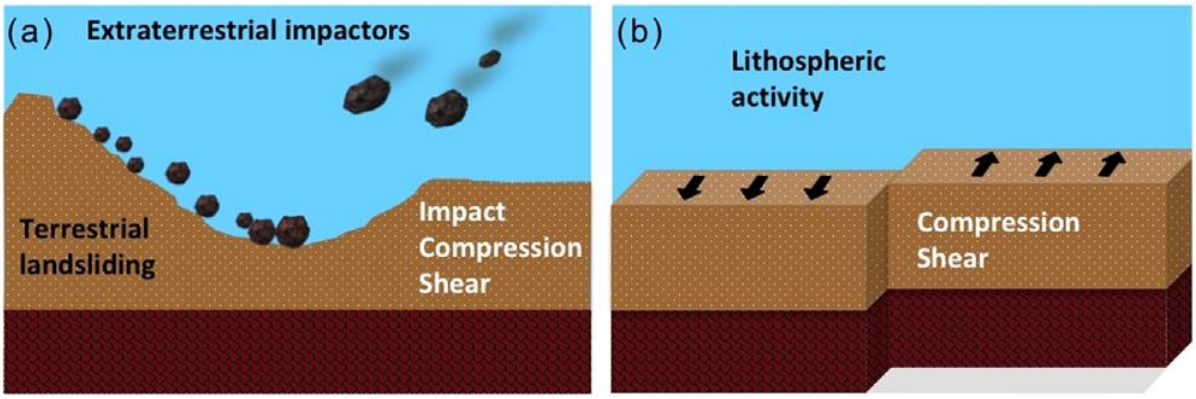

Figure 9. Prebiotic sources of mechanical energy might have included impacts, erosion, earthquakes, tectonics, and weathering. Adapted with permission from the reference.[23] Copyright 2021 John Wiley & Sons, Inc.

In another study, mechanical energy was used to drive the synthesis of monosaccharides.[139] A wide range of linear and branched monosaccharides including ribose was obtained after ball milling formaldehyde, glycolaldehyde, and glyceraldehyde together with mineral additives as catalysts. These results showed that the formose reaction route to sugars is operational and is kinetically accelerated under solid-state reaction conditions devoid of aqueous media. The role of minerals in mechanochemical formose reaction was further investigated in a separate study.[140] In contrast to the aqueous formose reaction pathway that yields decomposition products, the solid-state pathway enriches monosaccharides over time. The selectivity towards individual sugars was demonstrated to be fine-tuned using a particular mineral catalyst. Moreover, the mechanochemical formose pathway was shown to be independent of the atmospheric conditions and proceeds under cryogenic conditions, thus offering both terrestrial and extraterrestrial prebiotic routes towards sugars. In addition, mechanochemistry was used in a unified route towards purine and pyrimidine nucleosides.[141] The formation of critical *N*-glycosidic bond was achieved by either solid-state grinding and subsequent heating of *N*-isoxazolylurea, ribose, and boric acid at 95 °C, or by drying their aqueous solutions at the same temperature. Conversion of intermediate *N*-isoxazolyl-*N*'-ribosyl-ureas to pyrimidine nucleosides was achieved by the catalytic action of iron-sulfur minerals.

Overall, the use of mechanochemical reactions in prebiotic chemistry is still in its beginnings, likely influenced by until recently dominant hypothesis that life originated in hydrothermal vents (deep in the ocean).

§ 3. EXPERIMENTAL

3.1. Molecular recognition of nucleobases

3.1.1. General conditions

Nucleobases used in this study were purchased from Carbosynth. Ball milling experiments were conducted using InSolido Technologies vibratory ball mill and in 14 mL poly(methyl methacrylate) (PMMA) containers. Two 1.4 g stainless steel (SS) balls (7 mm in diameter) were used as milling media, reactions were conducted at 30 Hz frequency and for the duration of 2h. When performing liquid-assisted grinding (LAG) reactions, 50 μL of H_2O was added to overall 200 mg of the reaction mixture. Cryogenic ball milling experiments were conducted using the same abovementioned conditions. Jars were immersed in liquid nitrogen before the start of the experiment and liquid nitrogen was poured continuously for the duration of the reaction. Slurry reactions in water were performed by mixing equimolar amounts of reactants (overall mass of the reaction mixture was 100 mg) in 3 mL of H_2O at room temperature. Mixing was achieved in 10 mL round bottom flask using magnetic stirrer for several days. To grow single crystals of **2** and **10**, 10 mg of **1** and **4** was dissolved in 5 mL of methanol and left for solvent to evaporate at room temperature. Hydrothermal synthesis of **9** was performed by heating equimolar amounts of A and H at 120 °C overnight in a pressure-resistant quartz tube. For solid-state nuclear magnetic resonance (ssNMR) analysis, we used sample that was prepared by neat grinding (NG) A and H equimolar mixture in stainless steel jars with two 1.4 g stainless steel balls and subsequent heating to 300 °C. High-resolution PXRD patterns which were used for crystal structure solution were collected in a glass capillary on a Stoe Stadi-P laboratory diffractometer equipped with copper radiation (λ = 1.54059 Å). Single crystal measurements were performed on an Oxford Diffraction Xcalibur Nova R (λ = 1.54719 Å) equipped with an Oxford Instruments CryoJet liquid nitrogen cooling device. Program package CrysAlis PRO12 [CrysAlisPro, Rigaku Oxford Diffraction, 2018, version: 1.171.39.46, Rigaku Corporation, Oxford, UK] was used for data reduction and numerical absorption correction. The structures were solved and refined with SHELXL [G. M. Sheldrick, *Acta Cryst.* **A71** (2015) 3–8.]. Models were refined using the full-matrix least squares refinement; all non-hydrogen atoms were

refined anisotropically. Hydrogen atoms were located in a difference Fourier map and refined as a mixture of free restrained and riding entities. In **2**, crystallization water molecule (O3) is disordered over two positions (designated as A and B) with respective occupancies of 63 % and 37 %.

3.1.2. Variable-temperature synchrotron powder X-ray diffraction monitoring

In situ synchrotron PXRD monitoring of solid-state heating/cooling reactions was performed at P02.1 beamline at PETRA III, DESY. Glass capillaries of 0.5 mm width were filled with ball milled solid samples and were heated/cooled using Oxford Cryostream system. The latter was operating in 100-500 K temperature range. Distance between the sample and the detector was 2010 mm. X-rays of 0.20741 Å wavelength were used and X-ray diffraction pattern was collected on a spinning capillary. Each sample was thermally equilibrated at 300 K and was heated/cooled using a 5 K/min ramp. For each experiment, X-ray diffraction patterns were collected in 10 K steps. After reaching the desired temperature set-point, sample was thermally equilibrated, and X-ray diffraction pattern was collected. Exposure time for each collected diffraction pattern was 60 s for dark image followed by 60 s for real image. To obtain 2D plots, data were processed in MATLAB using an in-house developed script.

3.1.3. Crystal structure solution from powder X-ray diffraction data

For crystal structure solution from PXRD data, all calculations were performed using the program Topas.[142]

<u>The workflow for structure solution</u>

1) The scan was loaded in xy file format.

2) Data indexing was performed by manually selecting 15-20 reflections. This can also be done in an automatic fashion. However, since successful unit cell hits are highly dependent on the selected reflections, we generally relied on our own assessment.

3) Unit cells with the best values were selected for Pawley refinement. The objective of Pawley refinement was to identify the best unit cell candidate. If suggested unit cells had the same parameters but different symmetry, we generally proceeded with the space group having higher

symmetry. The main criterium for identifying the unit cell was the assessment of the number of molecules in the asymmetric unit. When this was done, the data was exported as INP file.

4) To prepare the complete INP file, the coordinates of the crystallizing molecule(s) were extracted from the database CIFs containing the same molecule(s) (alternatively, this can be done by creating a Z-matrix). This was performed by exporting the target molecule(s) as pdb file in Mercury software. The objective of assembling the INP file is to place the target molecule(s) in the unit cell after which simulated annealing in direct space was used to find the position, orientation, and meaningful interactions of the crystallizing molecule(s).

5) The general criteria to recognize that simulated annealing process has finished was when R_{wp} remained constant for a significant period. This was either hours or days. However, the final decision was made after visually inspecting the CIF file.

An example of INP file ready for simulated annealing is given in Figure 7.1.

Rietveld refinement of crystal structures

1) The background and unit cell parameters were refined first.

2) To refine the atom positions, we used geometry restraints (using the penalty function) on bond distances and angles as well as planarity restraints. The values of bond distances and angles were extracted from CIF file in Mercury.

3) For validation of the structural parameters, we used Mogul Geometry Check option in Mercury. If there were discrepancies with respect to the database mean values, the database mean values of each specific bond distance or angle were further used as starting points for the next iteration of the refinement.

4) Final CIFs were deposited to Cambridge Crystallographic Data Centre (CCDC).

An example of INP file that was used in final Rietveld refinement is given in Figures 7.2. and 7.3.

3.1.4. Other solid-state analytical techniques

ATR-FTIR spectra were recorded on PerkinElmer SpectrumTwo spectrometer equipped with a diamond cell in a 4000-400 cm^{-1} range. PXRD patterns were collected on PanAlytical Aeris diffractometer (CuKα, Ni filter) in Bragg-Bretano geometry, with the sample mounted on zero background silicon holder. DSC experiments were performed using Simultaneous Thermal Analyzer (STA) 6000 (PerkinElmer, Inc.) in platinum crucibles at 5 K/min heating rate under 20 mL/min nitrogen gas flow.

3.2. Peptide synthesis

3.2.1. General conditions

Glycine, 2,5-diketopiperazine (DKP), TiO_2, SiO_2, and Montmorillonite K10 were purchased from Sigma-Aldrich. Gly_2–Gly_4 and L-Ala-L-Ala were purchased from Bachem. L-Ala and glycolic acid were purchased from Carbosynth. Mica (MICA-V-4) grade was purchased from SPI. All milling experiments were conducted in InSolido Technologies vibratory ball mill. For the reactions with TiO_2, reactions were conducted at 30 Hz frequency and all reactions were performed for the duration of 18 h. Overall mass of reaction mixture was always around 400 mg (0.8 mmol Gly and 4 mmol TiO_2). There, we used 14 mL hardened stainless steel (SS) containers specifically designed for high temperature milling reactions and three 1 g SS balls (6 mm in diameter) as milling media. 25 μL of H_2O was used as the liquid additive in LAG reactions. For the reactions with SiO_2, they were conducted at different frequency and reaction time. Overall mass of reaction mixture was always 500 mg (1.33 mmol Gly and 6.66 mmol SiO_2). There, we used 28 mL hardened stainless steel (SS) containers specifically designed for high temperature milling reactions and two 1.4 g SS balls (7 mm in diameter) as milling media.

3.2.2. Ion-pairing high-performance liquid chromatography

Oligopeptides were analyzed by Ion-Pairing High-Performance Liquid Chromatography (IP-HPLC) on Agilent 1290 Infinity instrument equipped with a Phenomenex Luna C18 column (250 mm long, 4.6 mm diameter and 3 μm particle size). The mobile phase was an aqueous solution of 50 mM KH_2PO_4 and 7.2 mM sodium hexanesulfonate adjusted to pH 2.5 by the addition of 85 % H_3PO_4 (HPLC grade). The mobile phase was used isocratically with 1 mL/min flow rate. Sample injection volume of 1 μL was used, the total method time was 20 min, and

the column was equilibrated at 30 °C. The instrument was equipped with multi-wavelength UV-VIS detector set to record absorbance at 195, 200, 210 and 214 nm (with reference set at 360 nm). The analysis relied exclusively on the data collected at 195 nm. Before the first separation experiment, the column was equilibrated with the mobile phase for 24 h.

Individual standards preparation

2.5 mg of Gly, DKP, Gly_2, Gly_3, and Gly_4 was transferred to 1.5 mL HPLC vial and filled with mobile phase. Vials were stirred with IKA Vortex for 1 min to aid with dissolution.

Sample preparation

Around 40 mg of each milled sample was transferred into Falcon tube and filled with 1.5 mL of mobile phase. Each tube was stirred with IKA Vortex for 1 min to aid with dissolution of oligopeptides. After that, all tubes were centrifugated at 15000 rpm for 5 min. Then, 0.8 mL of the supernatant phase was transferred into a glass vial and analyzed by HPLC.

Calculation

Quantification of oligopeptides was based on measurement (for Gly-Gly_4) and extrapolation (for Gly_5-Gly_{11}) of molar responses. The standards were accurately weighed and analyzed. Obtained peak areas were divided by molar concentration to get molar response factor. The molar response ratio was plotted against oligomer length. Linear relationship was confirmed with obtained correlation coefficient of 0.9959. The relationship was extrapolated to glycine oligomers up to Gly_{11}. The obtained calibration results closely correlated with data reported previously.[82] From the obtained results (molar concentration of Gly oligomers in samples) the total yield percentage was calculated as the ratio of reaction yield to theoretical yield.

3.2.3. Mass spectrometry

LC/MS (Q-TOF)

Oligopeptides were analyzed by liquid chromatography (LC) on an Agilent 1290 Infinity interfaced to Quadrupole time-of-flight (Q-TOF) mass spectrometry (MS) (Model 6545) instrument equipped with a Waters Acquity BEH HILIC amide column (150 mm long, 2.1 mm diameter and 1.7 µm particle size). The mobile phase (MP) consisted of solvent A: ACN and B: 5 mM $HCOONH_4$ at pH 3.0 adjusted with HCOOH. Elution was started with isocratic condition for 2 min (10 % MP B) followed by gradient elution (2-20 min, 10 % MP B to 50 %

MP B). The flow was set to 0.4 mL/min. Sample injection volume of 3 µL was used, the total method time was 25 min, and the column was equilibrated at 30 °C. A dual AJS ESI ion source was operated in positive (+) ESI mode. The gas temperature (300 °C), drying gas (6 l/min), nebulizer (60 psi), capillary voltage (3000 V), fragmentor (100 V) and skimmer (65 V) parameters were optimized using glycine. The Q-TOF MS was operated in spectral mass range set at 40-1700 m/z and an acquisition rate of 1.4 spectra/s (acquisition time = 714 ms/spectrum). Internal mass tuning and calibration for the Q-TOF was preformed prior sample analysis. The instrument was calibrated in high-resolution mode (4 GHz, low mass range, 0-1700 m/z). A continuous two-point mass (121.050873 m/z and 922.009798 m/z) reference mass correction is further applied during data acquisition using reference solution.

Sample preparation

Samples prepared for IP-HPLC was subsequently diluted with acetonitrile (100 µL of sample and 900 µL of ACN) to match the starting mobile phase composition.

Calculation

Glycine and glycine oligomers were detected as the m/z equivalent to the $[M+H]^+$ ion of each analyte. From the obtained TIC chromatogram, extracted ion chromatograms were generated in m/z window of 3 units ($[M+H]^+$ of each oligomer to $[M+H]^+ + 3$). Obtained masses ($[M+H]^+$) were compared to theoretical masses and the difference is expressed in ppm. Glycine oligomers were confirmed with error lower than 3 ppm for the $[M+H]^+$ ions. Additionally, the isotopic distribution was verified for highest detected oligomer (Gly_{10}) in the sample where Gly_3 was milled with TiO_2 at 130 °C, with obtained error less than 2 ppm for three most prominent isotopes.

Direct MS (Q-TOF)

For selected samples direct analysis technique was chosen due to presumed lower solubility of alanine oligomers and simplicity. Samples of L-Ala milled with TiO_2 at 130 °C as well as Gly+L-Ala milled with TiO_2 at 100 °C were injected using Agilent 1290 Infinity instrument without the column. The samples were transferred into Q-TOF MS (Model 6545) instrument with mobile phase (10mM $HCOONH_4$ at pH 2.3 adjusted with HCOOH/acetonitrile, 50/50v/v) used only as a transfer media, with flow rate of 0.4mL/min. A dual AJS ESI ion source was operated in positive (+) ESI mode. The gas temperature (300 °C), drying gas (6 l/min), nebulizer

(60 psi), capillary voltage (3000 V), fragmentor (100 V) and skimmer (65 V) parameters were used. The Q-TOF MS was tuned and calibrated in extended dynamic range (2GHz) in 3200 m/z mass range mode. Spectral acquisition range was set from 50-950 m/z with an acquisition rate of 1.5 spectra/s (acquisition time =666 ms/spectrum). A continuous two-point mass (121.050873 m/z and 922.009798 m/z) reference mass correction was further applied during the data acquisition using reference solution.

Glycine and alanine oligomers were detected from the obtained accurate mass spectra (TIC) of the samples as the m/z equivalent to the $[M+H]^+$ ion of each analyte. From the obtained TIC chromatogram $[M+H]^+$ obtained exact masses ($[M+H]^+$) were compared to theoretical masses and the differences were expressed in ppm. All found oligomers were confirmed with error lower than 4 ppm for the $[M+H]^+$ ions.

LC/MS (Triple quadrupole)

Sample preparation

Around 40 mg of each milled sample was added to Falcon tube and filled with 1.5 mL of mobile phase. Each tube was stirred with IKA Vortex for 1 min. After that, all tubes were transferred to ultrasonic bath heated up to 35 °C and sonicated for 15 min. Finally, the content of each tube was filtrated thought 0.2µM pore size regenerated cellulose filter, transferred into autosampler vial and analyzed by Triple quadrupole LC/MS.

Oligomers were analyzed by hydrophilic interaction liquid chromatography (HILIC-HPLC) on Agilent 1290 Infinity interfaced to Triple quadrupole MS (Model 6470B) instrument equipped with Waters Acquity BEH HILIC amide column (150 mm long, 2.1 mm diameter and 1.7 µm particle size). The MF consisted of solvent A: 90 % ACN+10 % 200 mM $HCOONH_4$ at pH 3.0 and B: 20mM $HCOONH_4$ at pH 3.0. Elution was gradient, starting with 90 % MP A at 0 min to 80 % MP A at 10 min. The flow was set to 0.8 mL/min. Sample injection volume of 0.25 µL was used and the total method time was 16 min. The column was equilibrated at 30 °C. An AJS ESI ion source was operated in positive (+) ESI mode. The gas temperature (300 °C), gas flow (7 L/min), nebulizer (45 psi), capillary voltage (3500 V), fragmentor voltage (25 V), sheet gas temperature (400 V) and sheet gas flow (11 L/min). The Triple quadrupole MS was operated in SIM mode (monitoring $[M+H]^+$ ions of glycine and glycine oligomers to Gly_6) with Dwell time set to 50 ms.

MALDI TOF

For matrix-assisted laser desorption/ionization time-of-flight (MALDI TOF) analysis, we used 4800 Plus MALDI TOF/TOF Analzyer (Applied Biosystems, USA). Samples were dissolved in CHCA matrix. Ultra-pure water (Pliva d.d., Zagreb), acetonitrile HPLC grade (Merck), trifluoroacetic acid (Merck), and CHCA matrix (Merck) were used. CHCA matrix (5 mg/mL) was prepared in acetonitrile solution and 0.1 % trifluoroacetic acid (50/50, v/v) mixed with ultra-pure water. 2 mg of sample was centrifuged for 10 min in aqueous solution 1 mL, after which 2 µL of sample was mixed with 8 µL of CHCA matrix. 1 µL of the latter was transferred to MALDI holder. Samples were calibrated with near external calibration method with standards of azithromycin and domicile peaks resulting from the matrix ionization. Analysis of the samples was performed in both positive and negative modes.

§ 4. RESULTS AND DISCUSSION

4.1. Prebiotic emergence of molecular recognition on a nucleobase level

It was recently shown that model nucleobases, having a methyl group attached at the glycosidic nitrogen atom, selectively self-assemble from multi-component nucleobase solid mixtures.[143] Although equivalent experiments with canonical nucleobases failed to yield base-pairing, the idea that solid-state can provide an alternative reaction medium for chemical selection based on self-assembly served as a motivation for this study.

In this section, the objective was to explore molecular recognition properties in a pool of plausibly prebiotic nucleobases. These included canonical nucleobases: adenine (A), thymine (T), uracil (U), guanine (G), cytosine (C) and those nitrogen heterocycles that have been detected in meteorites or are present in contemporary biology; 2,6-diaminopurine (D), hypoxanthine (H), and xanthine (X). We studied their interactions in reaction conditions that cover all extremes of prebiotic water activity and elucidated their selective self-assembly processes. We used slurry reactions in water to mimic prebiotic bulk water activity, LAG with water to mimic prebiotic conditions of catalytic water presence, and solid-state heating reactions to mimic dry prebiotic environments. Variable-temperature synchrotron powder X-ray diffraction (VT-SPXRD) monitoring enabled insights into the nucleobase cocrystal formation mechanisms. For example, cocrystallization from an amorphous matrix or cocrystal hydrate to cocrystal anhydrate transformations upon wet/dry cycling.

4.1.1. 2,6-Diaminopurine and uracil

As a starting experiment, an equimolar amount of powdered D and U was milled together for 2 h. PXRD analysis revealed that the reaction product **1** was poorly crystalline (Figure 10). After ageing **1** in water vapor, as well as by LAG with the addition of catalytic amount of water, **1** is converted to a highly crystalline phase **2** that is absent of diffraction peaks corresponding to reactants (Figure 10). Furthermore, **2** can also be obtained by slurring equimolar amounts of untreated reactants in water at room temperature for six days (Figure 10).

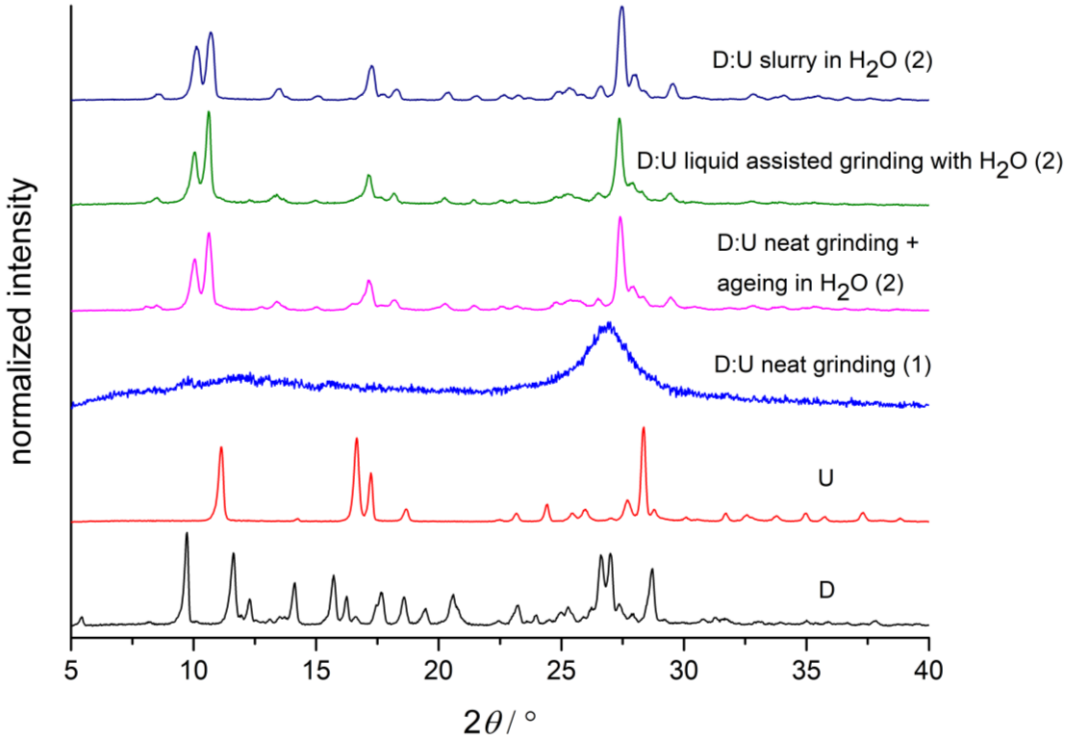

Figure 10. PXRD patterns for the D and U system.

We grew single crystals of **2** and elucidated its crystal structure by SCXRD (Figure 7.4 and Table 7.1).[a] **2** is a cocrystal channel hydrate and crystallizes in monoclinic *C*2/*c* space group. The asymmetric unit of **2** contains D and U assembled by three intermolecular Watson-Crick hydrogen bonds together with disordered water molecules (Figure 11a). Water content of **2** was determined by TGA and resulted in 5 % (Figure 7.5). Inside the crystal, D and U molecules form hydrogen-bonded layers (Figure 11b). Disordered water molecules fill structural voids and form channels along crystallographic *c*-axis (Figure 11c). Van der Waals interactions are acting between hydrogen-bonded layers that are distanced at 3.4 Å (Figure 12). Furthermore, disordered water molecules serve as vertical bridges between horizontal hydrogen-bonded layers by being involved in van der Waals interactions with D and U molecules within layers.

[a] The crystal structure of **2** was solved by collaborating co-author Dr. Krešimir Molčanov (Ruđer Bošković Institute).

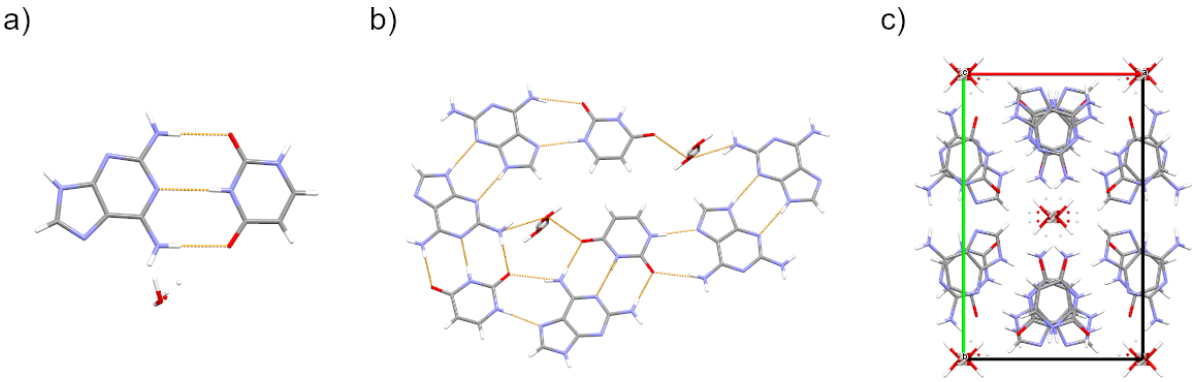

Figure 11. (a) The asymmetric unit of **2** contains D and U molecules assembled by Watson-Crick hydrogen-bonding as well as disordered water molecules (b) Hydrogen-bonded layers inside the crystal along crystallographic *b*-axis. (c) View from the crystallographic *c*-axis showing disordered water molecules that lie in the channel.

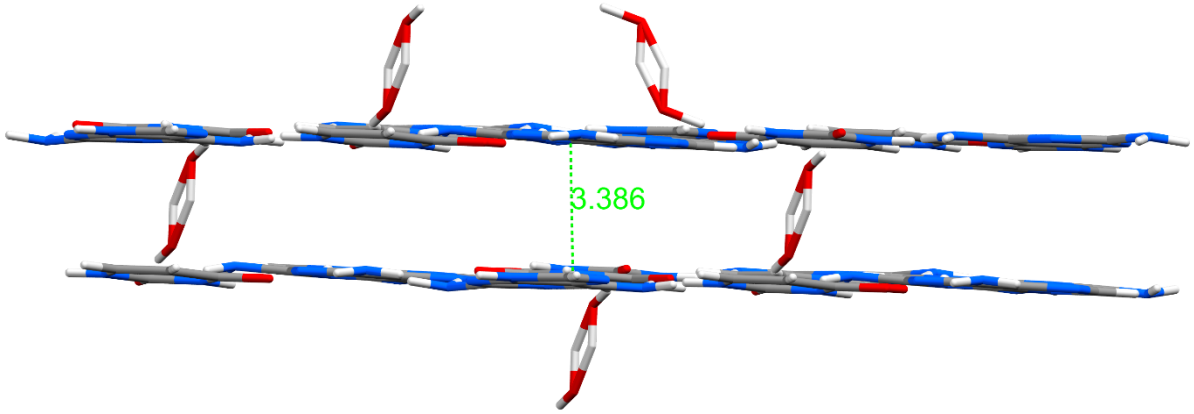

Figure 12. Distance in Å between hydrogen bonded layers of **2**.

The heating effect of wet/dry cycling on D and U base-pair was monitored in situ by VT-SPXRD (Figure 13). At around 400 K, the formation of new Bragg reflections started to appear and reached maximum intensities at around 430 K. After 430 K, this new phase **3** which is presumably a D:U cocrystal anhydrate, is the only one present in the reaction mixture until the end of monitoring at 490 K. We attempted to solve the crystal structure of **3** from PXRD

data but were not successful. The best Pawley fit was achieved by using $P\bar{1}$ space group but simulated annealing failed to provide any reasonable structural model (Figure 14).

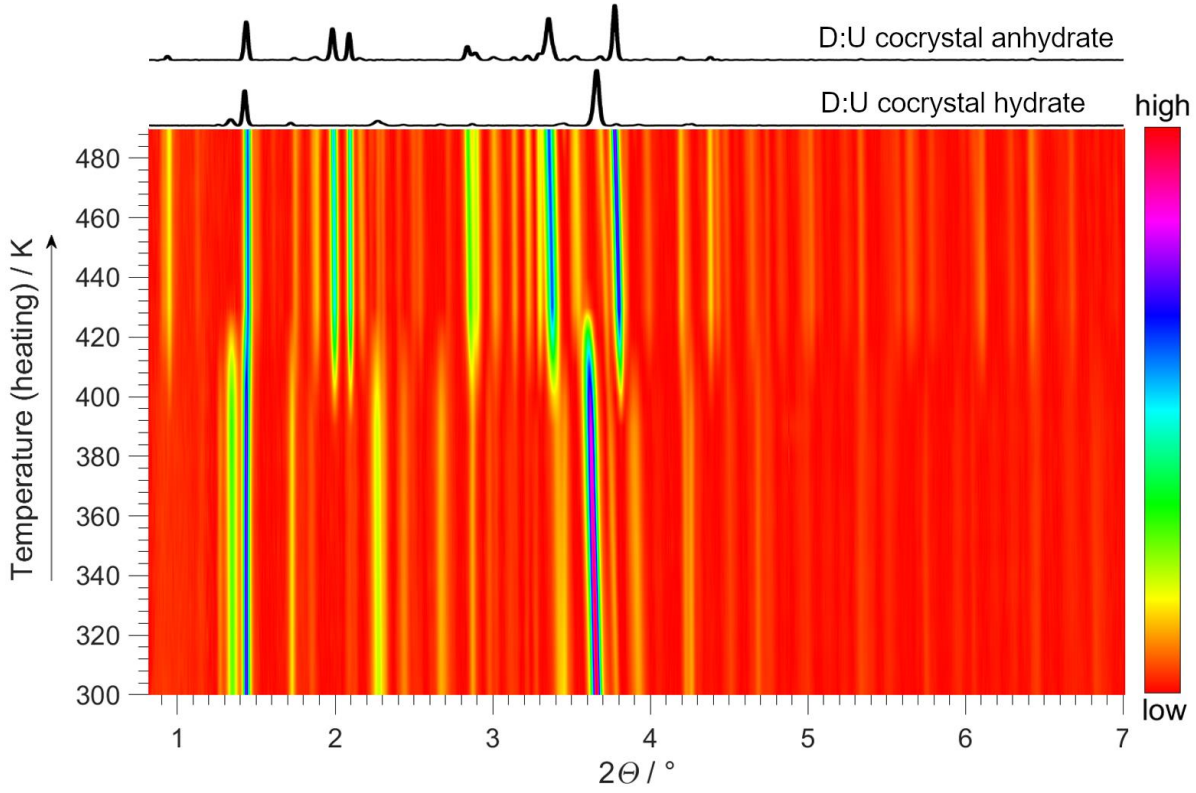

Figure 13. a) In situ SPXRD monitoring of heating **2** using synchrotron radiation (λ = 0.20741 Å). To obtain 2D plot, individual PXRD patterns were stacked upon each other. Note: some Bragg reflections are shifting towards lower 2θ values during heating due to thermal expansion of the unit cell.

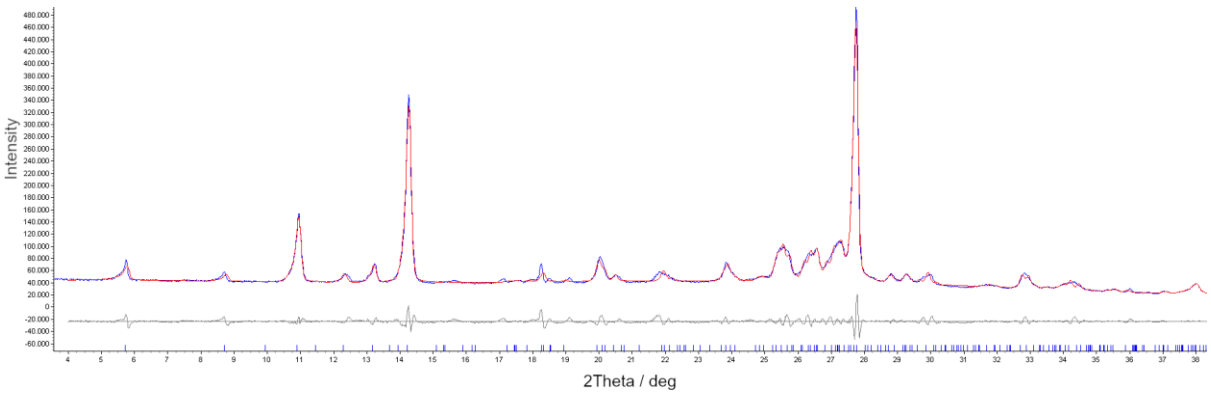

Figure 14. The best Pawley fit for **3** was achieved by using $P\bar{1}$ space group (a = 10.22 Å, b = 7.20 Å, c = 16.58 Å, α = 112.00°, β = 82.30°, γ = 96.33°, V = 1144.02 Å^3). Simulated annealing failed to provide any reasonable structural model.

To test whether the formation of D:U cocrystal hydrate proceeds under cryogenic conditions, cryogenic ball milling experiment was conducted. Rietveld refinement confirmed that after LAG an equimolar amount of D and U with H_2O for 30 min in the 80-100 K temperature range, D and U self-assembled as **2** (Figure 15).

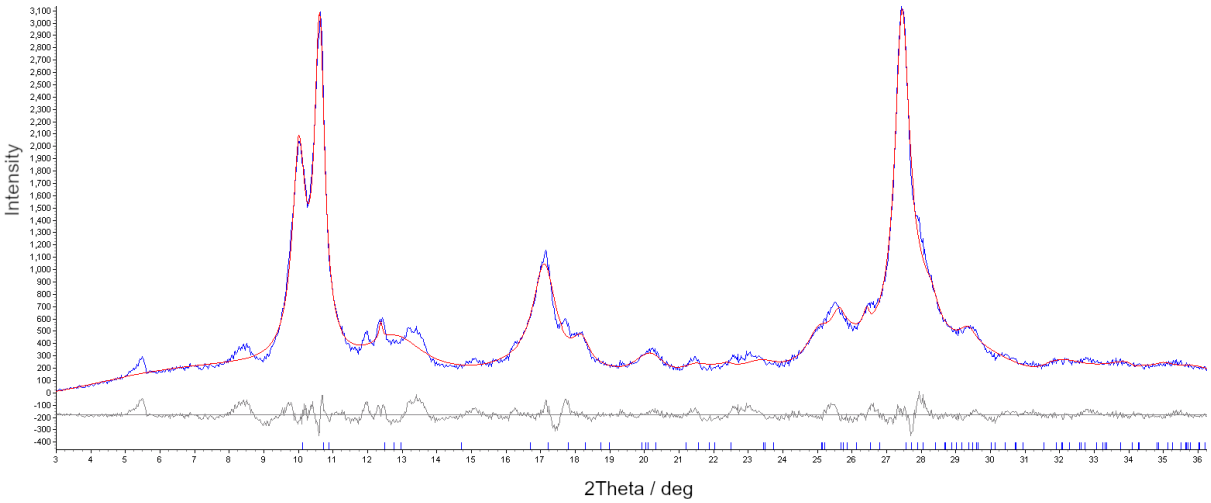

Figure 15. Rietveld refinement of LAG an equimolar amount of D and U with H_2O for 30 min in the 80-100 K temperature range. The resulting product is mainly composed out of **2**.

4.1.2. 2,6-Diaminopurine and thymine

After neat grinding an equimolar amount of powdered D and T, the resulting solid mixture was analyzed by PXRD and revealed that the reaction product **4** was poorly crystalline (Figure 16). After ageing **4** in water vapor, highly crystalline phase **5** absent of diffraction peaks corresponding to reactants was obtained and is presumably D:T cocrystal hydrate (Figure 16). Crystalline phase **5** can also be obtained by LAG equimolar amounts of D and T with the catalytic addition of water as well as by slurry mixing equimolar amounts of D and T in water for six days at room temperature.

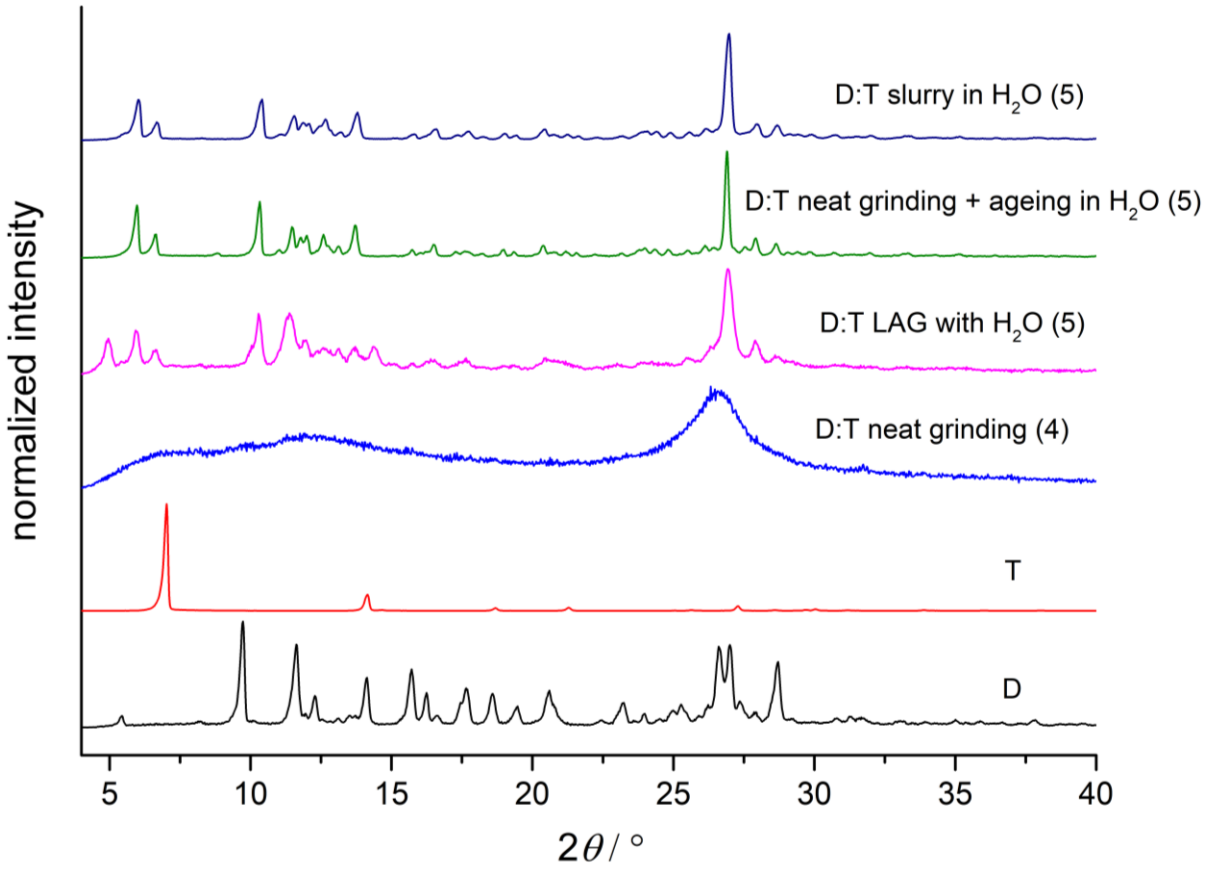

Figure 16. PXRD patterns for the D and T system.

Attempts to grow single crystals of **5** from aqueous solution failed and resulted in single crystals of D hydrate that is known from the literature.[144] On the other hand, after dissolving **4** in methanol, we obtained needle crystals and solved the crystal structure using SCXRD data (Figure 7.6 and Table 7.2).[b] This showed that D and T crystallized as a 1:2 stoichiometric cocrystal hydrate **10** in a monoclinic *Pc* space group. The asymmetric unit of **10** contains four symmetrically non-equivalent T and water molecules as well as two symmetrically non-equivalent D molecules (Figure 17). Crystal packing is characterized by hydrogen-bonded layers where each D molecule is connected to two distinct T molecules by three intermolecular Watson-Crick hydrogen bonds as well as to a third T molecule by $R_2^2(9)$ hydrogen bonding

[b] The crystal structure of **10** was solved by collaborating co-author Prof. Ivica Đilović (Chemistry Department, Faculty of Science, University of Zagreb).

motif (Figure 18a). Furthermore, horizontal hydrogen-bonded layers are vertically interconnected by hydrogen-bonded water molecules (Figure 18b).

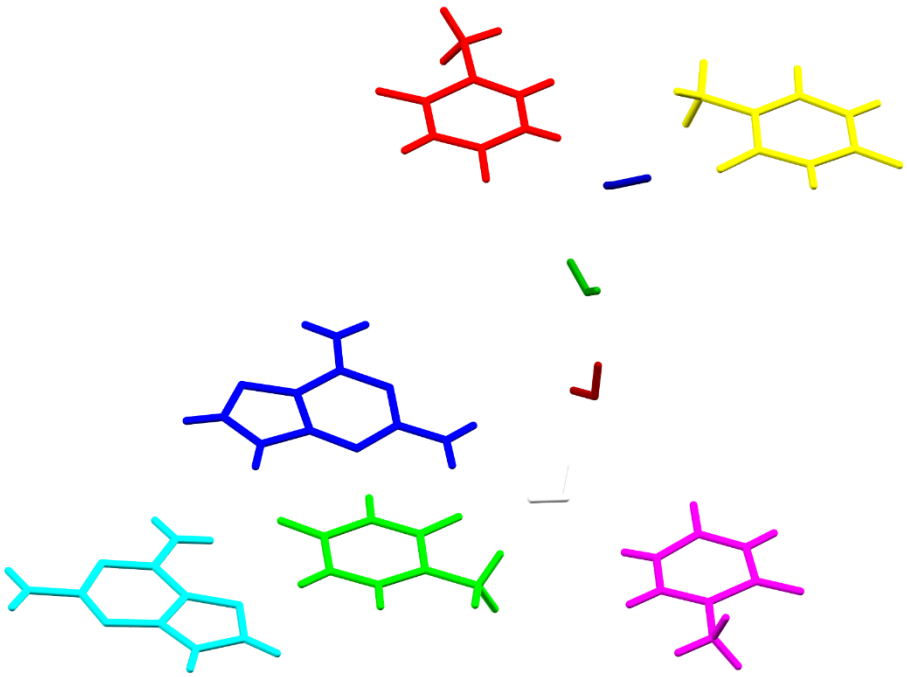

Figure 17. Asymmetric unit of **10** contains four symmetrically non-equivalent T and water molecules as well as two symmetrically non-equivalent D molecules.

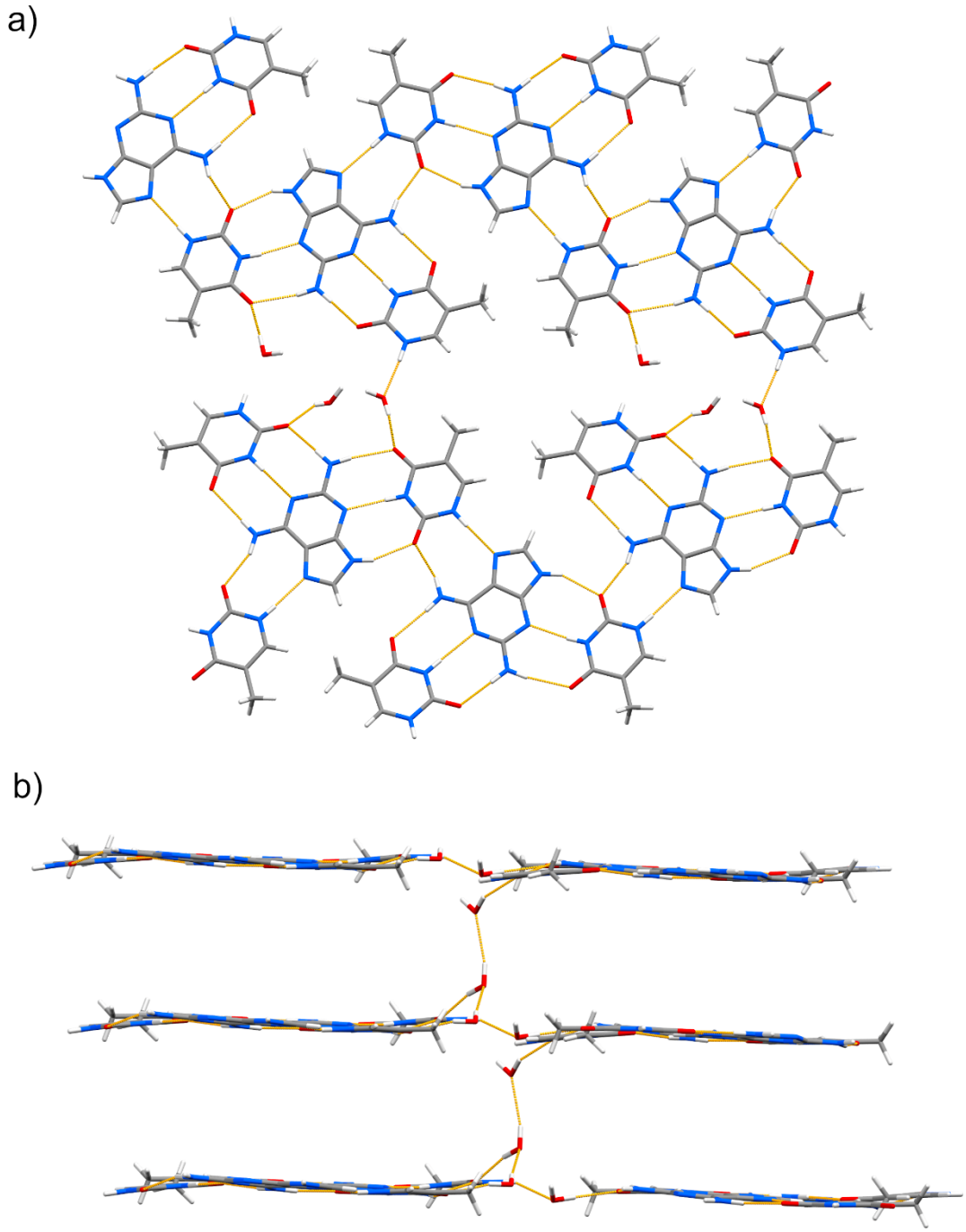

Figure 18. Crystal packing in **10**. a) Hydrogen-bonded layers viewed down the crystallographic *c*-axis. b) Horizontal hydrogen-bonded layers are vertically interconnected by hydrogen-bonded water molecules. View along crystallographic *b*-axis.

Heating of **4** was monitored in situ by VT-SPXRD (Figure 19a). At around 420 K several low diffraction peaks appear which are a result of gradual re-crystallization of the starting mixture. Next, at approximately 460 K crystallization of **6**, which is presumably D:T cocrystal anhydrate, is observed as evidenced by the appearance of new high intensity diffraction peaks. The appearance of **6**, which remains as the only crystalline phase inside the reaction mixture until the end of monitoring at 490 K, coincides with the crystallization peak at 460 K in DSC trace of neat ground equimolar mixture of D and T (Figure 19b).

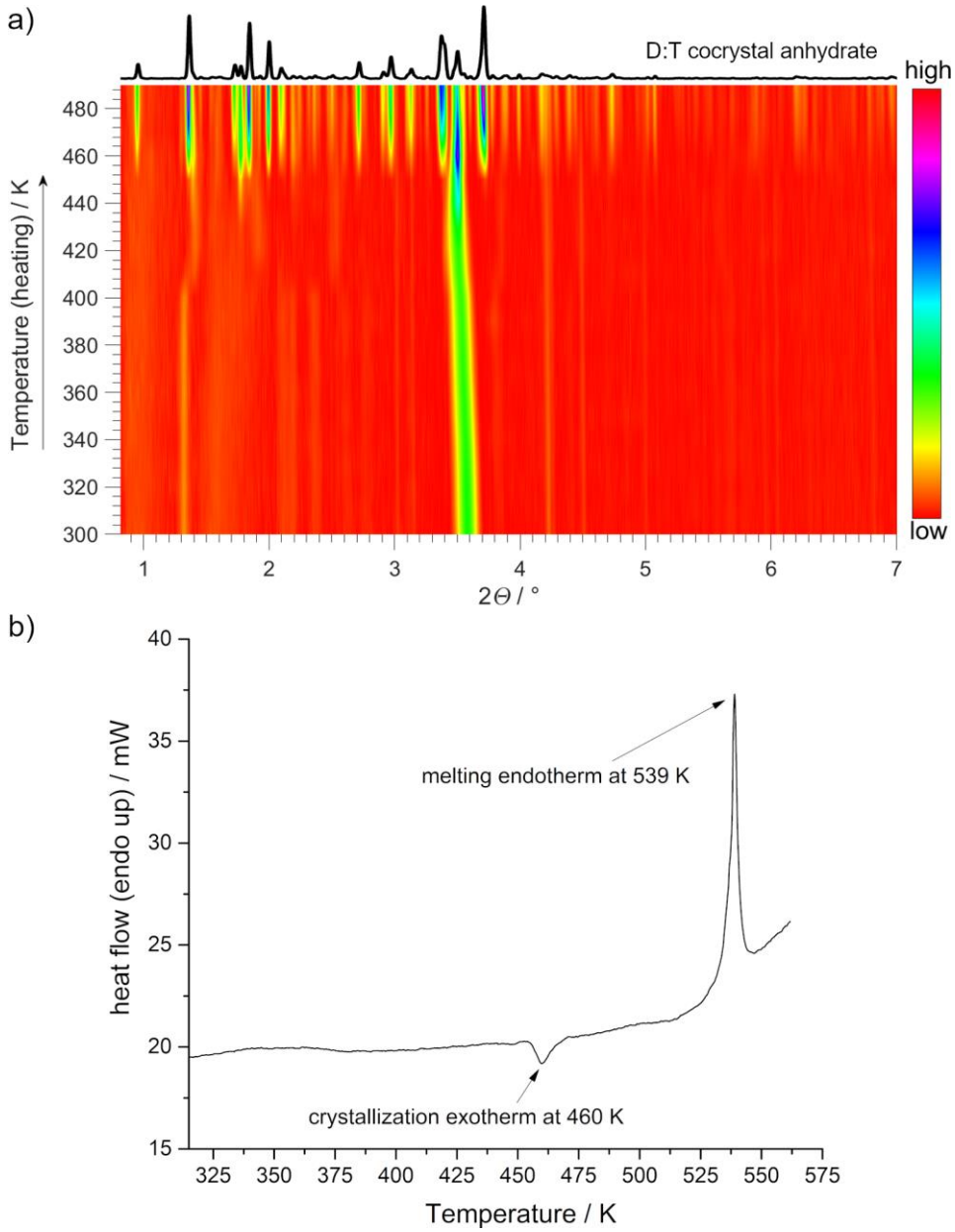

Figure 19. a) In situ SPXRD monitoring of heating **4** using synchrotron radiation (λ =0.20741 Å). To obtain 2D plot, individual PXRD patterns were stacked upon each other. b) DSC trace of a neat ground equimolar mixture of D and T.

Crystal structure of **6** was solved from the laboratory PXRD data (Figure 20 and Table 7.3). **6** is indeed a D:T cocrystal anhydrate and crystallizes in triclinic $P\bar{1}$ space group. The asymmetric unit of **6** contains two symmetrically non-equivalent D and T molecules, making eight molecules in the unit cell (Figure 21). Crystal packing is characterized by hydrogen-bonded layers. Each T molecule is connected to one D molecule by three intermolecular Watson-Crick hydrogen bonds as well as to a second D molecule by $R_2^2(9)$ hydrogen bonding motif (Figure 22). Hydrogen-bonded layers mutually interact through van der Waals interactions.

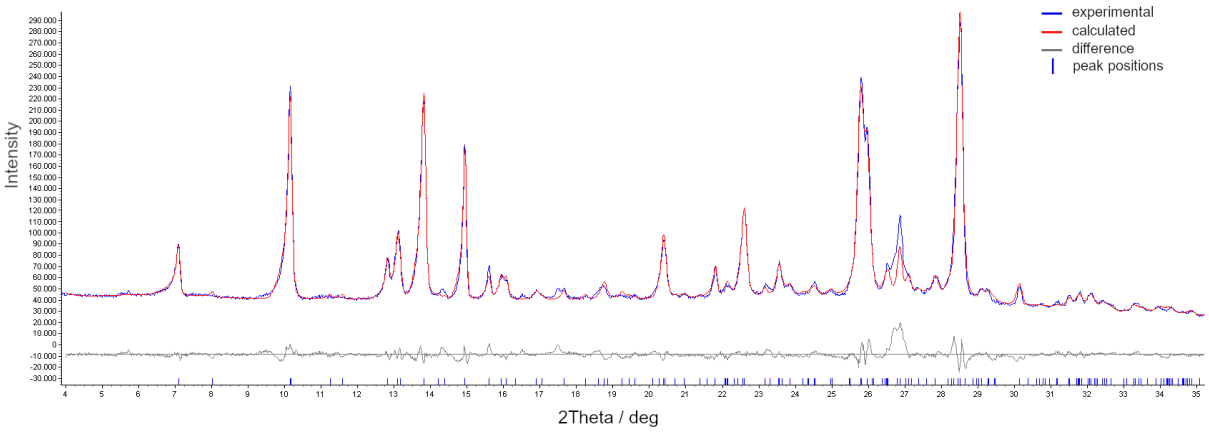

Figure 20. Rietveld fit of **6** crystal structure that was solved from the laboratory PXRD data (λ = 1.5406 Å).

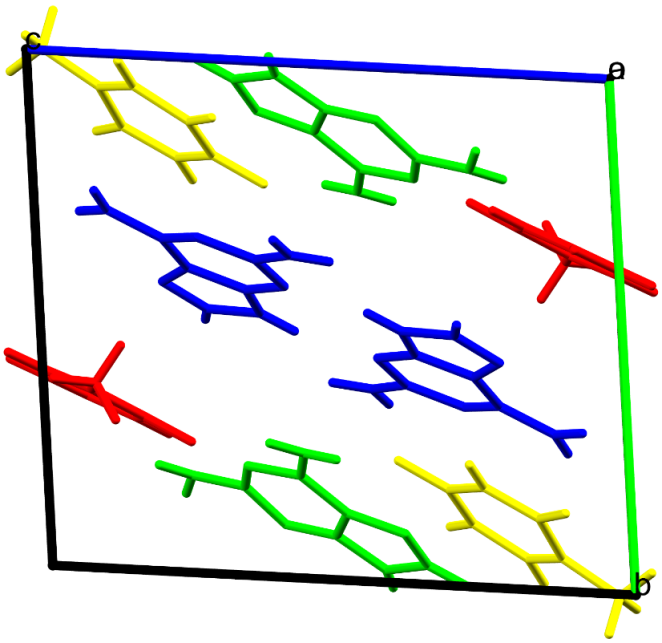

Figure 21. The unit cell of **6** is composed of four D and four T molecules.

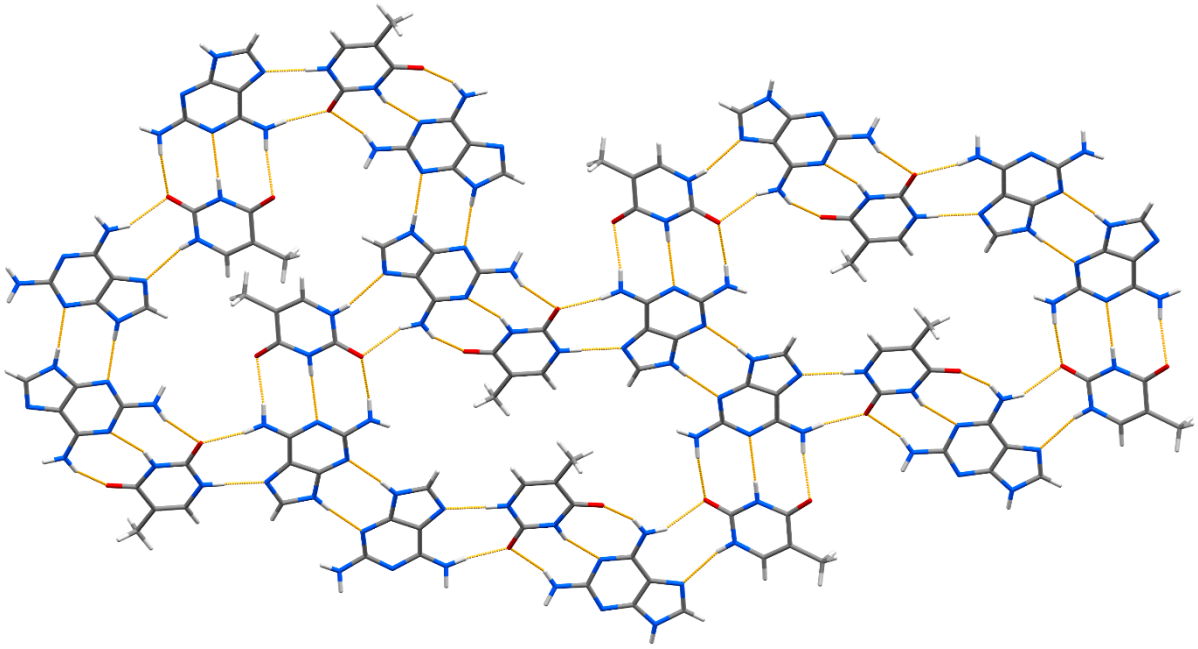

Figure 22. Hydrogen-bonded layers in **6** seen along crystallographic *b*-axis.

4.1.3. Adenine and hypoxanthine

After neat grinding an equimolar amount of powdered A and H, the resulting solid mixture was analyzed by PXRD and revealed that the reaction product **7** was poorly crystalline (Figure 23). Slurry mixing equimolar amounts of A and H in H_2O at 65 °C (Figure 7.7), LAG equimolar amounts of A and H with H_2O at 65 °C (Figure 7.8), and heating **7** in DSC experiment to 300 °C (Figure 7.9) resulted in the formation of a new crystalline phase **8** that was absent of diffraction peaks corresponding to reactants (Figure 23). Furthermore, hydrothermal synthesis of equimolar amounts of A and H resulted in the formation of distinct new crystalline phase **9** (Figure 23).

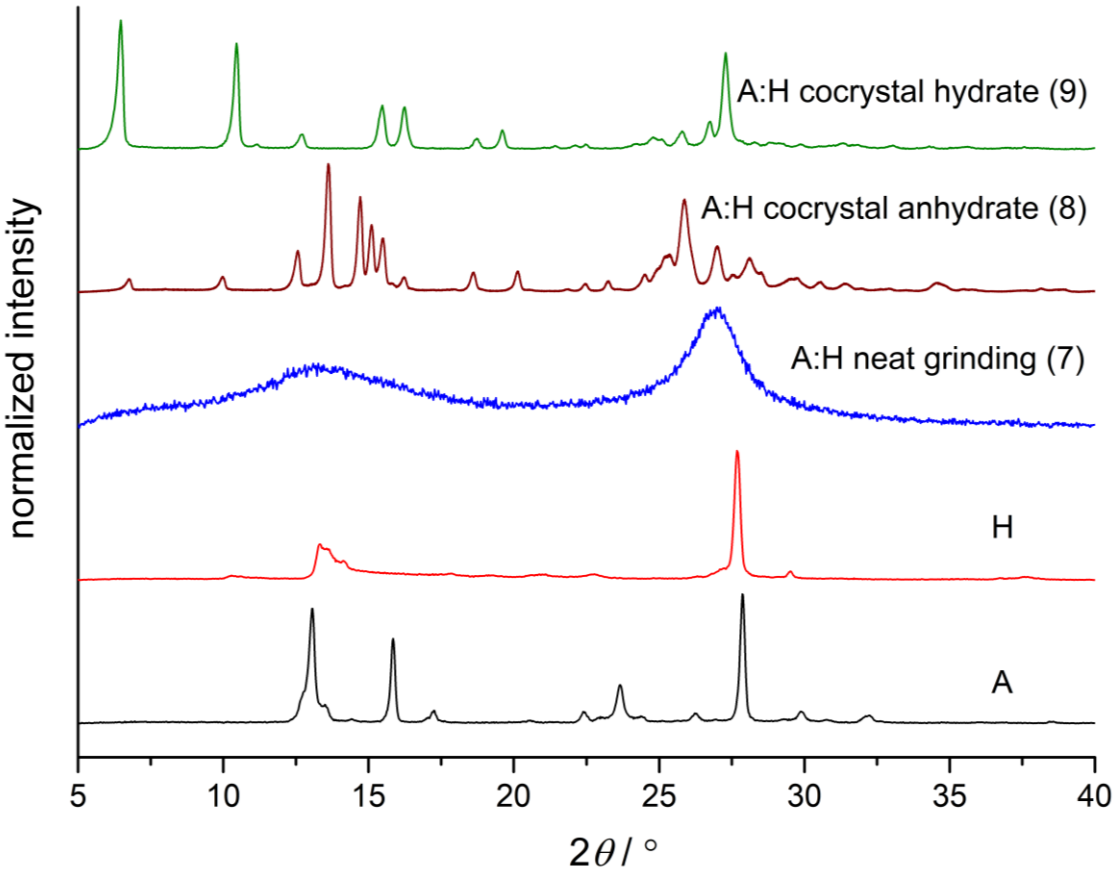

Figure 23. PXRD patterns for A and H system.

Since crystal structure solution of **8** from the PXRD data indicated that three different structural models were possible, we resolved this conundrum by the combination of crystallography, DFT computation, and solid-state nuclear magnetic resonance (ssNMR) analysis (Figure 24).[c] DFT computation showed that model A had by far the highest ground state energy (Figure 24a). The ground state energies of B and C were quite close one to another, although B still had notably lower energy. For model A, the calculated ^{13}C and ^{15}N chemical shifts did not match well with the experimentally obtained shifts (Figure 24b). This model was also excluded by the 2D DQ-SQ BABA (Back-to-Back) ^{1}H-^{1}H homonuclear correlation experiment (Figure 24d). However, the latter showed the presence of the diagonal cross peak at about 13.5 ppm in the ^{1}H dimension, a consequence of protons that resonate at this chemical

[c] DFT computational and ssNMR analysis was performed by collaborating co-authors Dr. Gregor Mali (National Institute of Chemistry, Ljubljana) and Dr. Andraž Krajnc (National Institute of Chemistry, Ljubljana).

shift being close in space (i.e., two equivalent protons in two neighboring molecules are quite close one to another). For model A, this was not the case since equivalent N-H protons that would be close in space were absent. This resulted in exclusion of model A. Regarding models B and C, both models predicted ^{1}H, ^{13}C and ^{15}N chemical shifts that agreed with the experimental ones reasonably well. However, in the ^{13}C-detected proton spin diffusion 2D experiment (Figure 24c), we would have expected a strong correlation peak between C2ads and H1ads for model C (highlighted by the blue circle). The peak would have had a similar intensity as the H1hyp-C2hyp cross peak. This was clearly not the case. On the other hand, the observed cross peaks agreed well with the proposed model B. Therefore, a combination of ssNMR experiments and DFT computation suggested that model B is the best model among the three models.

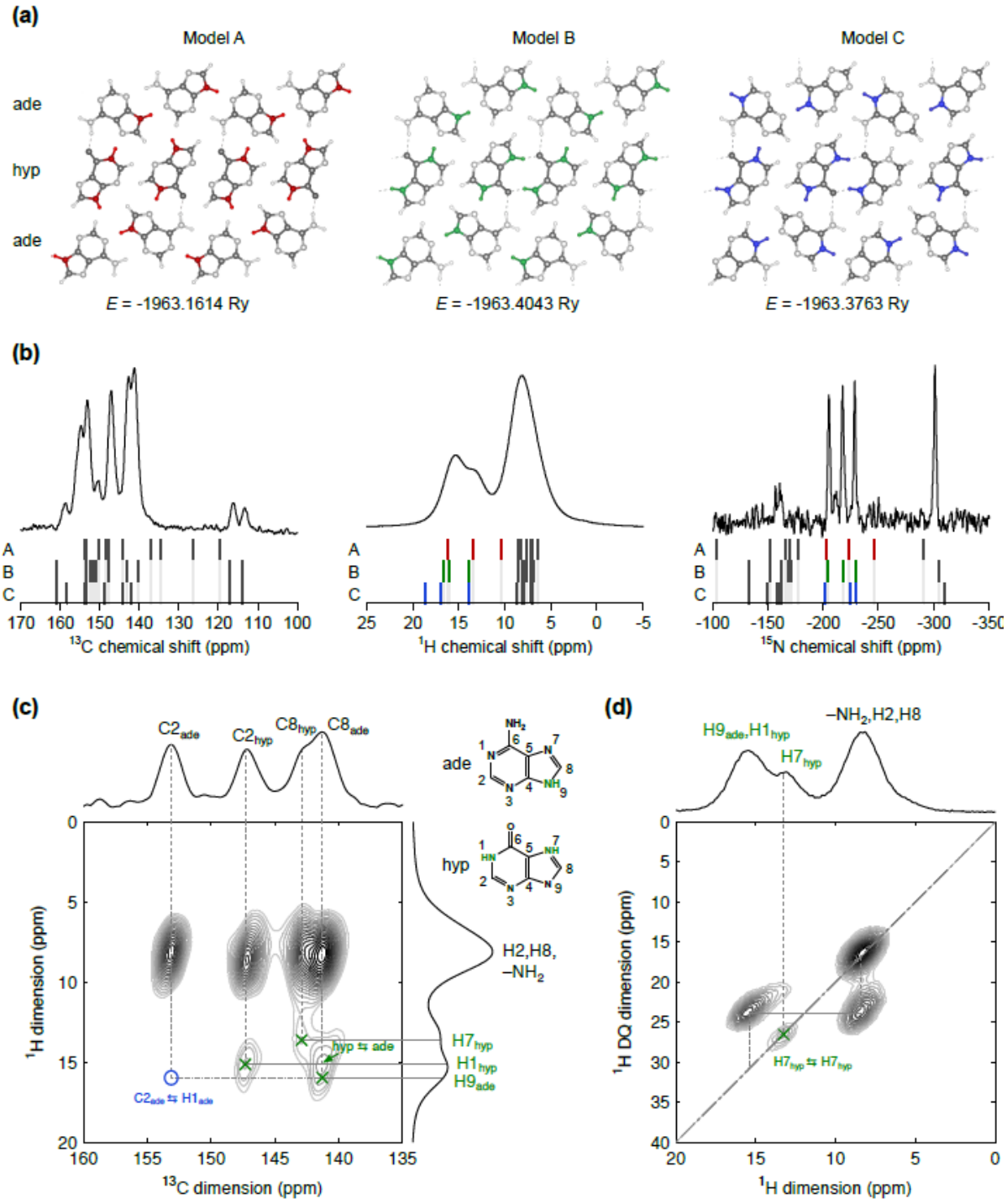

Figure 24. DFT computational and ssNMR analysis. a) Structural model A: adenine (tautomer H9) + hypoxanthine (tautomer H9), structural model B: adenine (tautomer H9) + hypoxanthine (tautomer H7), and structural model C: adenine (tautomer H1) + hypoxanthine (tautomer H7) with DFT ground state energies. b) ^{13}C, ^{1}H, and ^{15}N calculated and experimental chemical shifts. c) ^{13}C-detected proton spin diffusion 2D experiment. d) 2D DQ-SQ Back-to-Back ^{1}H-^{1}H homonuclear correlation experiment.

As a result, structural model B which included adenine in 9H tautomeric form and hypoxanthine in 7H tautomeric form was chosen for Rietveld refinement of **8** (Figure 25 and Table 7.4). **8** crystallizes in $P2_1/n$ space group with one adenine and one hypoxanthine molecules in the asymmetric unit and overall, eight molecules in the unit cell (Figure 26). Adenine and hypoxanthine molecules interact through Hoogsteen $R_2^2(9)$ hydrogen-bonding mode and crystal packing is characterized by hydrogen-bonded sheets (Figures 27 and 28). They are composed of homomeric adenine strands that hydrogen-bond with complementary homomeric hypoxanthine strands (Figure 29). Hypoxanthine molecules are distanced at 3.4 Å and are involved in parallel displaced π-π stacking interactions (Figure 30).

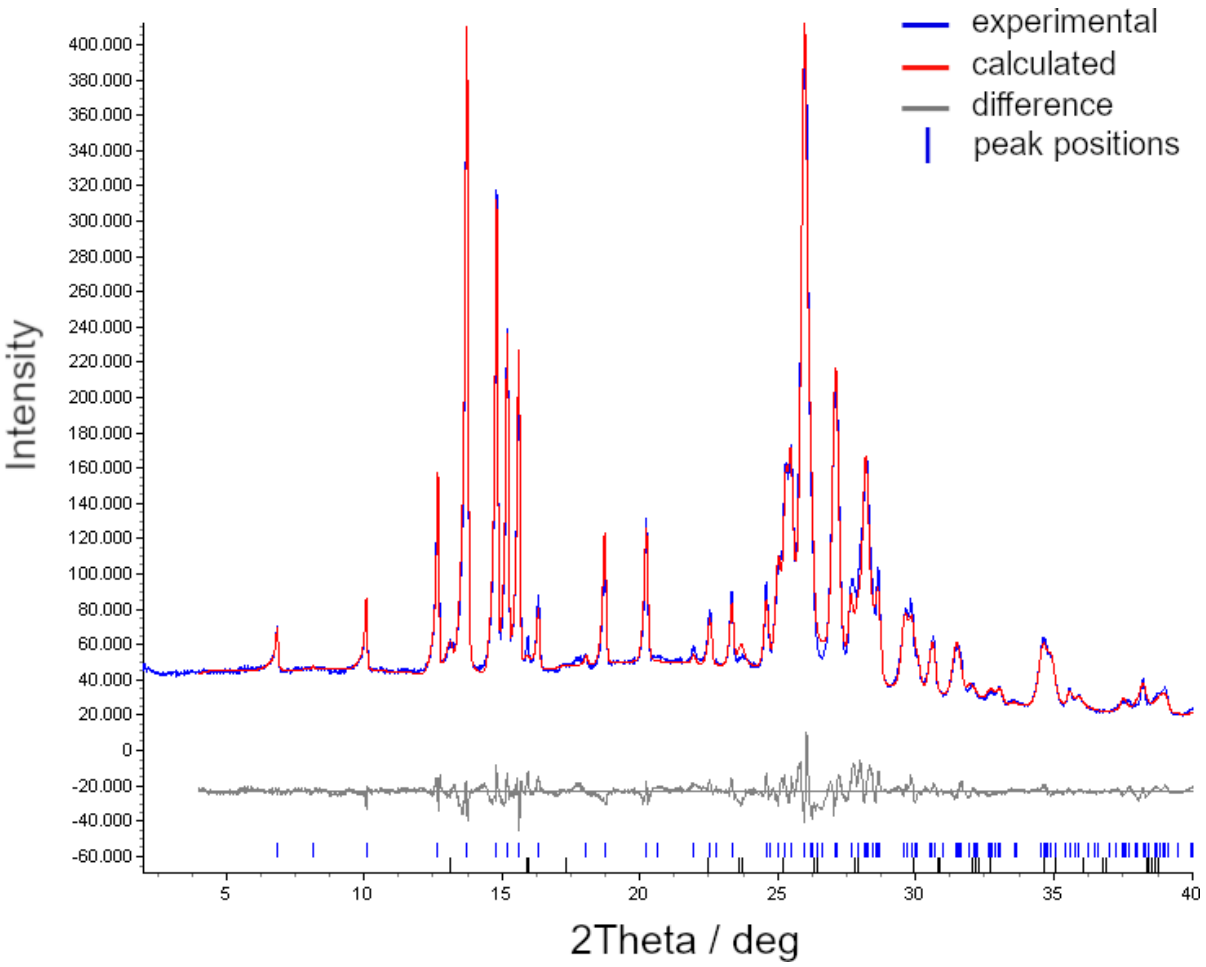

Figure 25. Rietveld fit of **8** crystal structure that was solved from the laboratory PXRD data (λ = 1.5406 Å) with inputs from DFT computation and ssNMR analysis. Note that experimental PXRD pattern contained small amounts of unreacted adenine which was modeled in Rietveld refinement.

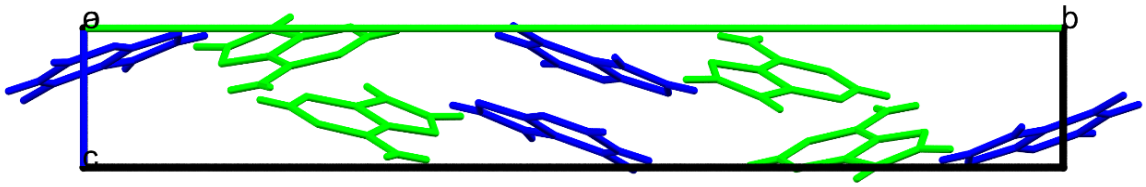

Figure 26. The unit cell of **8** viewed from crystallographic *a*-axis.

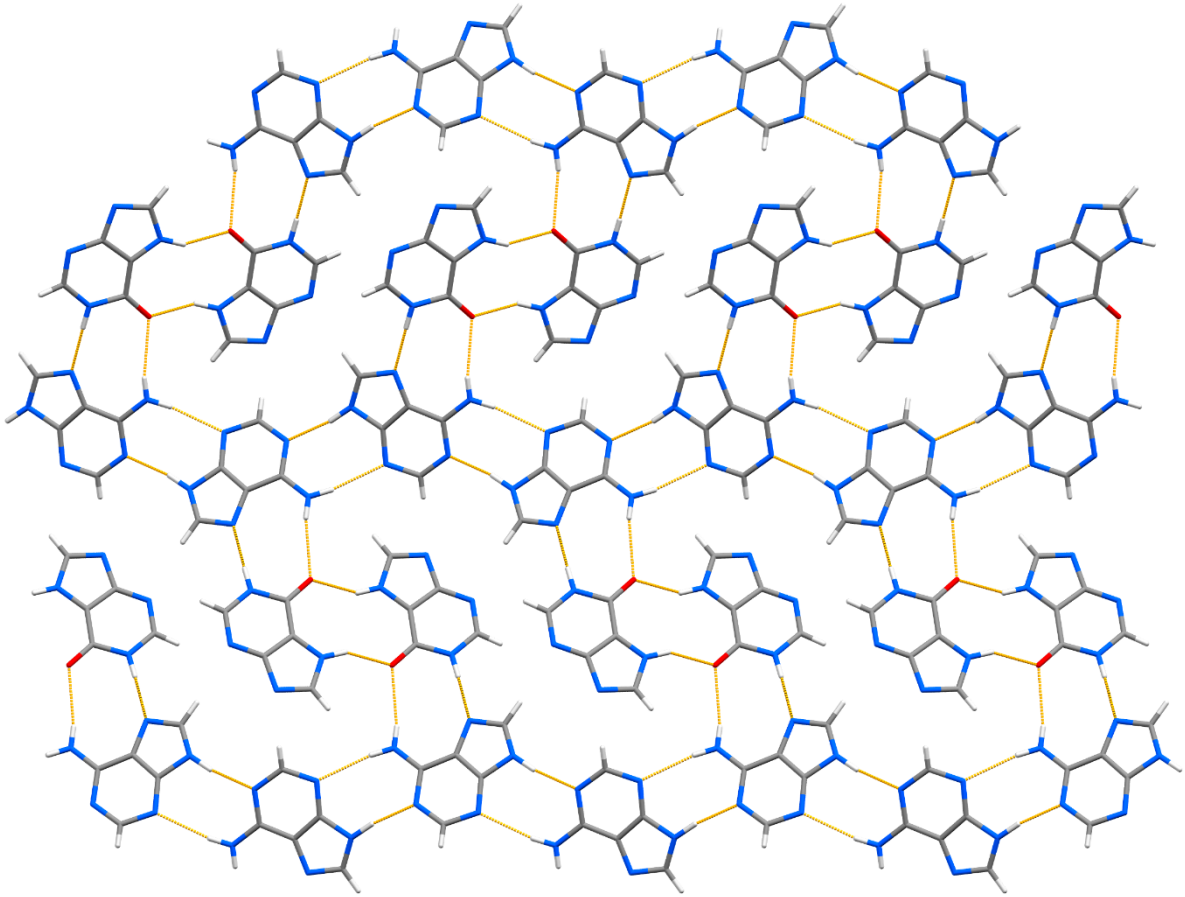

Figure 27. Top view of hydrogen-bonded sheets in **8** seen along crystallographic *c*-axis. Adenine and hypoxanthine molecules interact through Hoogsteen $R_2^2(9)$ hydrogen-bonding mode.

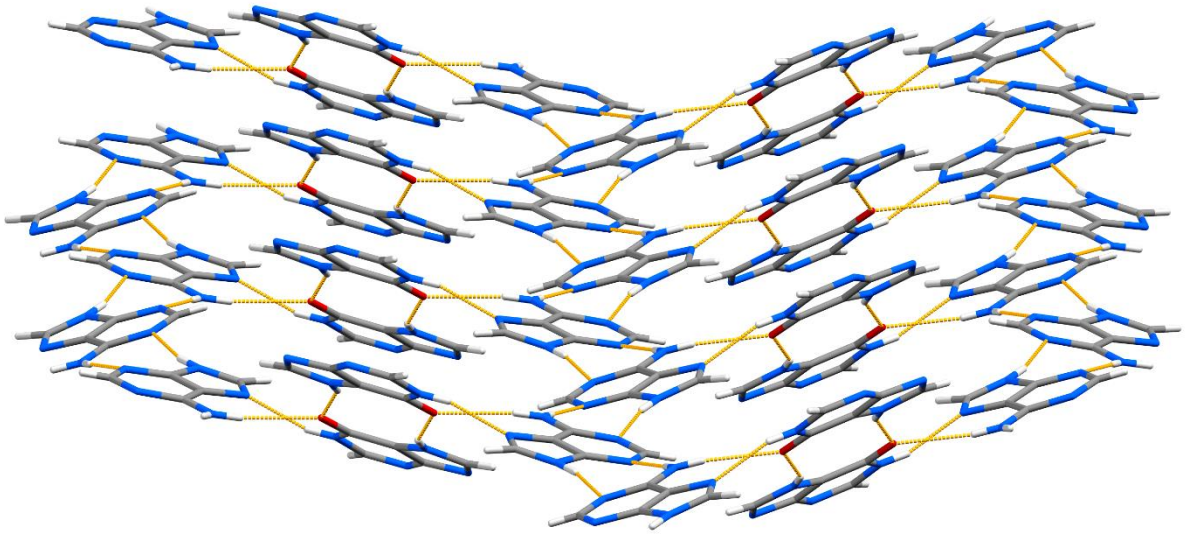

Figure 28. Side view of hydrogen-bonded sheets in **8** seen along crystallographic *a*-axis.

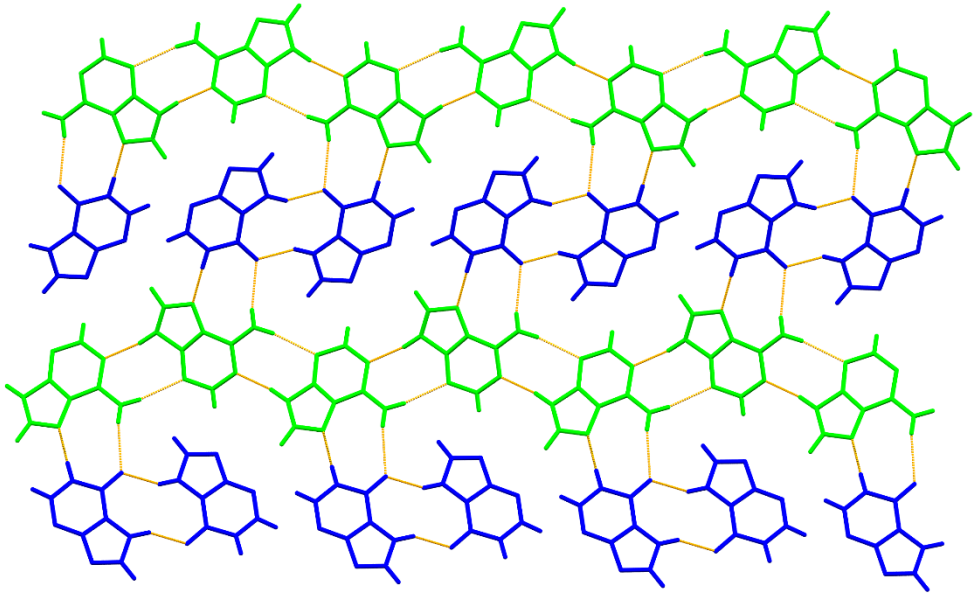

Figure 29. Sheets in **8** are composed of homomeric adenine strands that hydrogen-bond with complementary homomeric hypoxanthine strands. View along crystallographic *c*-axis.

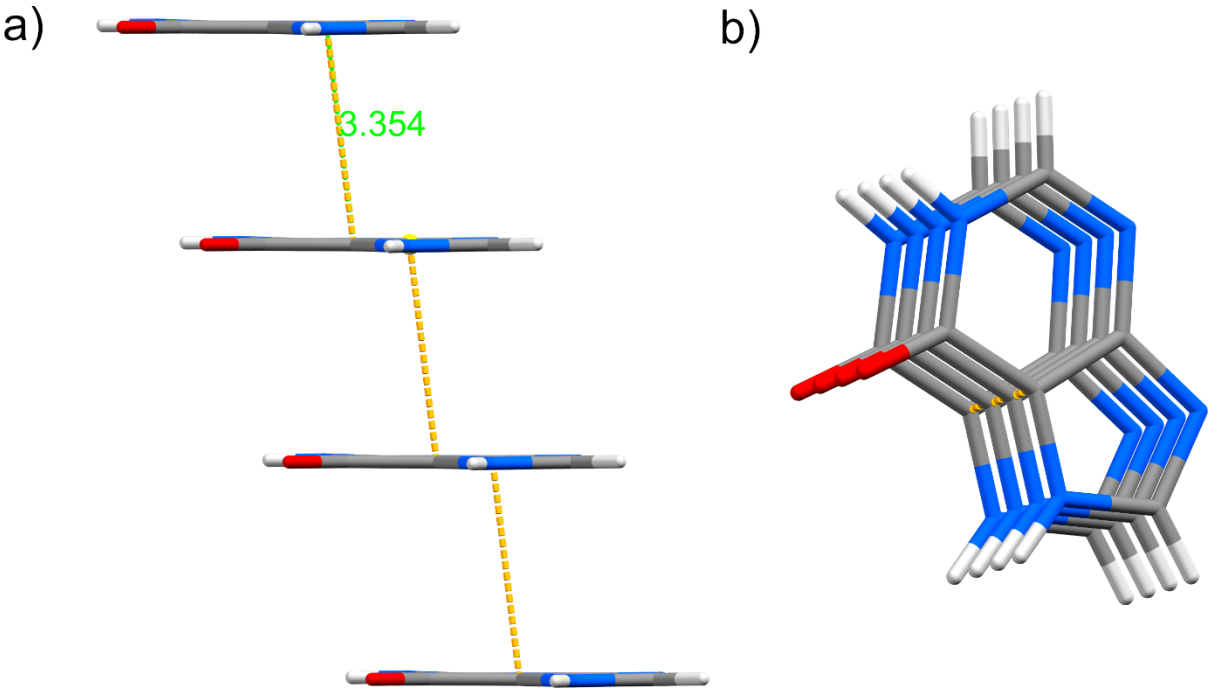

Figure 30. Hypoxanthine molecules are distanced at ca 3.4 Å and are involved in parallel displaced π-π stacking interactions: a) view from the crystallographic *a*-axis, b) view from the crystallographic *c*-axis.

To solve the crystal structure of **9** we used the identical structural model as in **8** (adenine in 9H tautomeric form and hypoxanthine in 7H tautomeric form) and solved it from the laboratory PXRD data (Figure 31 and Table 7.5). **9** is a channel hydrate and crystallizes in $P2_1/n$ space group with four adenine, four hypoxanthine, and eight water molecules in the unit cell (Figure 32). Hydrogen-bonded sheets are closely packed with water molecules filling structural voids and lying in the channel along crystallographic *c*-axis (Figure 33). Water molecules interact with the exposed C–H groups of adenine and hypoxanthine by van der Waals interactions. Each adenine molecule is involved in two $R_2^2(9)$ hydrogen bonding motifs and one $C(2)$ chain motif with hypoxanthine molecules, and vice versa (Figure 34). Adenine molecules are distanced at 3.2 Å and are involved in parallel displaced π-π stacking interactions (Figure 35).

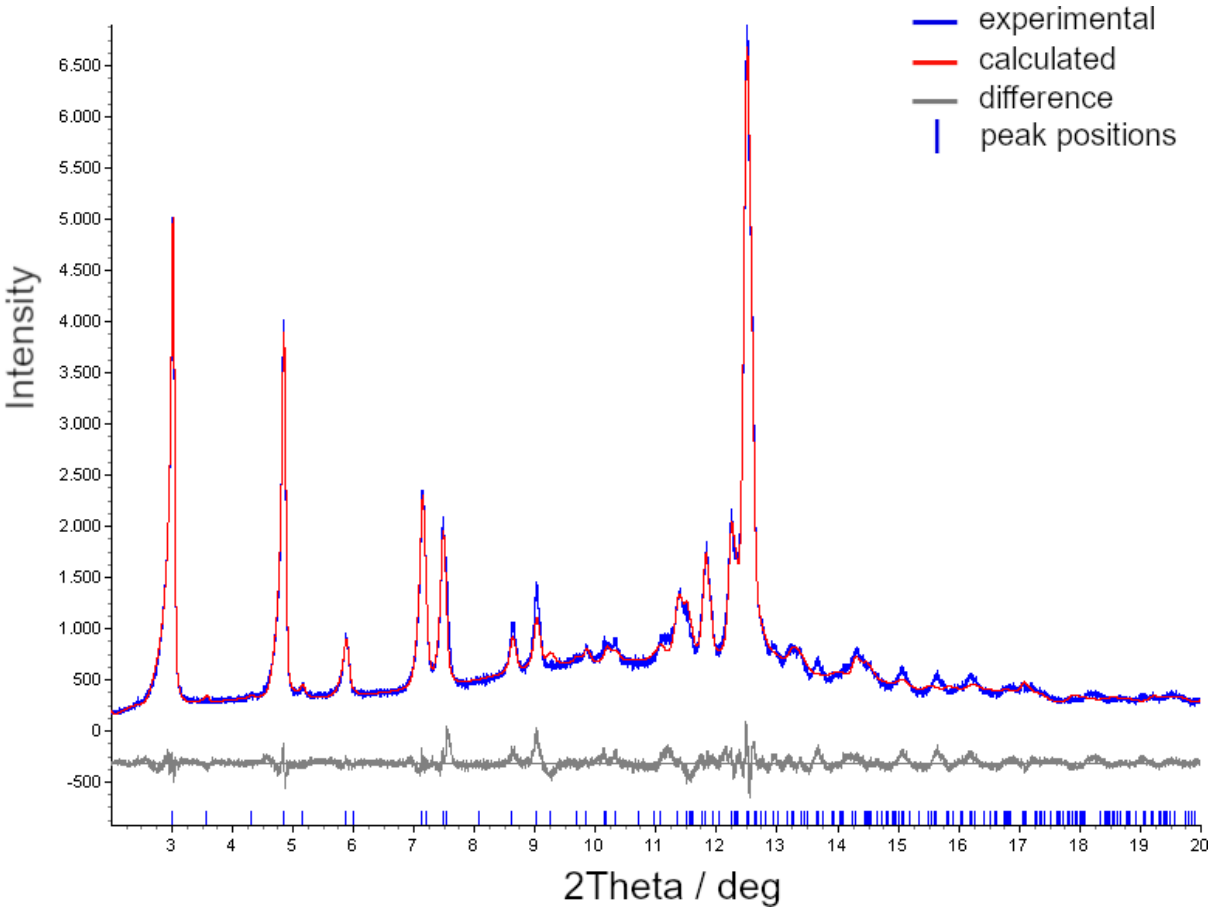

Figure 31. Rietveld refinement of **9** crystal structure that was solved from the laboratory PXRD data (λ = 0.7107 Å) using the same structural model as in **8** (adenine in 9H tautomeric form and hypoxanthine in 7H tautomeric form).

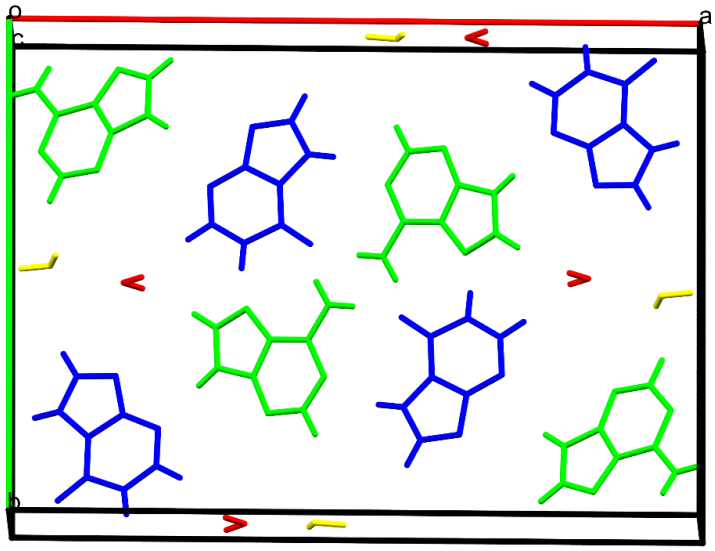

Figure 32. The unit cell of **9** is composed of four adenine, four hypoxanthine, and eight water molecules.

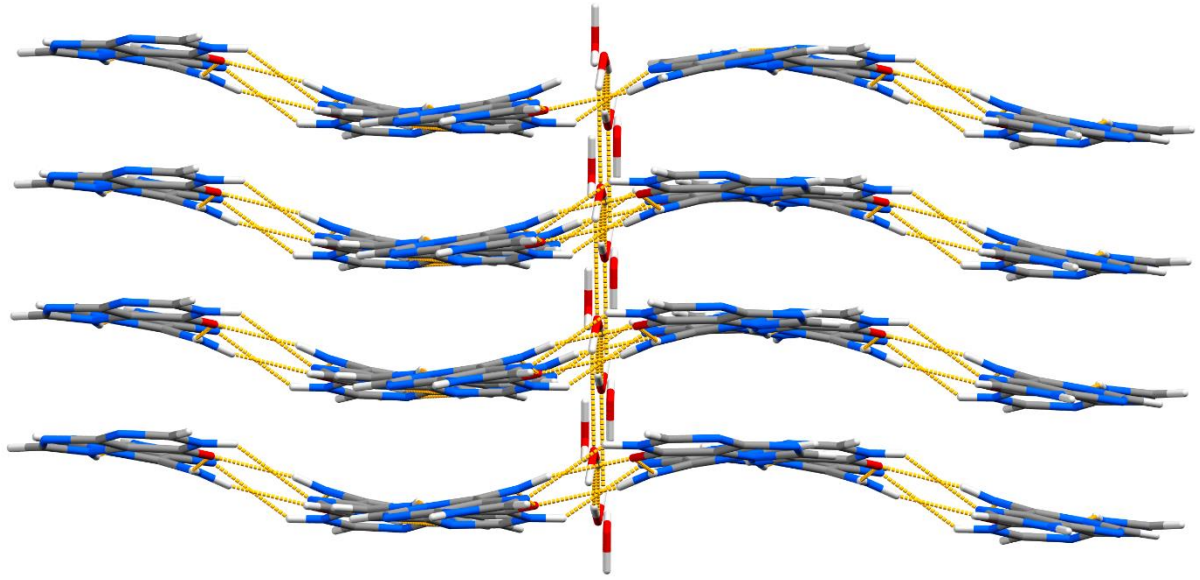

Figure 33. Hydrogen-bonded sheets in **9** are closely packed with water molecules filling structural voids and lying in the channel along crystallographic *c*-axis.

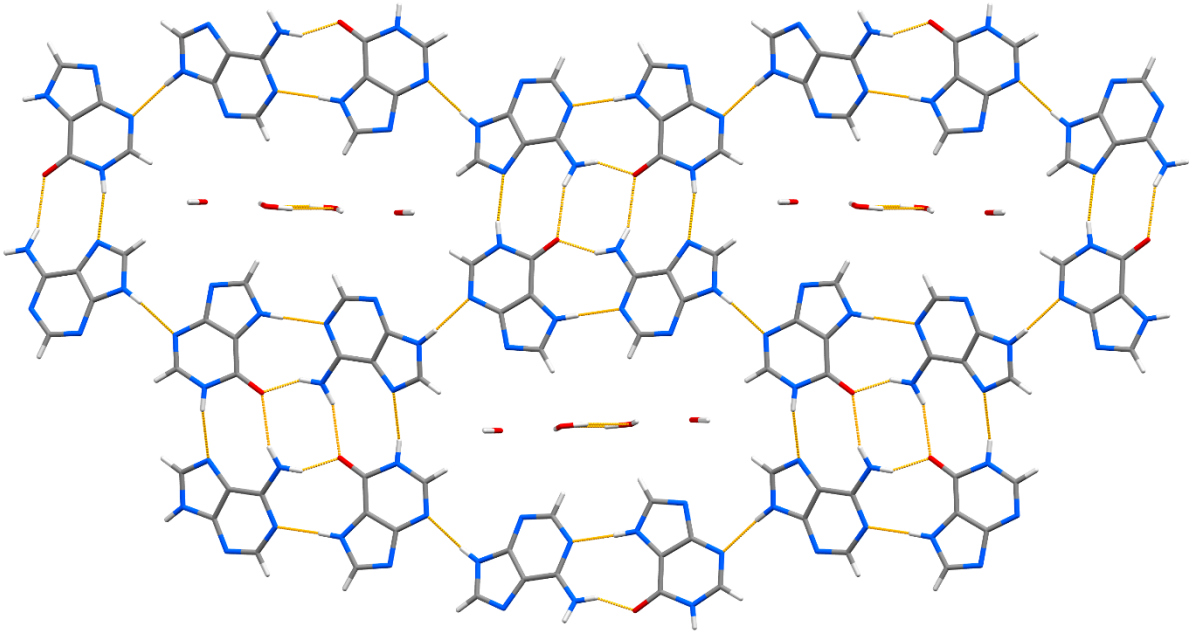

Figure 34. In hydrogen-bonded sheets of **9**, each adenine molecule is involved in two $R_2^2(9)$ hydrogen bonding motifs and one $C(2)$ chain motif with hypoxanthine molecules, and vice versa.

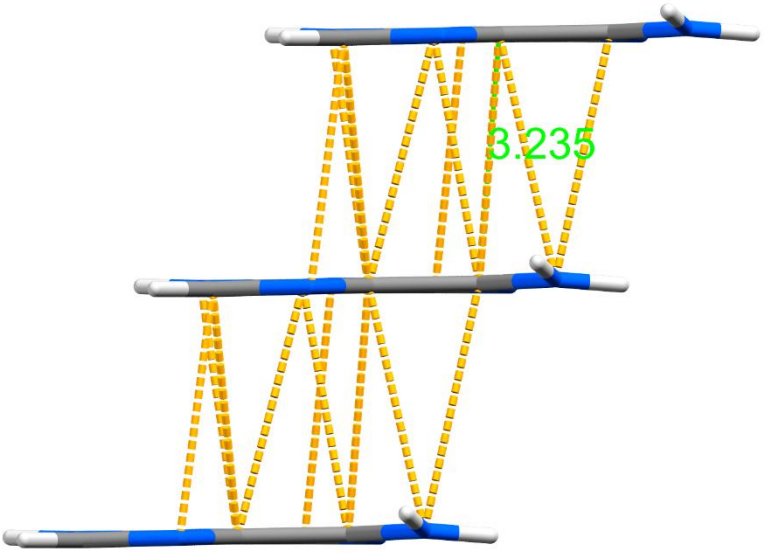

Figure 35. Adenine molecules are stacked at ca 3.2 Å and are involved in parallel displaced π-π stacking interactions.

4.1.4. Selectivity of base-pairing and other combinations

The summary of all base-pairing combinations that were investigated in this dissertation is given in Table 1 (Figures 7.10–7.13). D and U, D and T, as well as A and H base-pair in all reaction conditions of different prebiotic water activity. Furthermore, all of them base pair selectively when in the presence of competing nucleobases (Table 2, Figures 7.14-7.18).

Some of the reaction conditions needed for nucleobase self-assembly, such as catalytic presence of water/absence of water and cryogenic/elevated temperatures, might be of astrobiological significance. The individual nucleobases used in our study have been detected in meteorites and their extraterrestrial origin has been ascertained. Meteorites are asteroidal fragments that have fallen to the Earth's surface[78] and likely served as an exogenous source of prebiotic organic molecules during the time of the heavy bombardment.[9,16] Moreover, since meteorite composition depends on aqueous alteration, thermal metamorphism, or hydrous metamorphism, these post-accretion events might have presented opportunities for the self-assembly of nucleobases. The latter represents an example of chemical evolution of important biomolecules in laboratory conditions that mimic their extraterrestrial environments.

Table 1. Results of base-pairing experiments using different prebiotic water availability.

Reactants	Bulk water	Catalytic presence of water	Dry conditions
D, U	D:U cocrystal hydrate (**2**)	D:U cocrystal hydrate (**2**)	D:U cocrystal anhydrate (**3**)
D, T	D:T cocrystal hydrate (**5** and **10**)	D:T cocrystal hydrate (**5** and **10**)	D:T cocrystal anhydrate (**6**)
D, X	poorly crystalline, inconclusive	poorly crystalline, inconclusive	poorly crystalline, inconclusive
A, H	A:H cocrystal anhydrate (**8**) and hydrate (**9**)	A:H cocrystal anhydrate (**8**)	A:H cocrystal anhydrate (**8**)
C, H	physical mixture	physical mixture	physical mixture
U, H	physical mixture	physical mixture	physical mixture
G, U	physical mixture	physical mixture	physical mixture

Table 2. Results of selectivity experiments in the presence of competing nucleobases.

Reactants	Bulk water	Catalytic presence of water	Dry conditions
D, U, A	D:U cocrystal hydrate (**2**) + unreacted A	D:U cocrystal hydrate (**2**) + unreacted A	D:U cocrystal anhydrate (**3**) + unreacted A
D, T, A	D:T cocrystal hydrate (**5**) + unreacted A	D:T cocrystal hydrate (**5** and **10**) + unreacted A	D:T cocrystal anhydrate (**6**) + unreacted A
D, U, T	D:U cocrystal hydrate (**2**) + D:T cocrystal hydrate (**5**)	D:U cocrystal hydrate (**2**) + D:T cocrystal hydrate (**5** and **10**)	D:U cocrystal anhydrate (**3**) + D:T cocrystal anhydrate (**6**)
A, H, U	A:H cocrystal anhydrate (**8**) + unreacted U	A:H cocrystal anhydrate (**8**) + unreacted U	A:H cocrystal anhydrate (**8**) + unreacted U
A, H, T	A:H cocrystal anhydrate (**8**) + unreacted T	A:H cocrystal anhydrate (**8**) + unreacted T	A:H cocrystal anhydrate (**8**) + unreacted T

4.2. Mechanochemical prebiotic peptide synthesis

4.2.1. Amino acid oligomerization

Pristine Gly was milled at room temperature conditions and the product was analyzed by IP-HPLC (Figure 36a).[d] It showed that trace amount of 2,5-diketopiperazine (DKP) and Gly_2 have formed. After that, we milled Gly with five-fold molar excess of TiO_2. We used TiO_2 as a mineral additive because it is known to have good amino acid adsorption properties and is prebiotically relevant.[65] The excess of mineral additive was used to simulate prebiotic environments where it is regarded that organic material was present in limited amounts on mineral surfaces. Same is true for extraterrestrial matter, i.e., meteorites, albeit to a greater extreme. There, amino acids are present in ppm quantities in complex mineral matrix. In our case, milling Gly with TiO_2 at ambient conditions clearly showed the formation of Gly_2 and Gly_3, along with DKP and unreacted Gly (Figure 36b). Next, we conducted ball milling reactions with TiO_2 at temperatures ranging from ambient to 130 °C (Figure 37a-e), guided by recent literature examples for temperature limitation.[67,68] These have shown that increase in milling temperature results in the formation of higher Gly oligomers, along with the formation of DKP as the dominant component of the reaction mixture. The highest Gly oligomer was obtained by milling Gly with TiO_2 at 130 °C, Gly_{10} as detected by IP-HPLC (Figure 37e). The analysis by LC/MS (Q-TOF) confirmed that even Gly_{11} was obtained in the same reaction (Figure 38).[e]

[d] IP-HPLC analysis was performed jointly with collaborating co-author Mr. Saša Grubešić (Xellia Pharmaceuticals).
[e] LC/MS (Q-TOF) analysis was performed by collaborating co-author Mr. Saša Grubešić (Xellia Pharmaceuticals).

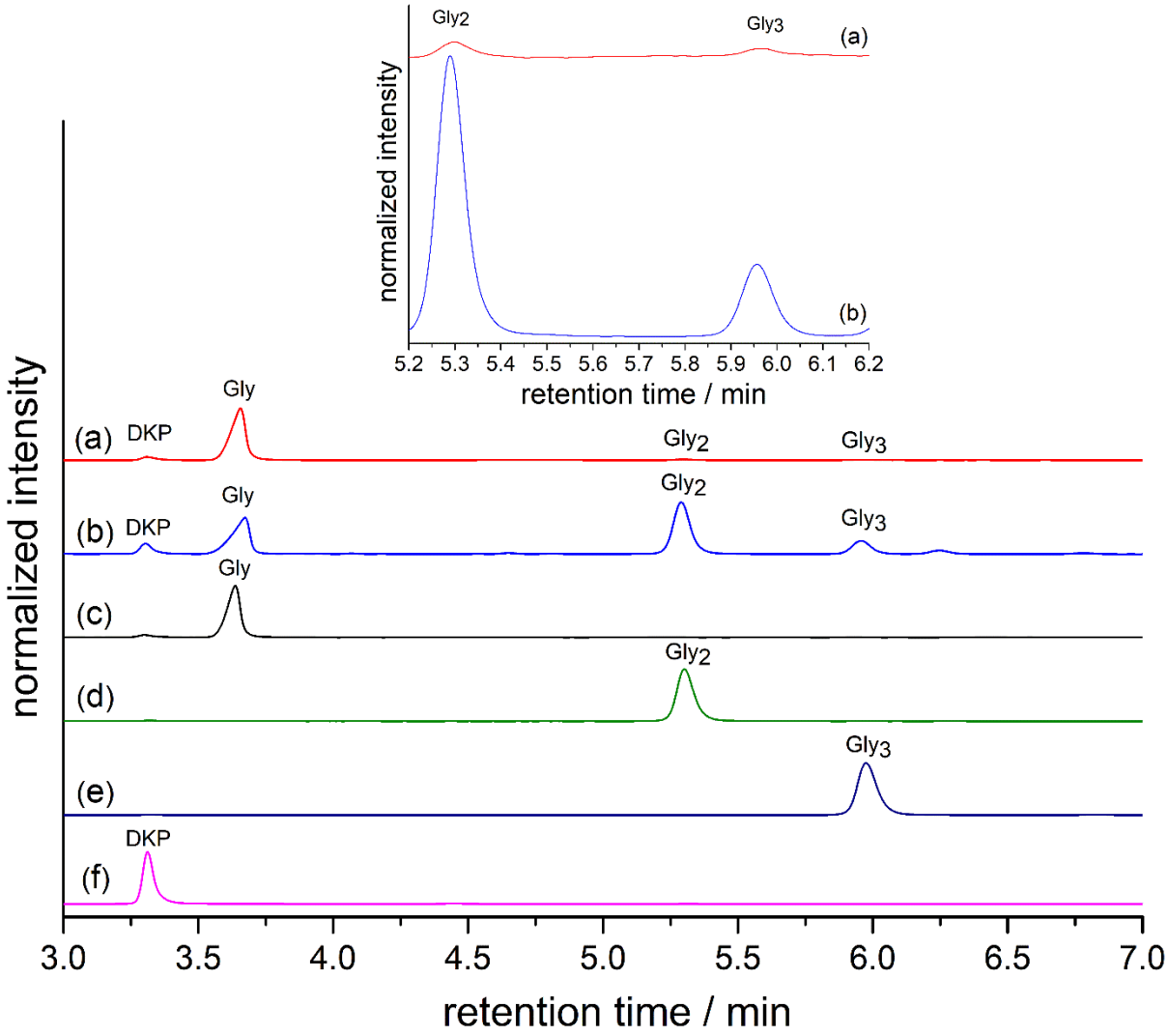

Figure 36. IP-HPLC chromatograms for: (a) pristine Gly milled at RT, (b) Gly milled with TiO_2 at RT, (c) Gly standard, (d) Gly_2 standard, (e) Gly_3 standard, and (f) DKP standard. The inset shows the zoomed region and reveals the presence of Gly_2 and Gly_3 in (a) and (b).

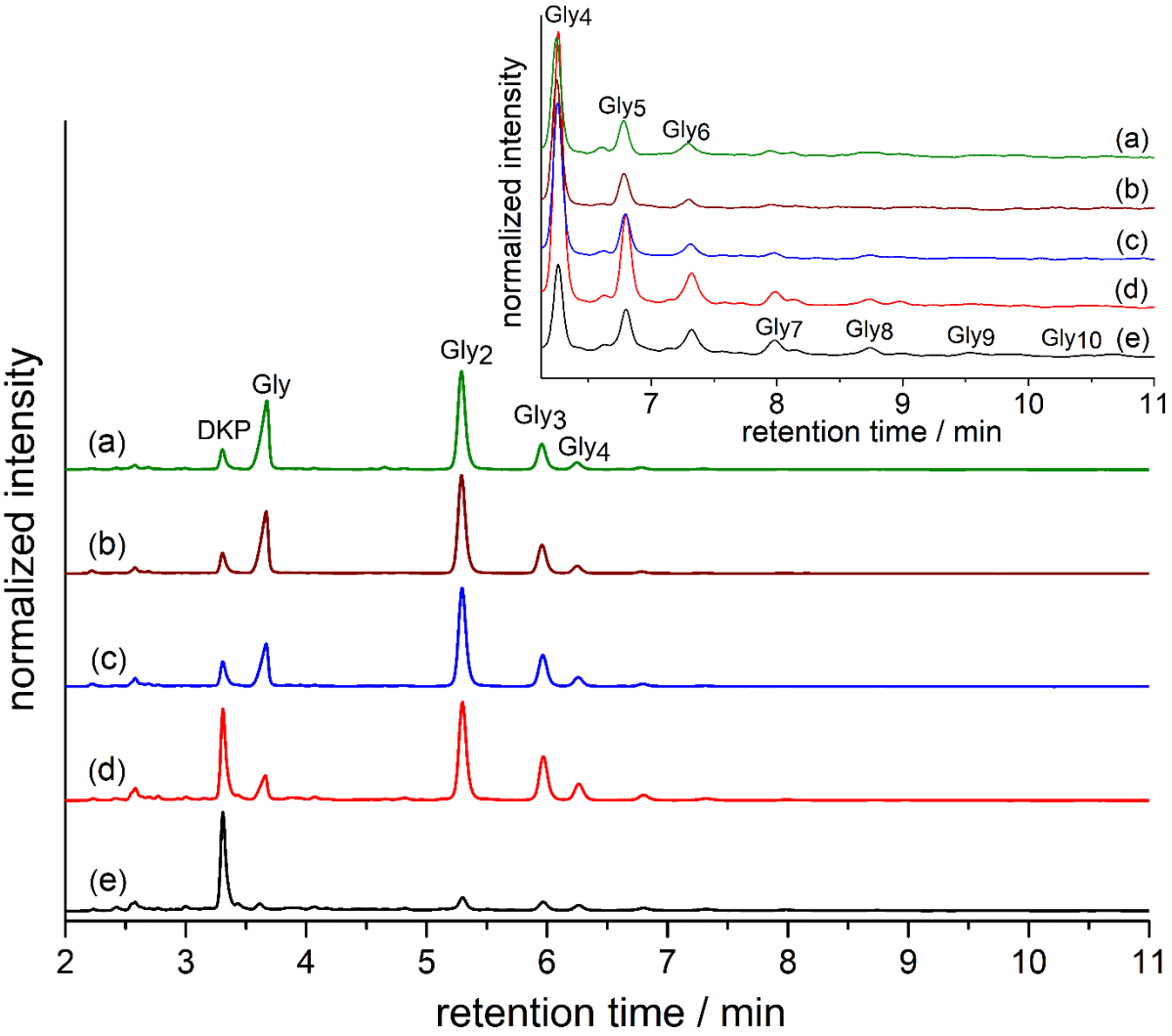

Figure 37. IP-HPLC chromatograms for Gly milled with TiO_2 at: (a) RT, (b) 40 °C, (c) 70 °C, (d) 100 °C, and (e) 130 °C. The inset shows the zoomed region of retention times in the 6.1–11 min range.

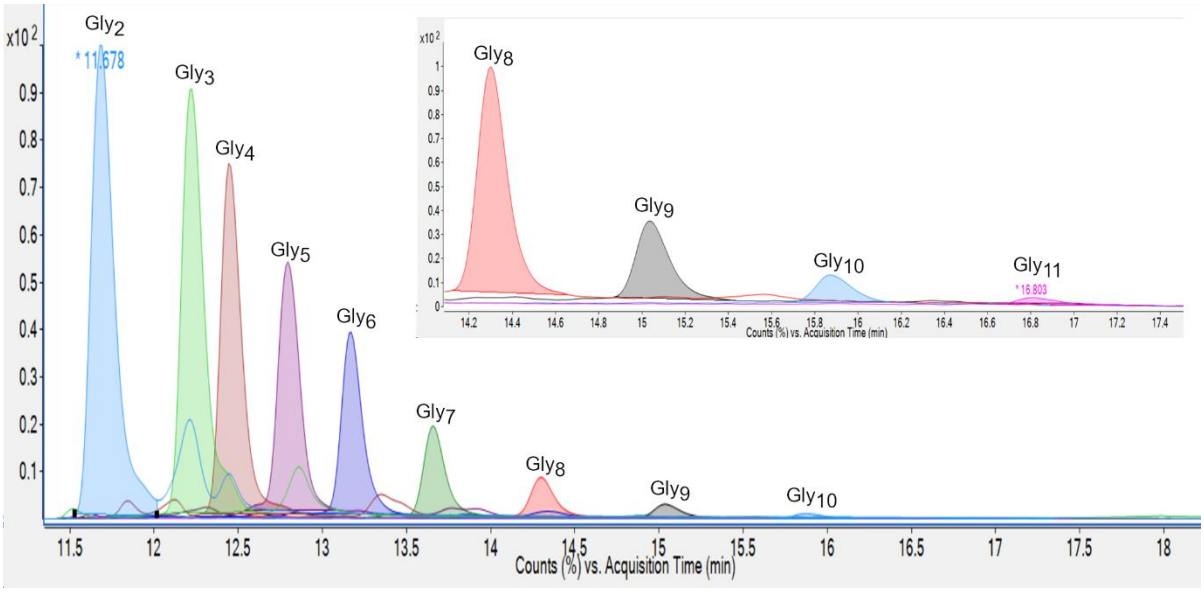

Figure 38. UPLC-MS analysis of Gly milled with TiO_2 at 130 °C. The inset shows the zoomed region in the acquisition time range from 14–17 min.

We also used IP-HPLC method to quantitatively determine the total yield of linear oligomers, according to the recent report (Table 3).[82] Total yield of linear oligomers was highest for Gly milled with TiO_2 at 100 °C, which had almost identical yield as Gly milled with TiO_2 at 70 °C. On one hand, these two temperatures showed that elevated temperature is needed to increase the conversion of Gly to linear oligomers. On the other hand, 130 °C exhibited lower linear Gly oligomer yield due to high temperature favoring Gly cyclization to DKP.

Table 3. Total yields of linear Gly oligomers for milling Gly with TiO_2 at different temperatures. [a] According to IP-HPLC. [b] According to UPLC-MS.

Ball milling temperature	**Longest detected oligomer[a]**	**Yield of linear oligomers**
ambient	Gly_6	6.7 %
40 °C	Gly_6	6.7 %
70 °C	Gly_8[b]	10.1 %
100 °C	Gly_8	10.2 %
130 °C	Gly_{11}[b]	8.5 %

Furthermore, we conducted experiments of milling Gly_2, Gly_3, and DKP with TiO_2 (Figure 39). These have shown that odd and even numbered oligomers have formed irrespective of the starting reactant, and that DKP is a productive intermediate in the mechanochemical prebiotic peptide bond formation.

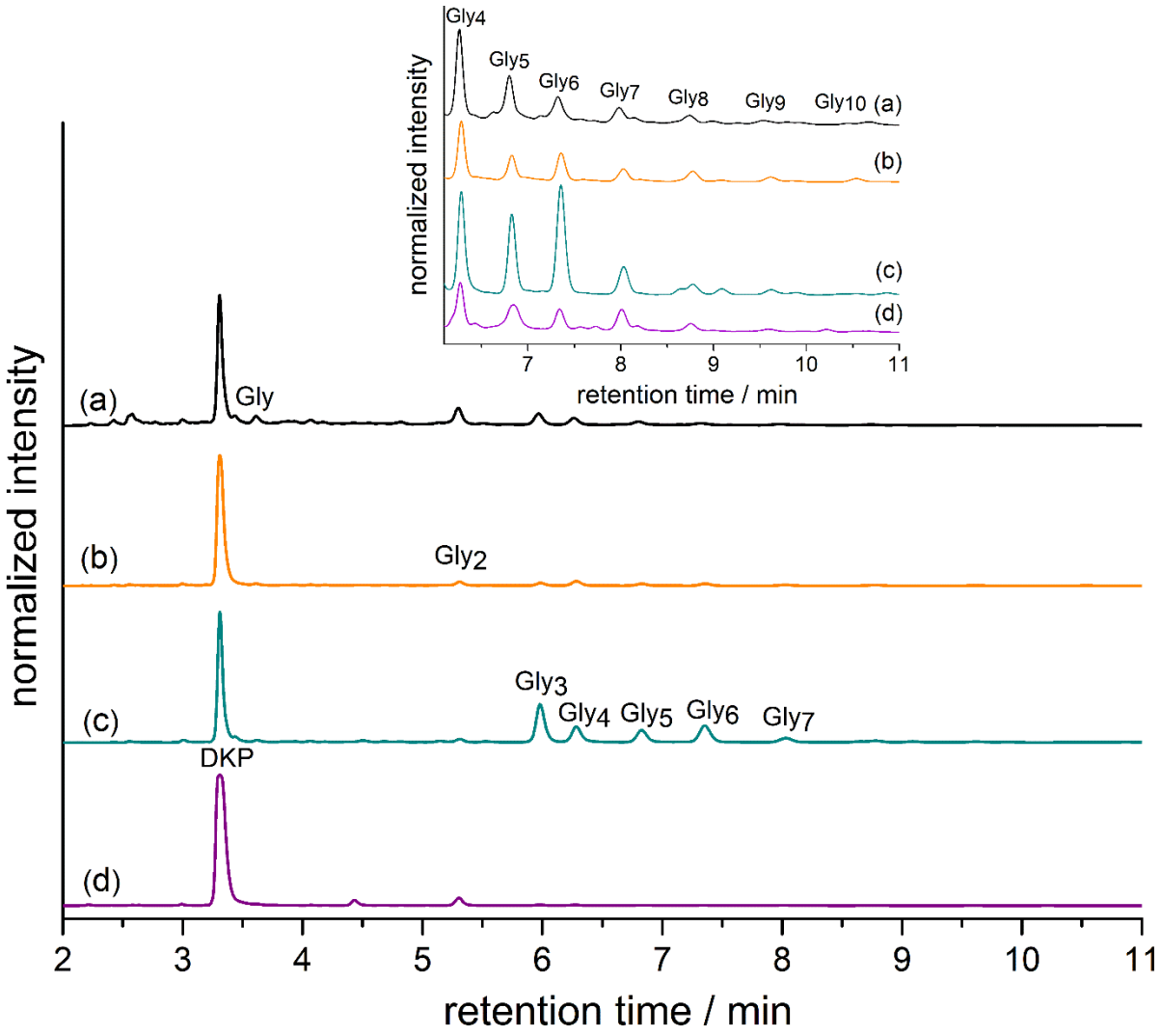

Figure 39. IP-HPLC chromatograms for: (a) Gly milled with TiO_2 at RT, (b) Gly_2 milled with TiO_2 at 130 °C, (c) Gly_3 milled with TiO_2 at 130 °C, and (d) DKP milled with TiO_2 at RT. The inset shows the zoomed region of retention times in the 6.1–11 min range.

Besides TiO_2, we tested other mineral additives, namely SiO_2, mica, and montmorillonite (Figures 7.19–7.21). All of them resulted in Gly oligomerization. LAG reactions with milling Gly in the presence of water at RT and 130 °C showed that mechanochemical prebiotic peptide bond formation proceeds in the presence of moisture (Figures 7.22 and 7.23).

To test if oligomers of other amino acids oligomerize under mechanochemical conditions, we milled L-alanine (Ala) with TiO_2 at 130 °C. The analysis by IP-HPLC showed

that Ala_2 formed (Figure 7.24), while direct MS analysis showed that oligomers up to Ala_5 formed (Table 4). After that, we conducted milling Gly and Ala with TiO_2 at 100 °C to see if hetero oligomers might form under mechanochemical conditions. Indeed, IP-HPLC indicated that might be the case (Figure 7.25), which was confirmed by direct MS analysis (Table 4).[f]

Table 4. Summary of each oligomer detected by the direct injection to MS (Q-TOF).

			Ala milled at 130°C		Gly+Ala milled at 100°C	
Oligomers	Elementary formulae	Theoretical $[M+H]^+$ (m/z)	Experimental $[M+H]^+$ (m/z)	Δ ppm	Experimental $[M+H]^+$ (m/z)	Δ ppm
Gly	$C_2H_5NO_2$	76,0393			76,0395	2,63
Gly_2	$C_4H_8N_2O_3$	133,0608			133,0611	2,25
Gly_3	$C_6H_{11}N_3O_4$	190,0822			190,0825	1,58
Gly_4	$C_8H_{14}N_4O_5$	247,1037				
Gly_5	$C_{10}H_{17}N_5O_6$	304,1252				
Gly_6	$C_{12}H_{20}N_6O_7$	361,1466				
Gly_7	$C_{14}H_{23}N_7O_8$	418,1681				
Gly_8	$C_{16}H_{26}N_8O_9$	475,1896				
Gly_9	$C_{18}H_{29}N_9O_{10}$	532,2110				
Gly_{10}	$C_{20}H_{32}N_{10}O_{11}$	589,2325				
Gly_{11}	$C_{22}H_{35}N_{11}O_{12}$	646,2539				
Gly_{12}	$C_{24}H_{38}N_{12}O_{13}$	703,2754				
DKP(Gly)	$C_4H_6N_2O_2$	115,0502				
Ala	$C_3H_7NO_2$	90,0550	90,0554	4,44	90,0554	4,44
Ala_2	$C_6H_{12}N_2O_3$	161,0921	161,0924	1,86	161,0924	1,86
Ala_3	$C_9H_{17}N_3O_4$	232,1292	232,1298	2,58		
Ala_4	$C_{12}H_{22}N_4O_5$	303,1663	303,1675	3,96		
Ala_5	$C_{15}H_{27}N_5O_6$	374,2034	374,205	4,28		
DKP(Ala)	$C_6H_{10}N_2O_2$	143,0815	143,0817	1,40		
Gly-Ala	$C_5H_{10}N_2O_3$	147,0764			147,0767	2,04
Ala-Ala-Gly	$C_8H_{15}N_3O_4$	218,1135			218,1136	0,46
Gly-Gly-Ala	$C_7H_{13}N_3O_4$	204,0979			204,098	0,49

[f] Direct MS Q-TOF analysis was performed by collaborating co-author Mr. Saša Grubešić (Xellia Pharmaceuticals).

We further set out to explore some of the critical reaction parameters that represent different geological and environmental scenarios to make them more universally plausible. These are the effects of mechanical loading and time. The transfer of mechanical load needed to initiate oligomerization is a crucial factor for mechanochemical peptide bond formation. Molecular collisions in ball mills are typically in the range of m/s,[145,146] and although they have much milder mechanical loading when compared to impact events which are in the km/s range, their relevance for prebiotic environments is hard to predict given the uncertainty and frequency of their appearance over time. On that note, subaerial geomorphological processes such as erosion and weathering are known to occur frequently and are capable of exchanging matter on the planetary surface.[147] Therefore, it is of interest to decouple the effect of solid-state mixing from that of mechanical activation in the solid-state.[148]

We simulated different mechanical loading that is transferred to the reaction mixture by using 5 Hz, 15 Hz, and 30 Hz milling frequencies. We propose that low milling frequency of 5 Hz might simulate the erosion of land and subaerial weathering. At 5 Hz, typical sound coming from the collision of milling balls is absent, the reaction mixture is gently shaken, and homogenization is limited. Intermediate milling frequency of 15 Hz might simulate the conditions that arise during earthquakes.[149] Highest milling frequency of 30 Hz might best simulate post-impact conditions of extraterrestrial bodies falling to Earth surface or rocks hitting the ground during heavy winds and volcanic eruptions. Our results show that the highest amounts of peptides are formed when milling is conducted at 30 Hz. The dependance of the total yield of linear Gly oligomers that are formed under these conditions on the reaction time are presented in Figure 40. The comparison of the effect of all three milling frequencies on mechanochemical peptide bond formation are summarized in Table 5 (Figures 7.26–7.37). IP-HPLC and LC/MS (Triple quadrupole)[g] analysis showed that 5 and 15 Hz frequencies enable peptide bond formation. The longest Gly oligomers were obtained by milling Gly with SiO_2 at room temperature for 48 h using 30 Hz frequency, Gly_8 (Table 5). Furthermore, milling Gly with TiO_2 using three stones for 18 h at 20 Hz and ambient temperature resulted in the formation of oligomers up to Gly_4 (Figure 41). Importantly, milling in N_2 and CO_2 atmospheres showed

[g] LC/MS (Triple quadrupole) analysis was performed by collaborating co-author Mr. Saša Grubešić (Xellia Pharmaceuticals).

that mechanochemical peptide bond formation proceeds in the presence of these two gases which were likely the main components of prebiotic Earth atmosphere,[150] although it gives a more complex mixture of products (Figures 7.39–7.42). For example, there is a high intensity peak at 6.1 min of retention time which we did not identify. Nevertheless, LC/MS (Triple quadrupole) analysis identified the presence of Gly_6 and Gly_5 when reaction was performed in N_2 (Figure 7.40) and CO_2 (Figure 7.42) atmospheres, respectively.

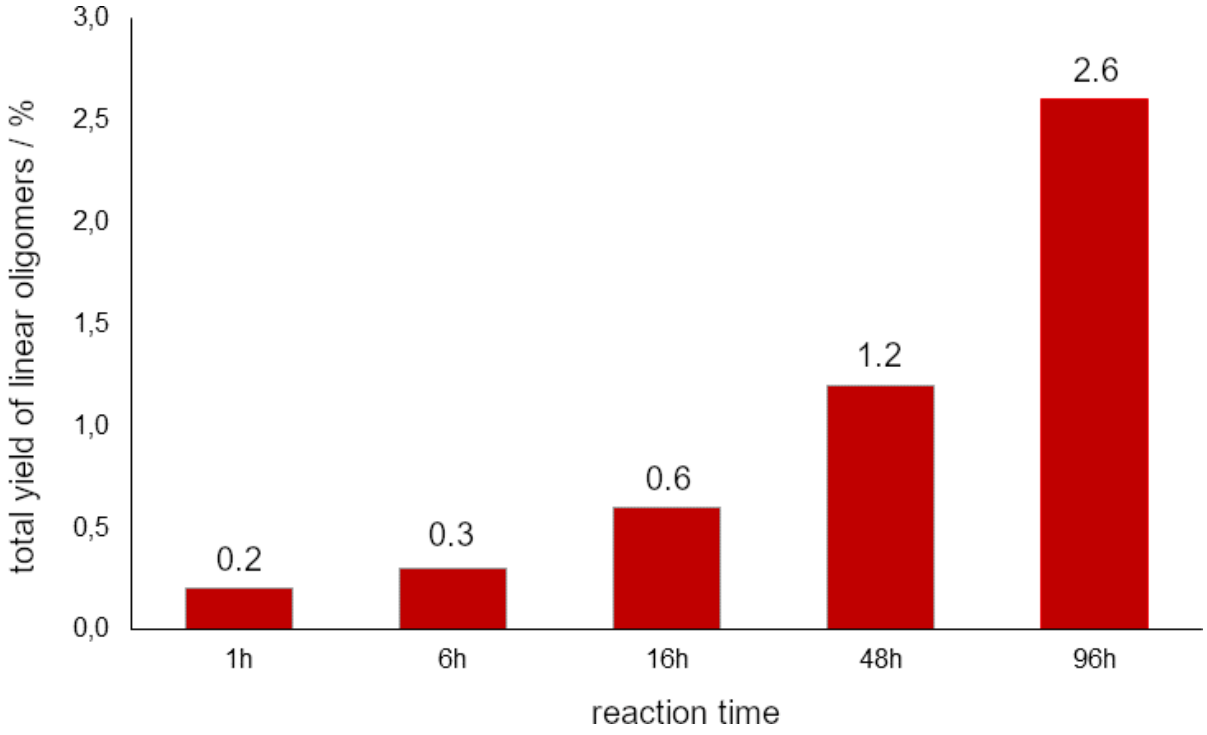

Figure 40. The effect of reaction time for milling Gly with SiO_2 at 40 °C on total yield of linear Gly oligomers.

Table 5. The longest glycine oligomers detected by: [a] IP-HPLC and [b] LC/MS (Triple quadrupole) for Gly milled with SiO_2 at room temperature.

	5 Hz	**15 Hz**	**30 Hz**
48 h	[a] Gly_3 [b] Gly_2	[a] Gly_4 [b] Gly_3	[a] Gly_8 [b] Gly_8
96 h	[a] Gly_2 [b] Gly_3	[a] Gly_3 [b] Gly_3	[a] Gly_7 [b] Gly_4

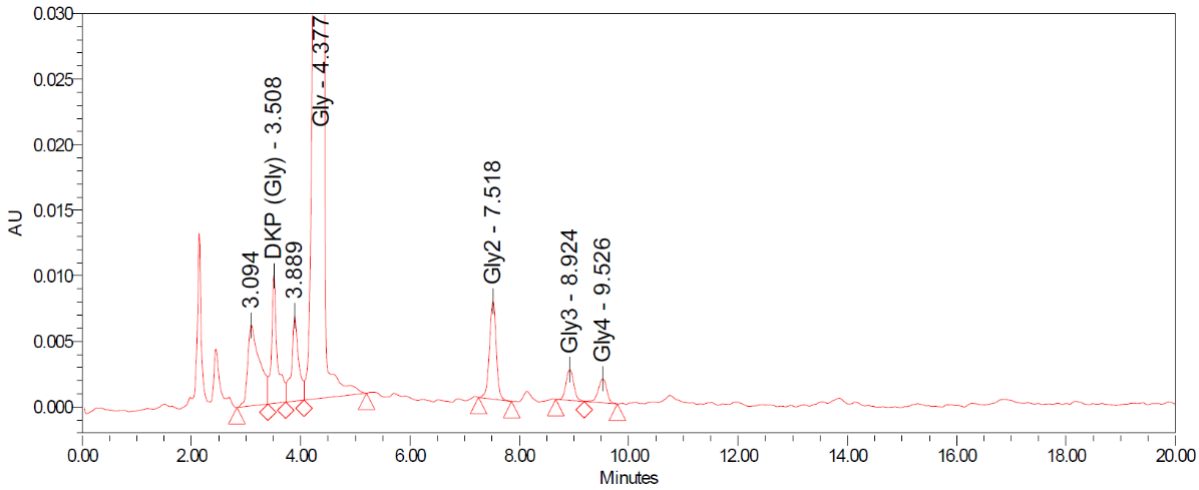

Figure 41. IP-HPLC chromatogram for milling Gly with TiO_2 using three hand-picked stones (Figure 7.38) for 18 h at 20 Hz and ambient temperature.

After studying the effects of mechanical loading, time, temperature, and atmosphere, we turned to oligomerization on complex mineral surfaces. We have selected different mineral classes including silicates (quartz, talc), sulfides (pyrite), sulphates (gypsum), phosphates (apatite), carbonates (calcite, magnesite), micas (muscovite), and zeolites (analcime, clinoptilolite) (Figure 42). The elemental composition of minerals was verified by X-ray fluorescence[h] (Table 6) and PXRD analysis combined with Rietveld analysis (Figures 7.43–7.52). Milling reactions were performed with the excess of minerals under the same molar ratio as those with TiO_2 and SiO_2, namely 5:1. We used 30 Hz frequency, two stainless steel balls as milling media, room temperature, and experiments were performed for 16 h. Reaction products were analyzed by IP-HPLC and LC/MS triple quadrupole (Figures 7.53–7.71). The results of these experiments are summarized in Table 7. Among all tested minerals, quartz yielded the longest oligomer detected by LC/MS triple quadrupole, Gly_7. Notably, unidentified peak at 6.3 min of retention time is present in all milling reactions with minerals alongside glycine

[h] XRF analysis of minerals elemental composition was performed by Dr. Ivan Nemet (Chemistry Department, Faculty of Science, University of Zagreb).

oligomers (same as in Figures 7.40 and 7.42 showing chromatograms for milling glycine with SiO_2 in N_2 and CO_2 atmospheres).

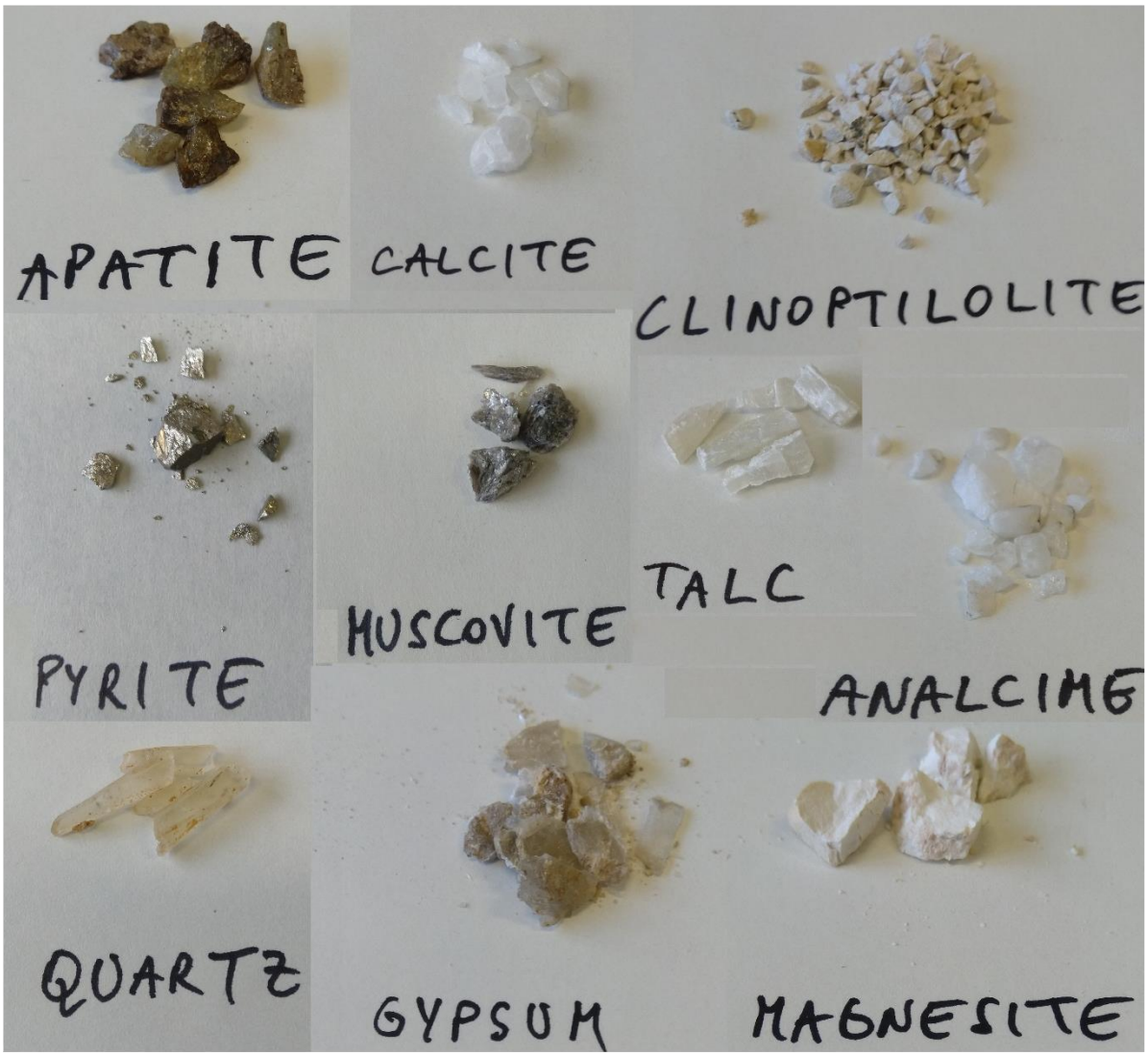

Figure 42. Mineral samples that were used in this study.

Table 6. Information about minerals origin and elemental composition as determined by X-ray fluorescence (XRF). Minerals were provided by Prof. Nenad Tomašić (Department of Geology, University of Zagreb).

Name	Formula	Method	*w* / %	Origin
analcime	$NaAlSi_2O_6(H_2O)$	XRF	SiO_2 55,26 Al_2O_3 22,40 FeO 0,07 CaO 3,12 MnO 0,06	Tirol, Austria
apatite	$Ca_5(PO_4)_3$	XRF	CaO 54,51 P_2O_5 39,40 FeO 0,196 MnO 0,06	Bancroft, Canada
calcite	$CaCO_3$	XRF	CaO 55,52 FeO 0,02 MnO 0,03 SiO_2 0,05	Trepča, Kosovo
clinoptilolite	$Ca_{1.9}Na_{1.76}K_{1.05}Mg_{0.17}Al_{6.72}Si_{29.2}O_{72}$	XRF	SiO_2 68,75 Al_2O_3 12,11 FeO 1,362 CaO 3,92 K 1,27 MnO 0,02	Serbia
gypsum	$CaSO_4$ x $2(H_2O)$	XRF	CaO 31,90 SO_3 46,1 SiO_2 0,14	Mt. Paderno, Italy
magnesite	$MgCO_3$	XRF	MgO 47,60 K_2O 0,06 MnO 0,07 FeO 0,02 SiO_2 0,03	Goleš, Kosovo
muscovite	$KAl_2(Si_3Al)O_{10}(OH,F)_2$	XRF	SiO_2 46,12 Al_2O_3 37,80 K_2O 10,50 MnO 0,02	Čanište, Macedonia
pyrite	FeS_2	XRF	Fe 45,61 S 52,22 SiO_2 0,05 MnO 0,06	Trepča, Kosovo
talc	$Mg_3Si_4O_{10}(OH)_2$	XRF	SiO_2 60,95 MgO 31,50 Fe_2O_3 1,163 CaO 0,222	Rabenwald, Austria
quartz	SiO_2	XRF	SiO_2 95,5	Trepča, Kosovo

Table 7. Summary of reactions where Gly was milled with different minerals.

Mineral	Longest detected oligomer by LC/MS (triple quadrupole)
analcime	Gly_2
apatite	Gly_3
calcite	Gly_4
clinoptilolite	Gly_2
gypsum	Gly_2
magnesite	Gly_4
muscovite	Gly_2
pyrite	Gly_3
talc	Gly_3
quartz	Gly_7

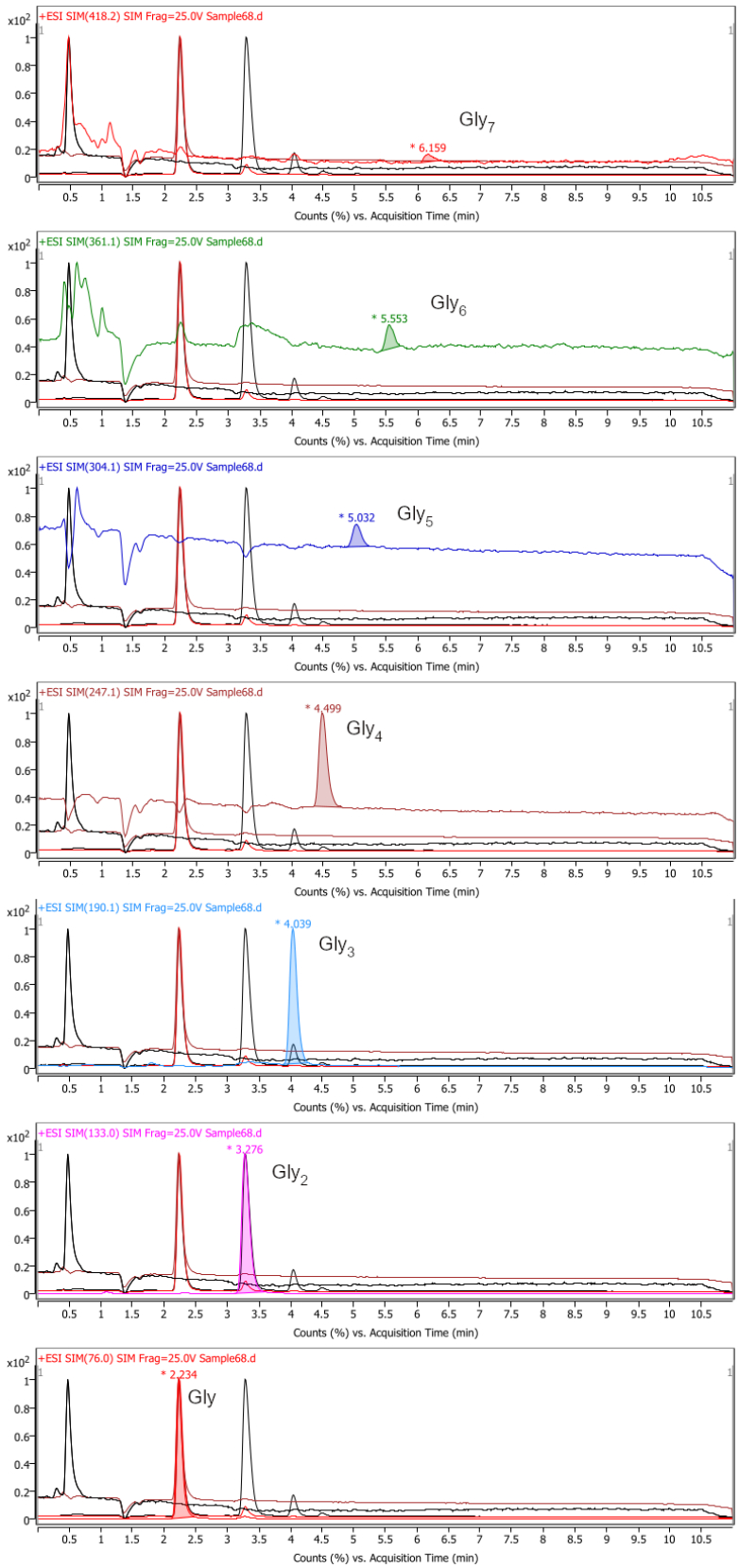

Figure 43. LC/MS (triple quadrupole) chromatograms for milling Gly with quartz at room temperature for 16 h and at 30 Hz.

4.2.2. Amino acid and α-hydroxy acid oligomerization

Together with amino acids, among the soluble organic material that is found in meteorites are also α-hydroxy acids. Therefore, investigation of their mutual interaction and the production of prebiotic oligomers is desired.[151,152] First, we milled glycolic acid (GA) with SiO_2 at room temperature, for 16 h, and at 30 Hz. The analysis by IP-HPLC showed several peaks which indicated that we might have synthesized ester bonds (Figure 7.72). Indeed, MALDI TOF analysis[i] confirmed that oligomers up to four glycolic acid residues have formed under these conditions (Figure 44 and Table 8).

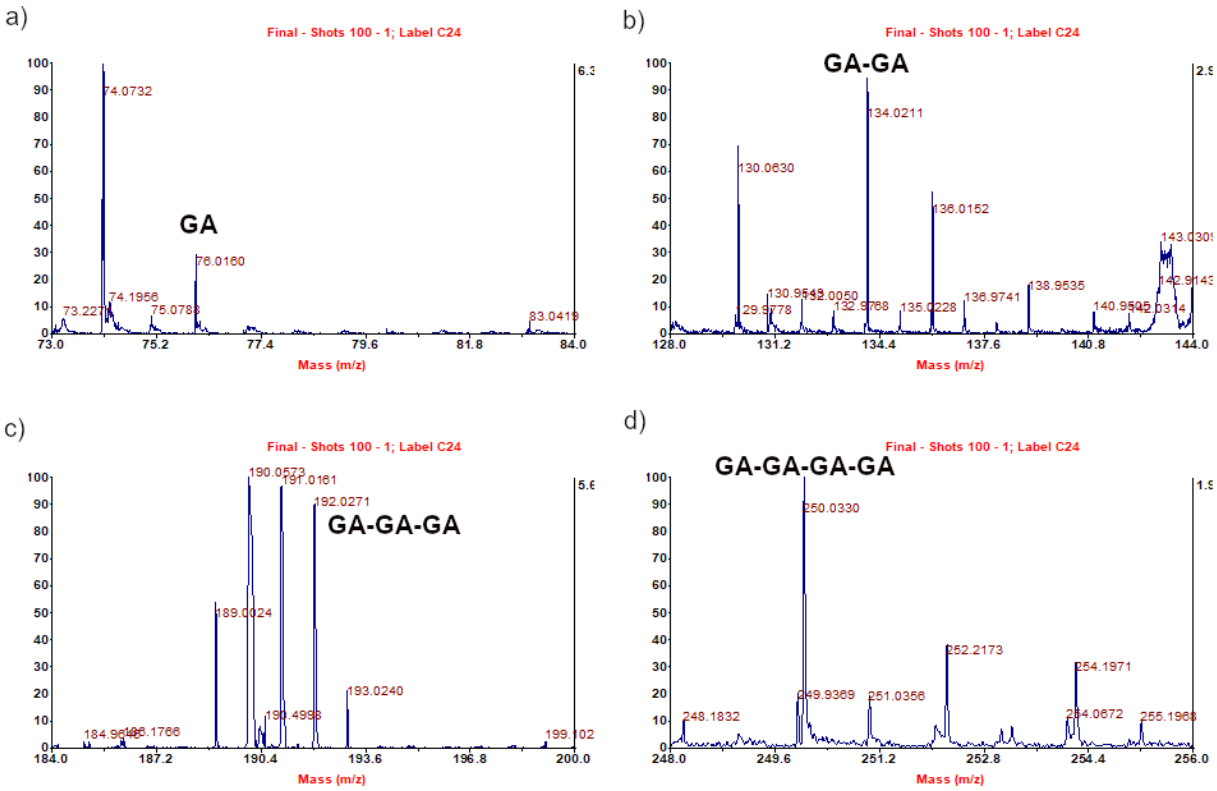

Figure 44. MALDI TOF spectra of the reaction product after milling glycolic acid with SiO_2 at room temperature, at 30 Hz, and for 16 h. Showing: a) glycolic acid, b) glycolic acid dimer, c) glycolic acid trimer, d) glycolic acid tetramer.

[i] MALDI TOF analysis was performed by collaborating co-author Dr. Mario Cindrić (Ruđer Bošković Institute).

Table 8. Summary of oligomers that were detected by MALDI TOF after milling glycolic acid with SiO_2 at room temperature, for 16 h, and at 30 Hz. There is no protonation due to the formation of oxazolone ion.

Product	Elementary formulae	Theoretical (m/z)	Experimental (m/z)	Δ ppm
1GA	$C_2H_4O_3$	76.0160	76.0160	0.00
2GA	$C_4H_6O_5$	134.0215	134.0211	-2.98
3GA	$C_6H_8O_7$	192.0270	192.0271	0.52
4GA	$C_8H_{10}O_9$	250.0325	250.0330	2.00

Encouraged by the formation of oligomers with ester bonds, glycine and glycolic acid were milled with SiO_2 at room temperature, for 16 h, and at 30 Hz. IP-HPLC chromatogram showed numerous peaks which indicated the rich reactivity under these reaction conditions (Figure 7.73). This was confirmed by MALDI TOF analysis which showed the formation of at least five different oligomers containing peptide and ester bonds (Figure 45). Furthermore, Table 9 contains detailed list of theoretical and experimentally observed molecular masses and Figure 46 shows molecular structures of products that were detected.

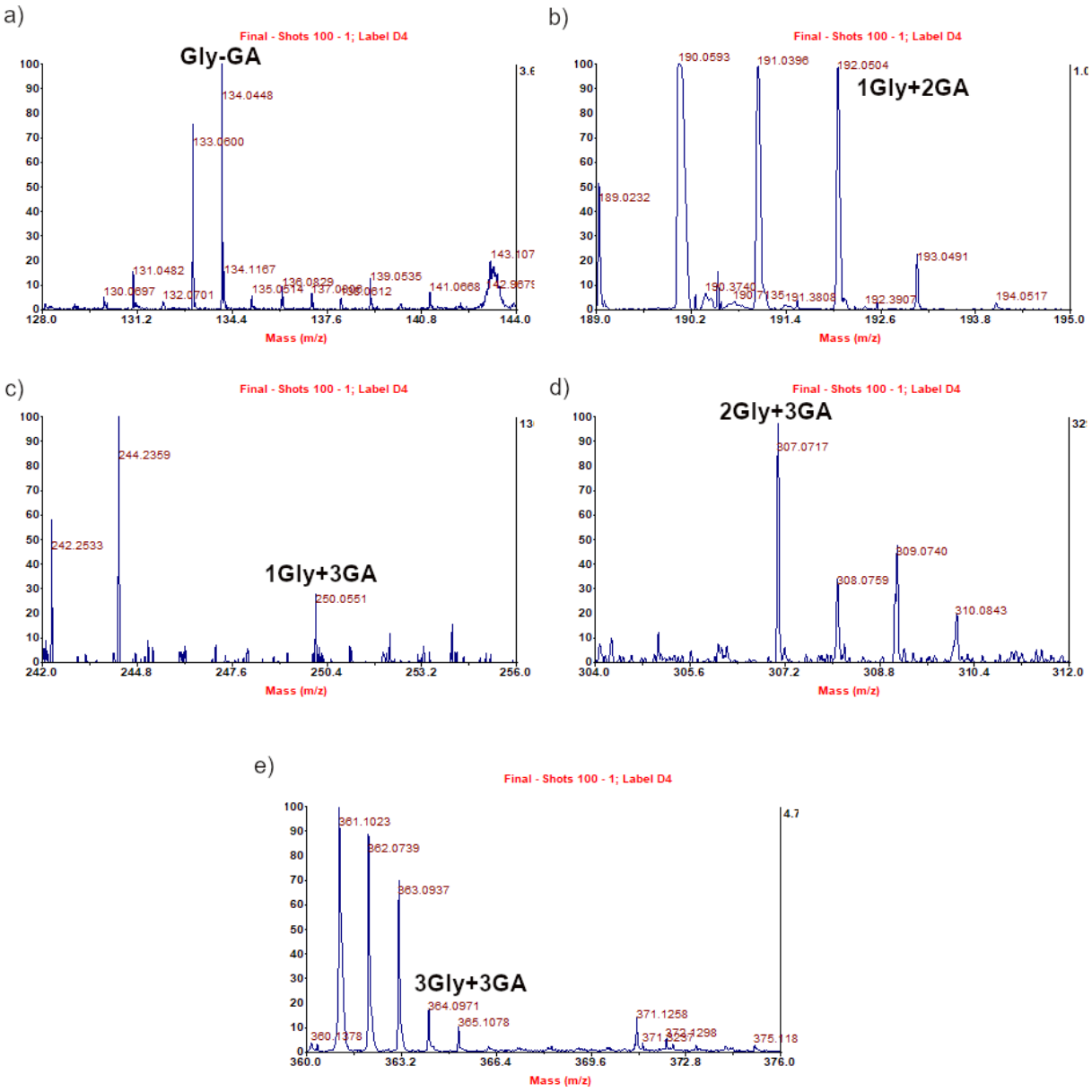

Figure 45. MALDI TOF spectra of the reaction product after milling glycine and glycolic acid with SiO_2 at room temperature, at 30 Hz, and for 16 h. Showing: a) glycine and glycolic acid dimer, b) glycine and glycolic acid trimer, c) glycine and glycolic acid tetramer, d) glycine and glycolic acid pentamer, e) glycine and glycolic acid hexamer.

Table 9. Summary of oligomers that were detected by MALDI TOF after milling glycine and glycolic acid with SiO_2 at room temperature, for 16 h, and at 30 Hz.

Product	Elementary formulae	Theoretical $[M+H]^+$ (m/z)	Experimental $[M+H]^+$ (m/z)	Δ ppm
1Gly+1GA	$C_4H_7NO_4$	134.0453	134.0448	-3.73
1Gly+2GA	$C_6H_9NO_6$	192.0503	192.0504	0.52
1Gly+3GA	$C_8H_{11}NO_8$	250.0558	250.0551	-2.80
2Gly+3GA	$C_{10}H_{14}N_2O_9$	307.0727	307.0717	-3.26
3Gly+3GA	$C_{12}H_{17}N_3O_{10}$	364.0987	364.0971	-4.39

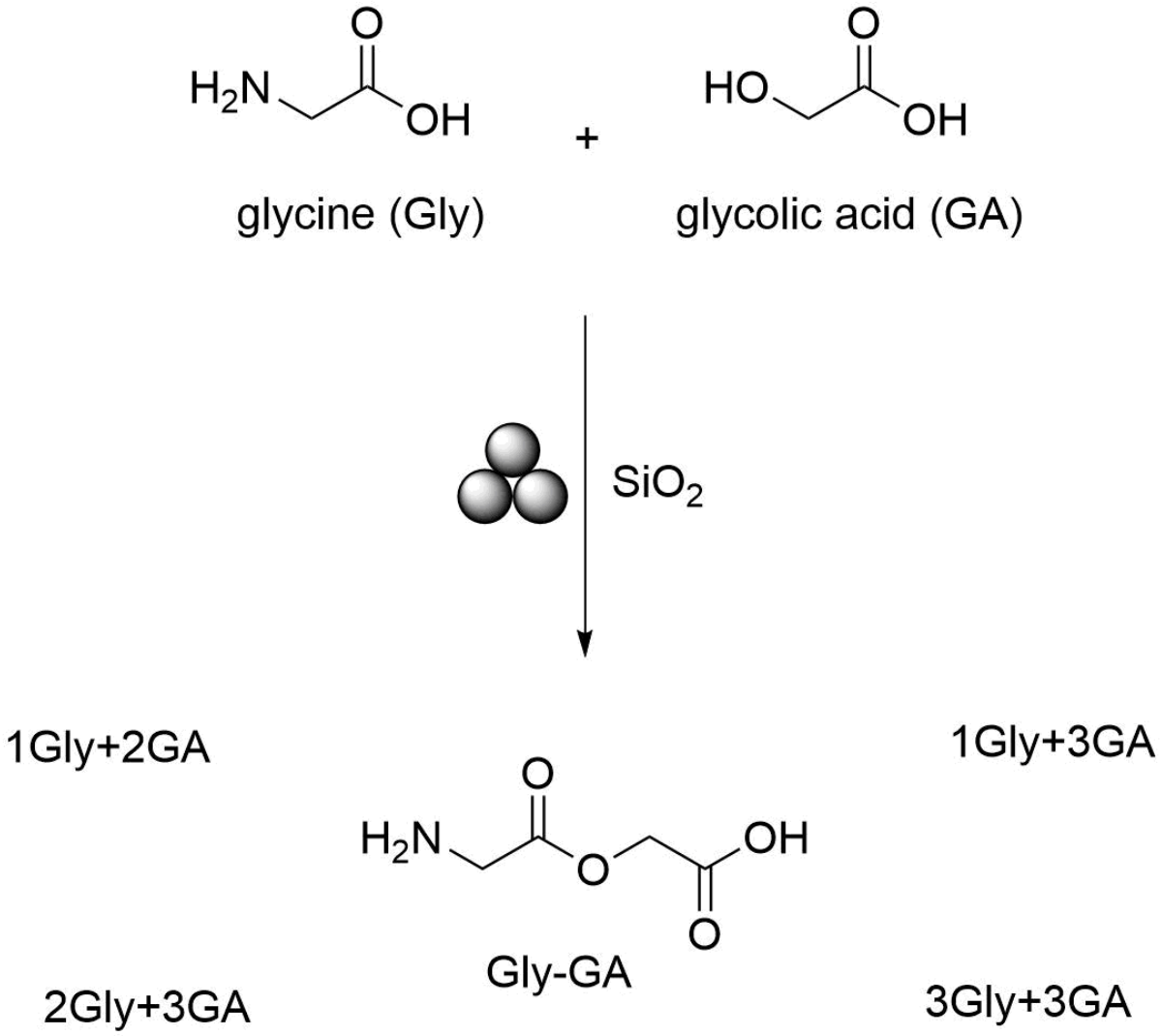

Figure 46. Reaction scheme for milling glycine and glycolic acid with SiO_2 at room temperature, for 16 h, and at 30 Hz. Represented oligomers were detected by MALDI TOF (Figure 45 and Table 9).

§ 5. CONCLUSION

Our results show that some of the molecular recognition patterns present in contemporary biochemistry are available on a nucleobase level under prebiotic conditions. Molecular recognition of 2,6-diaminopurine (D) and uracil (U), D and thymine (T), and adenine (A) and hypoxanthine (H) is selective and occurs under all reaction conditions that were tested: bulk water, catalytic presence of water, and in the absence of water (in the solid-state). The crystal structures of D:U cocrystal hydrate and D:T cocrystal hydrate were solved from the single-crystal X-ray diffraction data. The crystal structures of D:T cocrystal anhydrate, A:H cocrystal hydrate, and A:H cocrystal anhydrate were solved from the laboratory powder X-ray diffraction data. Interestingly, to self-assemble with A in both A:H cocrystal hydrate and anhydrate, H changes its tautomeric form (from naturally occurring 9H tautomer to 7H tautomer). Contrary to the literature fact that canonical nucleobase pairs (A:T, A:U, and G:C) do not self-assemble in water, we show that water molecules are enabling molecular recognition of D:U, D:T, and A:H nucleobase pairs. The latter is true even in the presence of competing nucleobases, e.g., U and T self-assemble with D in the presence of A, and A self-assembles with H in the presence of U and T. Since prebiotic synthetic pathways usually result in a mixture of nucleobases instead of a clean synthesis of a single nucleobase and multiple nucleobases are found in meteorites, it is reasonable to expect that they would have been in spatial proximity in an early Earth setting. For example, in impact craters where nucleobases would jointly diffuse from extraterrestrial bodies such as meteorites to the newly formed ponds. Therefore, it seems that supramolecular interactions between nucleobases might have been important for their chemical evolution. The influence of supramolecular interactions of organic building blocks of life on their chemical evolution is underrepresented in the literature. Most often, the propensity for the formation of covalent condensation bonds is taken as the criteria to assess whether a compound or a reaction pathway, is regarded as plausibly prebiotic. A synergistic approach based on the formation of covalent and non-covalent interactions may provide new insights into chemical evolution. For example, the existence of different tautomeric forms of nucleobases (or other organic building blocks) in the solid-state might offer different reactivity and selectivity in covalent bond-forming reactions such as glycosylation or phosphorylation. Besides, UV irradiation is considered to have been one of the main selection pressures on early Earth. Therefore,

supramolecular assemblies of nucleobases might have different photochemical properties when compared to individual nucleobases.

In the second part of the thesis, mechanical activation of amino acids with TiO_2 and SiO_2 enabled their oligomerization. For the analysis of the reaction products, we relied on ion-pairing high-performance liquid chromatography and mass spectrometry (Q-TOF, triple quadrupole, MALDI TOF). The longest glycine (Gly) oligomer that was obtained from the starting Gly was Gly_{11}. Milling Gly with TiO_2 (at 100 °C, 30 Hz, three 1 g stainless steel balls, for 18 h) resulted in the highest total yield of linear Gly oligomers, 10.2 %. Liquid-assisted grinding experiments with H_2O showed that mechanochemical peptide bond formation tolerates the presence of moisture. Reactions that started from Gly_2, Gly_3, and DKP showed that mechanochemical peptide bond formation is a dynamic process with simultaneous breaking and making of peptide bonds. Importantly, milling Gly and L-alanine (Ala) resulted in the formation of their hetero oligopeptides thus demonstrating that other amino acids can be incorporated into a growing peptide chain. It was shown that time, temperature, atmosphere, and mechanical loading are all important parameters for mechanochemical prebiotic peptide synthesis. Even milling at low frequencies and using stones as milling media resulted in the formation of peptides. It was also shown that glycolic acid (GA) oligomerizes under mechanochemical conditions. Milling GA with SiO_2 (at room temperature, 30 Hz, two 1.4 g stainless steel balls, for 16 h) resulted in the formation of an oligo-ester containing four residues of GA. Furthermore, milling Gly and GA with SiO_2 (at room temperature, 30 Hz, two 1.4 g stainless steel balls, for 16 h) resulted in the formation of at least five hetero oligomers containing peptide and ester bonds. This work demonstrated that the formation of peptides is operational under different geological and environmental scenarios in a solid-state reaction medium simulating dry prebiotic environment. These results may have implications for astrobiology research and reconsideration of extra-terrestrial conditions that can give rise to peptides. For example, weathering of amino acids in dry environments might result in the formation of peptides on mineral surfaces. Once formed, peptides might have catalyzed other important prebiotic reactions such as sugar syntheses. On the other hand, mechanochemical prebiotic peptide bond formation opens the possibility for the sustainable syntheses of modern peptides and amides by mechanochemistry.

§ 6. REFERENCES

1. L. E. Orgel, *Crit. Rev. Biochem. Mol. Biol.* **39** (2004) 99–123.
2. A. I. Oparin, *Proischogdenie Zhizni*, Moscovsky Robotchii, Moscow, 1924.
3. Eschenmoser, *Tetrahedron* **63** (2007) 12821−12844.
4. G. Wächtershäuser, *Microbiol Rev.* **52** (1988) 452–484.
5. Damer, D. Deamer, *Astrobiology* **20** (2020) 429–452.
6. S. Ross, D. Deamer, *Life* **6** (2016) 28.
7. M. Marshall, *Nature* **588** (2020) 210–213.
8. B. Damer, D. Deamer, *Life* **5** (2015) 872–887.
9. B. K. Pearce, R. E. Pudritz, D. A. Semenov, T. K. Henning, *Proc. Natl. Acad. Sci. U.S.A.* **114** (2017) 11327–11332.
10. M. J. V. Kranendonk, R. Baumgartner, T. Djokic, T. Ota, L. Steller, U. Garbe, E. Nakamura, *Astrobiology* **21** (2021) 39–59.
11. H. J. Cleaves II, A. Michalkova Scott, F. C. Hill, J. Leszczynski, N. Sahai, R. Hazen, *Chem. Soc. Rev.* **41** (2012) 5502–5525.
12. K. Kvenvolden, J. Lawless, K. Pering, E. Peterson, J. Flores, C. Ponnamperuma, I. R. Kaplan, C. Moore, *Nature* **228** (1970) 923–926.
13. M. P. Callahan, K. E. Smith, H. J. Cleaves II, J. Ruzicka, J. C. Stern, D. P. Glavin, C. H. House, J. P. Dworkin, *Proc. Natl. Acad. Sci. U.S.A.* **108** (2011) 13995–13998.
14. Y. Furukawa, Y. Chikaraishi, N. Ohkouchi, N. O. Ogawa, D. P. Glavin, J. P. Dworkin, C. Abe, T. Nakamura, *Proc. Natl. Acad. Sci. U.S.A.* **116** (2019) 24440–24445.
15. P. Schmitt-Kopplin, Z. Gabelica, R. D. Gougeon, A. Fekete, B. Kanawati, M. Harir, I. Gebefuegi, G. Eckel, N. Hertkorn, *Proc. Natl. Acad. Sci. U.S.A.* **107** (2010) 2763–2768.
16. Chyba, C. Sagan, *Nature* **355** (1992) 125–132.
17. A. J. Brearley, *Meteor. Early Sol. Syst. II* (2006) 587–624.
18. G. R. Huss, A. E. Rubin, J. N. Grossman, *Meteor. Early Sol. Syst. II* (2006) 567–586.
19. V. E. Hamilton, C. A. Goodrich, A. H. Treiman, H. C. Connolly Jr, M. E. Zolensky, M. H. Shaddad, *Nat Astron* **5** (2020) 350–355.
20. C. J. Bennett, C. Pirim, T. M. Orlando, *Chem. Rev.* **113** (2013) 9086−9150.
21. T. Okada *et al.*, *Nature* **579** (2020) 518–522.

22. A. A. Simon *et al.*, *Science* **370** (2020) eabc3522.
23. T. Stolar, S. Grubešić, N. Cindro, E. Meštrović, K. Užarević, J. G. Hernández, *Angew. Chem. Int. Ed.* **60** (2021) 12727–12731.
24. S. A. Benner, H.-J. Kim, M. A. Carrigan, *Acc. Chem. Res.* **45** (2012) 2025–2034.
25. J. D. Sutherland, *Nat Rev Chem* **1** (2017) 1–7.
26. M. Yadav, R. Kumar, R. Krishnamurthy, *Chem. Rev.* **120** (2020) 4766–4805.
27. M. A. Pasek, *Chem. Rev.* **120** (2020) 4690–4706.
28. Preiner, M. et al. *Life* **10** (2020) 1–25.
29. A. C. Rios, Y. Tor, *Isr. J. Chem.* **53** (2013) 1–15.
30. M. Fialho, T. P. Roche, N. V. Hud, *Chem. Rev.* **120** (2020) 4806–4830.
31. S. Miyakawa, H. J. Cleaves, S. L. Miller, *Orig Life Evol Biosph* **32** (2002) 209–218.
32. R. Saladino, C. Crestini, F. Ciciriello, G. Costanzo, E. Di Mauro, *Chem. Biodivers.* **4** (2007) 694–720.
33. C. Menor-Salván, M. R. Marín-Yaseli, *Chem. Soc. Rev.* **41** (2012) 5404–5415.
34. C. Menor-Salvan, M. R. Marín-Yaseli, *Chem. Eur. J.* **19** (2013) 6488–6497.
35. I. Jiménez, C. Gibard, R. Krishnamurthy, *Angew. Chem. Int. Ed.* **60** (2021) 2–11.
36. Z. Martins, *et al. Earth Planet. Sci. Lett.* **270** (2008) 130–136.
37. C. K. Materese, M. Nuevo, S. A. Sandford, *Astrobiology* **17** (2017) 761–770.
38. Y. Oba, Y. Takano, H. Naraoka, N. Watanabe, A. Kouchi, *Nat. Commun.* **10** (2019) 4413.
39. N. V. Hud, B. J. Cafferty, R. Krishnamurthy, L. D. Williams, *Chem. & Biol.* **20** (2013) 466–474.
40. P. O. P. Ts'o, *New York: Acad. Press.* **I** (1974) 453–584.
41. P. Cieplak, P. A. Kollman, *J. Am. Chem. Soc.* **110** (1988) 3734–3739.
42. E. Stofer, C. Chipot, R. Lavery, *J. Am. Chem. Soc.* **121** (1999) 9503–9508.
43. B. A. Anderson, K. Fauché, S. C. Karunakaran, J. R. Yerabolu, N. V. Hud, R. Krishnamurthy, *Chem. Eur. J.* **27** (2021) 4033-4042.
44. M. D. Kirnos, I. Y. Khudyakov, N. I. Alexandrushkina, B. F. Vanyushin, *Nature* **270** (1977) 369–370.
45. Y. Khudyakov, M. D. Kirnos, N. I. Alexandrushkina, B. F. Vanyushin, *Virology* **18** (1978) 8–18.

46. V. Pezo, F. Jaziri, P.-Y. Bourguignon, D. Louis, D. Jacobs-Sera, J. Rozenski, S. Pochet, P. Herdewijn, G. F. Hatfull, P.-A. Kaminski, P. Marliere, *Science* **372** (2021) 520–524.
47. D. Sleiman, P. S. Garcia, M. Lagune, J. Loc'h, A. Haouz, N. Taib, P. Röthlisberger, S. Gribaldo, P. Marlière, P. A. Kaminski, *Science* **372** (2021) 516–520.
48. Y. Zhou, X. Xu, Y. Wei, Y. Cheng, Y. Guo, I. Khudyakov, F. Liu, P. He, Z. Song, Z. Li, Y. Gao, E. L. Ang, H. Zhao, Y. Zhang, S. Zhao, *Science* **372** (2021) 512–516.
49. S. Reader, G. F. Joyce, *Nature* **420** (2002) 841–844.
50. F. H. C. Crick, *J. Mol. Biol.* **38** (1968) 367–379.
51. E. Orgel, *J. Mol. Biol.* **38** (1968) 381–393.
52. Groebke, J. Hunziker, W. Fraser, L. Peng, U. Diederichsen, K. Zimmermann, A. Holzner, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **81** (1998) 375–474.
53. B. D. Heuberger, C. Switzer, *ChemBioChem* **9** (2008) 2779–2783.
54. T. R. Battersby, M. Albalos, M. J. Friesenhahn, *Chem. & Biol.* **14** (2007) 525–531.
55. F. H. C. Crick, *J. Mol. Biol.* **19** (1966) 548–555.
56. J. M. Berg, J. L. Tymoczko, G. J. Gatto, L. Stryer, Biochemistry 8th Ed. 893–924 (2015), W. H. Freeman and Company, New York.
57. Danger, R. Plasson, R. Pascal, *Chem. Soc. Rev.* **41** (2012) 5416–5429.
58. Frenkel-Pinter, M. Samanta, G. Ashkenasy, L. J. Leman, *Chem. Rev.* **120** (2020) 4707–4765.
59. R. Cronin, C. B. Moore, *Science* **172** (1971) 1327–1329.
60. Altwegget *et al.*, *Sci. Adv.* **2** (2016) e160028.
61. S. Ioppolo, G. Fedoseev, K. J. Chuang, H. M. Cuppen, A. R. Clements, M. Jin, R. T. Garrod, D. Qasim, V. Kofman, E. F. van Dishoeck, H. Linnartz, *Nat Astron* **5** (2021)197–205.
62. Lange, F. Djago, B. Eddhif, Q. B. Remaury, A. Ruf, N. K. V. Leitner, L. L. S. Hendecourt, G. Danger, C. G. Rodier, S. Papot, P. Poinot, *J. Proteome Res.* **20** (2021) 1444–1450.
63. A. Shimoyama, R. Ogasawara, *Orig. Life Evol. Biosph.* **32** (2002) 165–179.
64. S. W. Fox, K. Harada *Science* **128** (1958) 1214.
65. G. Martra, C. Deiana, Y. Sakhno, I. Barberis, M. Fabbiani, M. Pazzi, M. Vincenti, *Angew. Chem. Int. Ed.* **53** (2014) 4671–4674.

66. J. G. Forsythe, S.-S. Yu, I. Mamajanov, M. A. Grover, R. Krishnamurthy, F. M. Fernández, N. V. Hud, *Angew. Chem. Int. Ed.* **54** (2015) 9871–9875.
67. Rodriguez-Garcia, A. J. Surman, G. J.T. Cooper, I. Suárez-Marina, Z. Hosni, M. P. Lee, L. Cronin, *Nat. Commun.* **6** (2015) 8385.
68. T. D. Campbell, R. Febrian, J. T. McCarthy, H. E. Kleinschmidt, J. G. Forsythe, P. J. Bracher, *Nat. Commun.* **10** (2019) 4508.
69. G. Schwendinger, B. M. Rode, *Anal Sci* **5** (1989) 411.
70. F. Sauer, M. Haas, C. Sydow, A. F. Siegle, C. A. Lauer, O. Trapp, *Nat Commun* **12** (2021) 7182.
71. J. Greenwald, M. P. Friedmann, R. Riek, *Angew. Chem. Int. Ed.* **55** (2016) 11609–11613.
72. P. Canavelli, S. Islam, M. W. Powner, *Nature* **571** (2019) 546–549.
73. J. P. Ferris, R. A. Hill Jr, R. Liu, L. E. Orgel, *Nature* **381** (1996) 59–61.
74. J. G. Blank, G. H. Miller, M. J. Ahrens, R. E. Winans, *Orig Life Evol Biosph* **31** (2001) 15–51.
75. H. Sugahara, K. Mimura, *Geochem. J.* **48** (2014) 51–62.
76. B. A. Steele, N. Goldman, I.-F. W. Kuo, M. P. Kroonblawd, *Chem. Sci.* **11** (2020) 7760–7771.
77. H. G. Hansma, *J. Theor. Biol.* **266** (2010) 175–188.
78. A. S. Burton, J. C. Stern, J. E. Elsila, D. P. Glavin, J. P. Dworkin, *Chem. Soc. Rev.* **41** (2012) 5459–5472.
79. S. Pizzarello, *Acc. Chem. Res.* **39** (2006) 231–237.
80. H. Sugahara, K. Mimura, *Icarus* **257** (2015) 103–112.
81. Bertrand, S. van der Gaast, F. Vilas, F. Hçrz, G. Haynes, A. Chabin, A. Brack, F. Westall, *Astrobiology* **9** (2009) 943–951.
82. T. D. Campbell, R. Febrian, H. E. Kleinschmidt, K. A. Smith, P. J. Bracher, *ACS Omega* **4** (2019) 12745–12752.
83. L. Takacs, *Chem. Soc. Rev.* **42** (2013) 7649–7659.
84. S. L. James, C. J. Adams, C. Bolm, D. Braga, P. Collier, T. Friščić, F. Grepioni, K. D. M. Harris, G. Hyett, W. Jones, A. Krebs, J. Mack, L. Maini, A. G. Orpen, I. P. Parkin, W. C. Shearouse, J. W. Steed, D. C. Waddell, *Chem. Soc. Rev.* **41** (2012) 413–447.

85. Baláž, M. Achimovičová, M. Baláž, P. Billik, A. Cherkezova-Zheleva, J. M. Criado, F. Delogu, E. Dutková, E. Gaffet, F. L. Gotor, R. Kumar, I. Mitov, T. Rojac, M. Senna, A. Streletskii, K. Wieczorek-Ciurowa, *Chem. Soc. Rev.* **42** (2013) 7571−7637.
86. T. Friščić, C. Mottillo, H. M. Titi, *Angew. Chem. Int. Ed.* **59** (2020) 1018–1029.
87. A. A. L. Michalchuk, E. V. Boldyreva, A. M. Belenguer, F. Emmerling, V. V. Boldyrev, *Front. Chem.* **9** (2021) 359.
88. S. Lukin, M. Tireli, T. Stolar, D. Barišić, M. V. Blanco, M. di Michiel, K. Užarević, I. Halasz, *J. Am. Chem. Soc.* **141** (2019) 1212–1216.
89. K. J. Ardila-Fierro, J. G. Hernández, *ChemSusChem* **14** (2021) 2145–2162.
90. D. Braga, L. Maini, F. Grepioni, *Chem. Soc. Rev.* **42** (2013) 7638–7648.
91. K. Užarević, I. Halasz, T. Friščić, *J. Phys. Chem. Lett.* **6** (2015) 4129−4140.
92. D. E. Crawford, J. Casaban, *Adv. Mater.* **28** (2016) 5747–5754.
93. R. Rightmire, T. P. Hanusa, *Dalton Trans.* **45** (2016) 2352–2362.
94. S. Garcia-Manyes, A. E. M. Beedle, *Nat. Rev. Chem.* **1** (2017) 0083.
95. B. H. Bowser, S. L. Craig, *Polym. Chem.* **9** (2018) 3583−3593.
96. J. L. Howard, Q. Cao, D. L. Browne, *Chem. Sci.* **9** (2018) 3080–3094.
97. J. Andersen, J. Mack, *Green Chem.* **20** (2018) 1435–1443.
98. C. Bolm, J. G. Hernández, *ChemSusChem* **11** (2018) 1410–1420.
99. V. Štrukil, *Synlett* **29** (2018) 1281–1288.
100. Eguaogie, J. S. Vyle, P. F. Conlon, M. A. Gîlea, Y. Liang, *Beilstein J. Org. Chem.* **14** (2018) 955–970.
101. D. Tan, F. García, *Chem. Soc. Rev.* **48** (2019) 2274–2292.
102. A. Beillard, X. Bantreil, T.-X. Métro, J. Martinez, F. Lamaty, *Chem. Rev.* **119** (2019) 7529–7609.
103. Perona, P. Hoyos, A. Farrán, M. J. Hernaiz, *Green Chem.* **20** (2020) 5559–5583.
104. Porcheddu, E. Colacino, L. De Luca, F. Delogu, *ACS Catal.* **10** (2020) 8344–8394.
105. M. Pérez-Venegas, E. Juaristi, *ACS Sustainable Chem. Eng.* **8** (2020) 8881–8893.
106. T. Stolar, K. Užarević, *CrystEngComm* **22** (2020) 4511–4525.
107. S. Kaabel, T. Friščić, K. Auclair, *ChemBioChem* **21** (2020) 742–758.
108. W. Pickhardt, S. Grätz, L. Borchardt, *Chem. Eur. J.* **26** (2020) 12903–12911.
109. K. Kubota, H. Ito, *Trends Chem.* **2** (2020) 1066–1081.
110. R.T. O'Neil, R. Boulatov, *Nat. Rev. Chem.* **5** (2021) 148–167.

111. A. P. Amrute, J. D. Bellis, M. Felderhoff, F. Schüth, *Chem. Eur. J.* **27** (2021) 6819–6847.

112. F. Hajiali, T. Jin, G. Yang, M. Santos, E. Lam, A. Moores, *ChemSusChem* **15** (2022) e202102535.

113. Bento, F. Luttringer, T. M. El Dine, N. Pétry, X. Bantreil, F. Lamaty, *Eur. J. Org. Chem.* (2022) e202101516.

114. N. Shan, F. Toda, W. Jones, *Chem. Commun.* (2002) 2372–2373.

115. T. Friščić, A. V. Trask, W. Jones, W. D. S. Motherwell, *Angew. Chem. Int. Ed.* **45** (2006) 7546–7550.

116. K. Užarević, V. Štrukil, C. Mottillo, P. A. Julien, A. Puškarić, T. Friščić, I. Halasz, *Cryst. Growth Des.* **16** (2016) 2342–2347.

117. H. Kulla, M. Wilke, F. Fischer, M. Rollig, C. Maierhofer, F. Emmerling, *Chem. Commun.* **53** (2017) 1664–1667.

118. S. Lukin, T. Stolar, I. Lončarić, I. Milanović, N. Biliškov, M. di Michiel, T. Friščić, I. Halasz *Inorg. Chem.* **59** (2020) 12200–12208.

119. K. Užarević, N. Ferdelji, T. Mrla, P. A. Julien, B. Halasz, T. Friščić, I. Halasz, *Chem. Sci.* **9** (2018) 2525–2532.

120. J. Andersen, J. Brunemann, J. Mack, *React. Chem. Eng.* **4** (2019) 1229–1236.

121. N. Cindro, M. Tireli, B. Karadeniz, T. Mrla, K. Užarević, *ACS Sustainable Chem. Eng.* **7** (2019) 16301–16309.

122. T. Seo, N. Toyoshima, K. Kubota, H. Ito, *J. Am. Chem. Soc.* **143** (2021) 6165–6175.

123. Takahashi, A. Hu, P. Gao, Y. Gao, Y. Pang, T. Seo, J. Jiang, S. Maeda, H. Takaya, K. Kubota, H. Ito, *Nat Commun* **12** (2021) 6691.

124. V. Štrukil, I. Sajko, *Chem. Commun.* **53** (2017) 9101–9104.

125. J. G. Hernández, *Beilstein J. Org. Chem.* **13** (2017) 1463–1469.

126. Bolm, J. G. Hernández, *Angew. Chem. Int. Ed.* **58** (2019) 3285–3299.

127. G.-F. Han, F. Li, Z.-W. Chen, C. Coppex, S.-J. Kim, H.-J. Noh, Z. Fu, Y. Lu, C. V. Singh, S. Siahrostami, Q. Jiang, J.-B. Baek, *Nat. Nanotechnol.* **16** (2021) 325–330.

128. T. Friščić, I. Halasz, P. J. Beldon, A. M. Belenguer, F. Adams, S. A. J. Kimber, V. Honkimäki, R. E. Dinnebier, *Nature Chem* **5** (2013) 66–73.

129. V. Ban, Y. Sadikin, M. Lange, N. Tumanov, Y. Filinchuk, R. Černý, N. Casati, *Anal. Chem.* **89** (2017) 13176–13181.

130. G. I. Lampronti, A. A. L. Michalchuk, P. P. Mazzeo, A. M. Belenguer, J. K. M. Sanders, A. Bacchi, F. Emmerling, *Nat Commun* **12** (2021) 6134.
131. Gracin, V. Štrukil, T. Friščić, I. Halasz, K. Užarević, *Angew. Chem. Int. Ed.* **53** (2014) 6193–6197.
132. Lukin, K. Užarević, I. Halasz, *Nat Protoc* **16** (2021) 3492–3521.
133. A. A. L. Michalchuk, F. Emmerling, *Angew. Chem. Int. Ed.* **61** (2022) e202117270.
134. K. Kubota, Y. Pang, A. Miura, H. Ito, *Science* **366** (2019) 1500–1504.
135. Schumacher, J. G. Hernández, C. Bolm, *Angew. Chem. Int. Ed.* **59** (2020) 16357–16360.
136. Huskić, C. B. Lennox, T. Friščić, *Green Chem.* **22** (2020) 5881–5901.
137. Bolm, R. Mocci, C. Schumacher, M. Turberg, F. Puccetti, J. G. Hernández, *Angew. Chem. Int. Ed.* **57** (2018) 2423–2426.
138. J. E. Elsila, N. M. Johnson, D. P. Glavin, J. C. Aponte, J. P. Dworkin, *Meteorit. Planet. Sci.* **56** (2021) 586–600.
139. S. Lamour, S. Pallman, M. Hass, O. Trapp, *Life* **9** (2019) 52.
140. M. Hass, S. Lamour, S. B. Christ, O. Trapp, *Commun. Chem.* **3** (2020) 140.
141. Becker, J. Feldmann, S. Wiedemann, H. Okamura, C. Schneider, K. Iwan, A. Crisp, M. Rossa, T. Amatov, T. Carell, *Science* **366** (2019) 76–82.
142. TOPAS, version 4.2; Bruker-AXS: Karlsruhe, Germany, 2009.
143. Stolar, S. Lukin, M. Etter, M. Rajić Linarić, K. Užarević, E. Meštrović, I. Halasz, *Chem. Commun.* **56** (2020) 13524–13527.
144. A. M. Atria, M. T. Garland, R. Baggio, *Acta Cryst.* **C66** (2010) o547.
145. E. Colacino, M. Carta, G. Pia, A. Porcheddu, P. C. Ricci, F. Delogu, *ACS Omega* **3** (2018) 9196 – 9209.
146. A.W. Tricker, G. Samaras, K. L. Hebisch, M. J. Realff, C. Sievers, *Chem. Eng. J.* **382** (2020) 122954.
147. P.A. Cawood, C.J. Hawkesworth, B. Dhuime, *GSA Bulletin* **125** (2013) 14–32.
148. L. Vugrin, M. Carta, S. Lukin, E. Meštrović, F. Delogu, I. Halasz, *ChemRxiv* (2022), DOI: 10.26434/chemrxiv-2022-cbp87
149. S. Lherminier, R. Planet, V. L. Vehel, G. Simon, L. Vanel, K. J. Måløy, O. Ramos, *Phys. Rev. Lett.* **122** (2019) 218501.
150. C. Catling, K. J. Zahnle, *Sci. Adv.* **6** (2020) eaax1420.

151. J. G. Forsythe, S.-S. Yu, I. Mamajanov, M. A. Grover, R. Krishnamurthy, F. M. Fernandez, N. V. Hud, *Angew. Chem. Int. Ed.* **54** (2015) 9871–9875.

152. A. D. McKee, M. Solano, A. Saydjari, C. J. Bennett, N. V. Hud, T. M. Orlando, *ChemBioChem* **19** (2018) 1913–1917.

153. Ferraris, D.W. Jones, J. Yerkess, *Z. Kristallogr.* **135** (1972) 240–252.

154. J. M. Hughes, M. Cameron, K. D. Crowley, *American Mineralogist* **74** (1989) 870–876.

155. L. Graf, *American Mineralogist* **46** (1961) 1283–1316.

156. K. Koyama, Y. Takeuchi, *Z. Kristallogr.* **145** (1977) 216–239.

157. J. C. A. Boeyens, V. V. H. Ichharam, *Z. Kristallogr.* **217** (2002) 9–10.

158. S. M. Richardson, J. W. Richardson, *American Mineralogist* **67** (1982) 69–75.

159. Bayliss, *American Mineralogist* **62** (1977) 1168–1172.

160. J. W. Gruner, *Z. Kristallogr.* **88** (1934) 412–419.

161. L. Levien, C. T. Prewitt, D. J. Weidner, *American Mineralogist* **65** (1980) 920–930.

§ 7. APPENDIX

```
macro PS  {  }
macro BGR {  }
macro RBR { @ }
macro RBT { @ }
macro REF {  }

no_LIMIT_warnings
'do_errors

Auto_T(100)

iters 1000000000
xdd "C:\Documents and Settings\Tomislav\Desktop\Stolar\Self-assembly\Experimental data\D starting\PMF\D-NG-240_hires_merged.xy"
r_exp 4.267 r_exp_dash 6.560 r_wp 13.448 r_wp_dash 20.676 r_p 10.494 r_p_dash 18.728 weighted_Durbin_Watson 0.120 gof 3.152
bkg BGR 228.4923695 -19.69729777 -55.15388878 24.77688085 31.84849651 5.964291148
start_X  4
finish_X  20
LP_Factor( 27)
Zero_Error(PS, 0.01515337979)
Rp 217.5
Rs 217.5
Simple_Axial_Model(PS, 3.827129369)
lam
ymin_on_ymax  0.001
la  0.6533 lo  0.709319 lh  0.2695
la  0.3467 lo  0.713574 lh  0.2795
x_calculation_step 0.002

str
CS_G(PS, 141.300175)
Strain_L(PS, 0.3083608399)
Strain_G(PS, 0.9368756804)
a PS 14.94544671
b PS 12.357385
c PS 6.755097821
r_bragg  4.30247054
phase_name "hkl_Phase"
space_group Pbca
MVW( 1201.146, 1247.57631, 100.000`)

scale @  4.42742893e-005` max 0.1
'view_structure
Out_CIF_STR(D_anhydrate_PMF.cif)

'2,6-diaminopurine
rigid
point_for_site   N3   ux 5.648   uy 5.955  uz 2.034
point_for_site   N4   ux 4.231   uy 4.377  uz 1.168
point_for_site   N5   ux 3.030   uy 2.424  uz 0.392
point_for_site   H10  ux 2.395   uy 2.796  uz -0.053
point_for_site   N6   ux 4.280   uy 0.812  uz 1.287
point_for_site   N7   ux 6.149   uy 3.770  uz 2.518
point_for_site   N8   ux 6.764   uy 1.582  uz 2.928
point_for_site   C5   ux 5.320   uy 4.656  uz 1.908
point_for_site   C6   ux 4.029   uy 3.054  uz 1.066
point_for_site   C7   ux 3.230   uy 1.086  uz 0.560
point_for_site   H9   ux 2.676   uy 0.435  uz 0.194
point_for_site   C8   ux 4.807   uy 2.053  uz 1.615
point_for_site   C9   ux 5.920   uy 2.455  uz 2.374
point_for_site   H6   ux 6.248   uy 6.189  uz 2.619
point_for_site   H5   ux 5.077   uy 6.568  uz 1.760
point_for_site   H7   ux 7.365   uy 1.853  uz 3.478
point_for_site   H8   ux 6.634   uy 0.733  uz 2.937

Rotate_about_axies( RBR  186.48150`, RBR -32.56154`, RBR -278.10057` )
Translate(RBT  51416.47372`, RBT  1108.52729` , RBT -9028.24471` )

prm B1  1.60395` min 0 max 12

'2,6-diaminopurine
site N3  num_posns  8 x   51416.90901 y  1108.95993 z  -9028.13074 occ N  1 beq  =B1;
site N4  num_posns  8 x   51416.79509 y  1108.83624 z  -9028.11405 occ N  1 beq  =B1;
site N5  num_posns  8 x   51416.65778 y  1108.73352 z  -9028.08603 occ N  1 beq  =B1;
site H10 num_posns  8 x   51416.67851 y  1108.66860 z  -9028.08670 occ H  1 beq  =B1;
site N6  num_posns  8 x   51416.55851 y  1108.86784 z  -9028.07469 occ N  1 beq  =B1;
site N7  num_posns  8 x   51416.76675 y  1109.02924 z  -9028.12005 occ N  1 beq  =B1;
site N8  num_posns  8 x   51416.62537 y  1109.10300 z  -9028.09106 occ N  1 beq  =B1;
site C5  num_posns  8 x   51416.82047 y  1108.94062 z  -9028.12295 occ C  1 beq  =B1;
site C6  num_posns  8 x   51416.70590 y  1108.82664 z  -9028.09887 occ C  1 beq  =B1;
site C7  num_posns  8 x   51416.57012 y  1108.76308 z  -9028.07208 occ C  1 beq  =B1;
site H9  num_posns  8 x   51416.52339 y  1108.71373 z  -9028.06168 occ H  1 beq  =B1;
site C8  num_posns  8 x   51416.64429 y  1108.90987 z  -9028.09084 occ C  1 beq  =B1;
site C9  num_posns  8 x   51416.67797 y  1109.01592 z  -9028.10192 occ C  1 beq  =B1;
site H6  num_posns  8 x   51416.92789 y  1109.02475 z  -9028.15875 occ H  1 beq  =B1;
site H5  num_posns  8 x   51416.94584 y  1108.90533 z  -9028.15090 occ H  1 beq  =B1;
site H7  num_posns  8 x   51416.64680 y  1109.16610 z  -9028.11518 occ H  1 beq  =B1;
site H8  num_posns  8 x   51416.56795 y  1109.10011 z  -9028.09058 occ H  1 beq  =B1;
```

Figure 7.1. An example of input file that was used for the crystal structure solution *via* simulated annealing.

```
macro PS  { @ }
macro BGR { @ }
macro RBR {  }
macro RBT {  }
macro REF { @ }

no_LIMIT_warnings
do_errors

'Auto_T(100)

iters 1000000000
xdd "C:\Documents and Settings\Tomislav\Desktop\Stolar\Self-assembly\Experimental data\D starting\PMF\D-NG-240_hires_merged.xy"
r_exp 5.806 r_exp_dash 10.879 r_wp 9.784 r_wp_dash 18.333 r_p 7.564 r_p_dash 17.718 weighted_Durbin_Watson 0.413 gof 1.685
bkg BGR 162.16077`_0.304336639 -70.2477652`_0.385290409 -20.8492304`_0.332689231 34.3419832`_0.318142462 -32.8780192`_0.306585502
'start_X  4
'finish_X  26
LP_Factor(90)
Zero_Error(PS, 0.01906`_0.00027)
Rp 217.5
Rs 217.5
Simple_Axial_Model(PS, 4.41109`_0.01691)
lam
ymin_on_ymax  0.001
la  0.6533 lo  0.709319 lh  0.2695
la  0.3467 lo  0.713574 lh  0.2795
x_calculation_step 0.002

str
CS_G(PS, 312.36483`_39.53892)
Strain_L(PS, 0.02753`_0.01404)
Strain_G(PS, 0.42073`_0.00812)
a PS  14.97675`_0.00064
b PS  12.38205`_0.00035
c PS  6.76696`_0.00015
r_bragg  4.42506285
phase_name "DAP anhydrate"
space_group Pbca
MVW( 1201.146, 1254.88447`_0.0697825878, 100.000`_0.000)

scale @  5.10703352e-005`_1.19e-007 max 0.1
'view_structure
Out CIF STR(DAP anhydrate.cif)

prm B1  2.12341`_0.04694 min 0 max 12

'2,6-diaminopurine
site N3  num_posns  8 x REF  0.90672`_0.00002 y REF  0.95805`_0.00004 z REF -0.14595`_0.00015 occ N  1 beq  =B1;
site N4  num_posns  8 x REF  0.79403`_0.00006 y REF  0.83551`_0.00005 z REF -0.11532`_0.00025 occ N  1 beq  =B1;
site N5  num_posns  8 x REF  0.65725`_0.00003 y REF  0.73391`_0.00004 z REF -0.08965`_0.00016 occ N  1 beq  =B1;
site H10 num_posns  8 x REF  0.67734`_0.00012 y REF  0.66919`_0.00015 z REF -0.09221`_0.00040 occ H  1 beq  =B1;
site N6  num_posns  8 x REF  0.56131`_0.00004 y REF  0.86672`_0.00006 z REF -0.07076`_0.00031 occ N  1 beq  =B1;
site N7  num_posns  8 x REF  0.77056`_0.00005 y REF  1.02736`_0.00004 z REF -0.11758`_0.00036 occ N  1 beq  =B1;
site N8  num_posns  8 x REF  0.62669`_0.00003 y REF  1.09950`_0.00004 z REF -0.09279`_0.00013 occ N  1 beq  =B1;
site C5  num_posns  8 x REF  0.82212`_0.00004 y REF  0.93907`_0.00004 z REF -0.11915`_0.00028 occ C  1 beq  =B1;
site C6  num_posns  8 x REF  0.70634`_0.00003 y REF  0.82450`_0.00004 z REF -0.09994`_0.00012 occ C  1 beq  =B1;
site C7  num_posns  8 x REF  0.56994`_0.00004 y REF  0.76386`_0.00004 z REF -0.07225`_0.00019 occ C  1 beq  =B1;
site H9  num_posns  8 x REF  0.51933`_0.00011 y REF  0.71412`_0.00017 z REF -0.06151`_0.00048 occ H  1 beq  =B1;
site C8  num_posns  8 x REF  0.64639`_0.00003 y REF  0.90552`_0.00004 z REF -0.08865`_0.00019 occ C  1 beq  =B1;
site C9  num_posns  8 x REF  0.68174`_0.00004 y REF  1.01193`_0.00004 z REF -0.09736`_0.00024 occ C  1 beq  =B1;
site H6  num_posns  8 x REF  0.92940`_0.00013 y REF  1.02325`_0.00016 z REF -0.13167`_0.00061 occ H  1 beq  =B1;
site H5  num_posns  8 x REF  0.94762`_0.00013 y REF  0.90384`_0.00015 z REF -0.17103`_0.00193 occ H  1 beq  =B1;
site H7  num_posns  8 x REF  0.64608`_0.00012 y REF  1.16376`_0.00014 z REF -0.08288`_0.00062 occ H  1 beq  =B1;
site H8  num_posns  8 x REF  0.56731`_0.00012 y REF  1.09463`_0.00015 z REF -0.07462`_0.00069 occ H  1 beq  =B1;
```

Figure 7.2. An example of input file that was used for final Rietveld refinement (part 1/2).

```
penalties_weighting_K1  100

Distance_Restrain( N3   C5  , 1.3689, 1.30134`_0.00077 , 0.02, 1)
Distance_Restrain( N3   H6  , 0.8791, 0.88117`_0.00213 , 0.02, 1)
Distance_Restrain( N3   H5  , 0.9076, 0.92443`_0.00260 , 0.02, 1)
Distance_Restrain( N4   C5  , 1.3647, 1.34978`_0.00084 , 0.02, 1)
Distance_Restrain( N4   C6  , 1.342, 1.32446`_0.00095  , 0.02, 1)
Distance_Restrain( N5   H10 , 0.8585, 0.85617`_0.00189 , 0.02, 1)
Distance_Restrain( N5   C6  , 1.3759, 1.34297`_0.00064 , 0.02, 1)
Distance_Restrain( N5   C7  , 1.348, 1.36428`_0.00068  , 0.02, 1)
Distance_Restrain( N6   C7  , 1.2961, 1.28020`_0.00093 , 0.02, 1)
Distance_Restrain( N6   C8  , 1.388, 1.36715`_0.00081  , 0.02, 1)
Distance_Restrain( N7   C5  , 1.3218, 1.33848`_0.00082 , 0.02, 1)
Distance_Restrain( N7   C9  , 1.3625, 1.35083`_0.00099 , 0.02, 1)
Distance_Restrain( N8   C9  , 1.3136, 1.36250`_0.00073 , 0.02, 1)
Distance_Restrain( N8   H7  , 0.8509, 0.84966`_0.00191 , 0.02, 1)
Distance_Restrain( N8   H8  , 0.8991, 0.89980`_0.00204 , 0.02, 1)
Distance_Restrain( C6   C8  , 1.388, 1.34847`_0.00066  , 0.02, 1)
Distance_Restrain( C7   H9  , 0.9736, 0.97935`_0.00190 , 0.02, 1)
Distance_Restrain( C8   C9  , 1.4441, 1.42119`_0.00074 , 0.02, 1)

Angle_Restrain(C5   N3  H6  ,120.07, 121.70099`_0.14371 , 3, 0.001)
Angle_Restrain(C5   N3  H5  ,121.93, 122.65714`_0.14584 , 3, 0.001)
Angle_Restrain(H6   N3  H5  ,114.65, 115.46404`_0.18308 , 3, 0.001)
Angle_Restrain(C5   N4  C6  ,115.25, 114.10088`_0.06237 , 3, 0.001)
Angle_Restrain(H10  N5  C6  ,126.42, 126.03873`_0.12397 , 3, 0.001)
Angle_Restrain(H10  N5  C7  ,125.76, 126.37069`_0.12585 , 3, 0.001)
Angle_Restrain(C6   N5  C7  ,107.81, 107.58702`_0.04357 , 3, 0.001)
Angle_Restrain(C7   N6  C8  ,106.84, 104.76191`_0.05851 , 3, 0.001)
Angle_Restrain(C5   N7  C9  ,117.66, 116.96296`_0.05517 , 3, 0.001)
Angle_Restrain(C9   N8  H7  ,122.67, 122.69860`_0.13027 , 3, 0.001)
Angle_Restrain(C9   N8  H8  ,122.38, 123.21869`_0.12796 , 3, 0.001)
Angle_Restrain(H7   N8  H8  ,112.92, 112.94041`_0.18514 , 3, 0.001)
Angle_Restrain(N3   C5  N4  ,119.19, 118.53509`_0.06267 , 3, 0.001)
Angle_Restrain(N3   C5  N7  ,114.61, 114.53959`_0.05926 , 3, 0.001)
Angle_Restrain(N4   C5  N7  ,126.17, 126.58024`_0.06519 , 3, 0.001)
Angle_Restrain(N4   C6  N5  ,129.07, 129.26325`_0.04934 , 3, 0.001)
Angle_Restrain(N4   C6  C8  ,126.53, 126.02251`_0.05036 , 3, 0.001)
Angle_Restrain(N5   C6  C8  ,105.683, 104.70967`_0.03997 , 3, 0.001)
Angle_Restrain(N5   C7  N6  ,111.92, 111.58448`_0.05317 , 3, 0.001)
Angle_Restrain(N5   C7  H9  ,125.15, 125.25939`_0.13129 , 3, 0.001)
Angle_Restrain(N6   C7  H9  ,122.93, 123.15567`_0.13368 , 3, 0.001)
Angle_Restrain(N6   C8  C6  ,112.20, 111.35608`_0.04904 , 3, 0.001)
Angle_Restrain(N6   C8  C9  ,132.067, 132.58332`_0.05282 , 3, 0.001)
Angle_Restrain(C6   C8  C9  ,116.757, 116.06057`_0.04651 , 3, 0.001)
Angle_Restrain(N7   C9  N8  ,118.15, 119.05389`_0.05441 , 3, 0.001)
Angle_Restrain(N7   C9  C8  ,121.47, 120.14365`_0.05211 , 3, 0.001)
Angle_Restrain(N8   C9  C8  ,120.37, 120.75929`_0.04930 , 3, 0.001)

Flatten(C7 H9 N5 H10 N6 C6 C8 C9 N4 N7 C5 N3, 0, 0.000729448925`_0.000529990844, 10)

Create_2Th_Ip_file(DAPanhydrate.pks)

prm p5  0.03046`_0.00130 min 0
spherical_harmonics_hkl sh5
sh_order  4 load sh_Cij_prm {
y00   !sh5_c00  1.00000
y20   sh5_c20   0.81573`_0.02834
y22p  sh5_c22p -0.42726`_0.03383
y40   sh5_c40  -0.39430`_0.03161
y42p  sh5_c42p -0.74314`_0.05395
y44p  sh5_c44p -0.12781`_0.01108}

lor_fwhm = sh5 p5 / Cos(Th);

xdd_out Plot-data-DAPanhydrate.txt load out_record out_fmt out_eqn
{
 " %11.5f  " = X;
 " %11.5f"   = Yobs;
 " %11.5f"   = Ycalc;
 " %11.5f\n" = Yobs-Ycalc;
 }
C_matrix_normalized
{ }
```

Figure 7.3. An example of input file that was used for final Rietveld refinement (part 2/2).

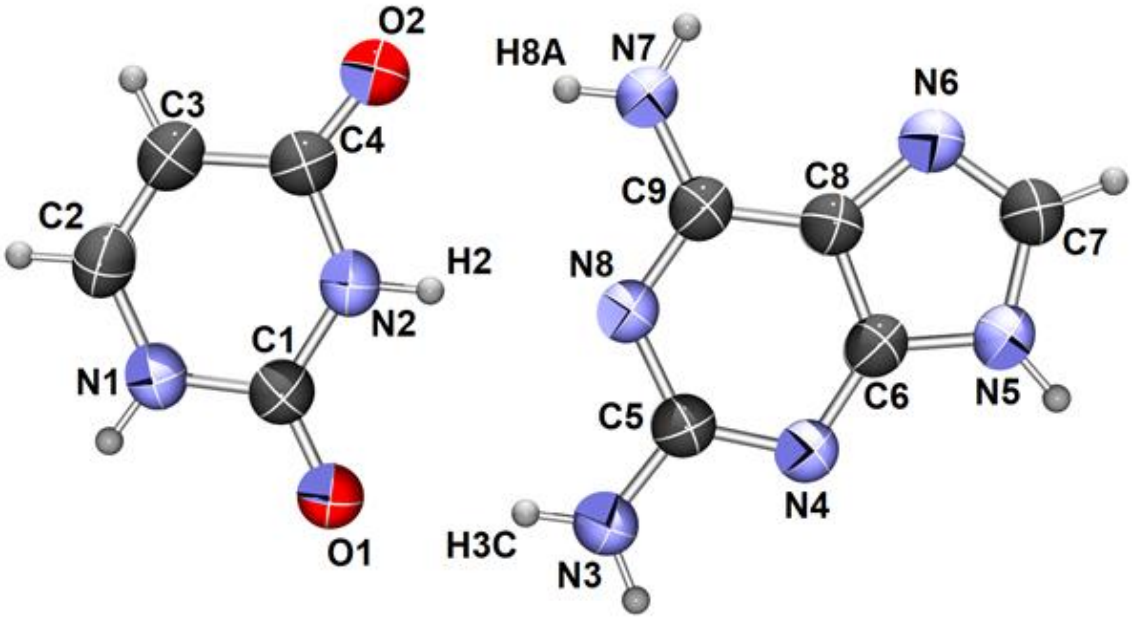

Figure 7.4. ORTEP-3 drawing of asymmetric unit of **2** (D:U cocrystal hydrate) with atom numbering scheme. Displacement ellipsoids are drawn for the probability of 50 % and hydrogen atoms are shown as spheres of arbitrary radii. Disordered water molecule has been omitted.

Table 7.1. Crystallographic, data collection and refinement data for single crystal X-ray diffraction data of **2** (D:U cocrystal hydrate).

Compound	**D:U cocrystal hydrate**
Empirical formula	$C_9H_{12}N_8O_3$
Formula wt. / g mol^{-1}	280.27
Colour	colourless
Crystal dimensions / mm	0.18 x 0.12 x 0.07
Space group	*C* 2/*c*
a / Å	10.6851(8)
b / Å	16.5477(10)
c / Å	14.2161(12)
α / °	90
β / °	102.930(8)
γ / °	90
Z	8
V / Å^3	2449.9(3)
D_{calc} / g cm^{-3}	1.520
λ / Å	1.54719 (Cu$K\alpha$)
μ / mm^{-1}	1.017
Θ range / °	5.02 – 76.84
T / K	293(2)
Diffractometer type	Xcalibur Nova
Range of *h*, *k*, *l*	$11 < h < 13$; $-20 < k < 17$; $-16 < l < 17$
Reflections collected	5096
Independent reflections	2498
Observed reflections ($I \geq 2\sigma$)	1785
Absorption correction	Multi-scan
T_{min}, T_{max}	0.4930; 1.0000
R_{int}	0.0380
R (F)	0.0673
R_w (F^2)	0.2133
Goodness of fit	1.012
H atom treatment	Mixed
No. of parameters	210
No. of restraints	12
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (eÅ^{-3})	0.278; - 0.244

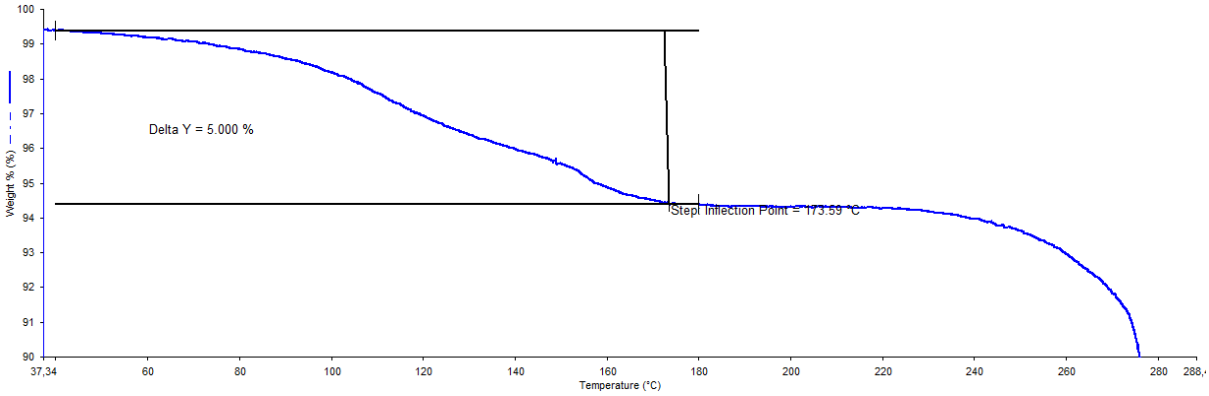

Figure 7.5. TGA curve of **2** (D:U cocrystal hydrate).

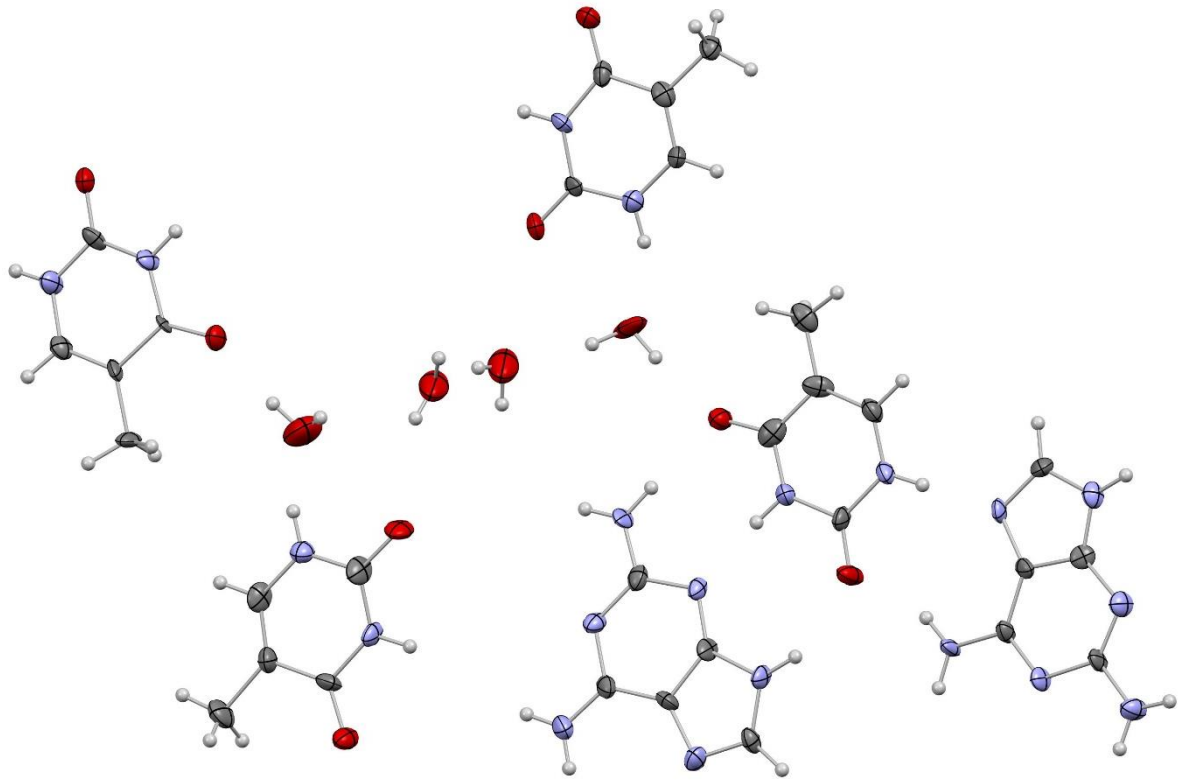

Figure 7.6. ORTEP-3 drawing of asymmetric unit of **10** (D:T 1:2 cocrystal hydrate). Displacement ellipsoids are drawn for the probability of 50 % and hydrogen atoms are shown as spheres of arbitrary radii.

Table 7.2. Crystallographic, data collection and refinement data for single crystal X-ray diffraction data of **10** (D:T 1:2 cocrystal hydrate).

Empirical formula	$C_{30}H_{44}N_{20}O_{12}$
Formula weight	876.85
Temperature/K	169.98(10)
Crystal system	monoclinic
Space group	Pc
a/Å	17.3694(5)
b/Å	16.9996(4)
c/Å	6.6652(2)
α/°	90
β/°	93.240(3)
γ/°	90
Volume/Å^3	1964.91(9)
Z	2
ρ_{calc}g/cm^3	1.482
μ/mm^{-1}	1.001
F(000)	920.0
Crystal size/mm^3	? × ? × ?
Radiation	Cu Kα (λ = 1.54184)
2Θ range for data collection/°	7.282 to 132.944
Index ranges	$-20 \leq h \leq 20$, $0 \leq k \leq 20$, $-7 \leq l \leq 7$
Reflections collected	6826
Independent reflections	6826 [R_{int} = ?, R_{sigma} = 0.1163]
Data/restraints/parameters	6826/0/574
Goodness-of-fit on F^2	1.059
Final R indexes [I>=2σ (I)]	R_1 = 0.0783, wR_2 = 0.2256
Final R indexes [all data]	R_1 = 0.1039, wR_2 = 0.2405
Largest diff. peak/hole / e Å^{-3}	0.57/-0.29
Flack parameter	0.1(4)

Table 7.3. Crystallographic data for **6** (D:T cocrystal anhydrate).

molecular formula	$C_{10}\ H_{12}\ N_8\ O_2$
formula weight (g mol^{-1})	276.28
space group	$P\bar{1}$
a (Å)	8.77213 (67)
b (Å)	11.07778 (52)
c (Å)	12.6355 (11)
α (deg)	95.7660 (37)
β (deg)	82.2760 (41)
γ (deg)	89.5282 (32)
V ($Å^3$)	1210.27 (15)
Z	4
temperature (K)	295
radiation type	Cu Kα
wavelength (Å)	1.5406
R_p	0.032
R_{wp}	0.047
GOF	7.61
$R(F^2)$	0.020

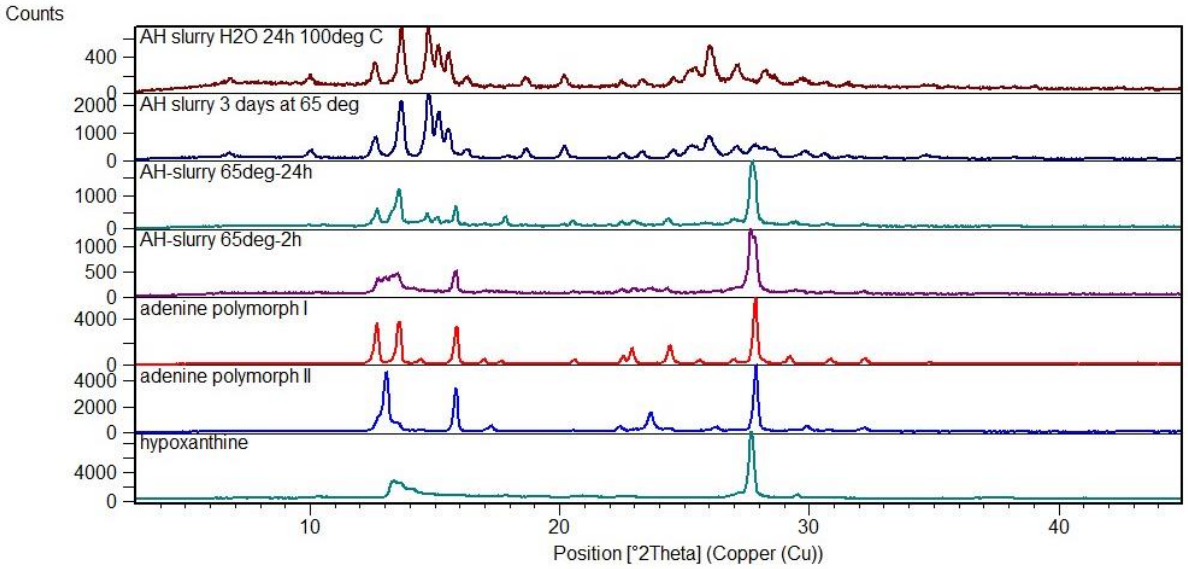

Figure 7.7. PXRD patterns of slurrying A and H in water using different reaction conditions.

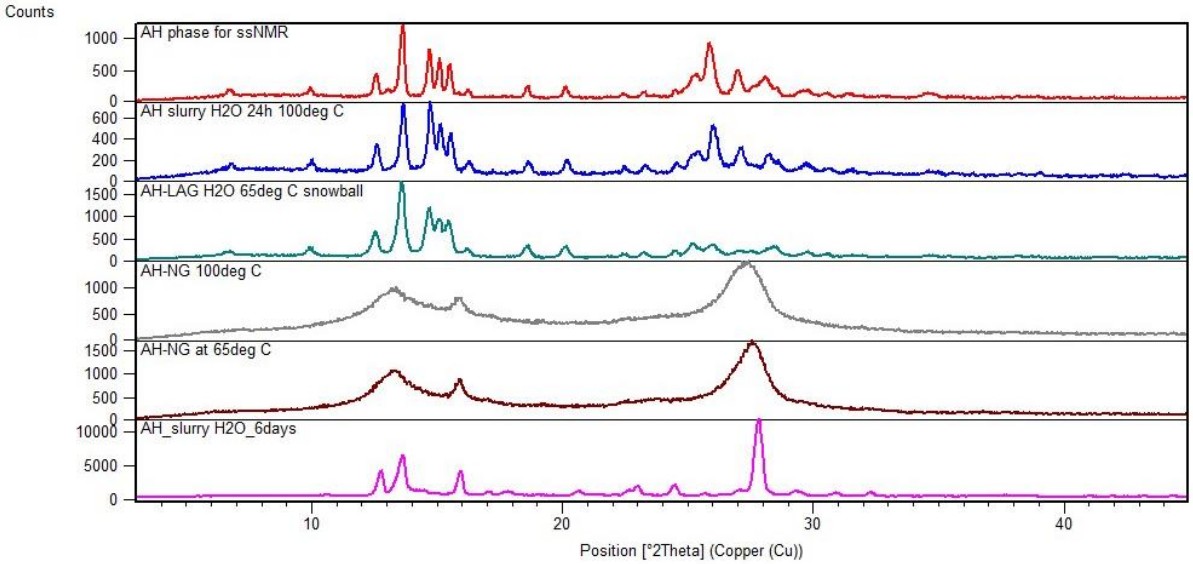

Figure 7.8. PXRD patterns of NG and LAG A and H at different temperatures.

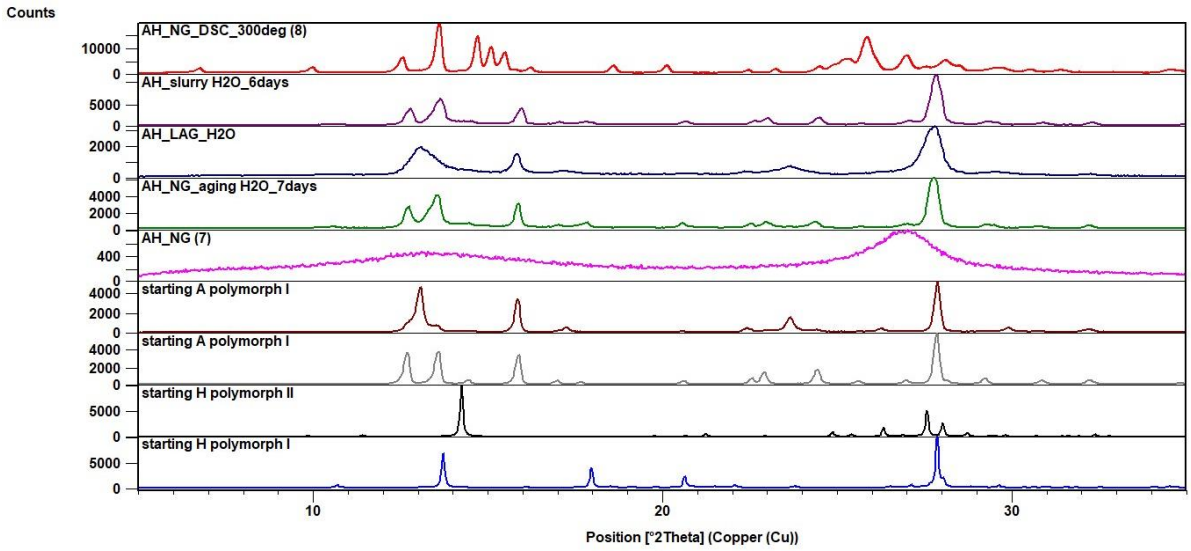

Figure 7.9. PXRD patterns of A and H system including DSC heating at 300 °C.

Table 7.4. Crystallographic data for **8** (A:H cocrystal anhydrate).

molecular formula	C_{10} H_9 N_9 O
formula weight (g mol^{-1})	271.26
space group	$P2_1/n$
a (Å)	12.00132 (53)
b (Å)	25.7944 (12)
c (Å)	3.64521 (11)
α (deg)	90
β (deg)	85.3997 (38)
γ (deg)	90
V ($Å^3$)	1124.800 (79)
Z	4
temperature (K)	295
radiation type	Cu Kα
wavelength (Å)	1.5406
R_p	0.034
R_{wp}	0.045
GOF	7.87
$R(F^2)$	0.019

Table 7.5. Crystallographic data for **9** (A:H cocrystal hydrate).

molecular formula	C_{10} H_{13} N_9 O_3
formula weight (g mol^{-1})	307.29
space group	$P2_1/n$
a (Å)	22.8501 (16)
b (Å)	16.8210 (11)
c (Å)	3.60152 (19)
α (deg)	90
β (deg)	86.7812 (70)
γ (deg)	90
V ($Å^3$)	1382.10 (15)
Z	4
temperature (K)	295
radiation type	Mo Kα
wavelength (Å)	0.71073
R_p	0.052
R_{wp}	0.069
GOF	1.81
$R(F^2)$	0.017

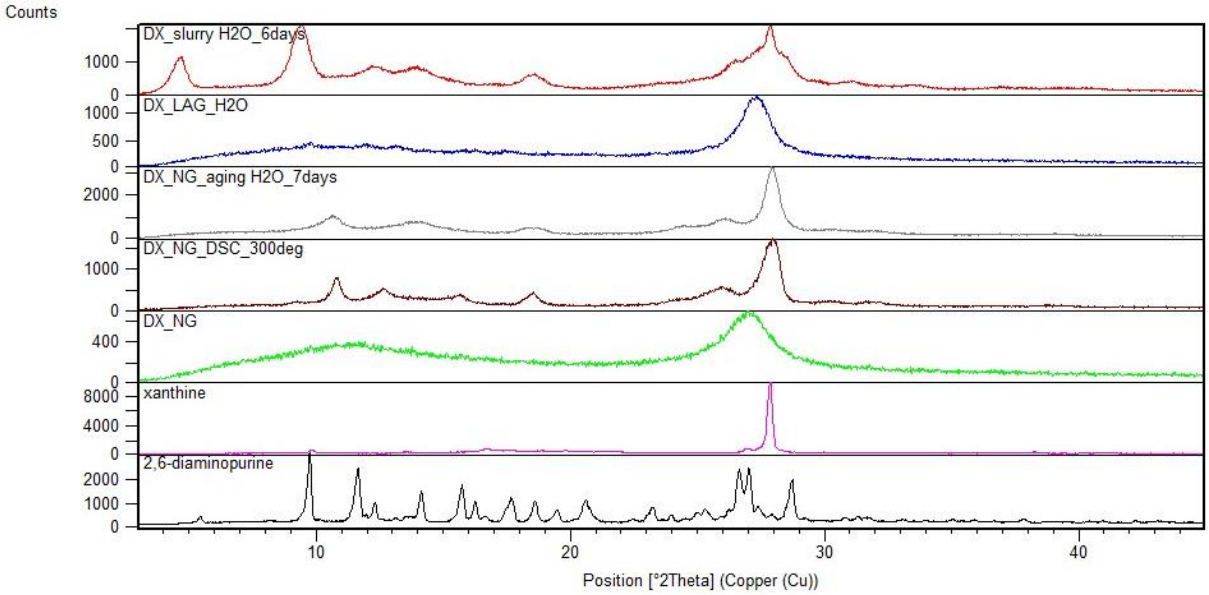

Figure 7.10. PXRD patterns for D and X system.

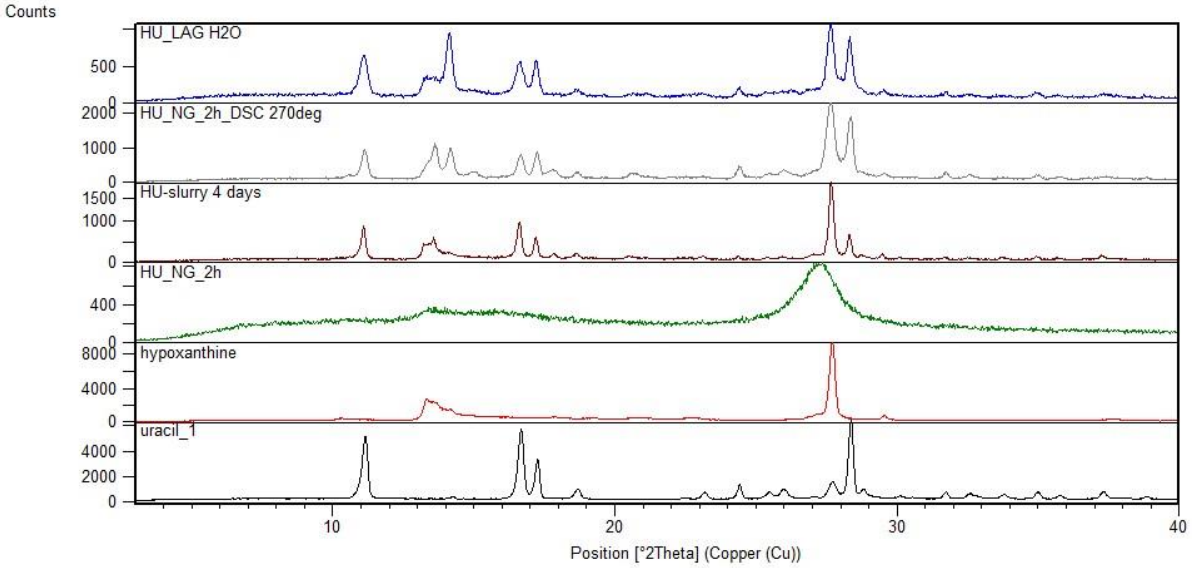

Figure 7.11. PXRD patterns for H and U system.

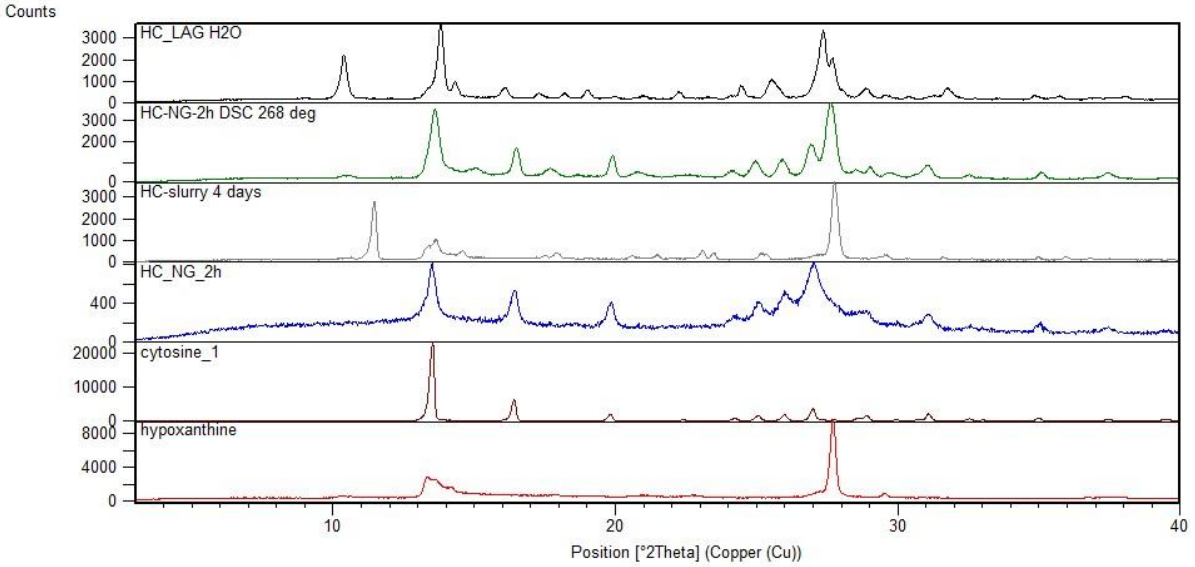

Figure 7.12. PXRD patterns for H and C system.

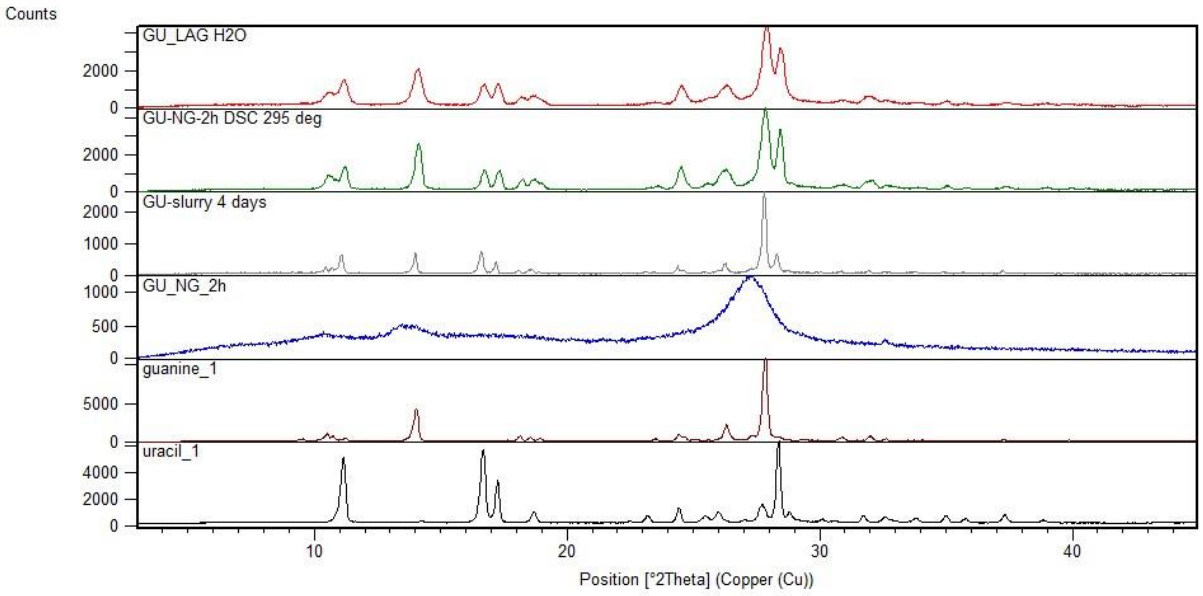

Figure 7.13. PXRD patterns for G and U system.

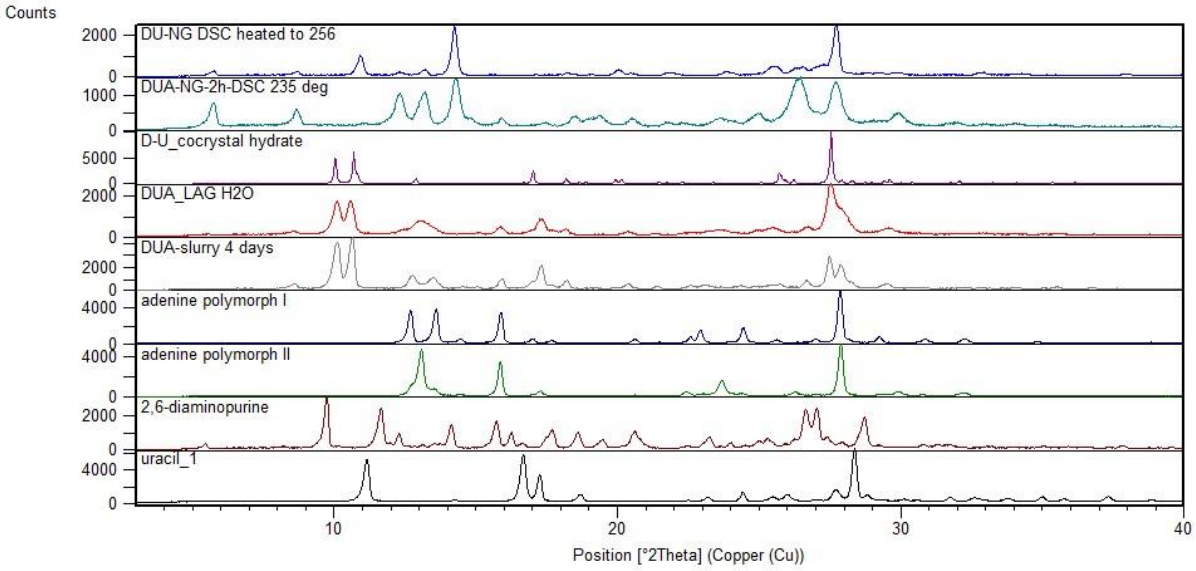

Figure 7.14. PXRD patterns for D, U, and A system.

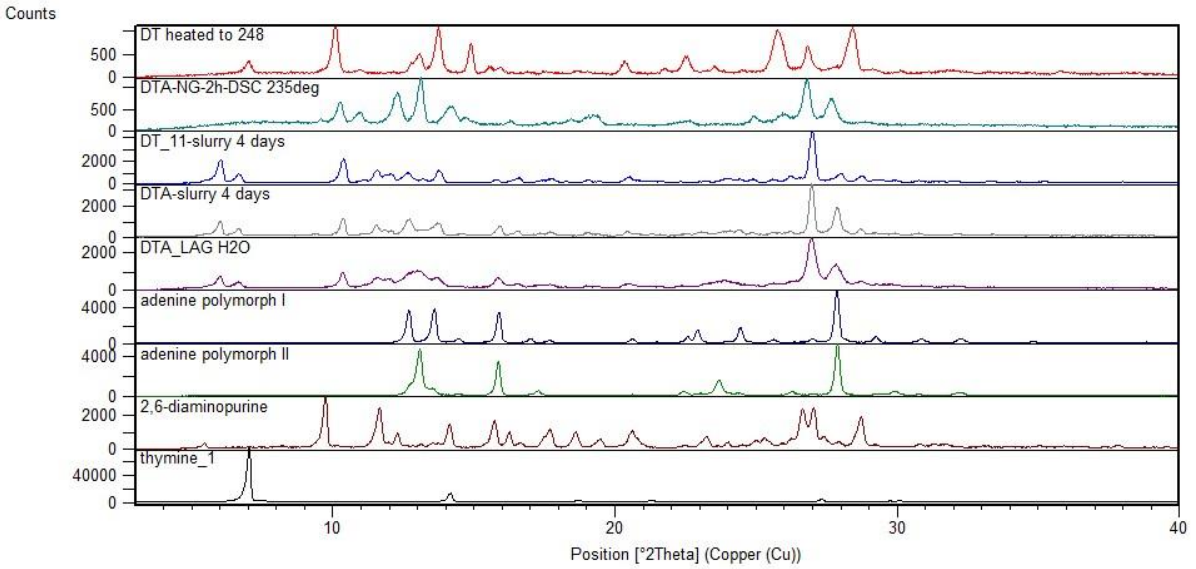

Figure 7.15. PXRD patterns for D, T, and A system.

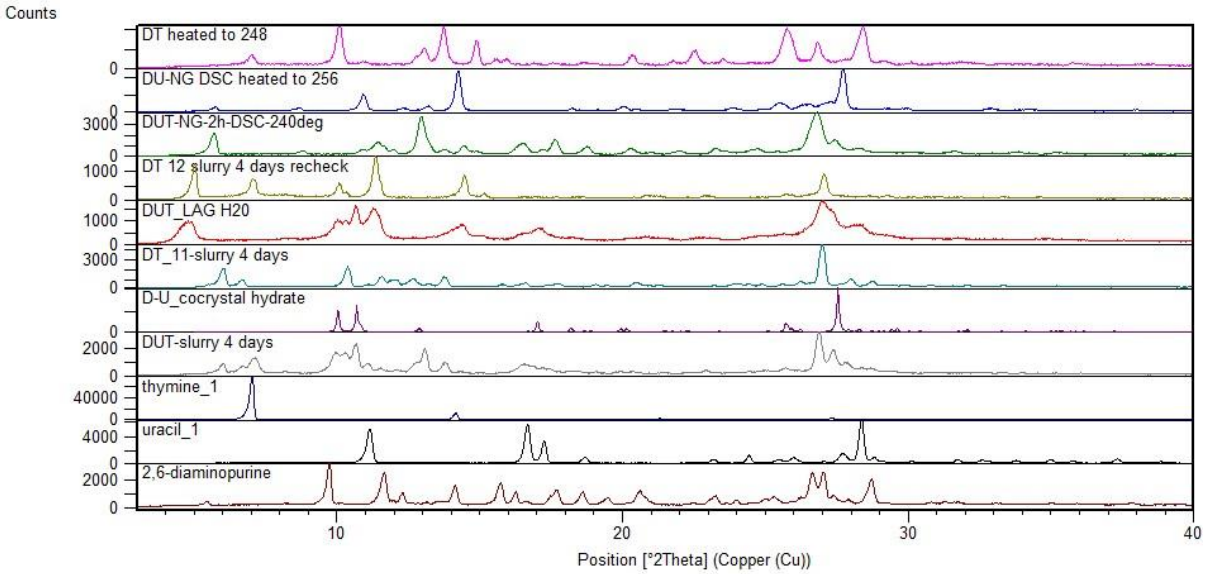

Figure 7.16. PXRD patterns for D, U, and T system.

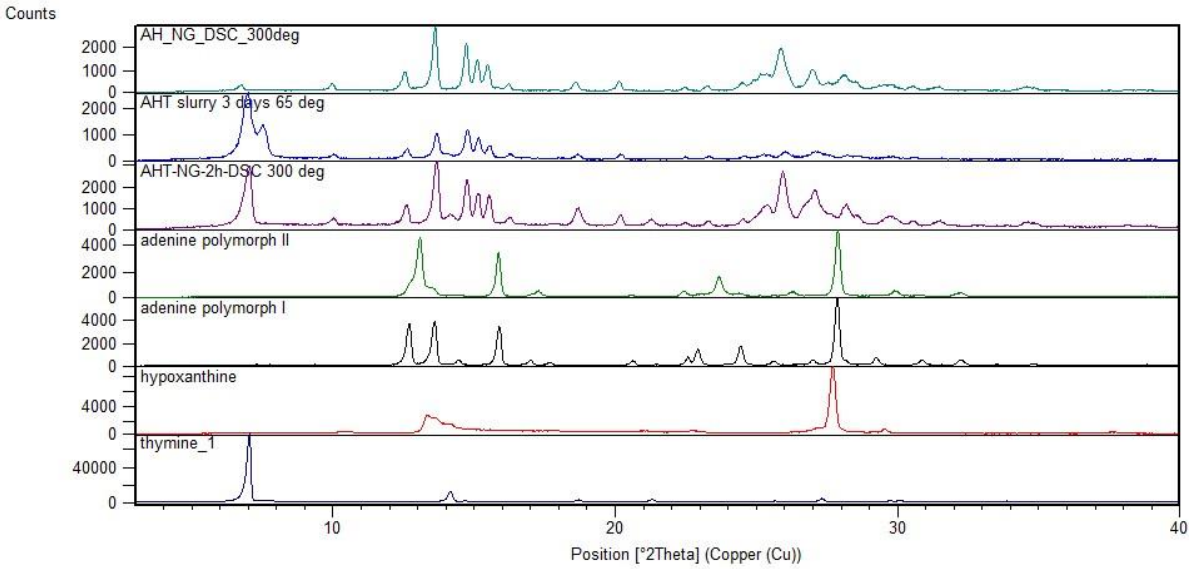

Figure 7.17. PXRD patterns for A, H, and T system.

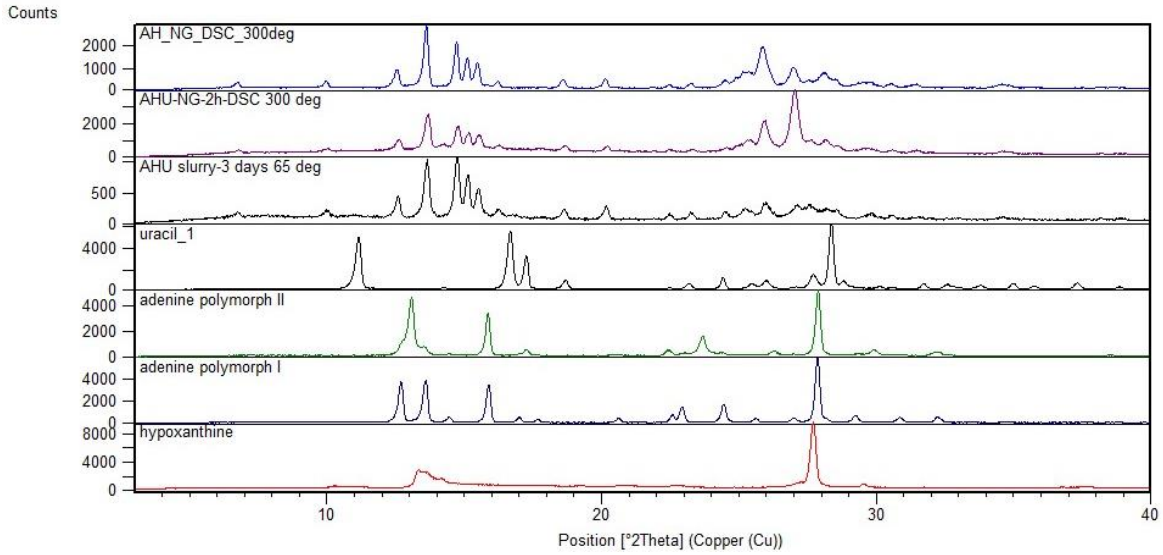

Figure 7.18. PXRD patterns for A, H, and U system.

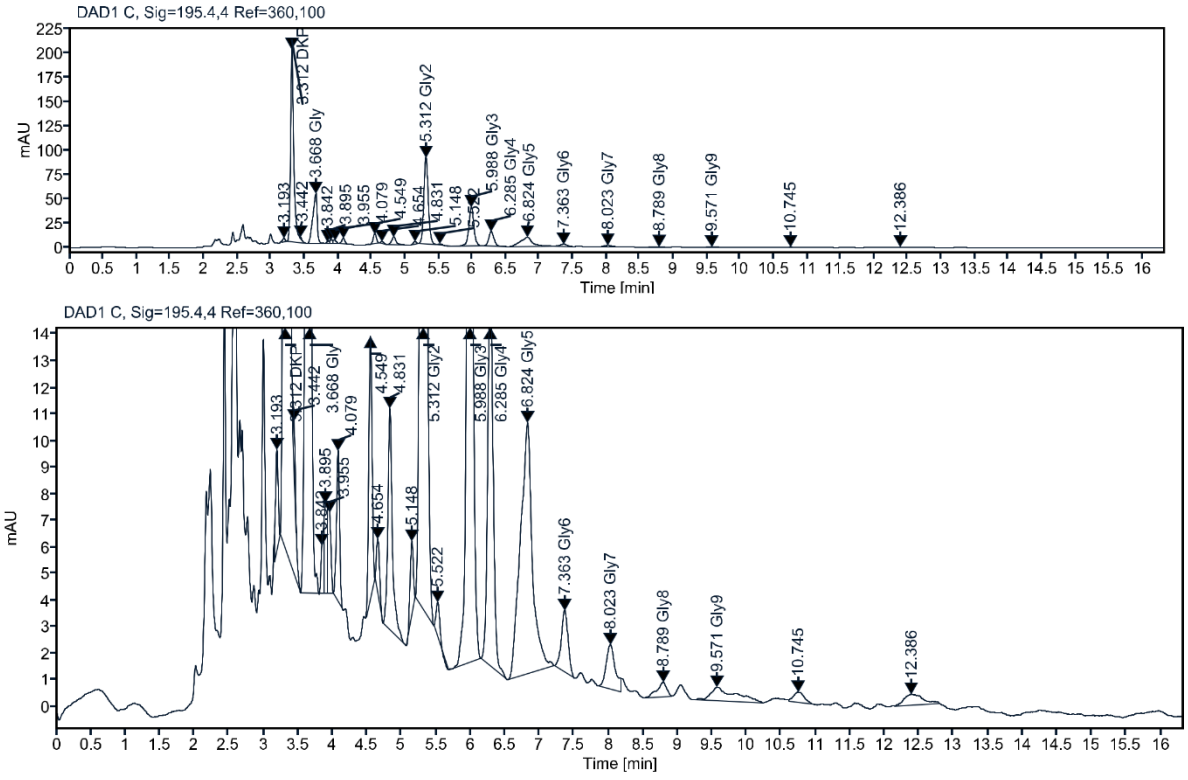

Figure 7.19. IP-HPLC chromatogram for milling Gly with SiO_2 at 100 °C.

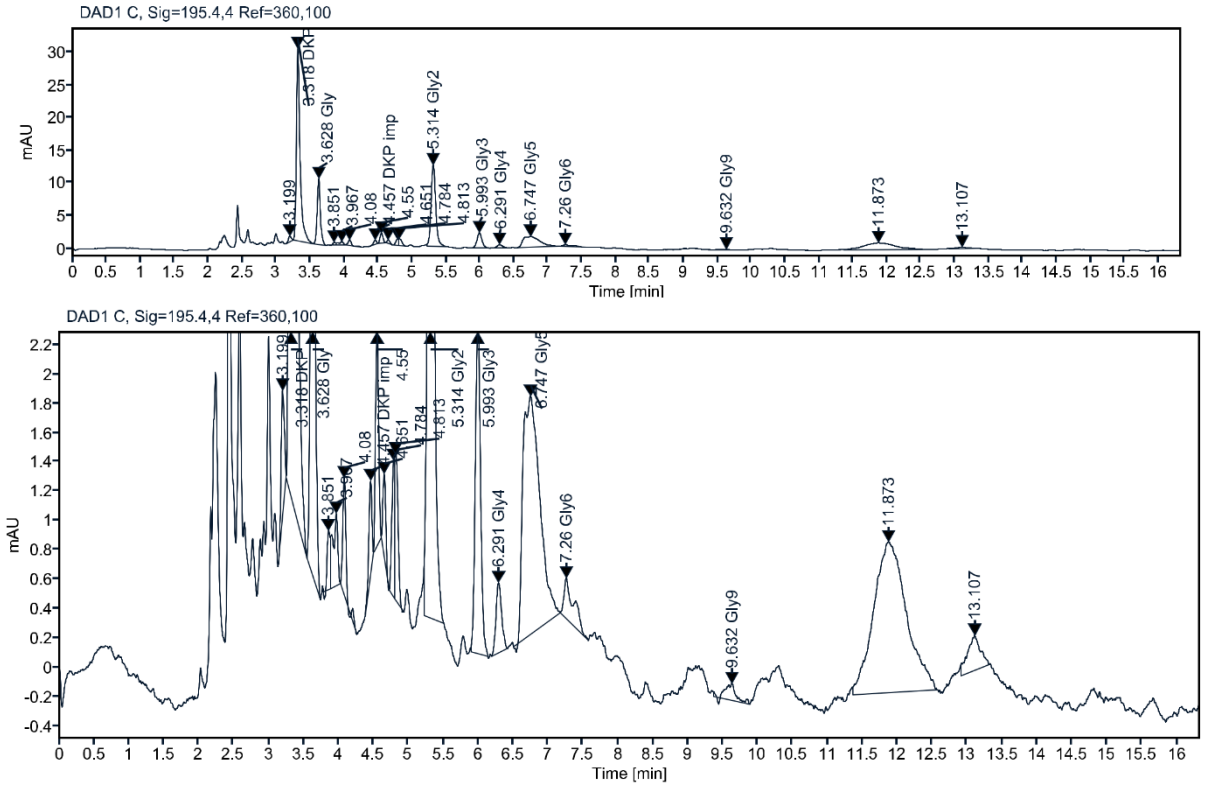

Figure 7.20. IP-HPLC chromatogram for milling Gly with montmorillonite at 100 °C.

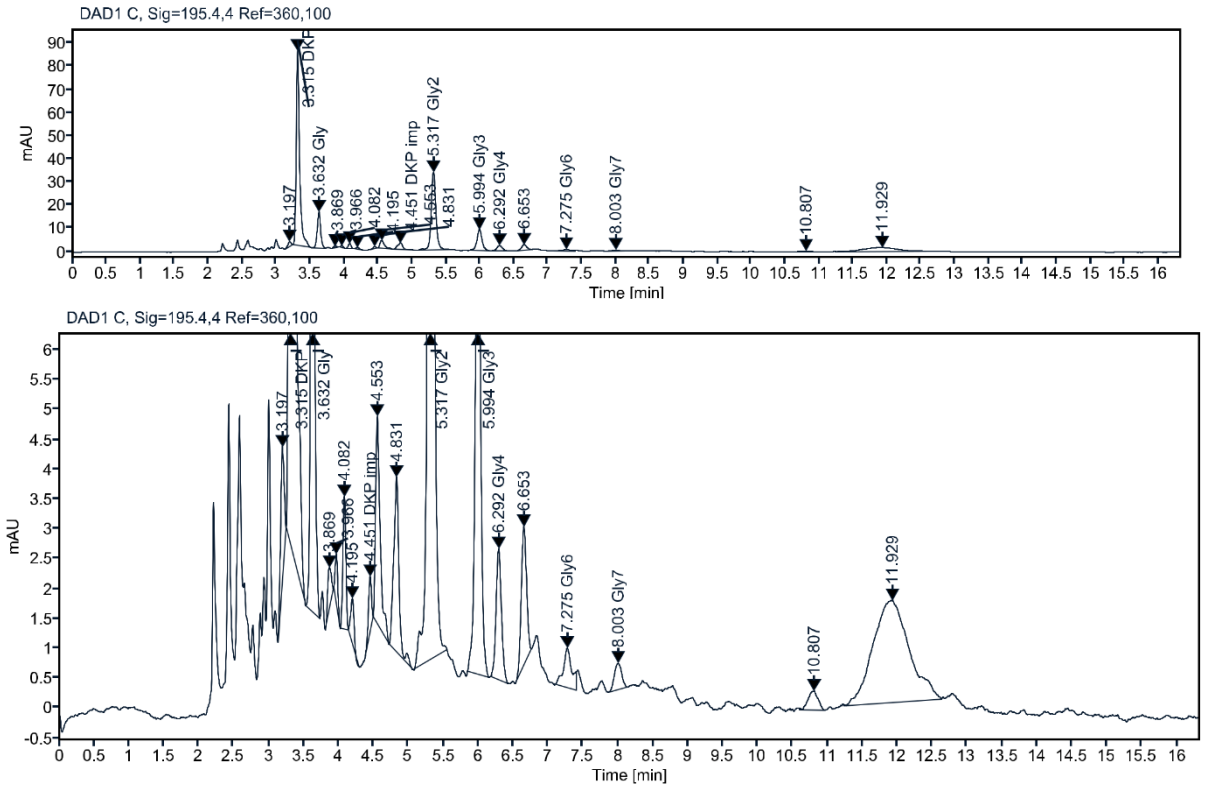

Figure 7.21. IP-HPLC chromatogram for milling Gly with mica at 100 °C.

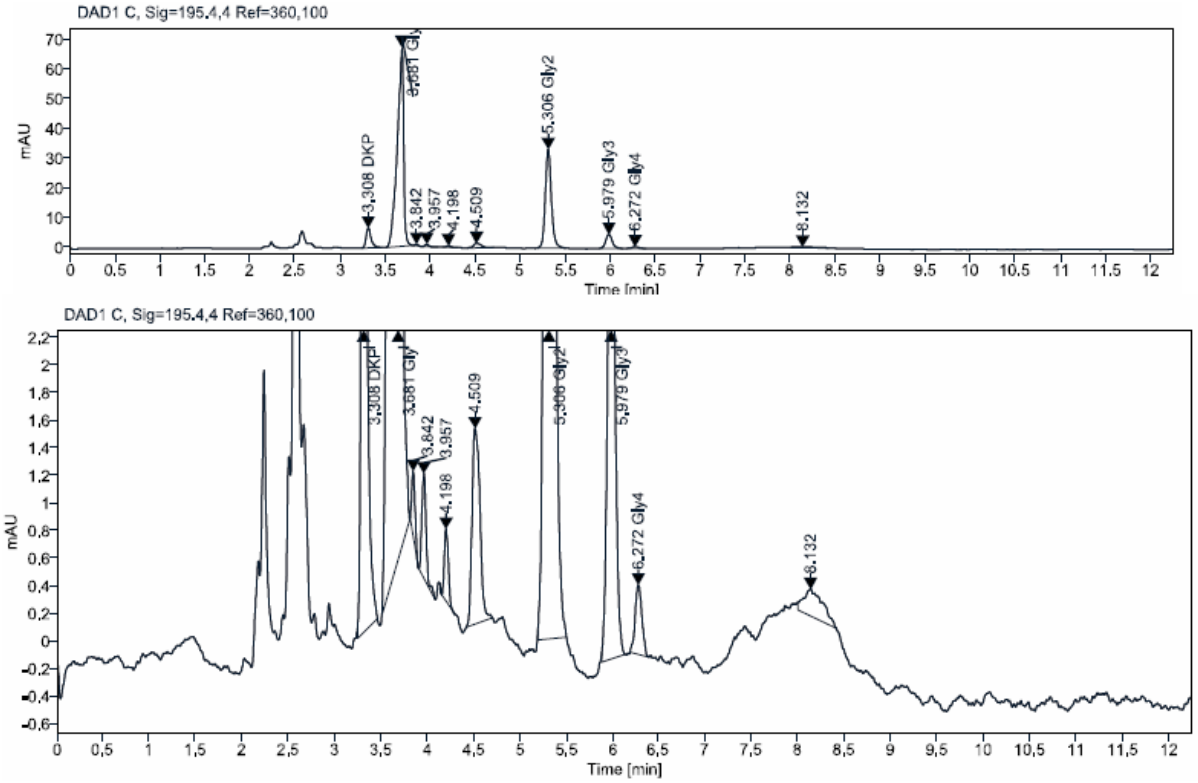

Figure 7.22. IP-HPLC chromatogram for milling Gly with TiO_2 at RT in the presence of water.

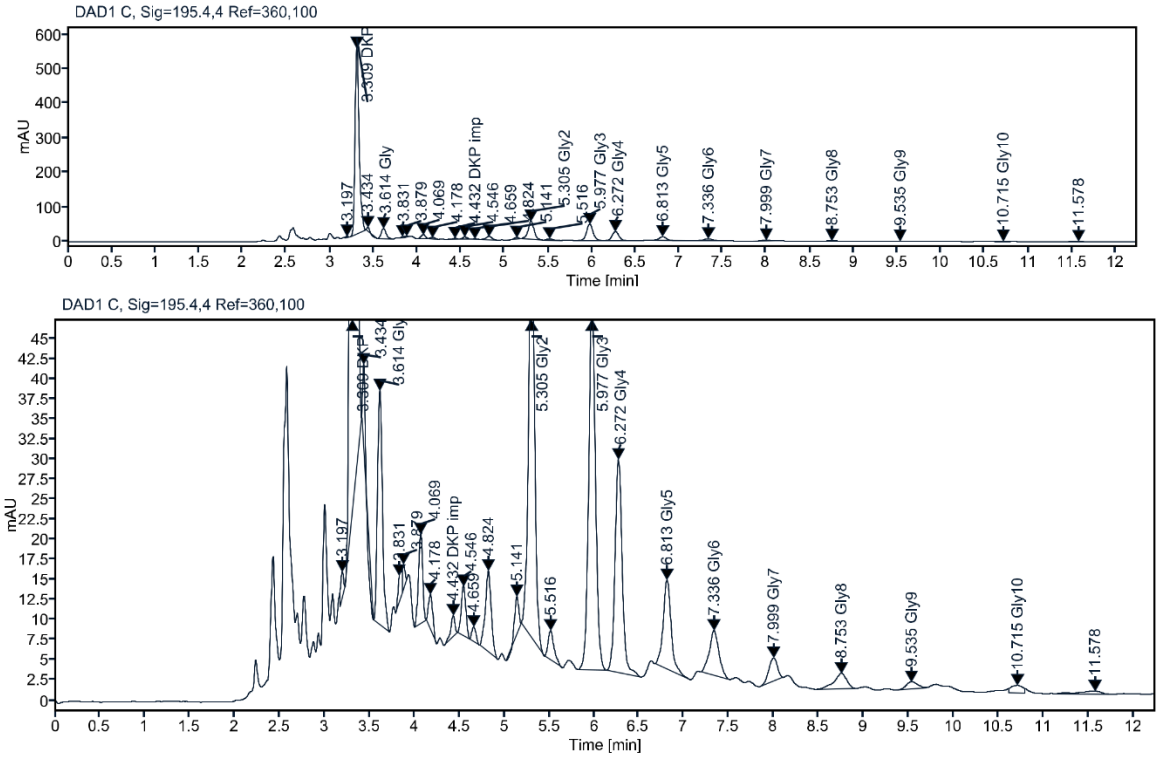

Figure 7.23. IP-HPLC chromatogram for milling Gly with TiO_2 at 130 °C in the presence of water.

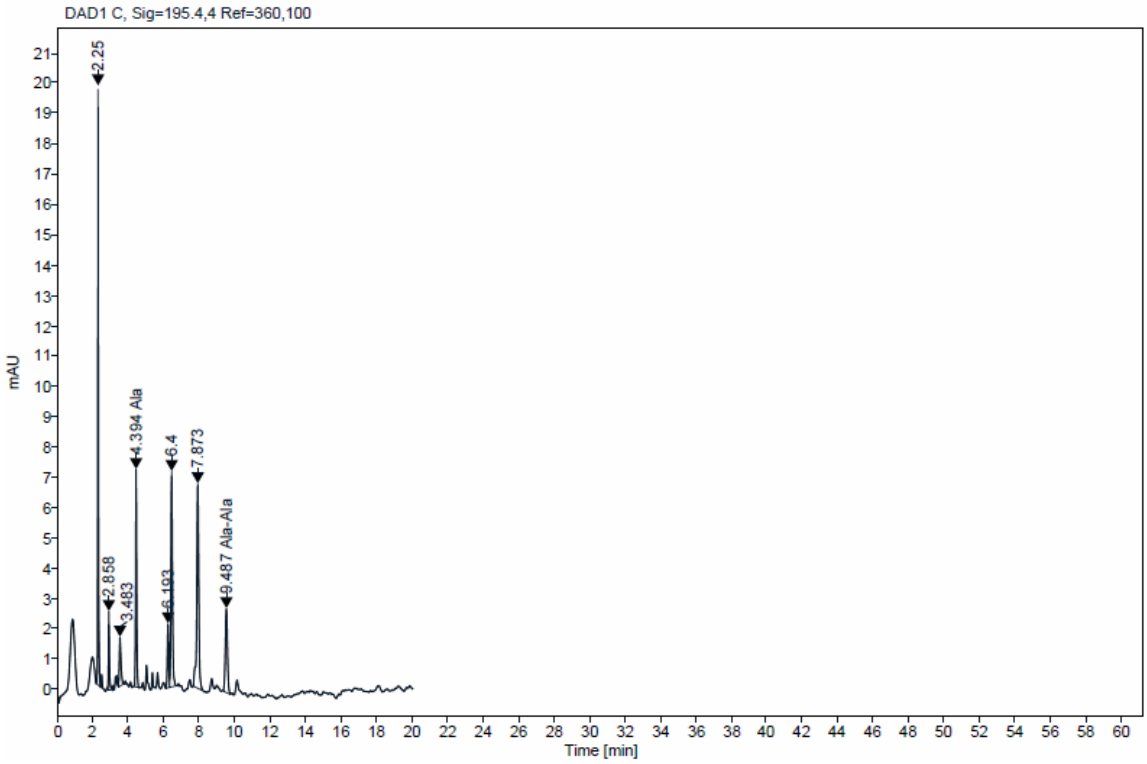

Figure 7.24. IP-HPLC chromatogram for milling Ala with TiO_2 at 130 °C.

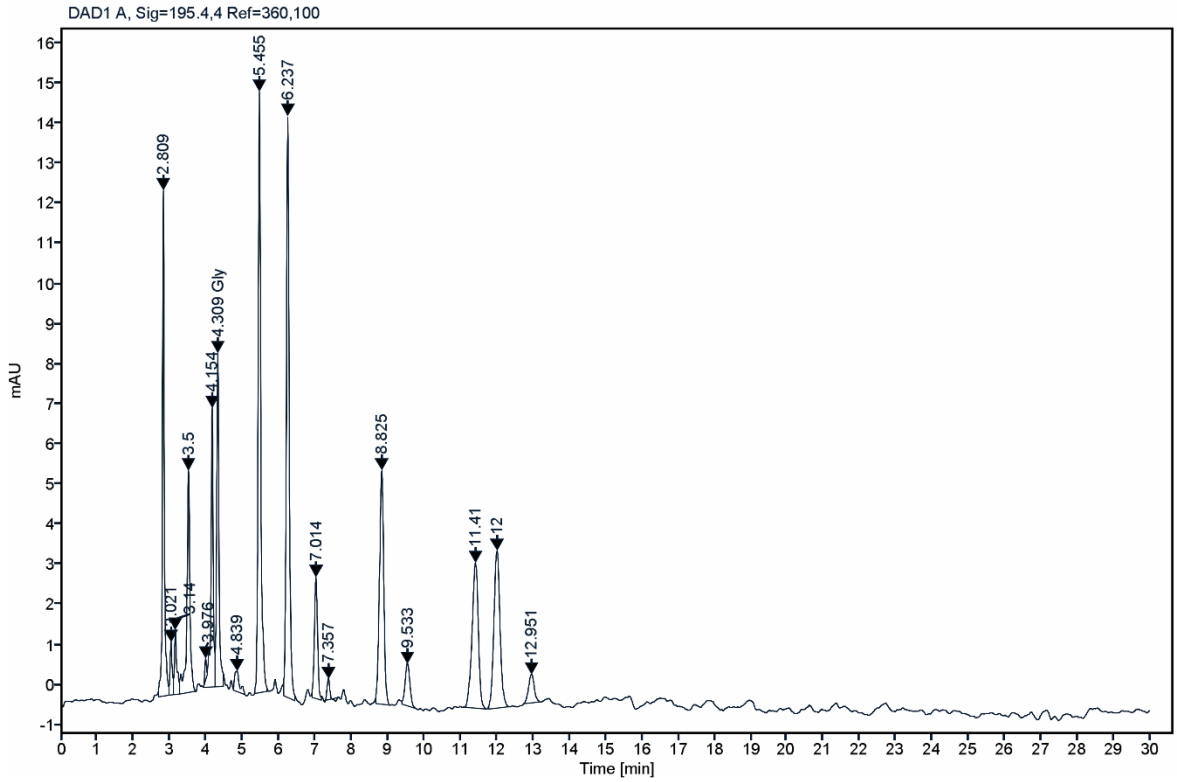

Figure 7.25. IP-HPLC chromatogram for milling Gly and Ala with TiO_2 at 100 °C.

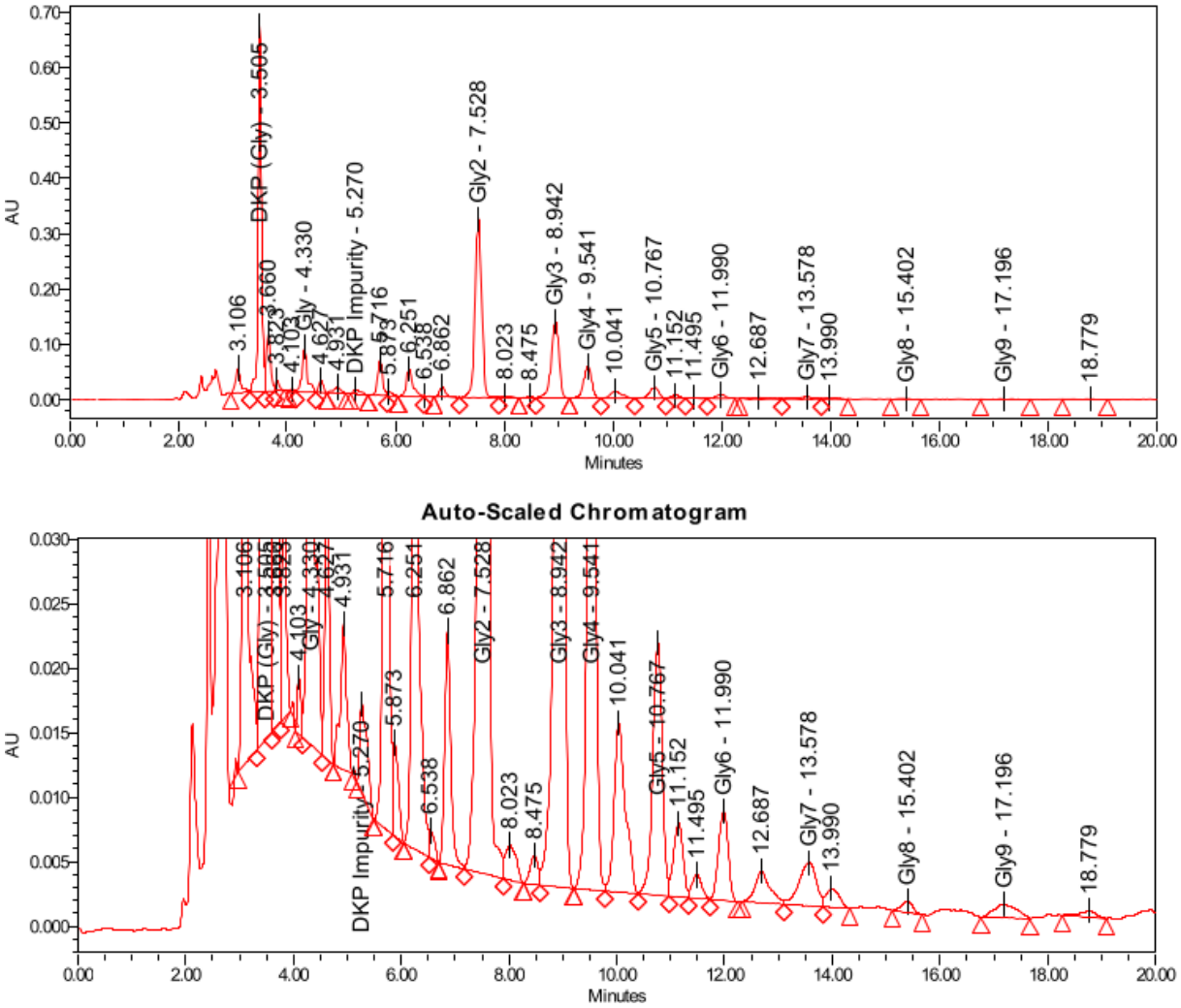

Figure 7.26. IP-HPLC chromatogram for milling Gly with SiO_2 at room temperature for 96 h and at 30 Hz.

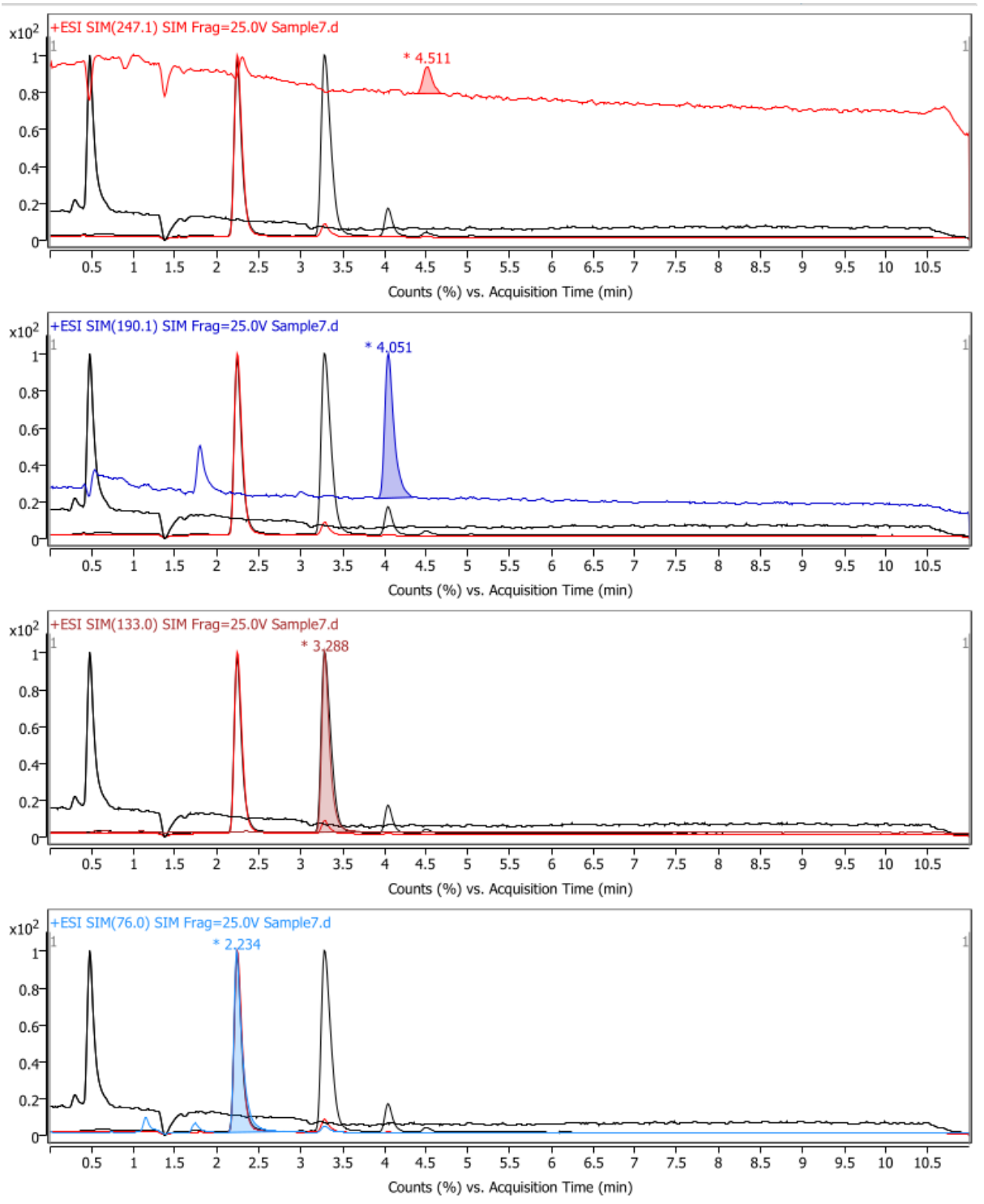

Figure 7.27. LC/MS (Triple quadrupole) chromatograms for milling Gly with SiO_2 at room temperature for 96 h and at 30 Hz.

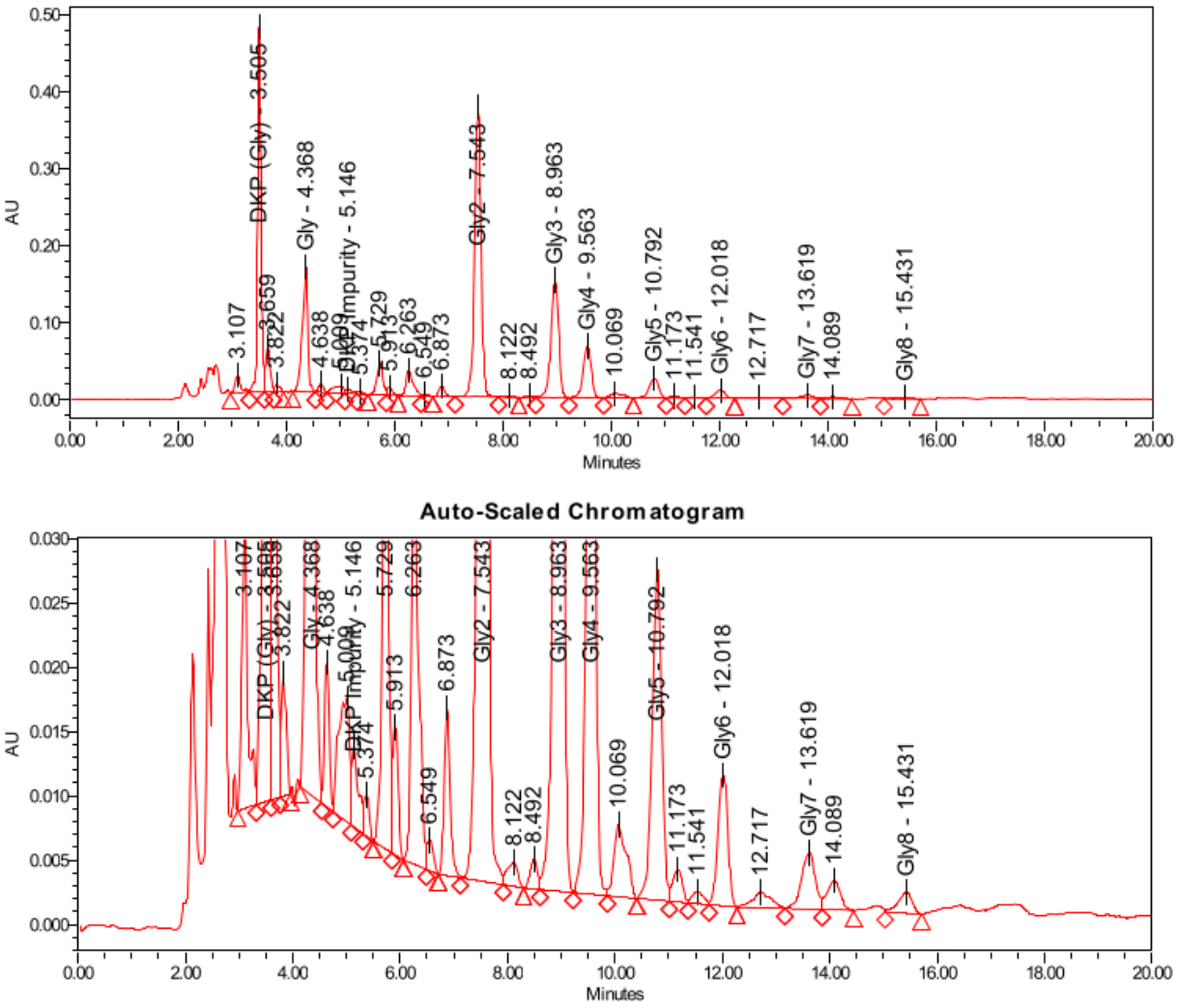

Figure 7.28. IP-HPLC chromatogram for milling Gly with SiO_2 at room temperature for 48 h and at 30 Hz.

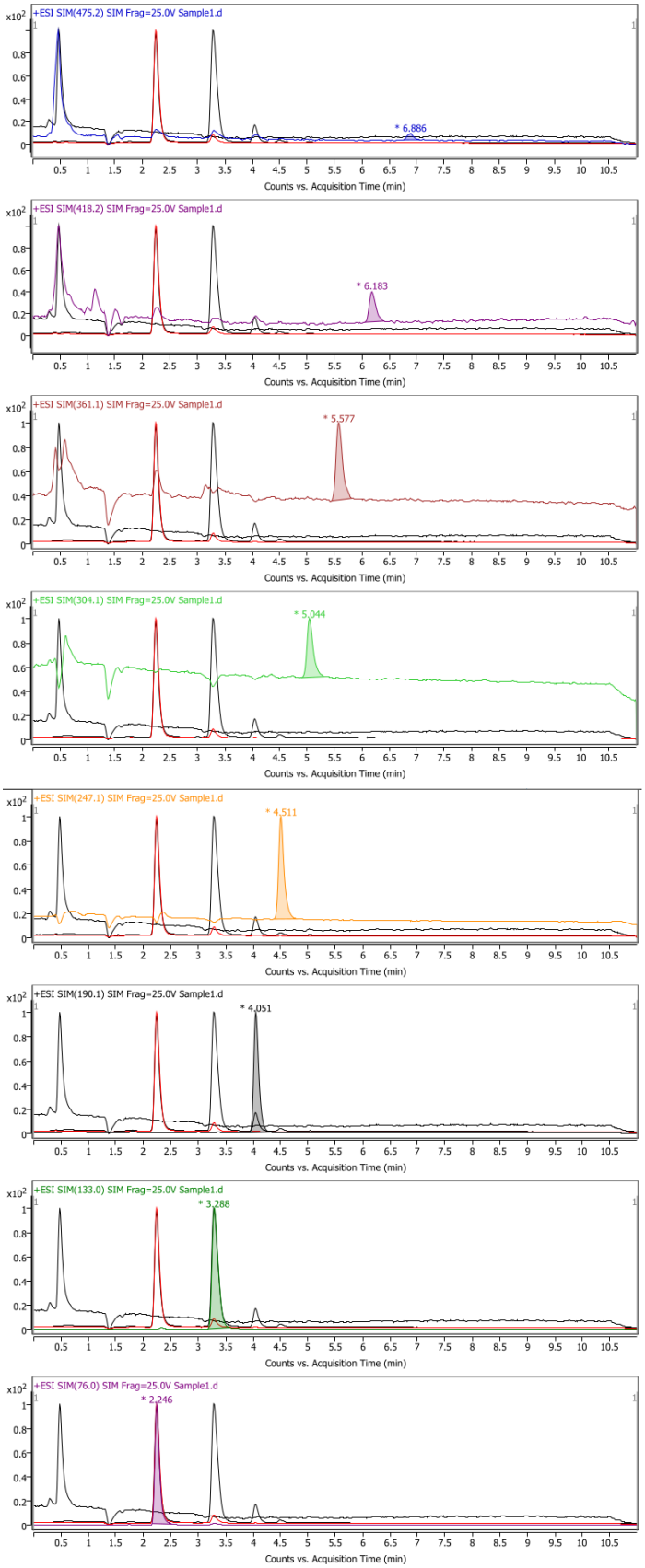

Figure 7.29. LC/MS (Triple quadrupole) chromatograms for milling Gly with SiO_2 at room temperature for 48 h and at 30 Hz.

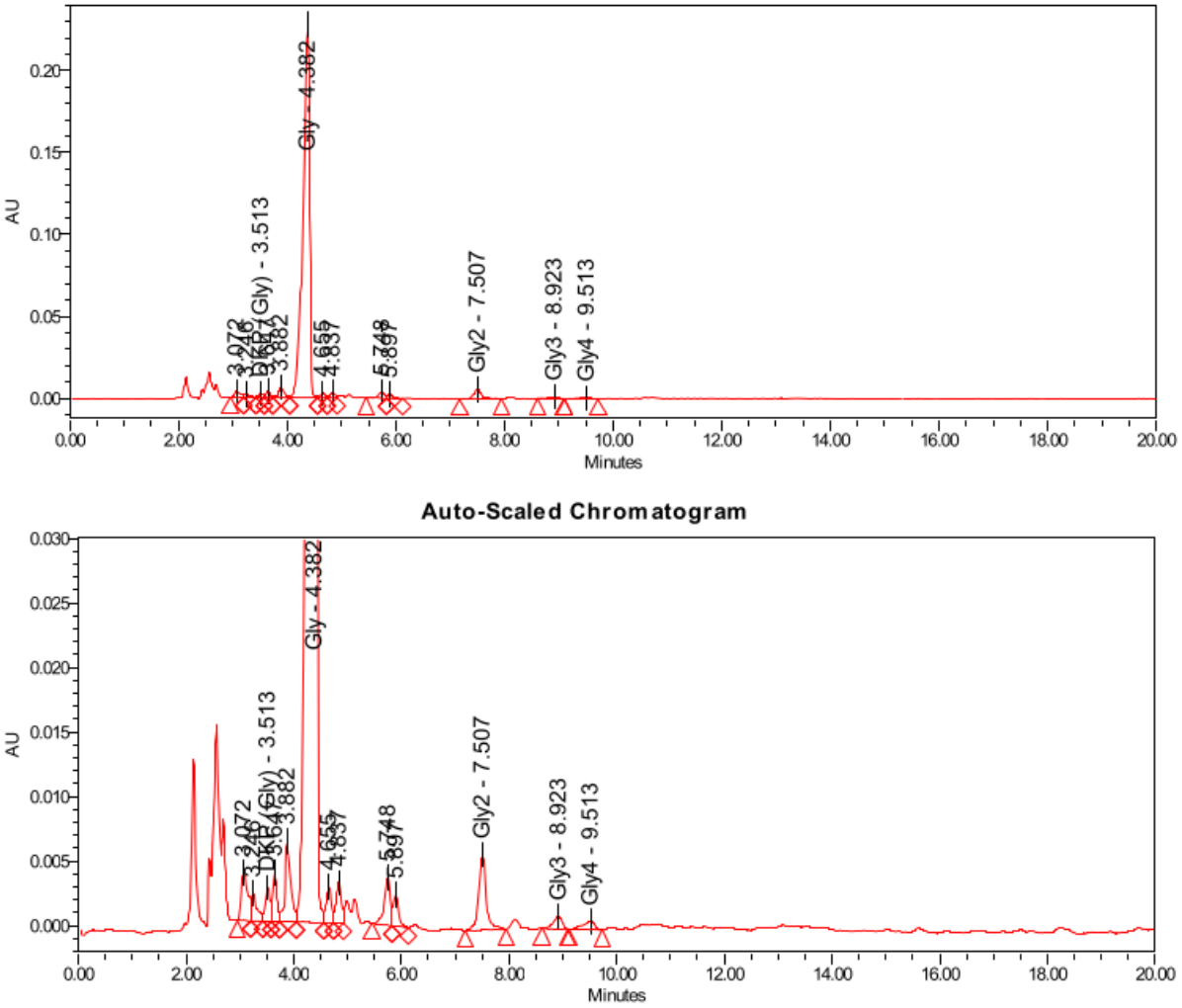

Figure 7.30. IP-HPLC chromatogram for milling Gly with SiO_2 at room temperature for 96 h and at 15 Hz.

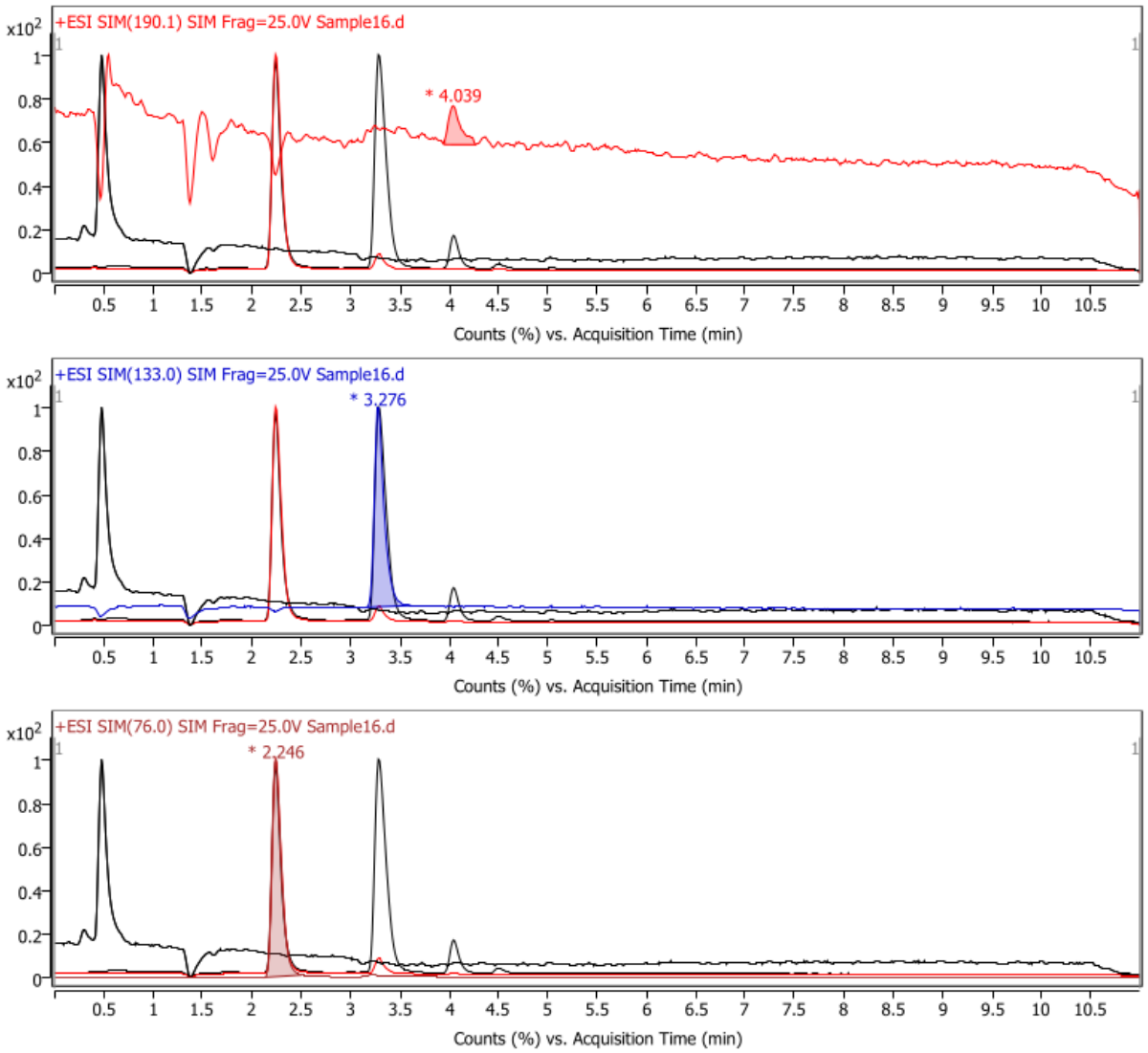

Figure 7.31. LC/MS (Triple quadrupole) chromatograms for milling Gly with SiO_2 at room temperature for 96 h and at 15 Hz.

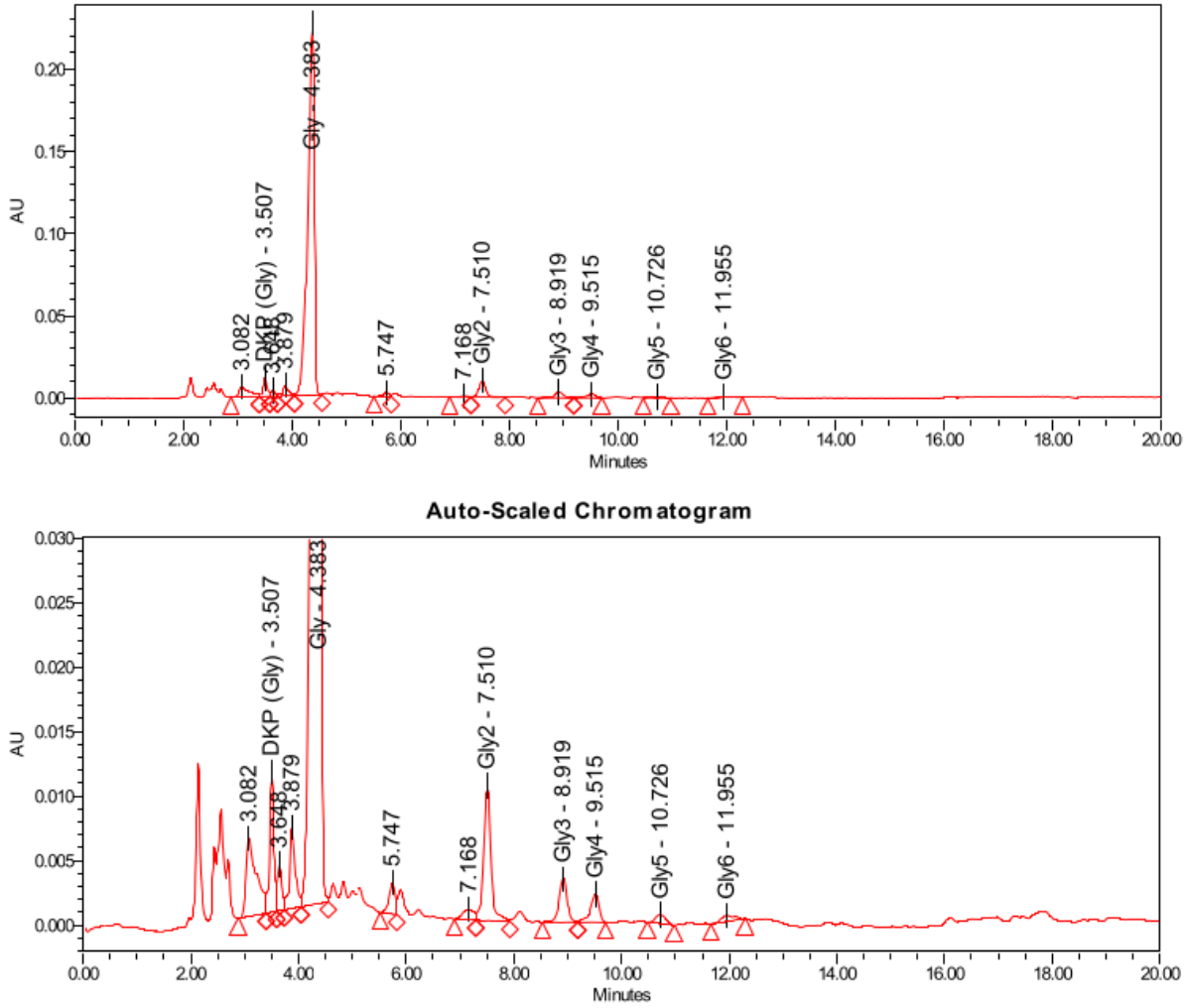

Figure 7.32. IP-HPLC chromatogram for milling Gly with SiO_2 at room temperature for 48 h and at 15 Hz.

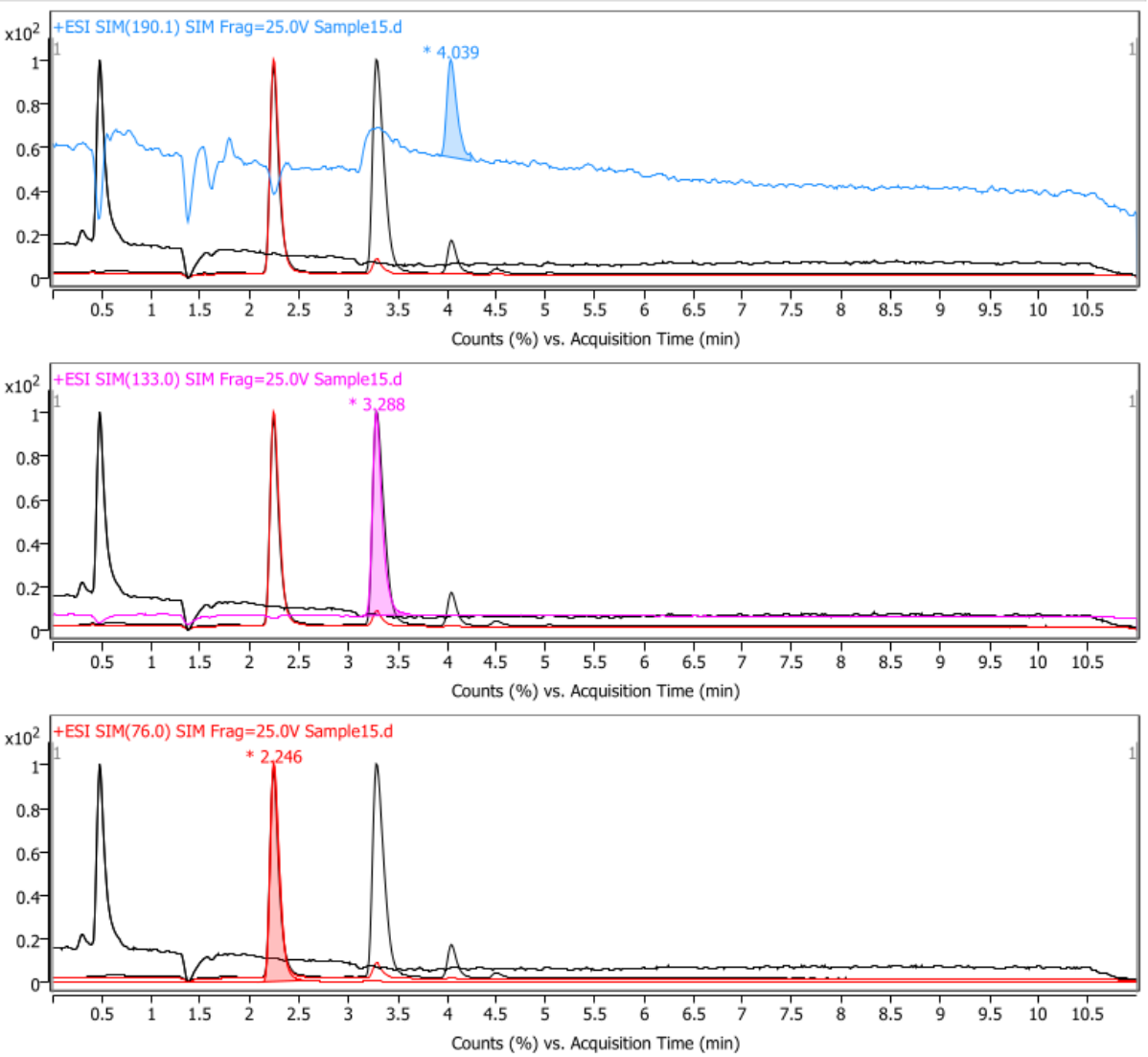

Figure 7.33. LC/MS (Triple quadrupole) chromatograms for milling Gly with SiO_2 at room temperature for 48 h and at 15 Hz.

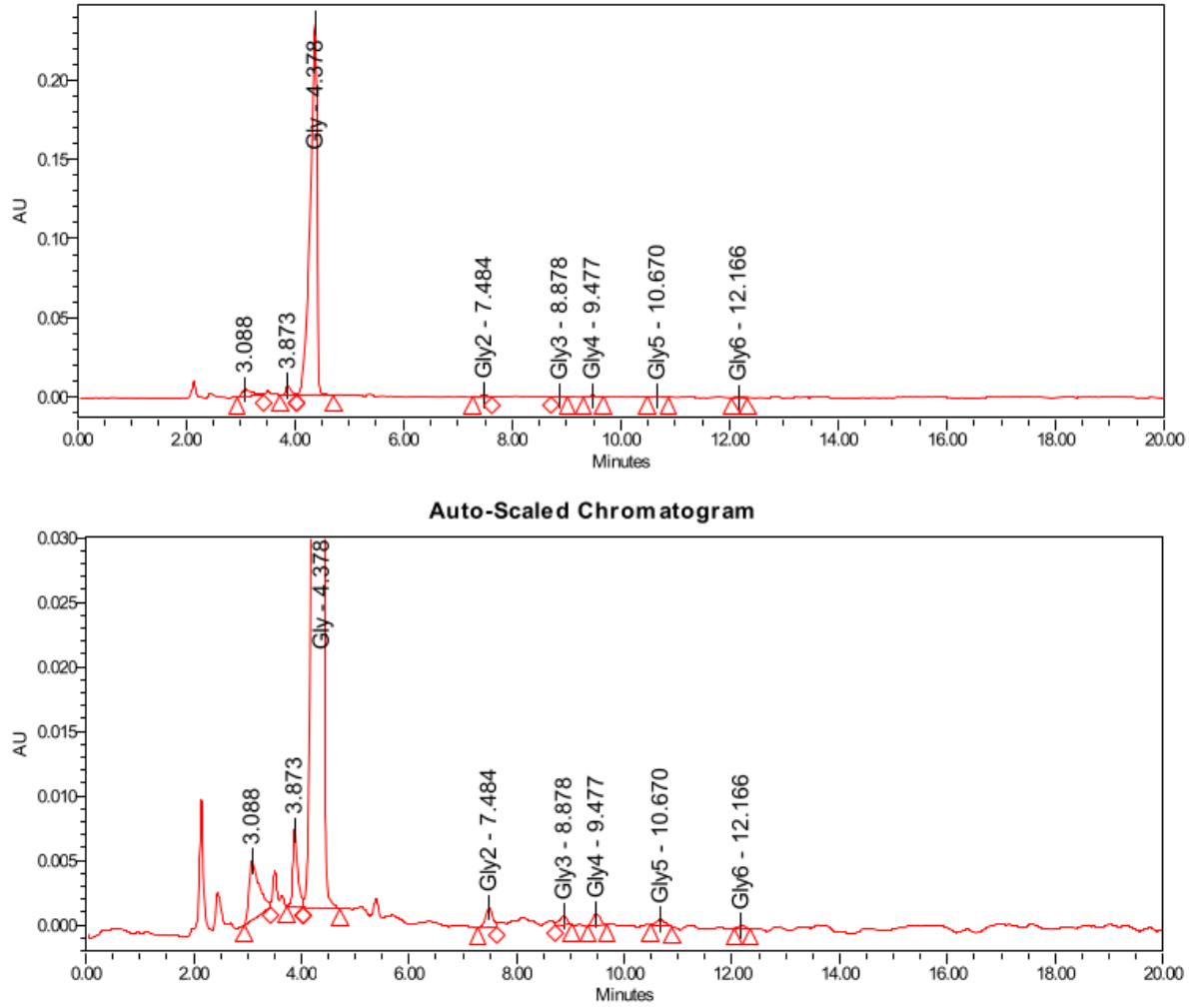

Figure 7.34. IP-HPLC chromatogram for milling Gly with SiO_2 at room temperature for 96 h and at 5 Hz.

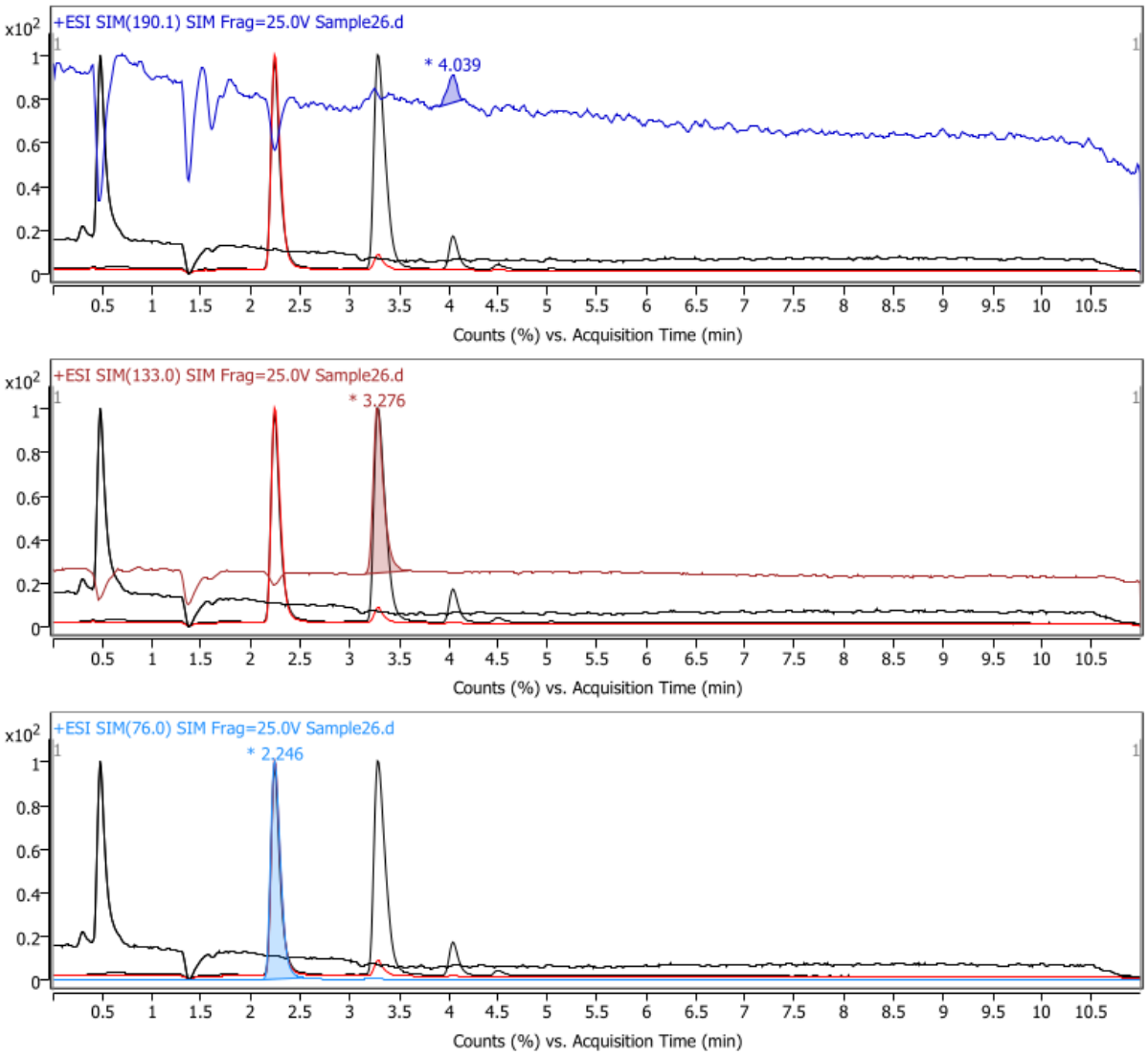

Figure 7.35. LC/MS (Triple quadrupole) chromatograms for milling Gly with SiO_2 at room temperature for 96 h and at 5 Hz.

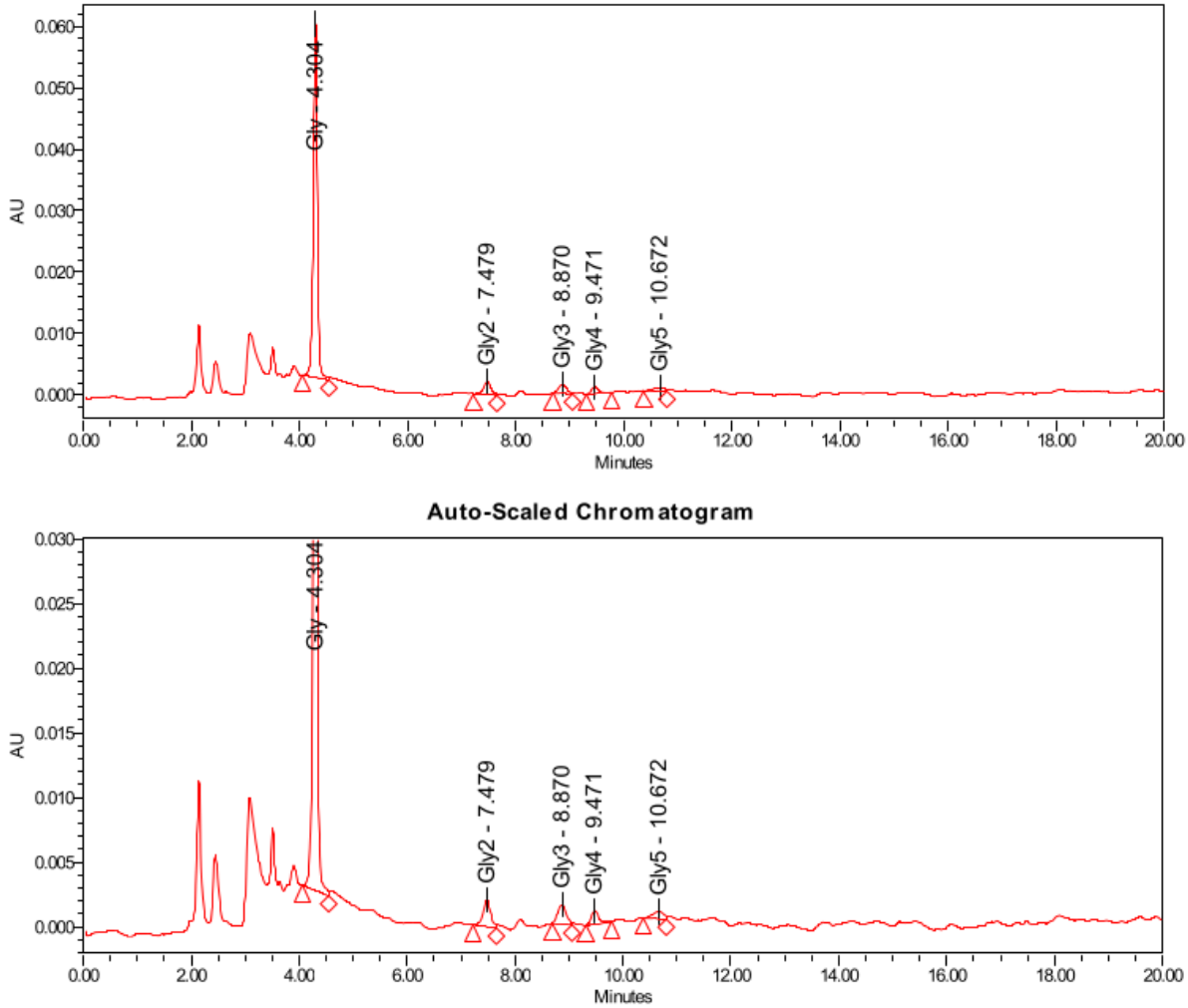

Figure 7.36. IP-HPLC chromatogram for milling Gly with SiO_2 at room temperature for 48 h and at 5 Hz.

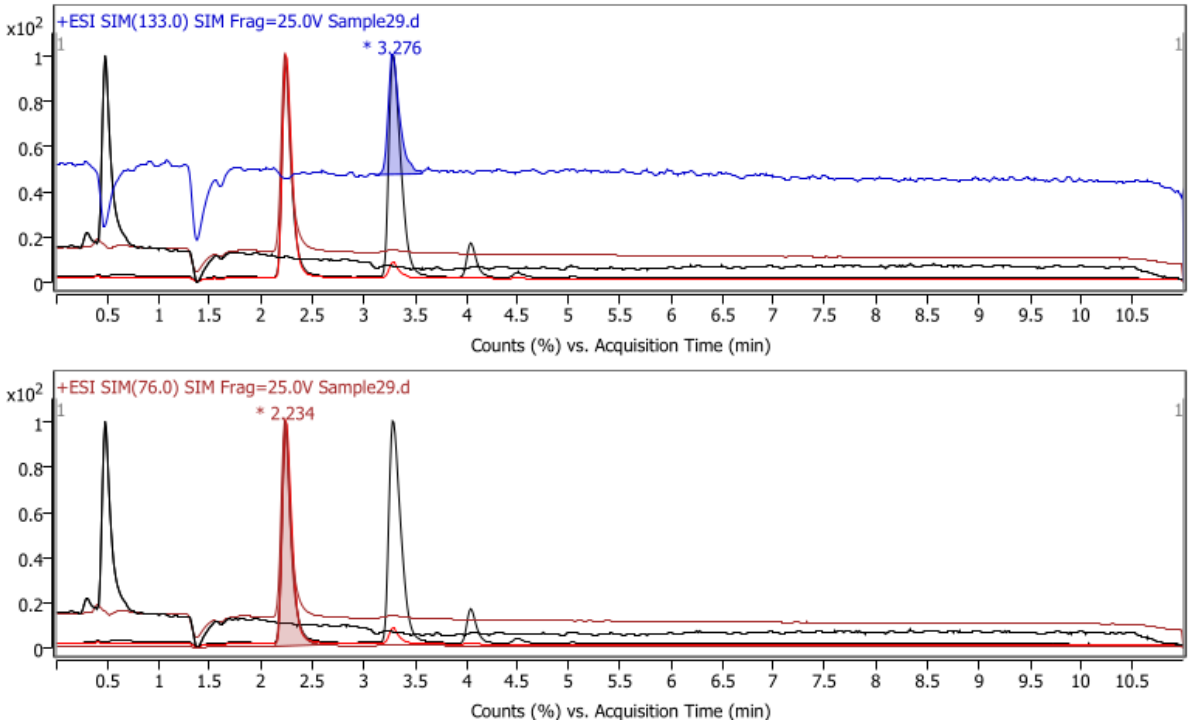

Figure 7.37. LC/MS (Triple quadrupole) chromatograms for milling Gly with SiO_2 at room temperature for 48 h and at 5 Hz.

Figure 7.38. Photograph showing three stones picked up in front of Ruđer Bošković Institute by Prof. José G. Hernández**.** These stones were used as a milling media for milling Gly with TiO_2 for 18h at 20 Hz and ambient temperature. Gly oligomers up to Gly_4 were formed in this way.

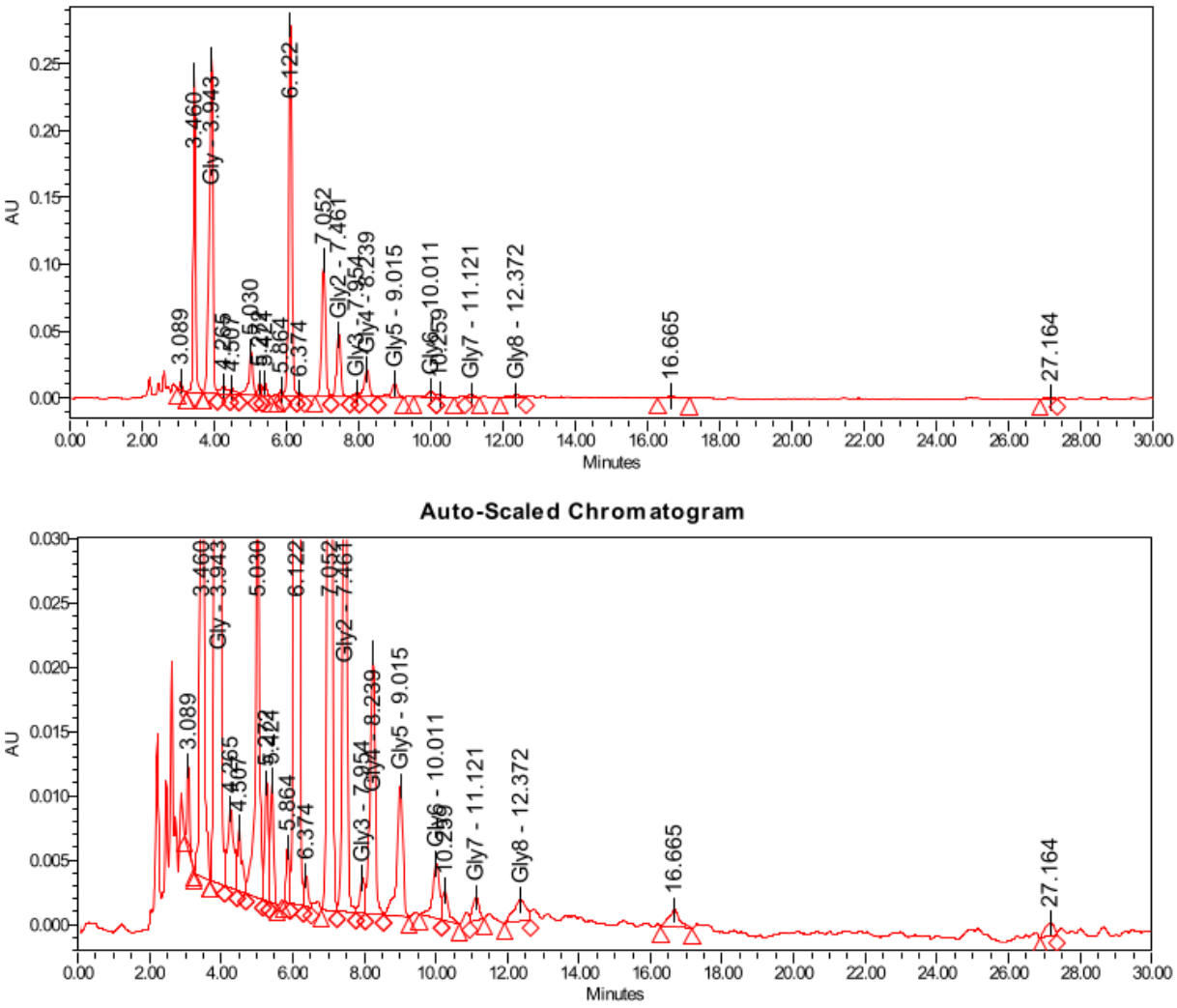

Figure 7.39. IP-HPLC chromatogram for milling Gly with SiO_2 in N_2 atmosphere at room temperature for 16 h and at 30 Hz. Milling under N_2 atmosphere was achieved by loading the jar in glovebox under N_2 atmosphere, closing it tightly and opening only before the chromatographic and mass spectrometry analysis.

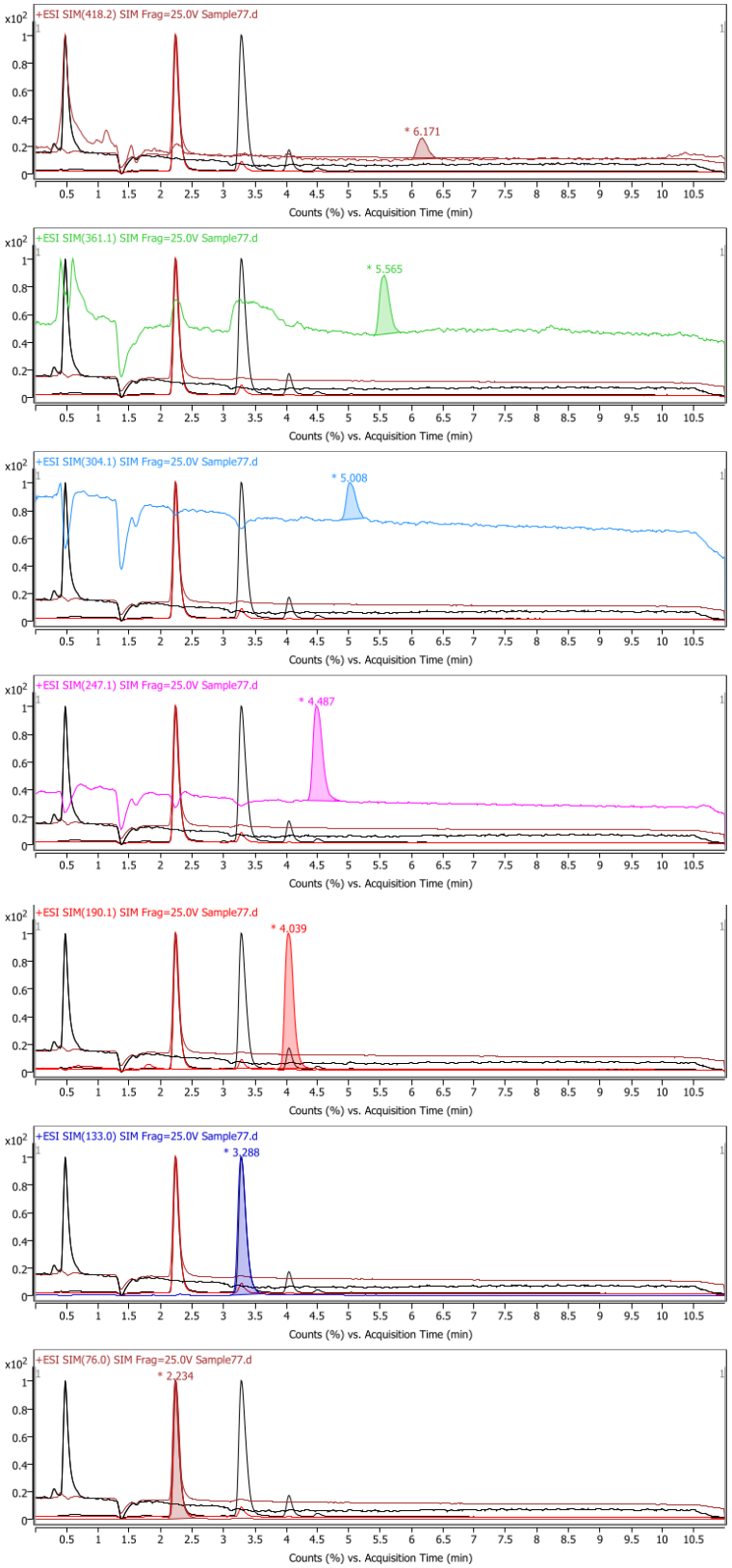

Figure 7.40. LC/MS (Triple quadrupole) chromatograms for milling Gly with SiO_2 in N_2 atmosphere at room temperature for 16 h and at 30 Hz. Milling under N_2 atmosphere was achieved by loading the jar in glovebox under N_2 atmosphere, closing it tightly and opening only before the chromatographic and mass spectrometry analysis.

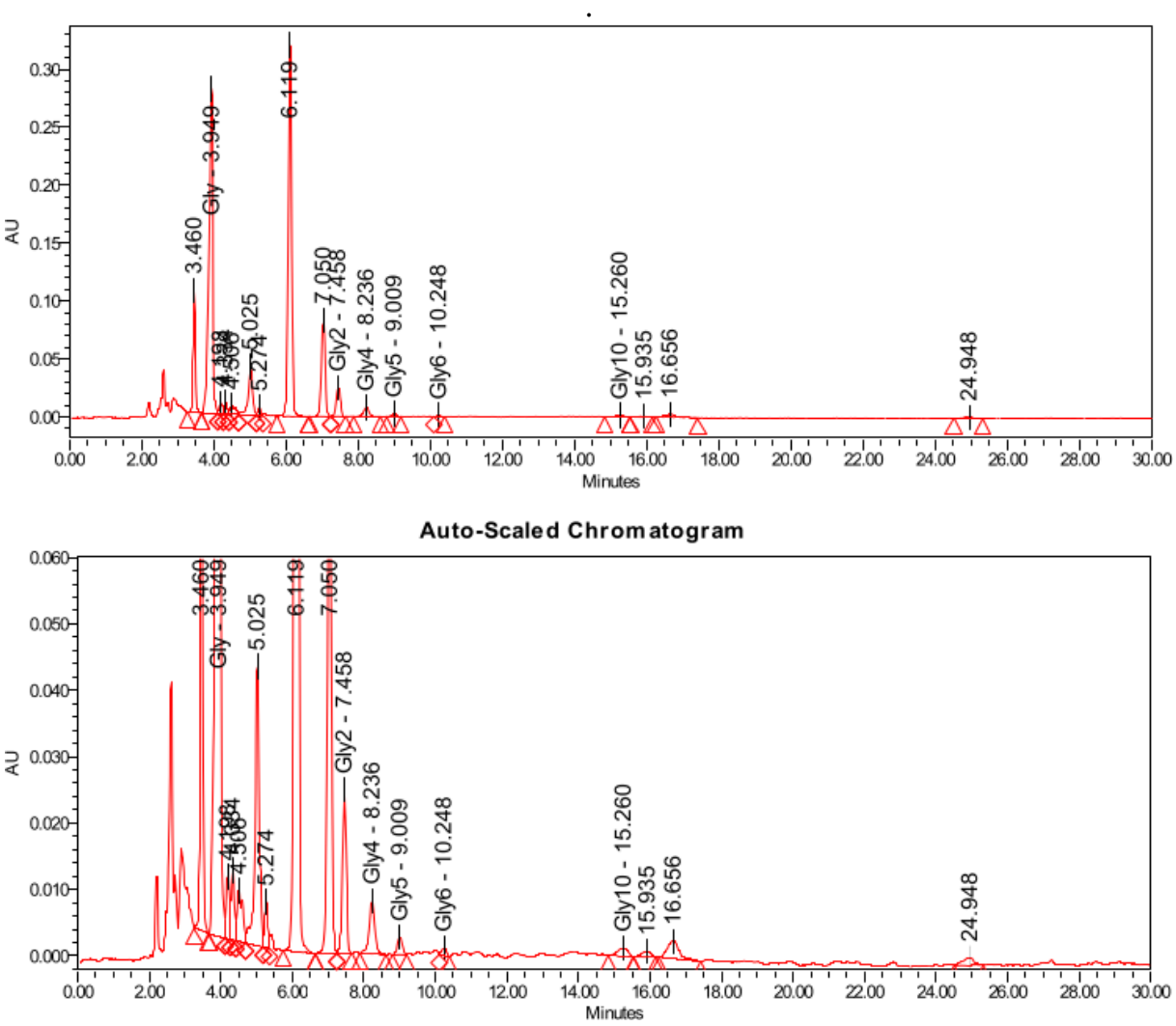

Figure 7.41. IP-HPLC chromatogram for milling Gly with SiO_2 in CO_2 atmosphere at room temperature for 16 h and at 30 Hz. Milling under CO_2 atmosphere was achieved by loading the jar containing the reaction mixture in glovebox with dry ice (solid CO_2) to the top, waiting for 30 s and closing it tightly. The jar was opened only before the chromatographic and mass spectrometry analysis. Note: due to high pressure inside, the jar had to be opened with special equipment and slowly vented.

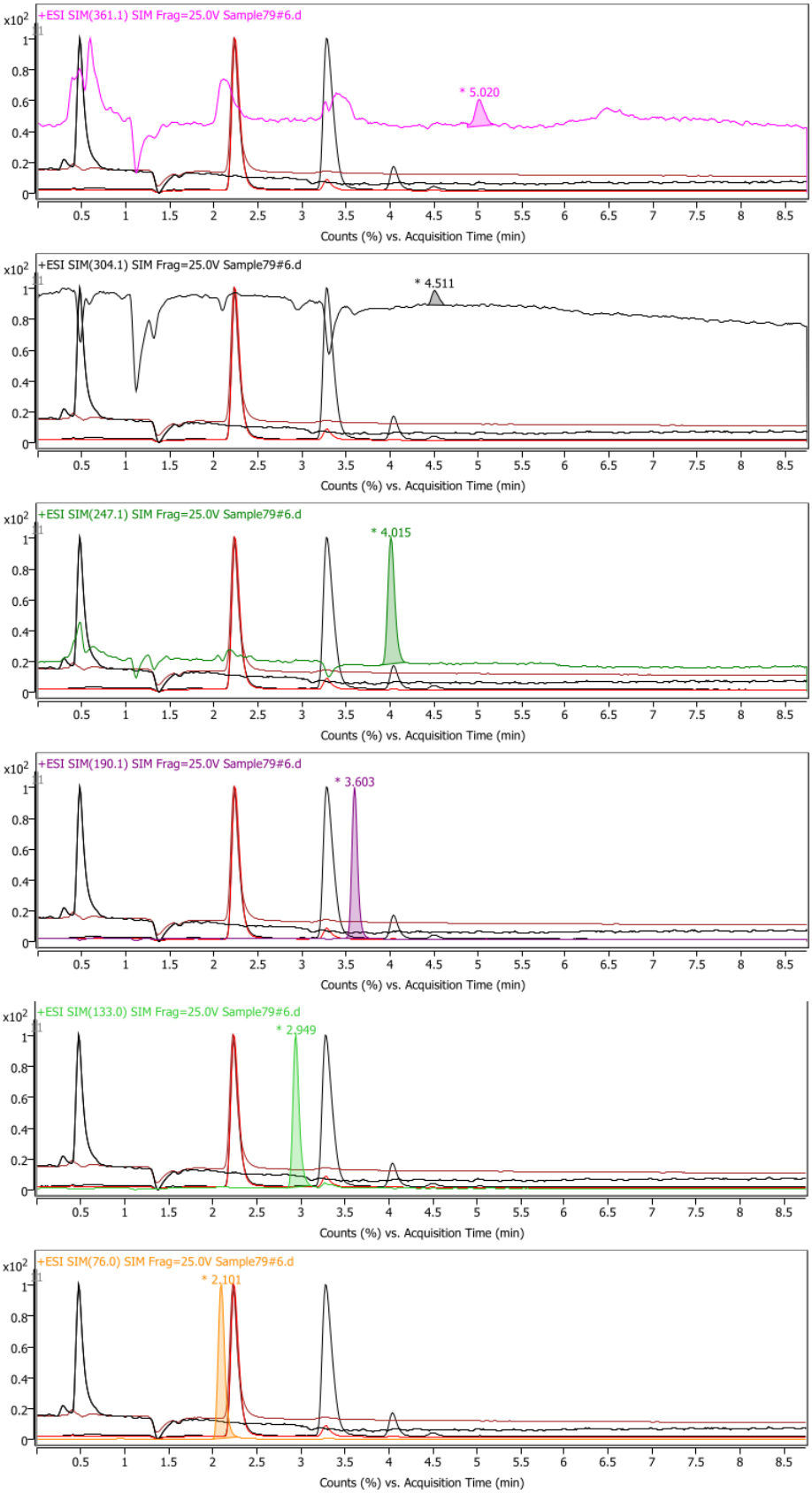

Figure 7.42. LC/MS (Triple quadrupole) chromatograms for milling Gly with SiO_2 in CO_2 atmosphere at room temperature for 16 h and at 30 Hz.

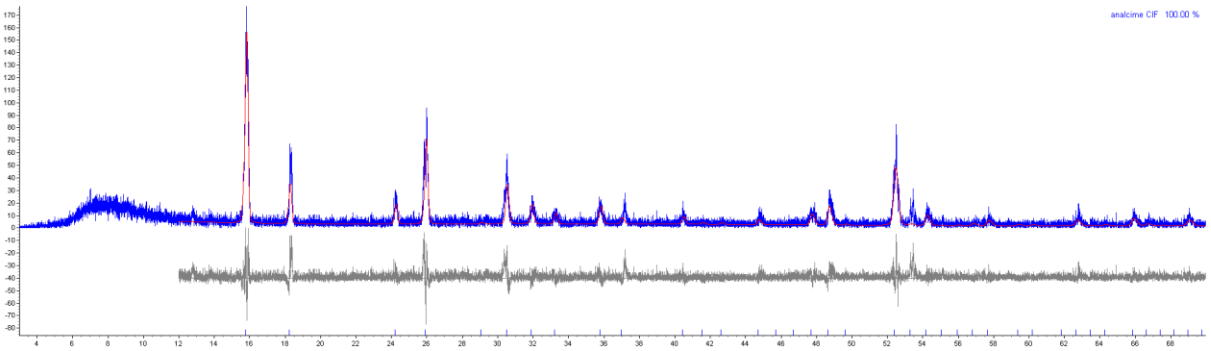

Figure 7.43. Rietveld refinement of analcime PXRD pattern and analcime CIF taken from the literature.[153]

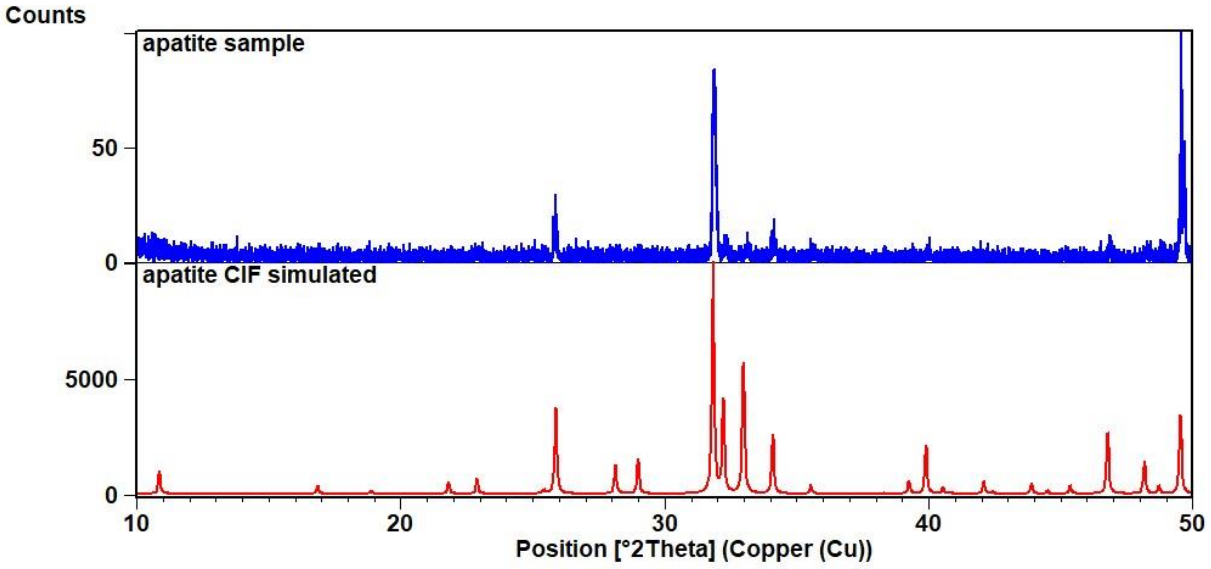

Figure 7.44. PXRD comparison of apatite sample and simulated from the CIF.[154]

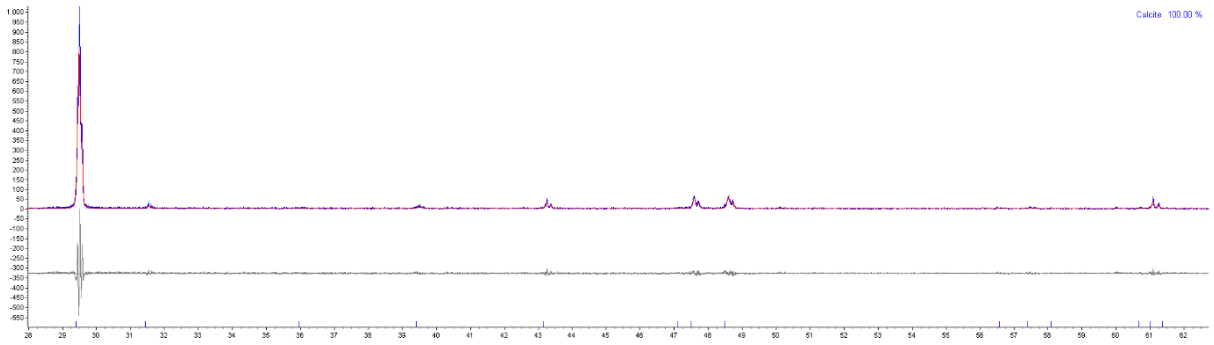

Figure 7.45. Rietveld refinement of calcite PXRD pattern and calcite CIF taken from the literature.[155]

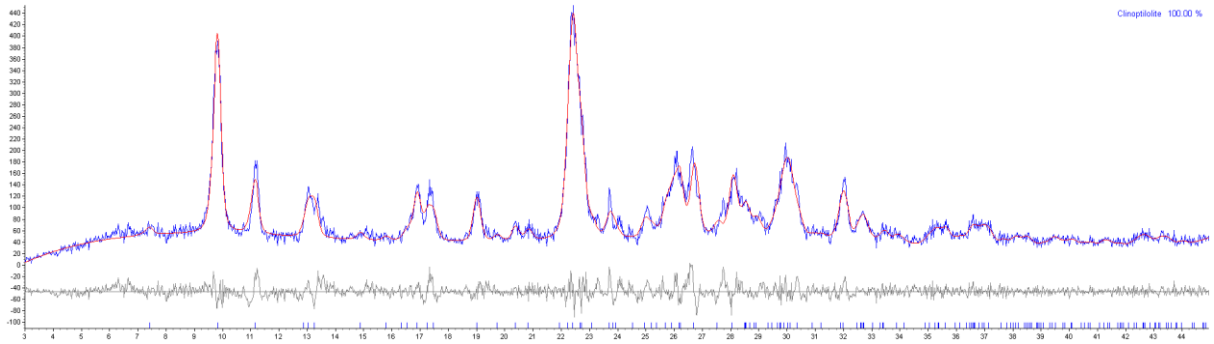

Figure 7.46. Rietveld refinement of clinoptilolite PXRD pattern and clinoptilolite CIF taken from the literature.[156]

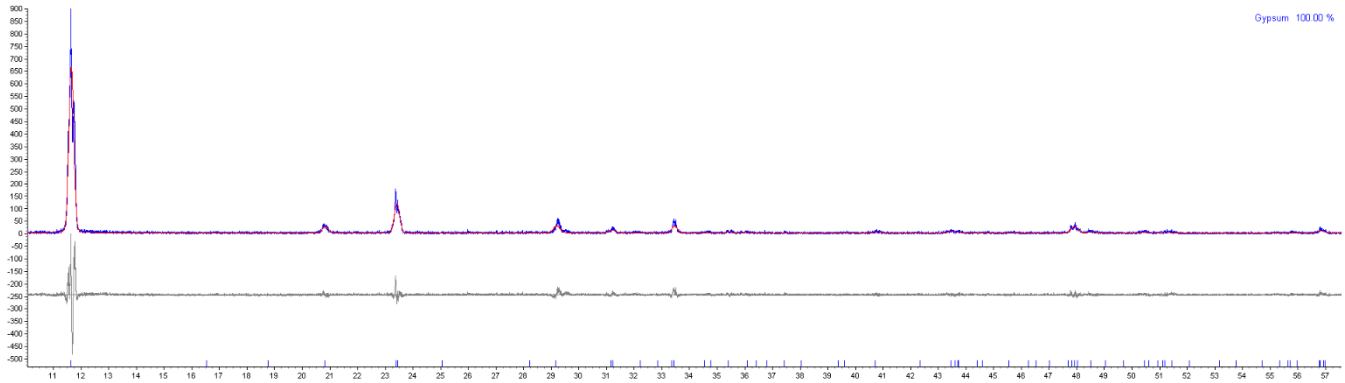

Figure 7.47. Rietveld refinement of gypsum PXRD pattern and gypsum CIF taken from the literature.[157]

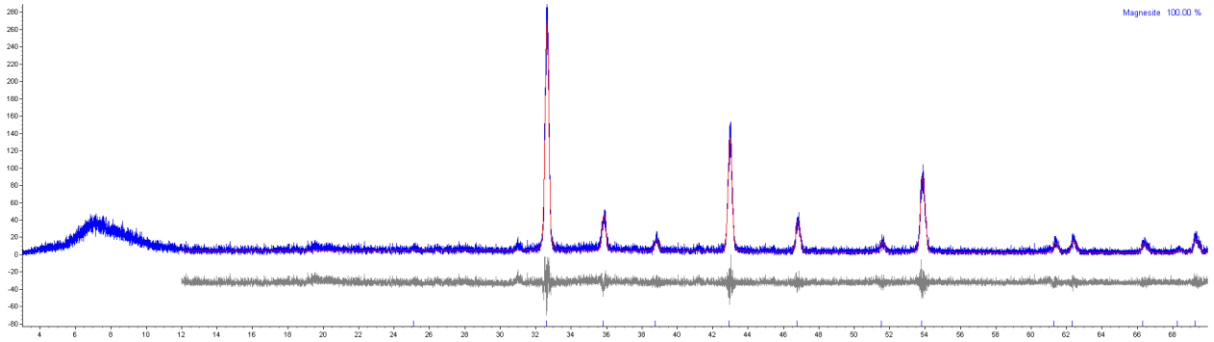

Figure 7.48. Rietveld refinement of magnesite PXRD pattern and magnesite CIF taken from the literature.[155]

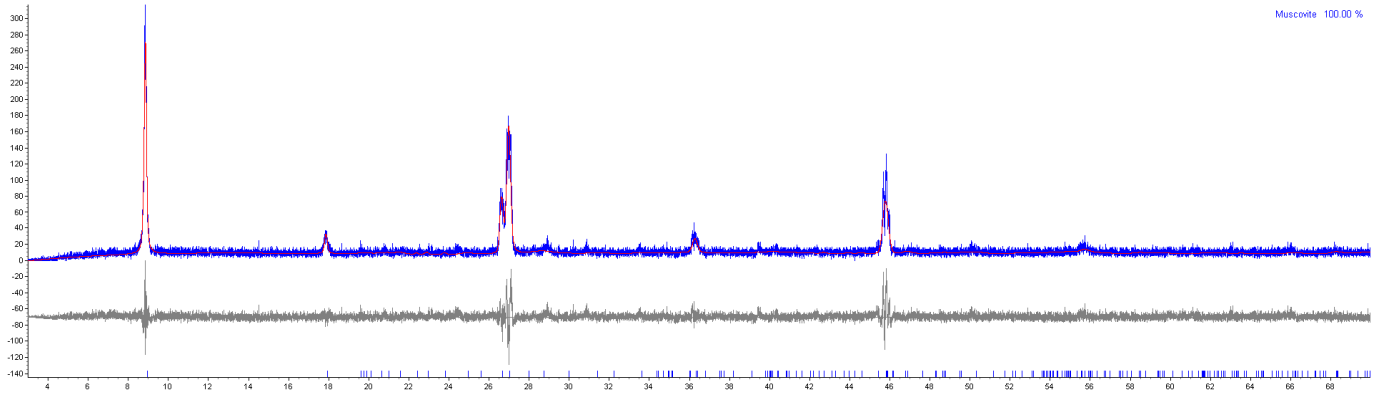

Figure 7.49. Rietveld refinement of muskovite PXRD pattern and muskovite CIF taken from the literature.[158]

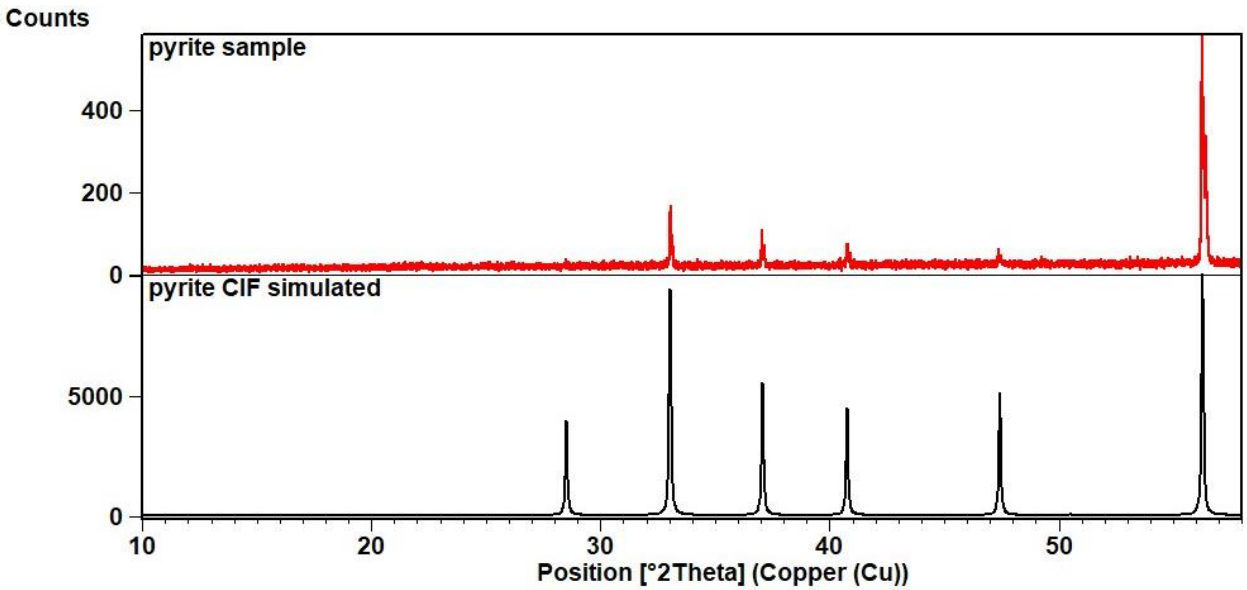

Figure 7.50. PXRD comparison of pyrite sample and simulated from the CIF.[159]

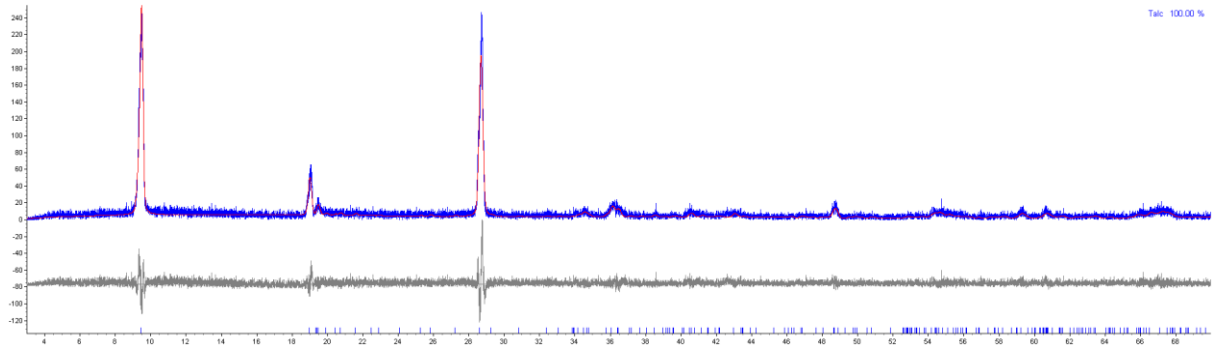

Figure 7.51. Rietveld refinement of talc PXRD pattern and talc CIF taken from the literature.[160]

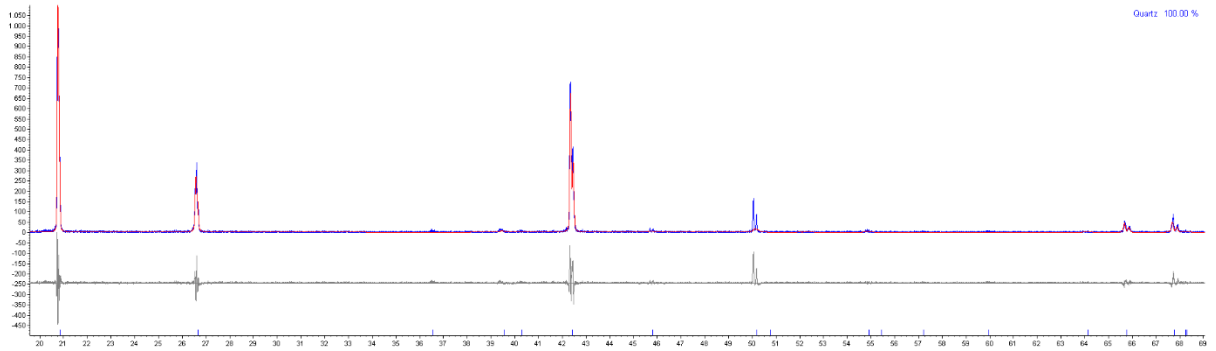

Figure 7.52. Rietveld refinement of quartz PXRD pattern and quartz CIF taken from the literature.[161]

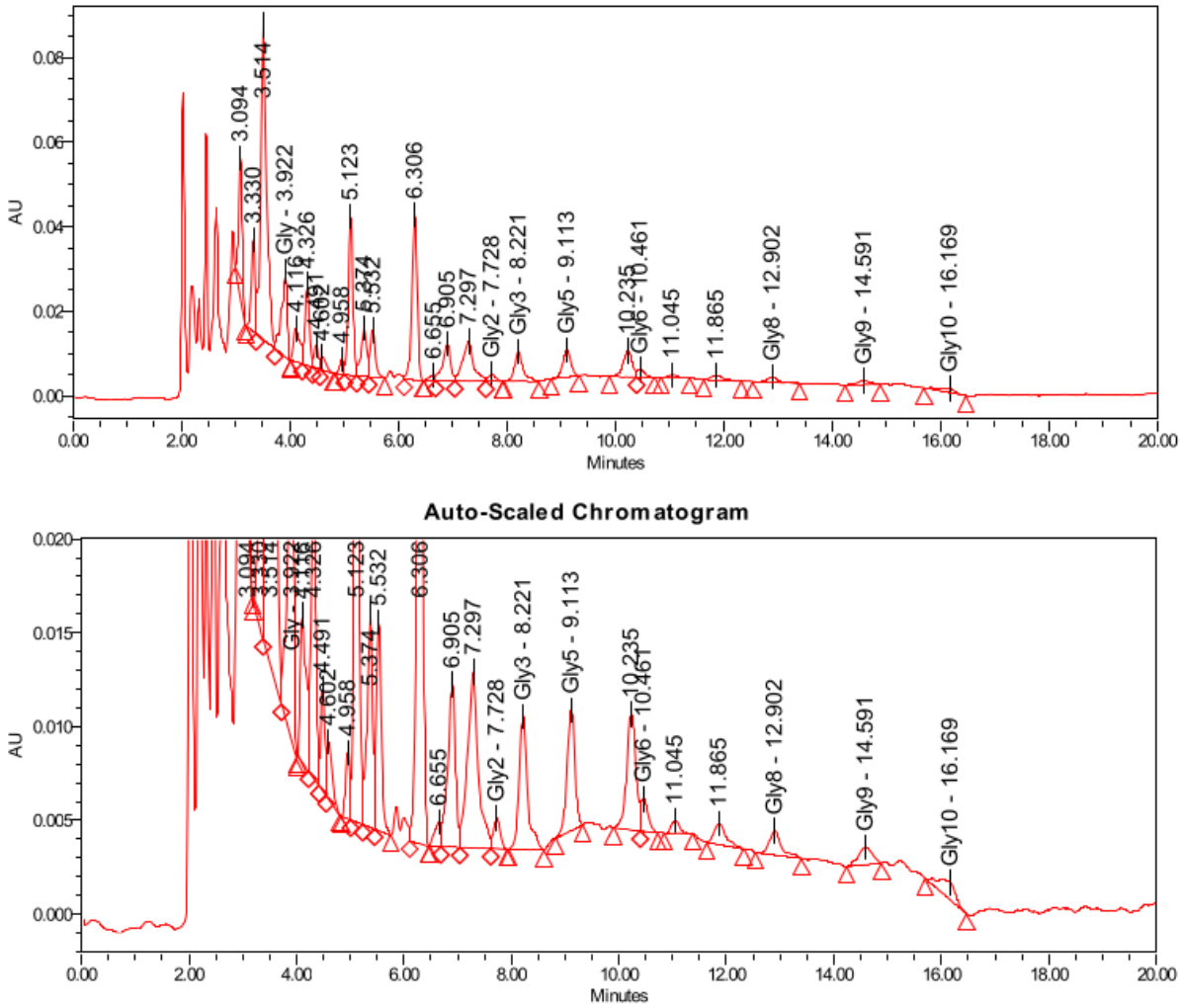

Figure 7.53. IP-HPLC chromatogram for milling Gly with pyrite at room temperature for 16 h and at 30 Hz.

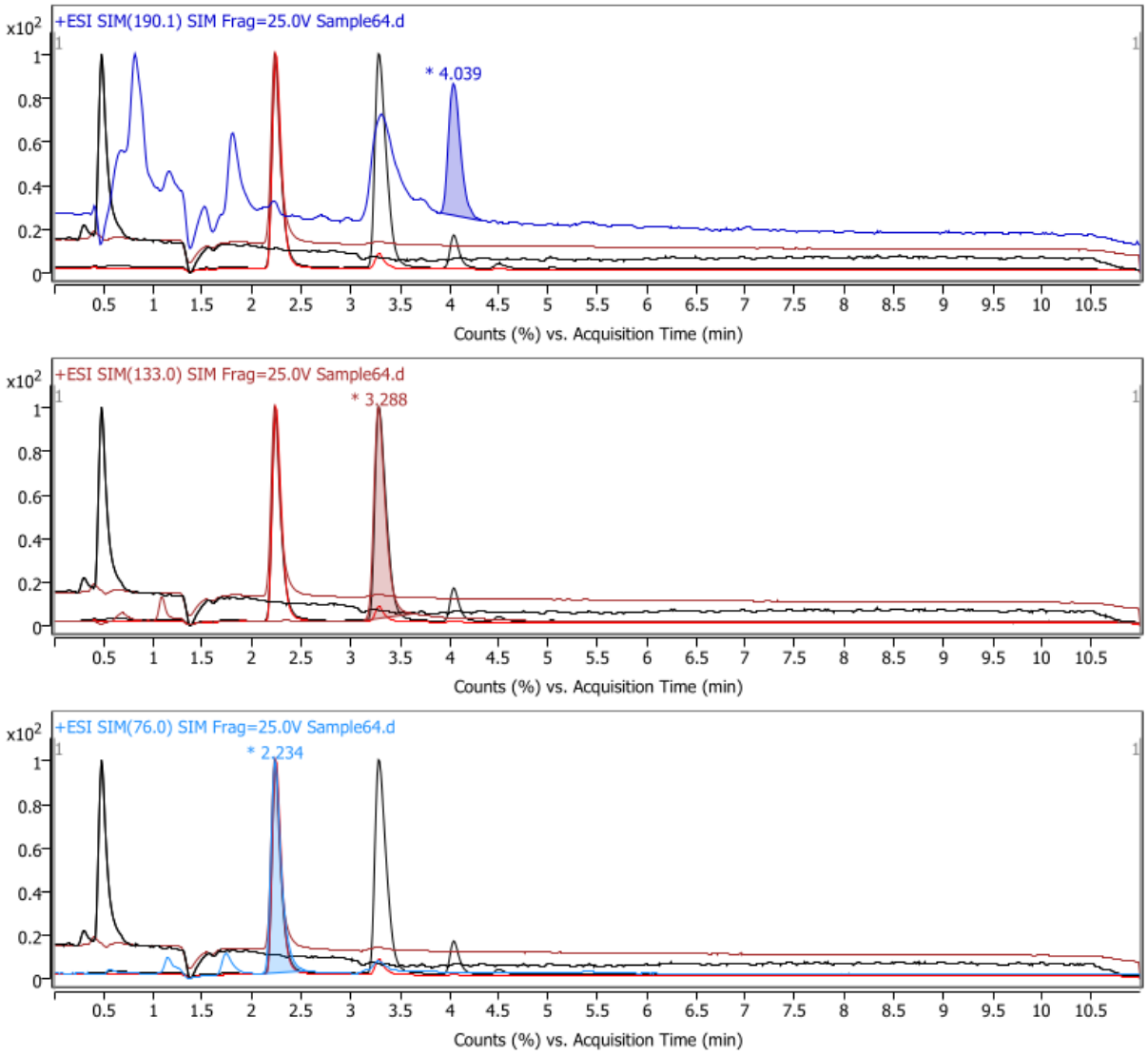

Figure 7.54. LC/MS (Triple quadrupole) chromatograms for milling Gly with pyrite at room temperature for 16 h and at 30 Hz.

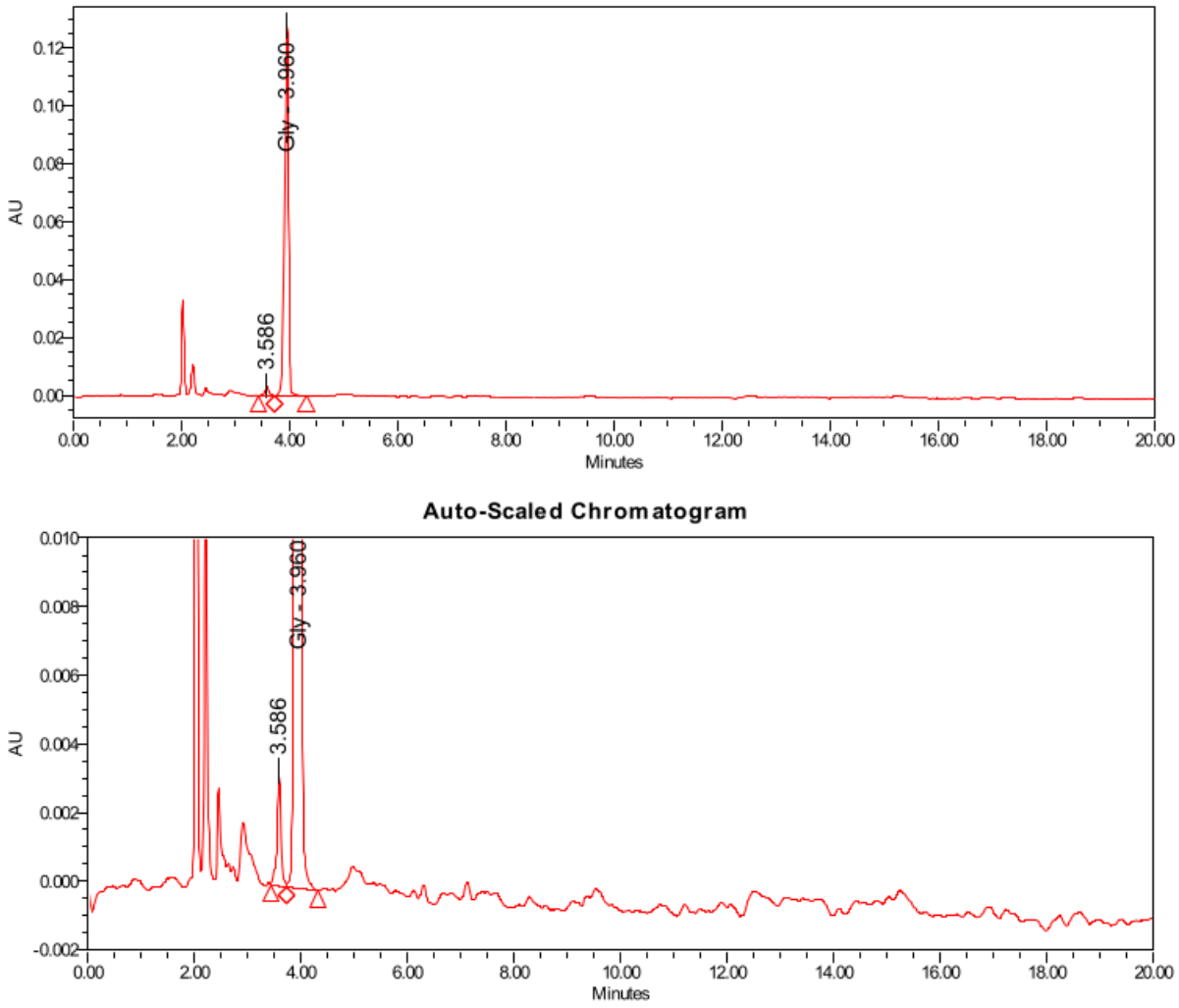

Figure 7.55. IP-HPLC chromatogram for milling Gly with gypsum at room temperature for 16 h and at 30 Hz.

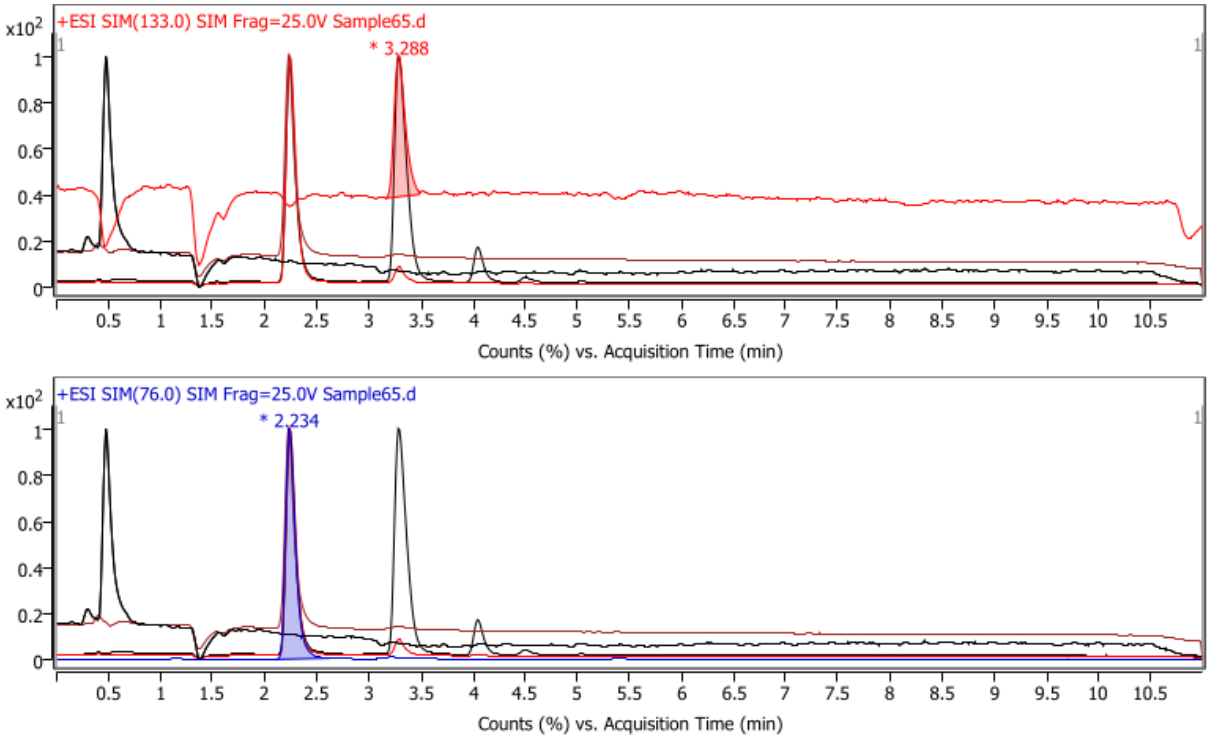

Figure 7.56. LC/MS (Triple quadrupole) chromatograms for milling Gly with gypsum at room temperature for 16 h and at 30 Hz.

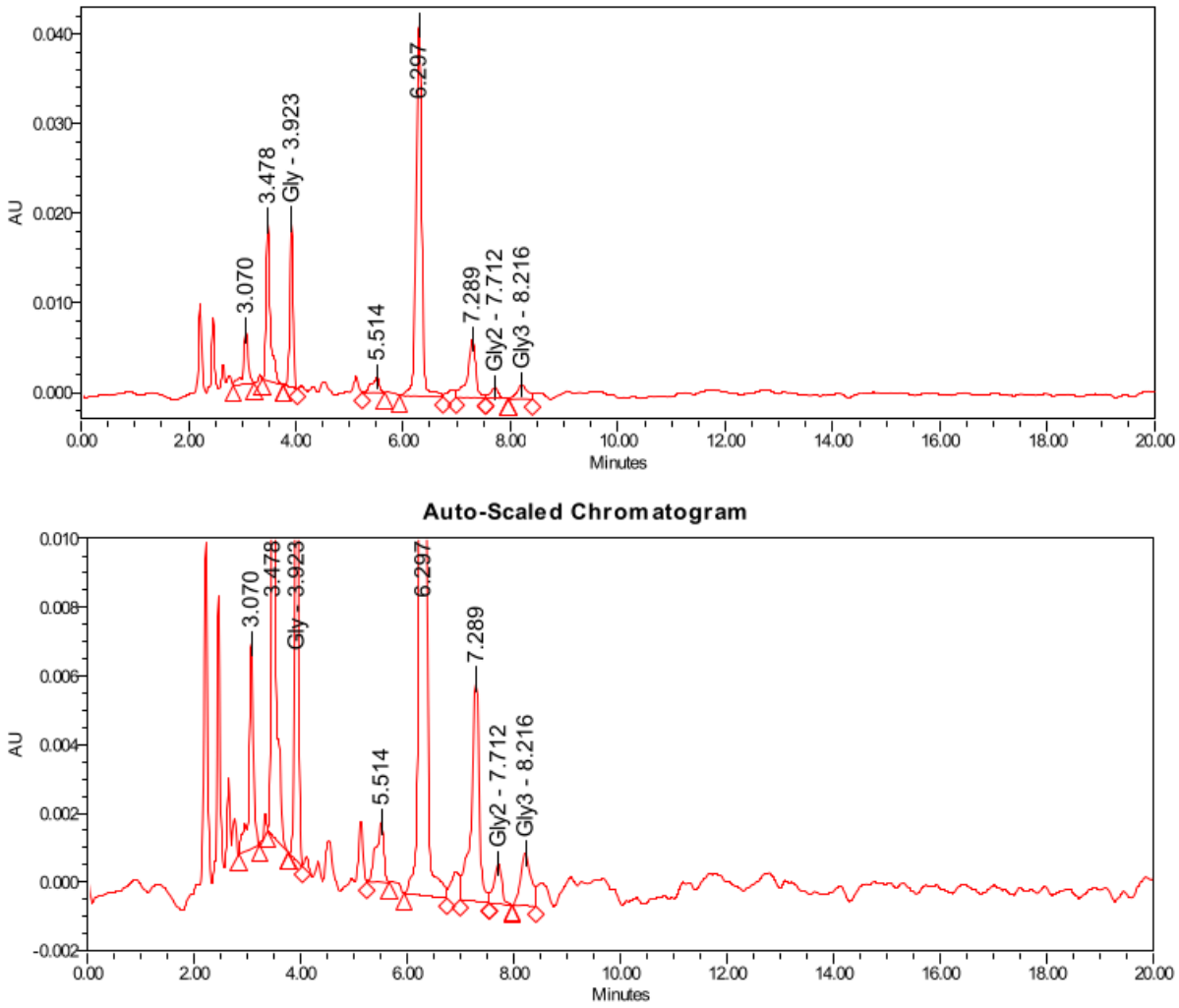

Figure 7.57. IP-HPLC chromatogram for milling Gly with talc at room temperature for 16 h and at 30 Hz.

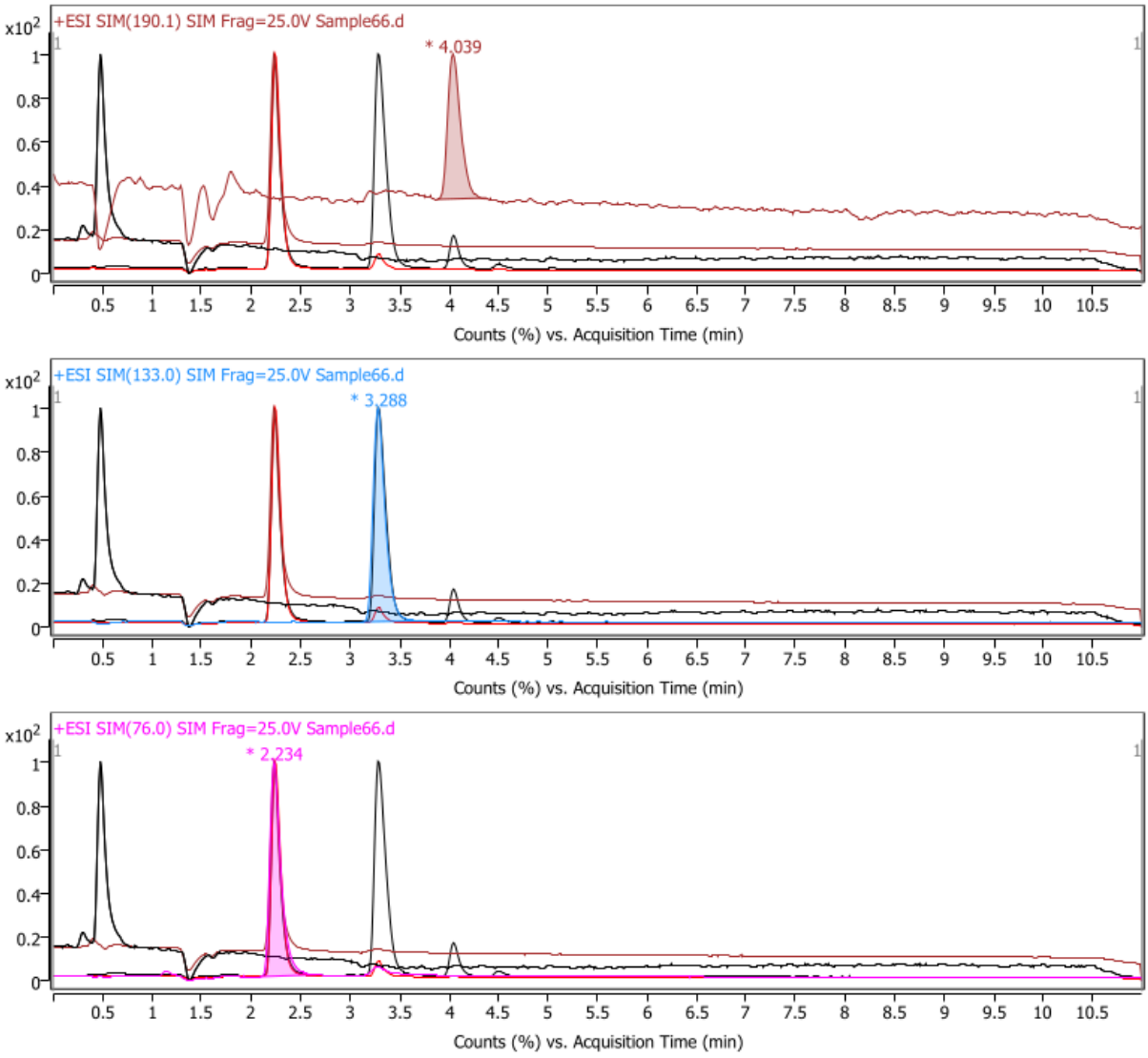

Figure 7.58. LC/MS (Triple quadrupole) chromatograms for milling Gly with talc at room temperature for 16 h and at 30 Hz.

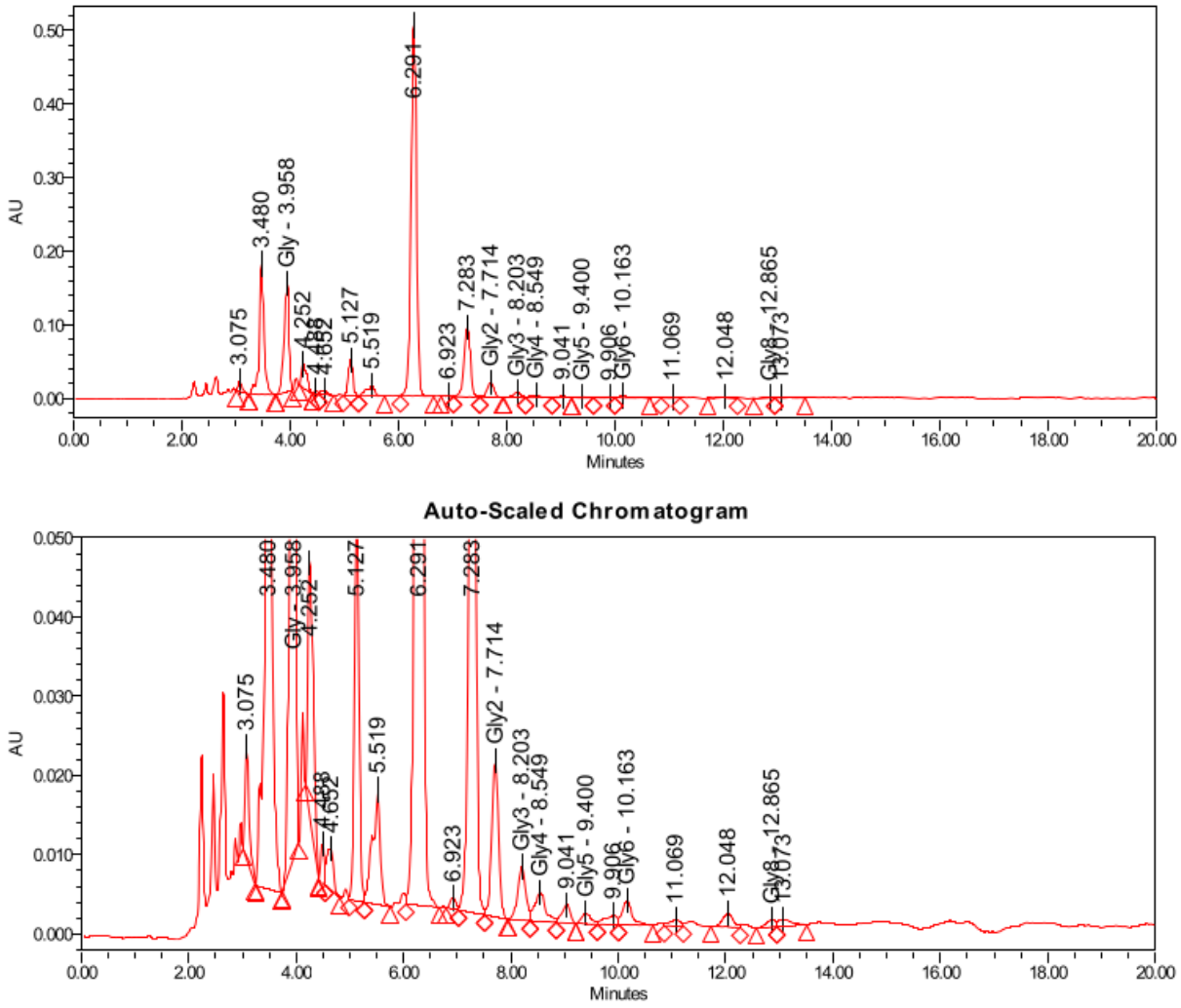

Figure 7.59. IP-HPLC chromatogram for milling Gly with magnesite at room temperature for 16 h and at 30 Hz.

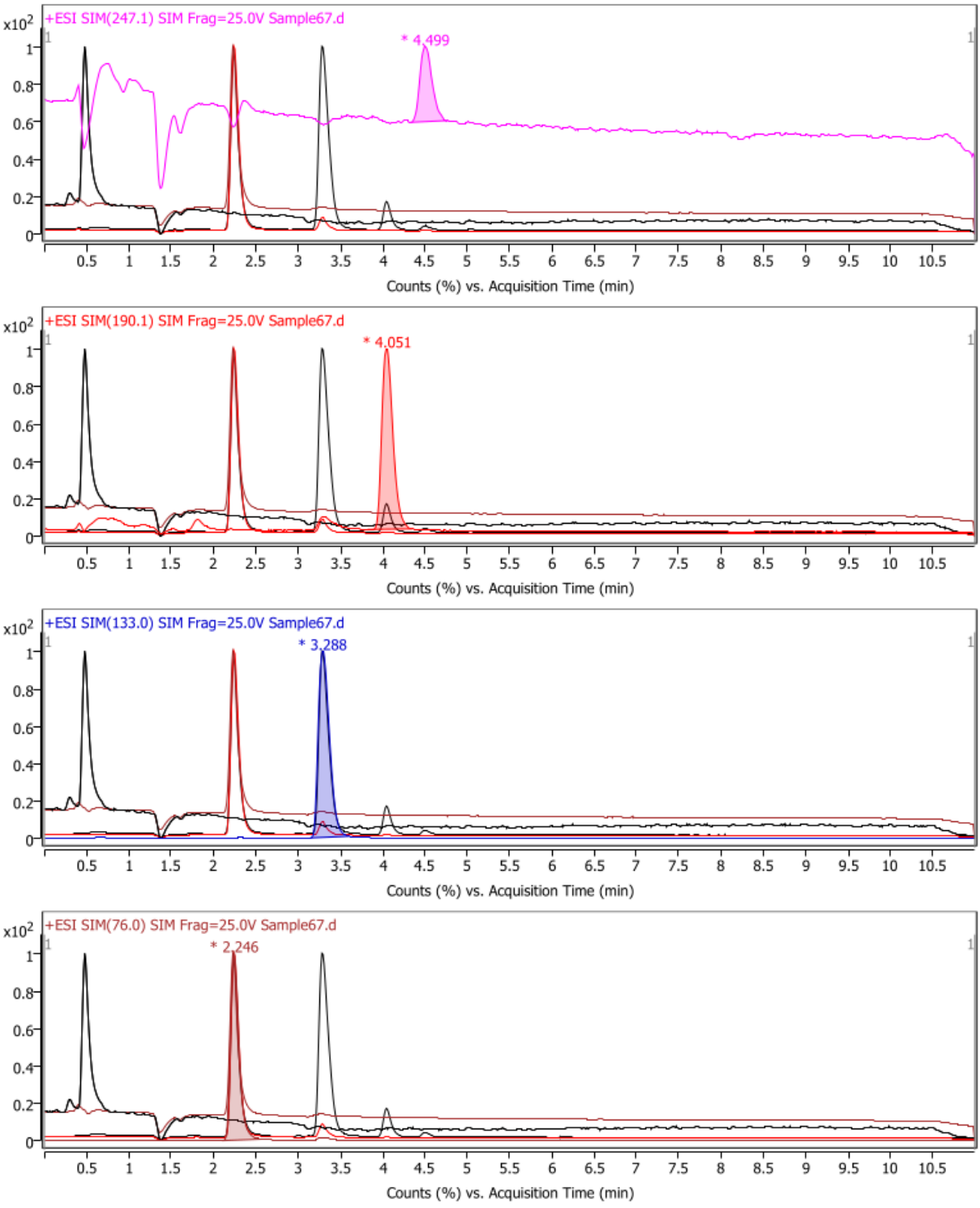

Figure 7.60. LC/MS (Triple quadrupole) chromatograms for milling Gly with magnesite at room temperature for 16 h and at 30 Hz.

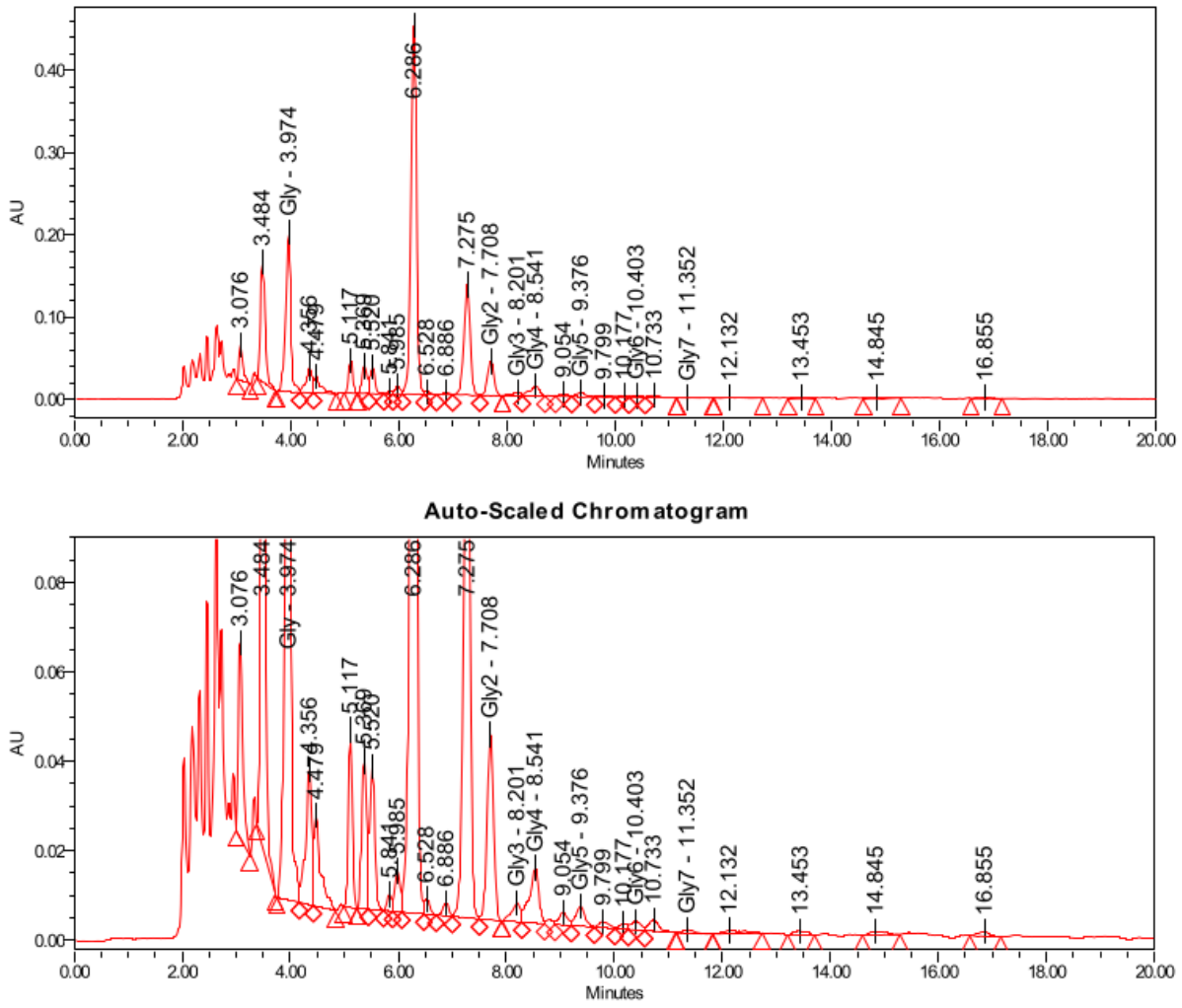

Figure 7.61. IP-HPLC chromatogram for milling Gly with quartz at room temperature for 16 h and at 30 Hz.

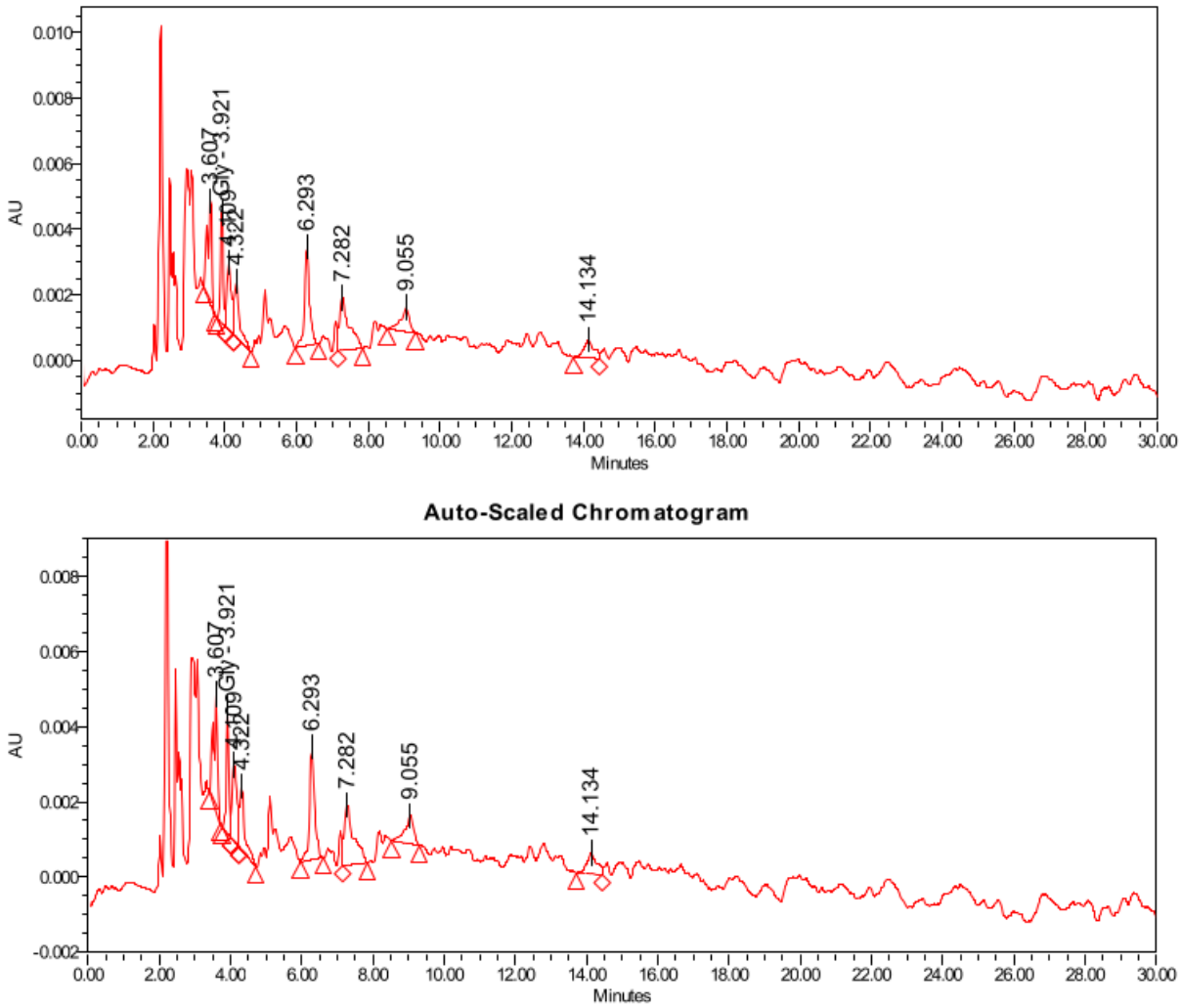

Figure 7.62. IP-HPLC chromatogram for milling Gly with clinoptilolite at room temperature for 16 h and at 30 Hz.

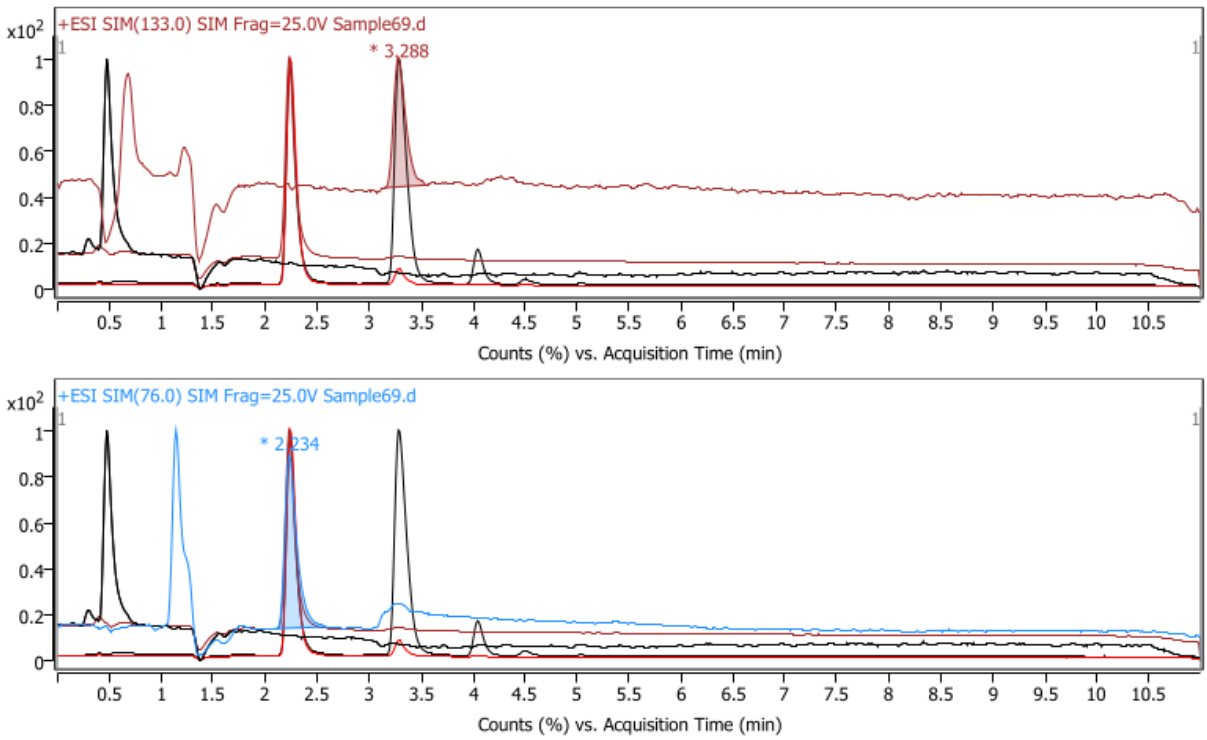

Figure 7.63. LC/MS (Triple quadrupole) chromatograms for milling Gly with clinoptilolite at room temperature for 16 h and at 30 Hz.

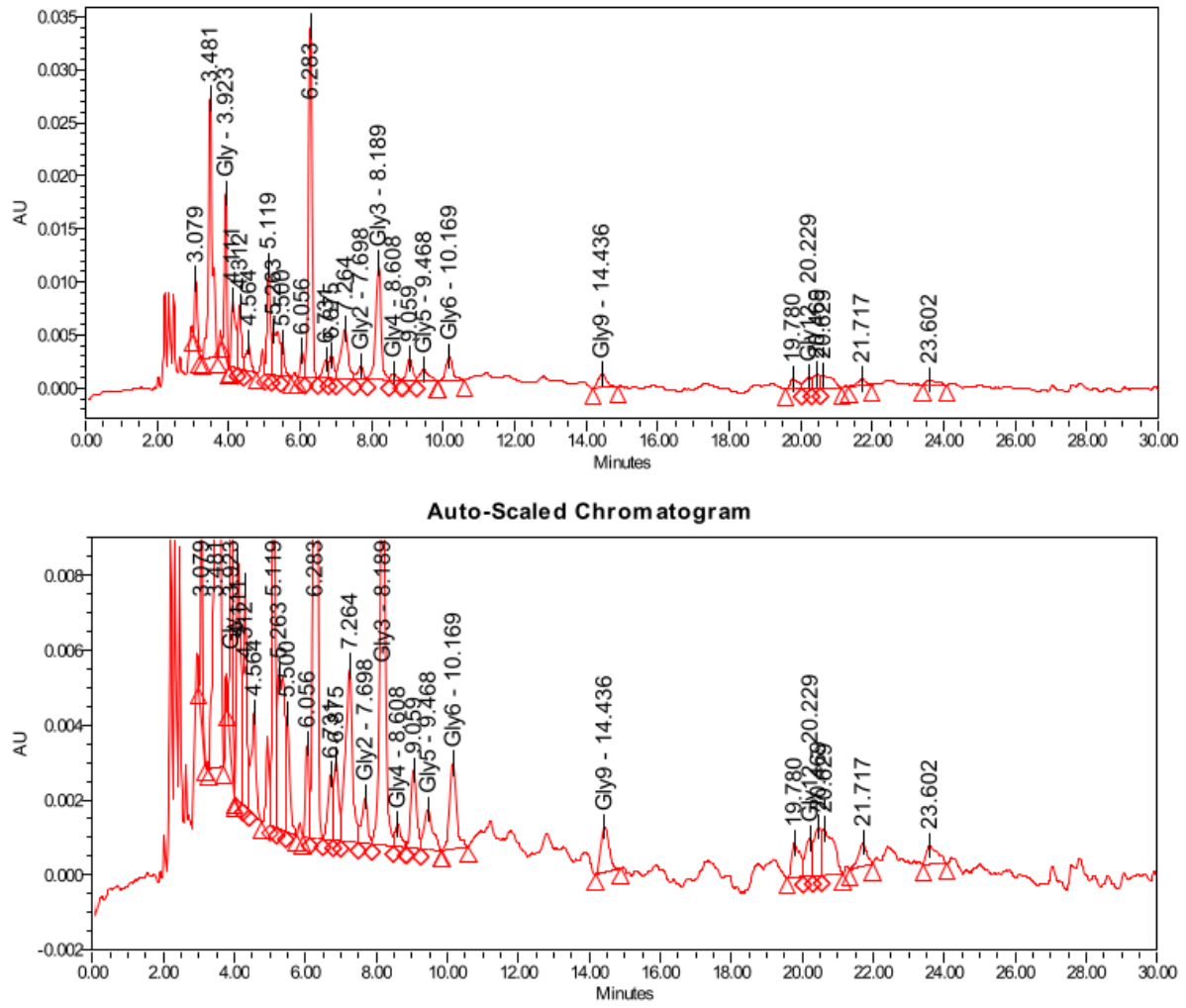

Figure 7.64. IP-HPLC chromatogram for milling Gly with apatite at room temperature for 16 h and at 30 Hz.

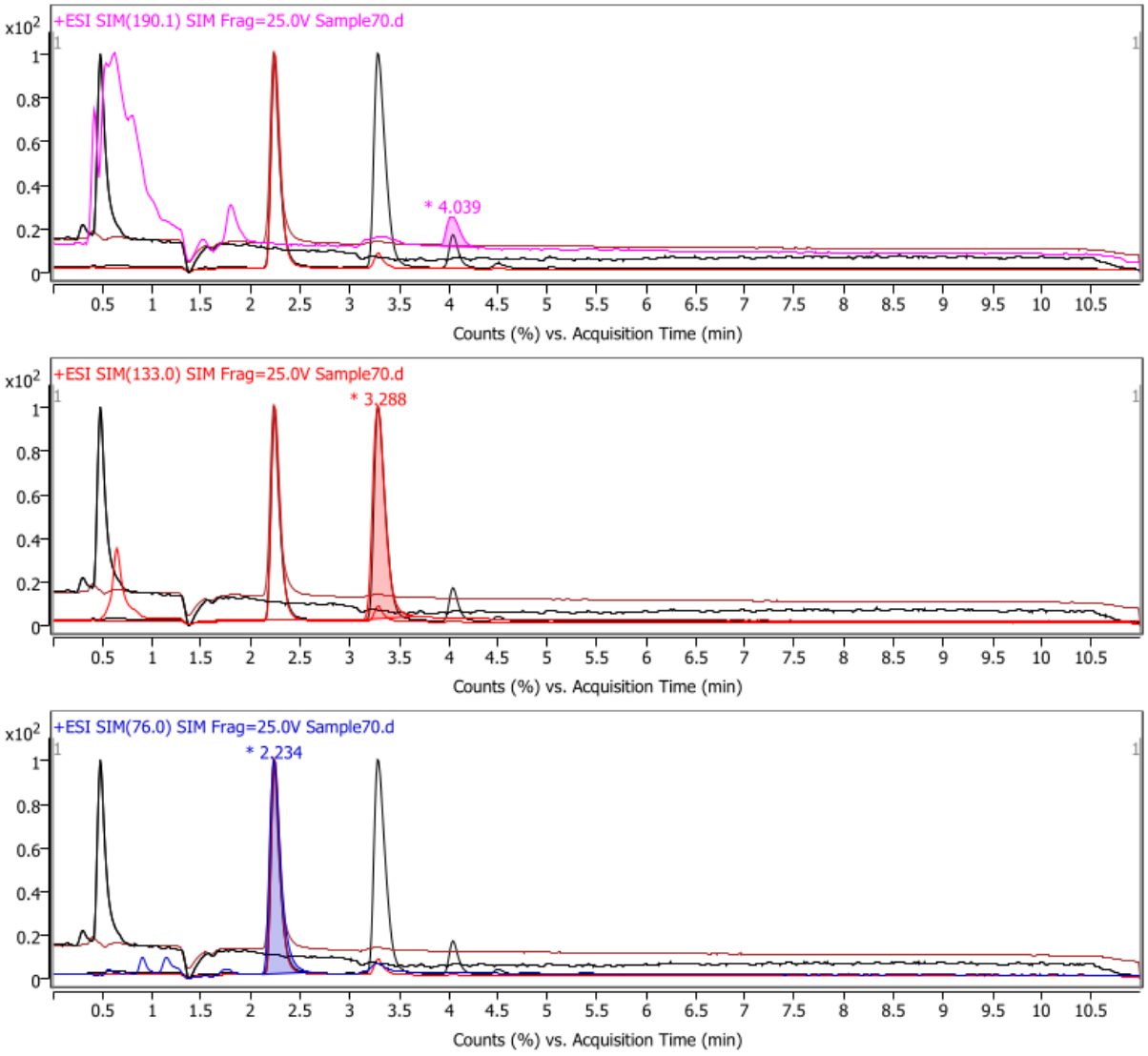

Figure 7.65. LC/MS (Triple quadrupole) chromatograms for milling Gly with apatite at room temperature for 16 h and at 30 Hz.

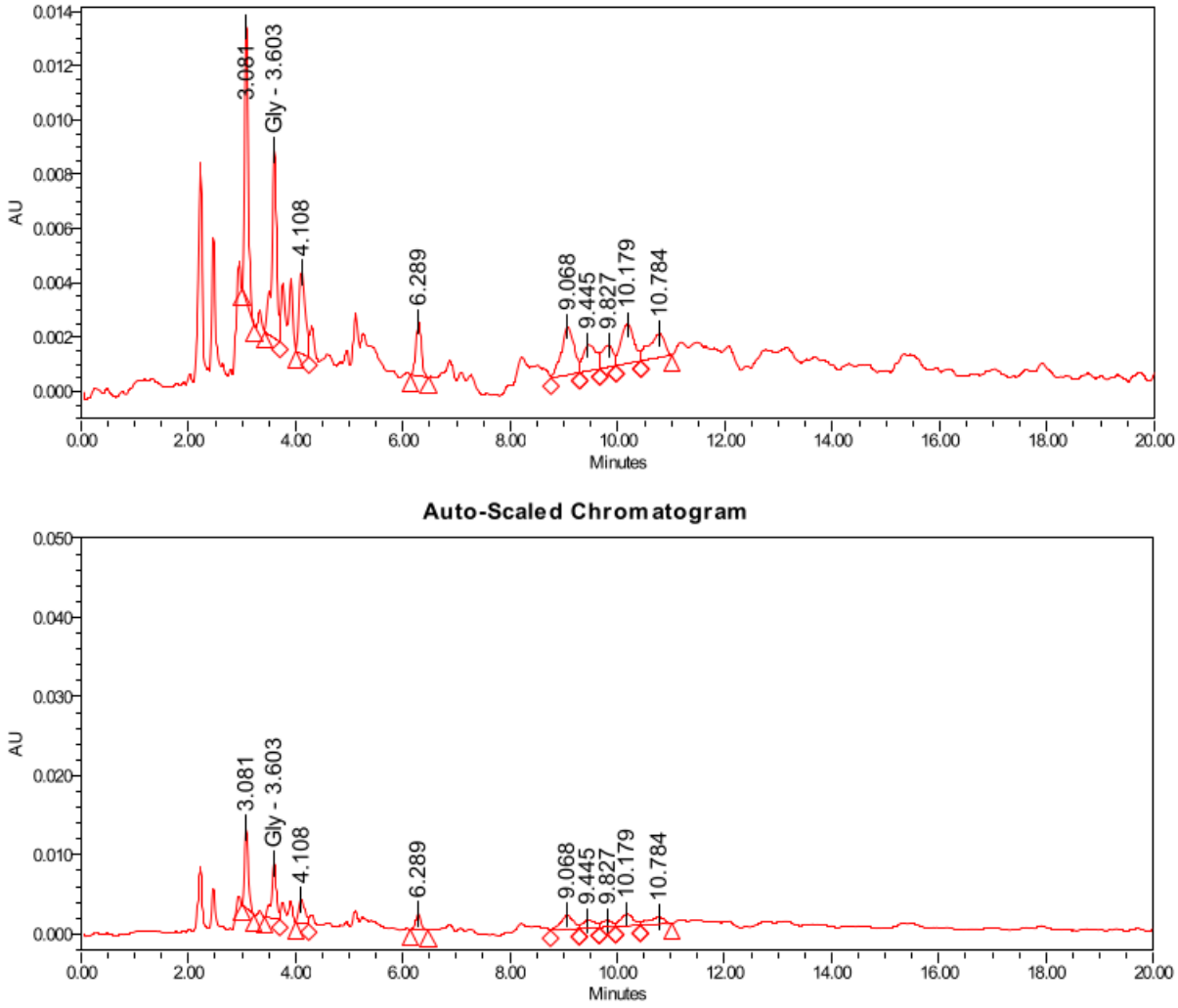

Figure 7.66. IP-HPLC chromatogram for milling Gly with analcime at room temperature for 16 h and at 30 Hz.

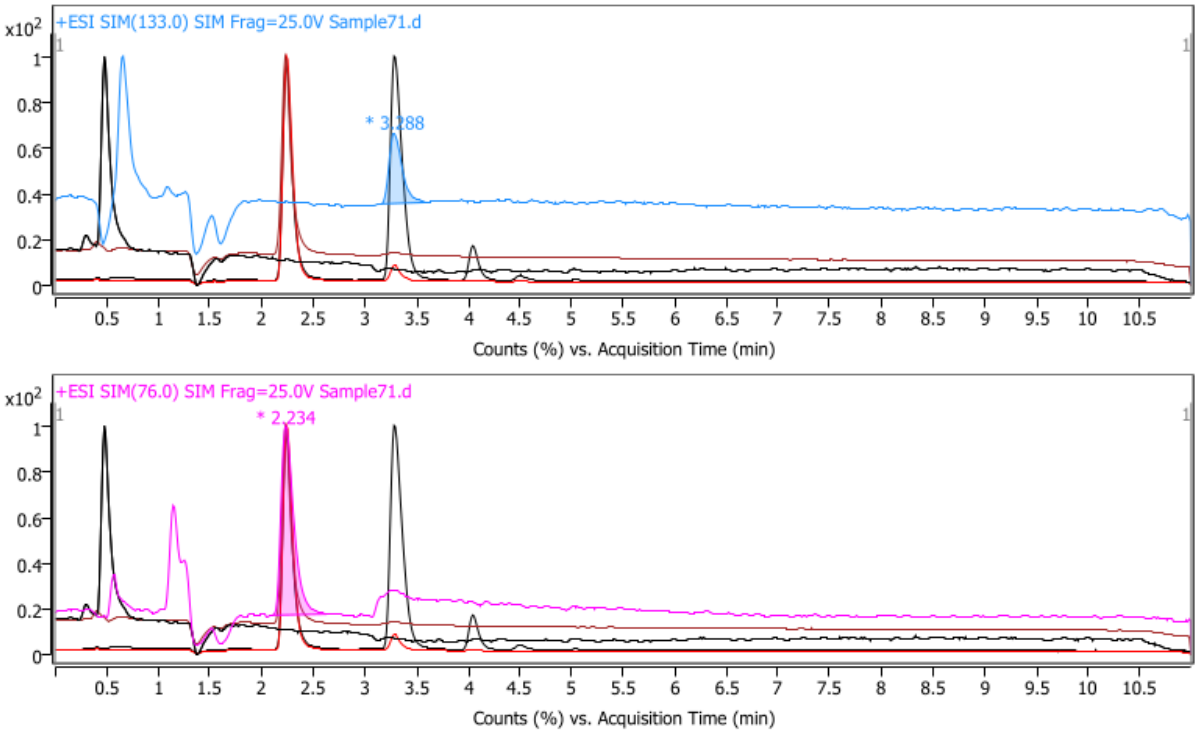

Figure 7.67. LC/MS (Triple quadrupole) chromatograms for milling Gly with analcime at room temperature for 16 h and at 30 Hz.

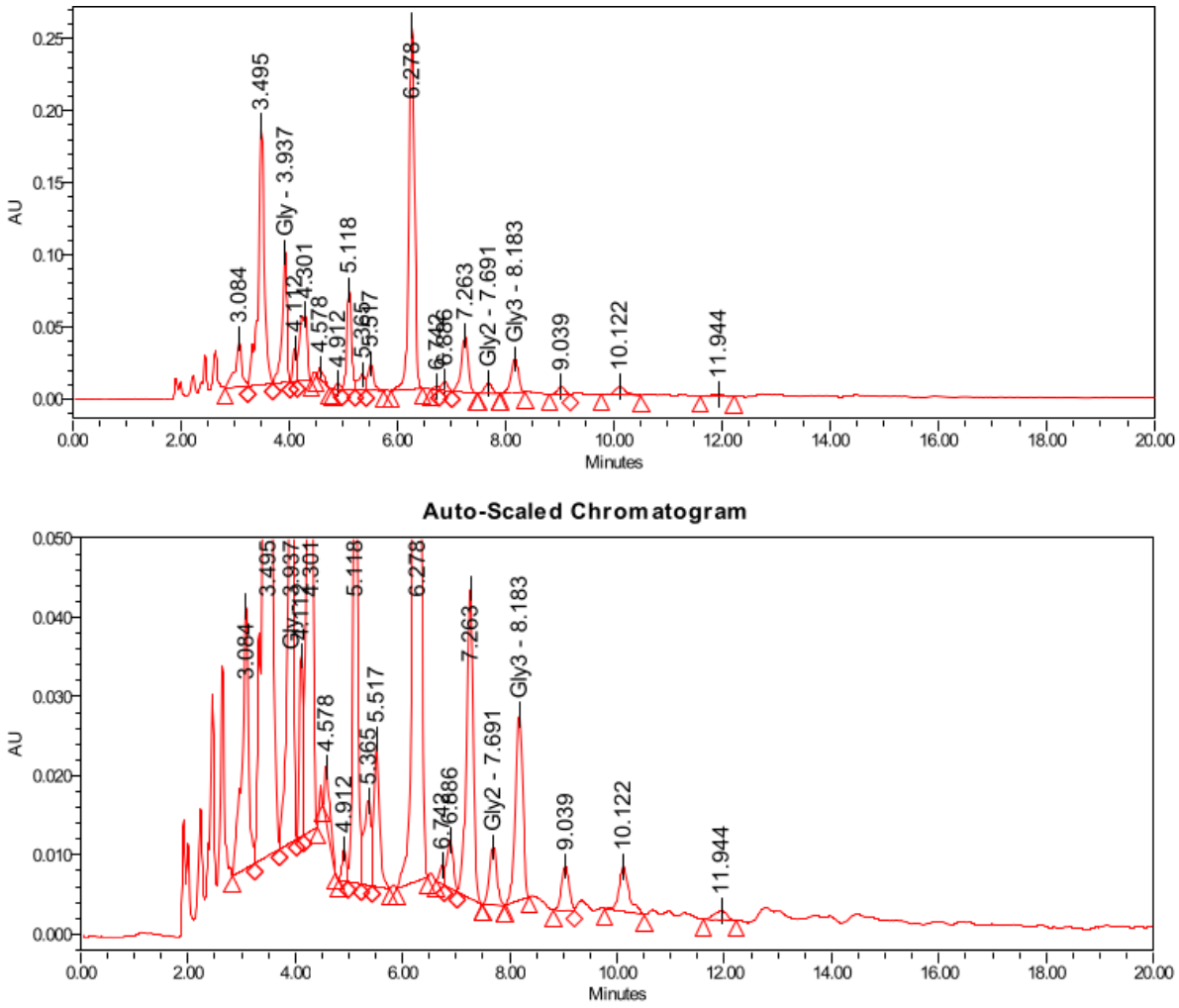

Figure 7.68. IP-HPLC chromatogram for milling Gly with calcite at room temperature for 16 h and at 30 Hz.

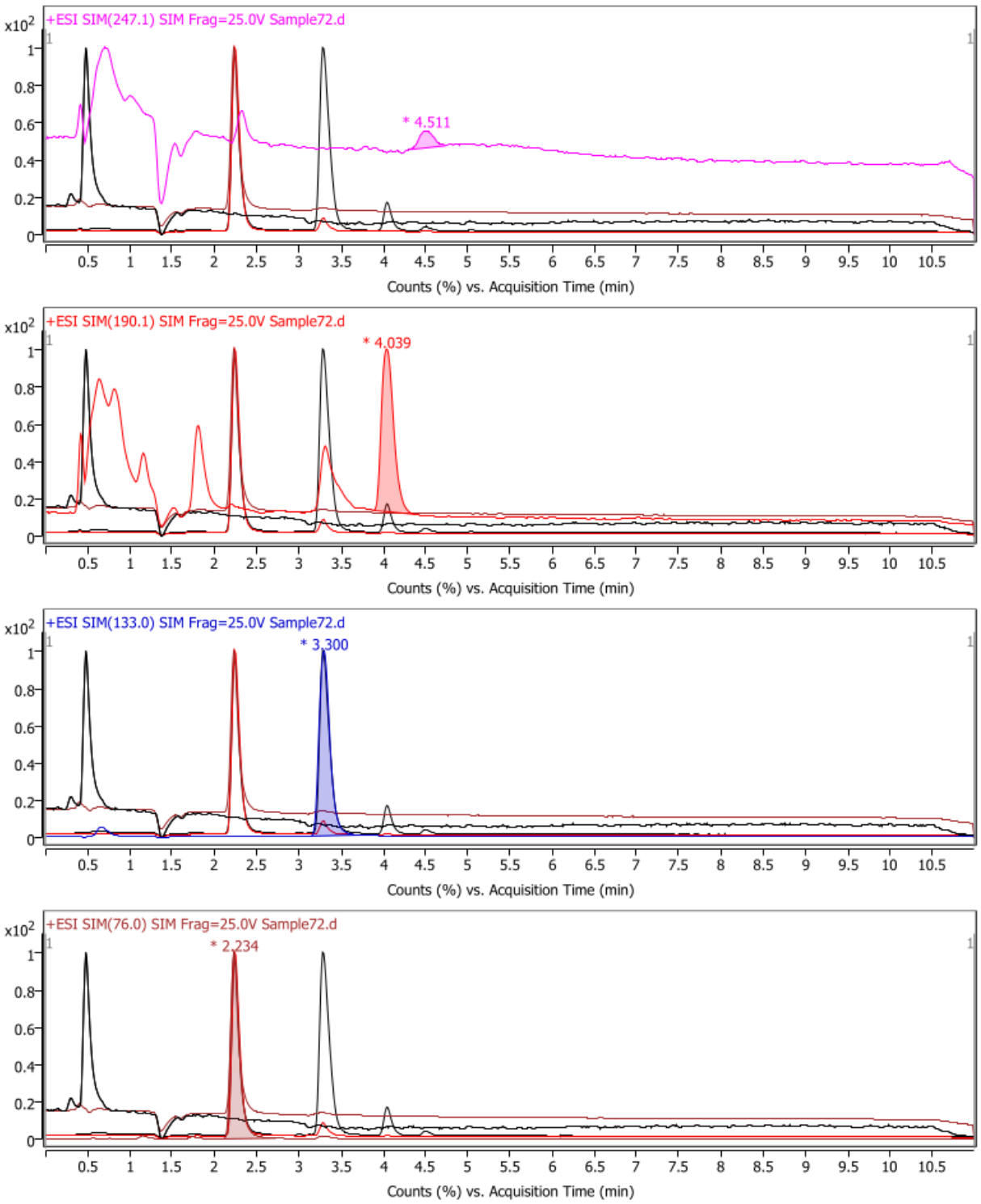

Figure 7.69. LC/MS (Triple quadrupole) chromatograms for milling Gly with calcite at room temperature for 16 h and at 30 Hz.

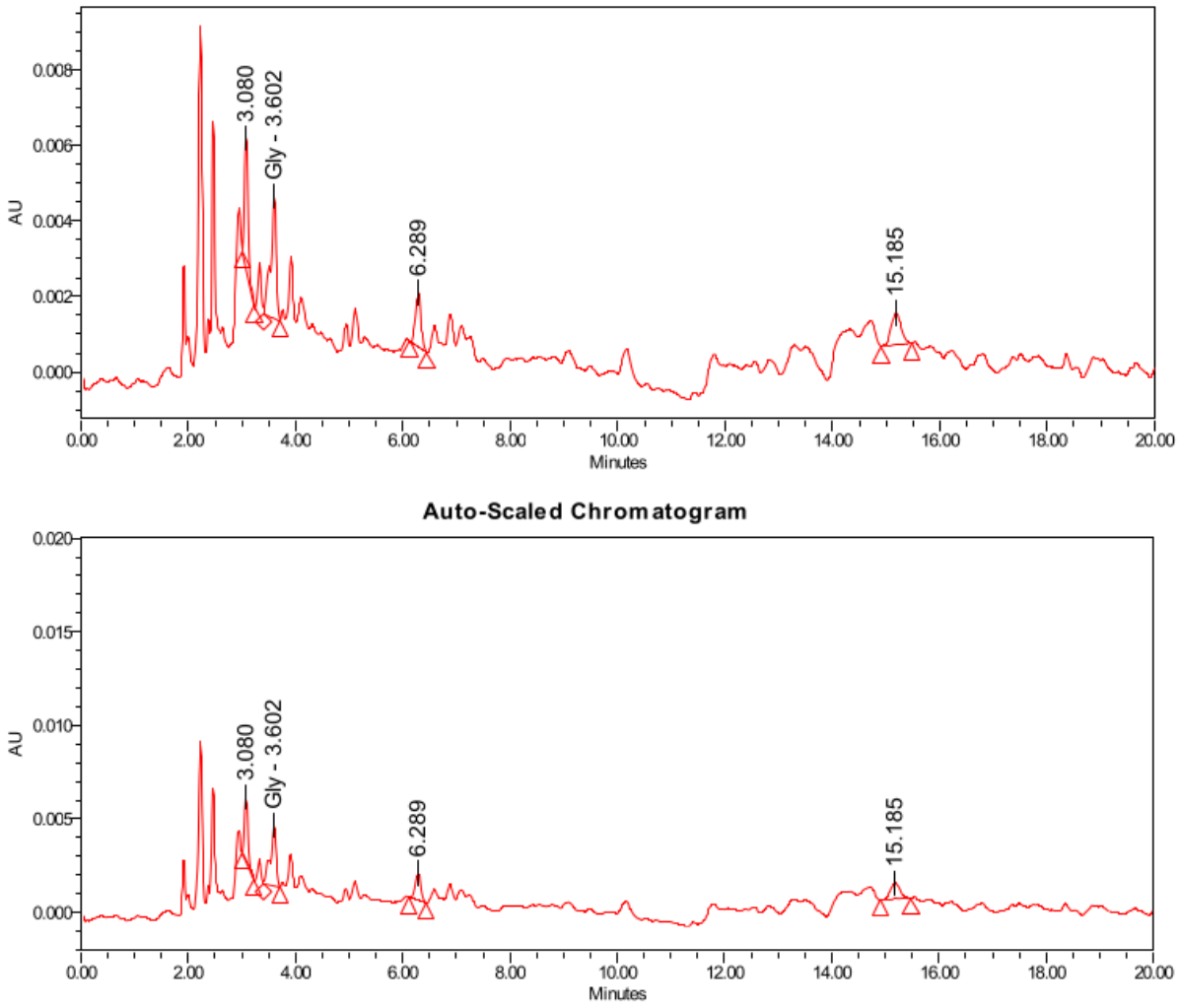

Figure 7.70. IP-HPLC chromatogram for milling Gly with muskovite at room temperature for 16 h and at 30 Hz.

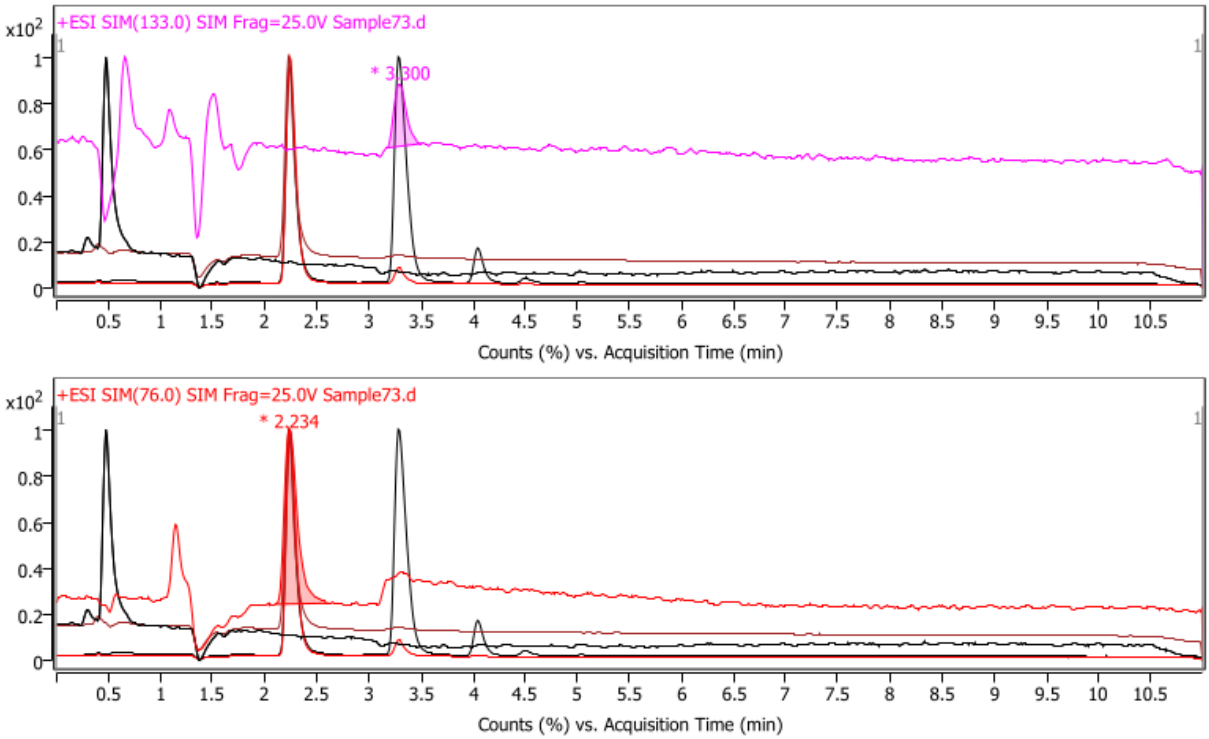

Figure 7.71. LC/MS (Triple quadrupole) chromatograms for milling Gly with muskovite at room temperature for 16 h and at 30 Hz.

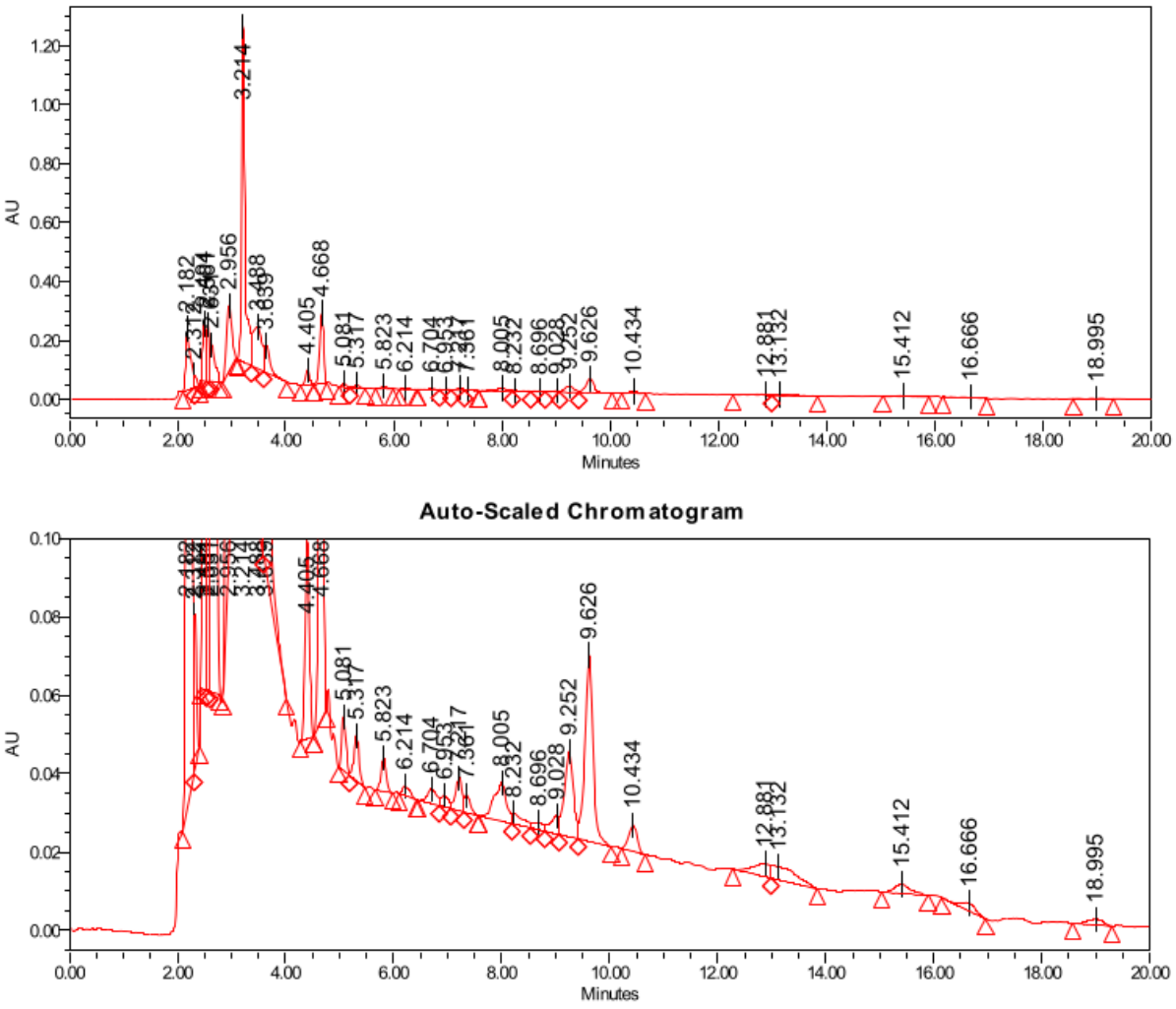

Figure 7.72. IP-HPLC chromatogram for milling glycolic acid with SiO_2 at room temperature for 16 h and at 30 Hz.

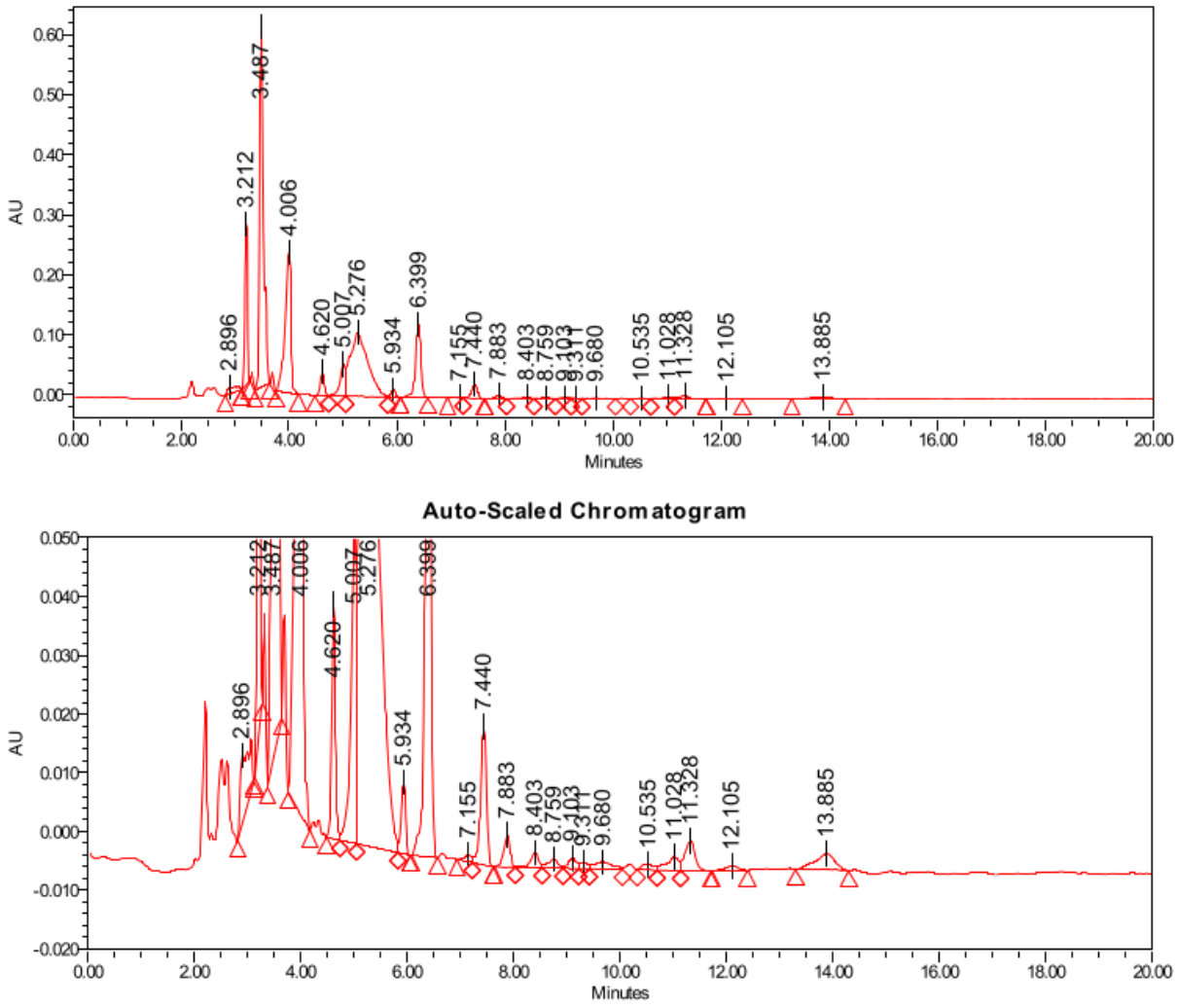

Figure 7.73. IP-HPLC chromatogram for milling glycine and glycolic acid with SiO_2 at room temperature for 16 h and at 30 Hz.

§ 8. CURRICULUM VITAE

Personal details:

Email:	tomislav.stolar@gmail.com
ORCID:	0000-0002-9824-4462
LinkedIn/Twitter:	Tomislav Stolar/@StolarTomislav

Working experience:

21/02/2018–Present	**Research assistant**, Laboratory for Applied and Sustainable Chemistry, Ruđer Bošković Institute, Zagreb (Croatia)
01/02/2016–31/07/2016	**External associate**, PLIVA Croatia, R&D – Teva Active Pharmaceutical Ingredients Physical Characterization Department, Zagreb (Croatia)
May 2013–Jan 2018	**Scientific research on a voluntary basis**, Ruđer Bošković Institute, Zagreb (Croatia)

Education:

Tomislav obtained his Bachelor's (2014.) and Master's degrees (2016.) from the Department of Chemistry, Faculty of Science, University of Zagreb (Croatia). He is currently enrolled in PhD studies at the same institution.

Publications (five selected publications, 16 in total, eight as first author, one as corresponding author, 486 citations on Google Scholar):

1) T. Stolar *et al.*, *Angew. Chem. Int. Ed.* **2021**, *60*, 12727–12731.

2) T. Stolar *et al.*, *ACS Appl. Mater. Interfaces* **2021**, *13*, 3070–3077.

3) T. Stolar *et al.*, *Chem. Commun.* **2020**, *56*, 13524–13527.

4) T. Stolar, K. Užarević, *CrystEngComm* **2020**, *22*, 4511–4525.

5) T. Stolar *et al.*, *ACS Sustainable Chem. Eng.* **2019**, *7*, 7102–7110.

Selected conferences & meetings

3) T. Stolar *et al.*, "The Effects of Geological and Environmental Conditions on the Rise of Prebiotic Peptides on Mineral Surfaces", 2022 Astrobiology Science Conference –**oral online presentation (11 + 4 min)**

2) T. Stolar *et al.*, "Mechanochemical prebiotic peptide bond formation in the absence of water", AbGradCon online event, 2021 – **oral presentation (12 + 3 min)**

1) T. Stolar *et al.*, "Importance of solid state in prebiotic chemistry studies", Solid State Science & Research 2019, Zagreb, Croatia - **award for distinguished oral presentation for young researcher (15 min)**

Professional services, independence, and leadership

Peer-reviewing for *ChemSusChem* and *ACS Sustainable Chemistry & Engineering* journals.

Projects as PI: CERIC-ERIC project 20212028 'SS-NMR for probing hydrogen-bonding interactions in nucleobase cocrystals'. 168 hours collection at the 600 MHz solid-state NMR spectrometer MAGIC, the Slovenian NMR Center, Ljubljana, Slovenia.

Voted representative of the Department of Physical Chemistry graduate students and post-docs (35 employees) at Ruđer Bošković Institute's "Council of graduate students and post-docs" as well as "Council of the Department of Physical Chemistry" from 2019.–2021.