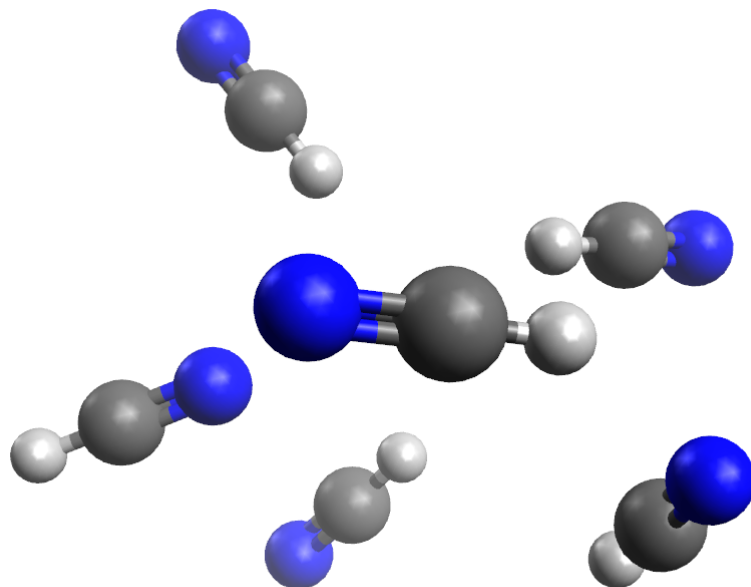




CHALMERS



En studie i potentiella reaktionsvägar för HCN-baserade molekyler

Kandidatarbete inom prebiotisk kemi

Kurskod KBTX10-22-09

Studenter

Jacob Bergström
Jonas Carlsson
Erika Danielsson
Natanael Lennartsson

Handledare

Martin Rahm
Marco Cappelletti
Fernando Izquierdo Ruiz

INSTITUTIONEN FÖR KEMI OCH KEMITEKNIK
AVDELNINGEN FÖR KEMI OCH BIOKEMI

CHALMERS TEKNISKA HÖGSKOLA
Göteborg, Sverige 2022
www.chalmers.se

Abstract

A study in potential reaction pathways of HCN-based molecules

This project delves into the reaction pathways that a HCN tetramer can follow when reacting with additional HCN molecules as well as its anion CN^- . Using DFT based computational chemistry we calculate the activation energies and kinetics of several nucleophilic attacks to determine how likely they are to occur. Further analysis of an intramolecular cyclisation is performed where the effects of various side groups are studied. Among the cyclisation reactions we found one reaction containing only HCN with a very low reaction barrier.

Sammandrag

Under det här projektet har vi studerat hur en tetramer av HCN kan reagera med andra HCN-molekyler och dess anjon CN^- . Genom att använda DFT-baserade beräkningar har energibarriärer och kinetik för ett antal nukleofila attacker studerats för att bestämma hur sannolika de är. Vi har även studerat sex olika sidogrupperns inverkan på en intramolekylär cykliseringsreaktion. Vi har funnit en HCN-tetramer som kan cykliseras i en reaktion med låg barriär.

Begrepp

Density Functional Theory, DFT - En kvantkemisk modell som används vid beräkningskemi för att simulera och optimera system av atomer.

Heterocyklisk förening - En cyklisk organisk förening med minst en annan atom än kol i ringen.

HOMO - Förkortning för "Highest Occupied Molecular Orbital".

Isomer - En molekyl som har samma summaformel men har olika struktur, är isomerer till varandra. Till exempel så har alla HCN-tetramerer summaformeln $\text{H}_4\text{C}_4\text{N}_4$ men kan se ut på många olika sätt.

LUMO - Förkortning för "Lowest Unoccupied Molecular Orbital".

Självcyklisering - En cykliseringsreaktion där en reaktant bildar en ringstruktur med sig själv.

Stereoisomer - Isomer som har samma ordning, eller samma bindningar mellan atomer, men olika orienteringar och vinklar mellan dessa.

Transition state, TS - Engelskans ord för övergångstillstånd.

Övergångstillstånd - Steget mellan reaktanter och produkter i en kemisk reaktion där bindningar uppstår och/eller bryts.

Innehåll

1	Inledning	1
2	Teori och metod	2
2.1	Nukleofila och elektrofila attacker	2
2.2	Nomenklatur	2
2.2.1	Molekylen M4	3
2.2.2	Studerad cyklisering	3
2.3	Kvantkemi	4
2.3.1	Orbitalteori	5
2.4	Kinetikberäkningar	5
3	Resultat och diskussion	7
3.1	Undersökning av olika nukleofila attacker 4	7
3.1.1	Halveringstider för M4-reaktionerna	9
3.2	Vidare undersökning av cykliseringsreaktionen	9
3.2.1	Halveringstider för cyklisering	11
3.2.2	Molekylorbitalanalys	12
3.2.3	Analys av laddningar	13
4	Slutsats	15
4.1	Framtidsutsikter	15
	Erkännanden	16
	Referenser	16
	Bilagor	
I	Härledning av uttryck för halveringstiden	20
II	Energier för undersökta M4-stereoisomerer	21
III	Cykliseringsenergier för reaktant- och produktkomplex	22
IV	Halveringstider för samtliga undersökta cykliseringsreaktioner	23
V	Molekylorbitalanalys	24

1 Inledning

Genom historien har människan ställt sig frågan om hur liv uppstod på jorden. Att förstå hur detta gick till kan hjälpa oss att förstå mer av universums och jordens historia. Det är också av betydelse för att hitta liv utanför jorden och att veta hur och var de kemiska grundstenar som behövs uppstår.[1]

Bland dessa grundstenar finns det ett starkt intresse av att förstå hur nukleobaser och heterocykliska ringar har bildats, detta då de har liknande strukturer som de molekyler som bygger upp DNA och RNA. Nukleobaser och heterocykliska ringar har påträffats både i meteoriter och laboratoriesituationer vilket visar på att dessa molekyler kan bildas abiotiskt.[2][3][4][5] Bildandet av nukleobaser kan därför vara en av nyckelpunkterna till hur liv uppstod på den tidiga jorden, samt att det även skulle kunna ske ute i universum.

För att hitta ett svar på frågan om hur dessa molekyler uppstått finns många tillvägagångssätt. Många tidigare studier har utgått ifrån en blandning av H_2O och HCN. Det finns belägg för att det är en rimlig utgångspunkt då H_2O innehåller syre vilket är nödvändigt för bildandet av alla de primära nukleobaserna förutom adenin.[6] Reaktionerna som bildar nukleobaserna har beskrivits använda produkten av HCN och H_2O , formamid (CHONH_2). Dock har de beräknade energibarriärerna för dessa reaktioner varit relativt höga då de har uppgått till över 30 kcal/mol. [7][6] Därför finns det ett intresse att undersöka reaktioner med lägre energibarriärer, till exempel reaktioner där HCN reagerar med sig själv som har uppskattats ha lägre barriärer; cirka 20 kcal/mol.[8] Att reaktionerna har lägre barriärer innebär att de skulle kunna ske spontant under mildare förhållanden så som vid lägre temperaturer.

Redan på 60-talet visade experimentella observationer att HCN kan genomgå polymeriseringsreaktioner för att bland annat bilda biologiskt intressanta molekyler som adenin.[9][8] I dagsläget saknas det kunskap om exakt vilka reaktionsvägar som HCN kan ta för att bilda dessa molekyler, med enbart ett fåtal reaktionsvägar som har hittats.[7][10][11] Ökad kunskap om dessa reaktionsvägar kan vara ett steg på vägen att hitta insikter om livets uppkomst; HCN har visats existera på många platser ute i universum, bland annat i månen Titans atmosfär, i protostjärnan GL 2591, på ett flertal kometer samt i det interstellära mediet.[1][2][12][13]

För att öka den nuvarande kunskapen har detta projekt genom kvantkemiska beräkningar undersökt ett antal reaktioner där HCN och dess anjon CN^- har reagerat med en HCN-tetramer. En av reaktionerna är en självcyklisering som studerades vidare med ett urval av olika sidogrupper i olika positioner.

2 Teori och metod

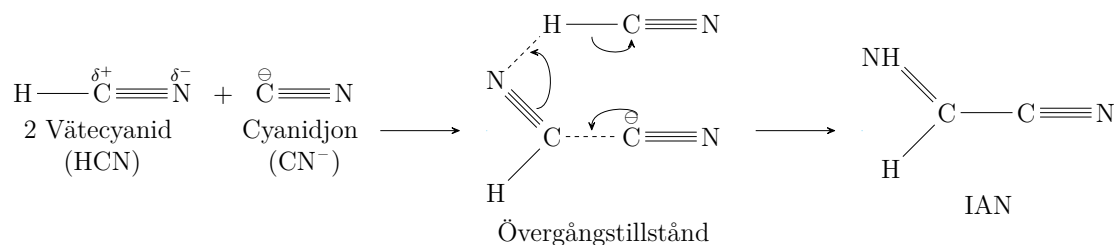
I detta projekt så har vi med kvantkemiska beräkningar studerat kemiska reaktioner. Till stöd för detta så har vi använt programmen Chemcraft[14] och Gaussian 16[15] som har körts genom datorklustret Vera på Chalmers som har tillhandahållits av SNIC via C3SE. Chemcraft användes för att producera indata för beräkningarna och för att visualisera resultaten och Gaussian 16 användes för att utföra alla beräkningar. Grunden för beräkningarna är *Density Functional Theory*, DFT, vars detaljer förklaras nedan.

2.1 Nukleofila och elektrofila attacker

För reaktioner där molekyler slås samman finns två typer av attacker. Den ena varianten är den elektrofila attacken där en positivt laddad del av en molekyl attackerar en negativt laddad del och skapar en bindning. Den elektrofila attacken kommer dock inte studeras här, då de enda reaktionerna som studeras i arbetet är nukleofila attacker.

En nukleofil attack illustreras i Figur 1 där två vätecyanidmolekyler, HCN, och en cyanidjon, CN^- reagerar för att bilda iminoacetonitril, IAN. [16]

Reaktionsmekanismen inleds med att den negativt laddade nukleofilen attackerar den positivt laddade elektrofilen. Det medför att elektroner förflyttas, vilket visas med pilarna i övergångstillståndet. De streckade linjerna anger var bindningar kommer uppstå till följd av elektronförflyttningarna.



Figur 1: Demonstration av nukleofil attack av CN^- på en HCN-molekyl med omkringliggande HCN, pilarna visar elektronnörelserna.

Då grunden för dessa reaktioner är laddningsskillnaden mellan nukleofilen och elektrofilen, så utfördes en analys av de partiella laddningarna på det elektrofila kolet och det nukleofila kvävet som deltog i den studerade cykliseringsreaktionen.

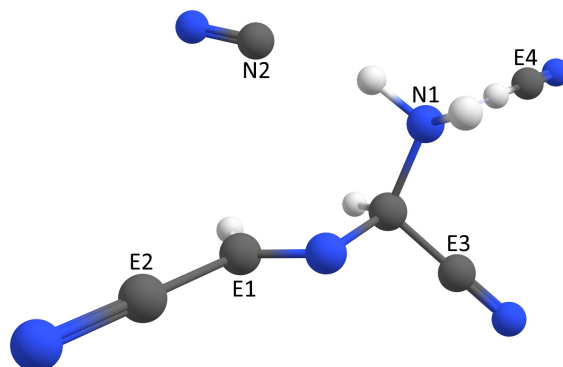
2.2 Nomenklatur

För att enklare kunna behandla individuella reaktioner så döptes de antingen efter vilka atomer som agerade nukleofil och elektrofil, eller efter vilka sidogrupper som en molekyl inkluderade.

2.2.1 Molekylen M4

Utgångspunkt för studien så var tetramer av HCN, härefter kallad M4, vars struktur visas i Figur 2. Den valdes då den bestod av flera intressanta funktionella grupper som möjliggjorde många olika nukleofila och elektrofila attacker.

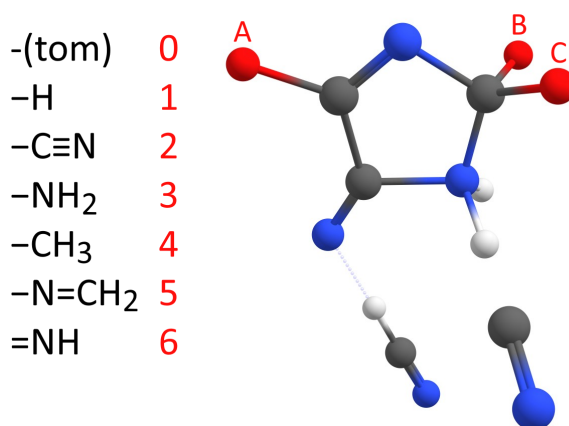
De studerade reaktionerna namngavs efter vilka delar av reaktanterna som agerade nukleofil eller elektrofil, där $Ni, i = 1, 2$ betecknar nukleofilen och $Ej, j = 1, 2, 3$ betecknar elektrofilen. Denna nomenklatur visas i Figur 2.



Figur 2: Reaktionskomplexet M4, CN^- och HCN, med utskriven nomenklatur för vilka atomer som kan agera nukleofiler, N, och elektrofiler, E.

2.2.2 Studerad cyklisering

Efter att de nukleofila attackerna hade studerats så studerades N1E2 ytterligare. Tre sidogrupper byttes ut enligt det schema som visas i Figur 3. Den ursprungliga produkten från N1E2 reaktionen blir med denna nomenklatur 112, med väten i positionerna A och B, och en cyanidgrupp i position C.



Figur 3: Övergångstillståndet för cykliseringen N1E2, där de tre sidogrupperna nu är rödmarkerade och namngivna A, B, och C, successivt byts ut med grupperna 1-6 till vänster i figuren.

2.3 Kvantkemi

Under vissa förhållande så är det inte möjligt att undersöka molekyler och reaktioner i experimentellt i laboratorium. Om detta inte går, på grund av hälsorisk, svårighet att få tag i ämnet eller andra anledningar så behövs ett annat sätt att undersöka olika ämnens egenskaper. Ett vanligt sätt är att med datorprogram som tillämpar olika teorier och modeller simulera molekyler och dess egenskaper, exempelvis energier och laddningar. Praktiska och beräkningsexperiment är dock inte ömsesidigt uteslutande och kan komplementera varandra.

För att få fram egenskaper för ett system av atomer kan Schrödingers elektronekvation användas

$$\hat{H}\Psi = E\Psi \quad (1)$$

där E är systemets totala energi, Ψ är flerelektronsvågfunktionen och $\hat{H}$ är Hamiltonianen. Dessa funktioner är dock väldigt komplicerade och behöver förenklas för att kunna lösas. Den första förenkling som görs är Born-Oppenheimer-approximationen, som säger att elektroners och atomkärnornas rörelser och vågfunktion kan separeras på grund av storleksskillnaden mellan partiklarna.[17] Hamiltonianen kommer därav enbart behöva beskriva elektronerna i systemet och skrivs

$$\hat{H} = \hat{T} + \hat{V} + \hat{U} \quad (2)$$

$\hat{T}$, $\hat{V}$ och $\hat{U}$ är en elektrons kinetiska energi, dess kärnattraktion, respektive repulsion från alla andra elektroner. En av de första metoderna som utvecklades för att uppskatta ekvationen var Hartree-Fock-metoden som används för att reducera flerpartikelsystemet till flera enklare enpartikelsystem. Denna och följande metoder är i grunden en uppskattning, men ger oftast resultat som stämmer tillräckligt bra överens med verkligheten.[18]

Det bevisades att lösningen till flerpartikelsystemet kan fås genom en funktional som beror av elektrondensiteten, men vad denna funktional är går inte att veta exakt. Det är användandet av elektrondensiteten som är grunden till DFT. Ett viktigt tillägg till DFT kom från Kohn och Sham, *KS*, som kombinerade Hartree-Fock-metoden tillsammans med elektrondensiteten genom att återinföra orbitaler i den tidiga DFT:n. I deras approximation används en fiktiv uppsättning av elektroner som inte interagerar med varandra och som definieras ha samma densitet som de verkliga elektronerna.[18] Detta gör att energin för systemet kan skrivas som

$$E = T_S + U + E_{XC} \quad (3)$$

där T_S är KS:s kinetiska energi, U är systemets coulombenergi och E_{XC} som står för *Exchange-Correlation Energy* som då innehåller den del av energin som försvinner i KS-approximationen.[19] Den svåra delen blir då att beräkna E_{XC} . Kohn-Shams motsvarighet till vågfunktioner, Ψ , vilket är KS-orbitaler som vid beräkningar approximeras av basset.

Generellt är ett basset en uppsättning funktioner som kombineras till andra nya funktioner. För de här kvantkemiska beräkningarna representerar basseten

orbitaler. Olika basset varierar bland annat i storlek och därmed hur väl de approximerar orbitaler, men större basset ökar även tiden beräkningar tar. Basset är oftast inte specialiserade för specifika atomer utan väljs utifrån önskad noggrannhet och beräkningstid.

För KS-approximationen används också en funktional approximerad från atomsystem med kända energier för att utföra energiberäkningarna.[18]. Funktionaler är till skillnad från basseten framtagna för att beskriva specifika typer av system, till exempel så finns det funktionaler som är mer anpassade för organiska molekyler och funktionaler som passar övergångsmetaller bättre. Det innebär att valet av funktional är systemberoende, då det inte finns några framtagna funktionaler som fungerar bra för alla typer av atomer.

Med hjälp av denna teori så har två typer av beräkningar utförts: Enkla geometrioptimeringar där energierna för en molekyl eller ett molekylkomplex minimerades, och beräkningar av övergångstillstånd för en reaktionsmekanism.

För reaktionerna som studerades utfördes tre beräkningar: En som optimerade reaktionskomplexet, en som letade efter reaktionens övergångstillstånd och en som optimerade produktkomplexet. Resultaten från dessa beräkningar är optimerade molekylära geometrier samt energier för de tre fallen. För att kunna få fram höjder på reaktionernas energibarriärer och produktenergier så utgick vi från en referens vars energi subtraherades från de beräknade värdena.

Beräkningarna använde funktionalen B3LYP och bassetet 6-31+G(d,p). B3LYP är den mest använda funktionalen för DFT och 6-31+G(d,p) är ett lämpligt basset att använda tillsammans med den för att simulera organiska molekyler.[20][21] De utfördes i en implicit lösning av vatten med en dielektricitetskonstant satt till 144.8 för att efterlikna en HCN-lösning. Det inkluderades även en explicit HCN-molekyl och en CN^- -jon som båda kunde agera som nukleofil, elektrofil eller katalysator.

2.3.1 Orbitalteori

Molekyler har flera orbitaler som är en kombination av de ingående atomernas elektronorbitaler. En del av dem är ockuperade av elektroner, andra är tomma. Den högsta ockuperade kallas "HOMO", *Highest Occupied Molecular Orbital*, och den lägsta icke ockuperade kallas "LUMO", *Lowest Unoccupied Molecular Orbital*. Vid en nukleofil attack kommer elektroner från den nukleofila reaktantens HOMO lägga sig i den elektrofila reaktantens LUMO. Generellt så betyder en lägre skillnad i HOMO- och LUMO-energierna en lägre energibarriär för reaktionen.[22] För att försöka förklara skillnaderna i barriärerna mellan sidogrupperna i cykliseringsreaktionerna så analyserades HOMO och LUMO för alla sidogrupper.

2.4 Kinetikberäkningar

För att avgöra hur fort och hur troligt det är att en viss reaktion sker så räcker det inte med att studera reaktionsbarriärer och energiskillnader mellan produkter och reaktanter. Även om en produkt tycks vara termodynamiskt stabil så kan reaktionshastigheten vara väldigt låg, vilket gör att reaktionen inte kommer att ske

i så stor utsträckning. Kinetikberäkningar är därav en viktig del för att kunna bedöma under vilka temperaturförhållanden en reaktion är intressant.

För en första ordningens reaktion där bakåtreaktionen kan försummas så beskrivs hastighetsekvationen av

$$r = k[C_1] \quad (4)$$

där r är reaktionshastigheten, k är reaktionens hastighetskonstant och $[C_1]$ är koncentrationen av det ämne C_1 som reagerar. Förbrukningen av C_1 ges då av

$$\frac{d[C_1]}{dt} = -r = -k[C_1] \quad (5)$$

Genom att separera variablerna i (5) och integrera kan ett uttryck för halveringstiden härledas till ekvation 6. För härledning, se Bilaga I.

$$t_{1/2} = \frac{-\ln \frac{1}{2}}{k} \quad (6)$$

För att komma åt hastighetskonstanten så kan Eyrings ekvation användas [17]

$$k = \left(\frac{\kappa k_b T}{h} \right) e^{\frac{-\Delta G^\ddagger}{RT}} \quad (7)$$

där k_b är Boltzmanns konstant, T temperaturen i K, h Plancks konstant, $\Delta G^\ddagger$ reaktionens beräknade energibarriärer och R den allmänna gaskonstanten. Konstanten κ representerar sannolikheten att övergångstillståndet inte går över i produkterna.

Det får antas att alla övergångstillstånd övergår i produkt och att κ därför kan antas vara lika med 1. Med detta antagande kan Eyrings ekvation, (7), användas för att få fram halveringstiden tillsammans med ekvation (6) för att tydligare se hur höjderna på de olika barriärerna påverkar verkligheten.

Halveringstiderna beräknades vid olika temperaturer som kan ses som intressanta utifrån ett prebiotiskt perspektiv. Dessa inkluderar Titans genomsnittliga yttemperatur (90 K)[23], den lägsta och högst yttemperaturen på kometen 67P (150 K och 230 K)[24], 0 °C, ungefär minimal och maximal temperatur på hydrotermiska källor (293 K och 623 K)[25], undre och övre uppskattningar av temperaturen i arkeiska haven (303 K och 343 K)[26] och den högsta temperaturen där liv har upptäckts (393 K)[27].

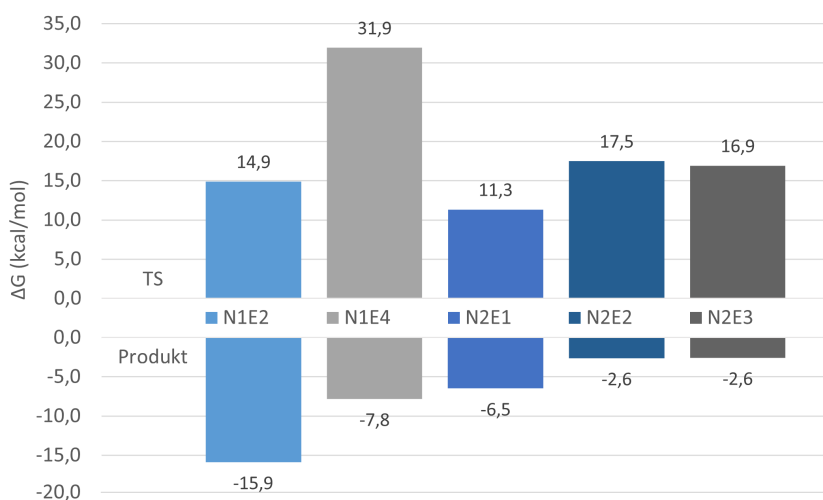
3 Resultat och diskussion

För att få en gemensam utgångspunkt för alla reaktioner så baserades dem på den mest stabila stereoisomeren av M4. Denna valdes utifrån fyra olika optimerade stereoisomerer. Se Bilaga II för alla stereoisomerernas energier. Orientering och visualisering av den mest stabila stereoisomeren återfinns tillsammans med en CN^- -jon, och en HCN-molekyl i Figur 2.

3.1 Undersökning av olika nukleofila attacker 4

I de första reaktionerna som studerades bestod reaktantkomplexet av en HCN-molekyl och en CN^- -jon som reagerade med M4.

De beräknade energierna för övergångstillstånd och produkter för reaktionerna $Ni, i = 1, 2, Ej, j = 1, 2, 3$ visualiseras i Figur 4. Inledningsvis kan det fastslås att samtliga produkter har negativa energier relativt reaktionskomplexet och därav är stabila.



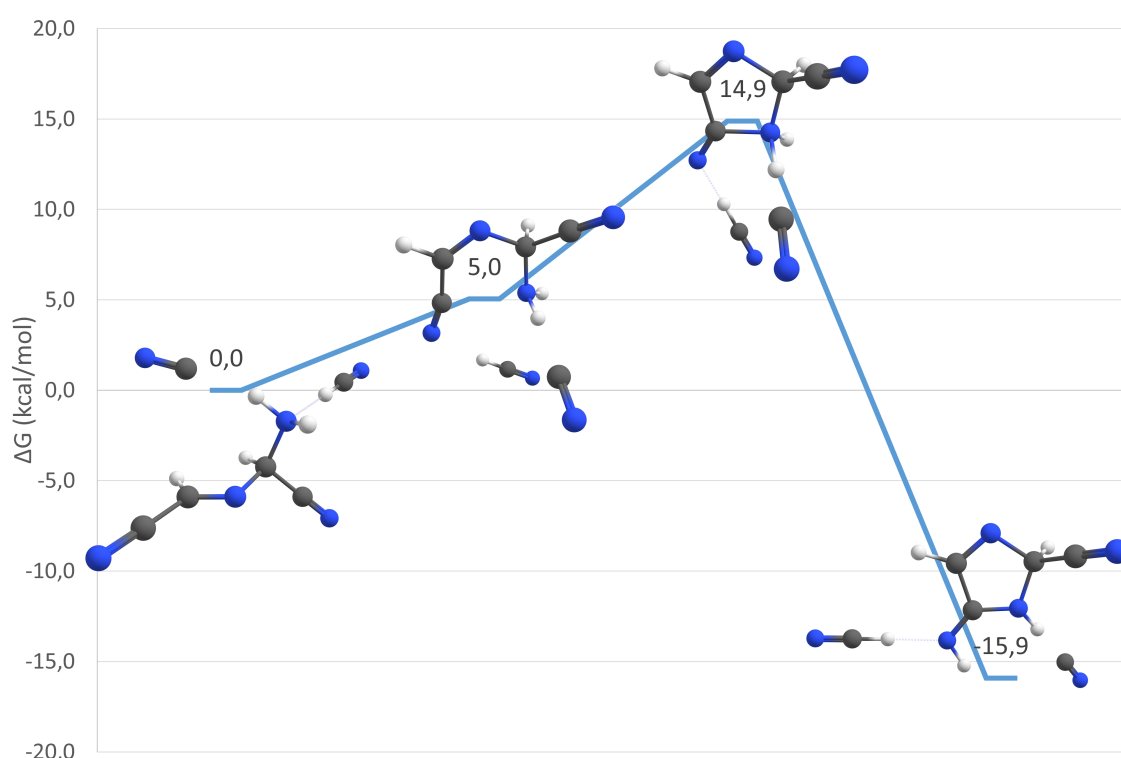
Figur 4: Energier för övergångstillstånd, här kallat TS, och produktkomplex relativt reaktionskomplexet.

I Figur 4 går det att se att reaktionerna N2E2 och N2E3 har samma produktenergier och mycket närliggande energibarriärer. Det är förväntat då de två reaktionerna är mycket lika: En cyanidjon gör en nukleofil attack på en cyanogrupp. Skillnaden är omgivningen till kolet som cyanogruppen sitter på. Från detta kan vi konstatera att samma typer av reaktioner har termodynamiskt lika egenskaper, även om de sker på olika delar av en molekyl.

Reaktionen N1E4 har den högsta barriären (31,9 kcal/mol) enligt Figur 4, vilket är betydligt högre jämfört med de övriga reaktionerna. Framför allt är barriären mer än dubbelt så hög som N1E2 (14,9 kcal/mol) trots att de är av samma typ av reaktion: En amin som attackerar ett kol i en cyanogrupp. Skillnaden mellan

reaktionerna skulle kunna bero på att övergångstillståndet i N1E2 skapar en ringstruktur som kan vara stabilare än det linjära övergångstillståndet i N1E4.

Den mest stabila produkten fås av reaktionen N1E2, vars produktenergi ligger på -15,9 kcal/mol. Detta är inte förvånande då ringformade molekyler bestående av fem till sex atomer ofta är mycket stabila. Den har även näst lägst barriär på 14,9 kcal/mol. Eftersom beloppet av produktens energi är större än energibarriären skulle denna reaktion teoretiskt kunna vara självdrivande, förutsatt ett slutet system och att endast denna reaktion kan ske. Reaktionsvägen för N1E2 visas i Figur 5. Utöver reaktions-, produktkomplex och övergångstillstånd visas även den stereoisomer av M4 som är mest lik övergångstillståndet, ett så kallat förreaktionskomplex. Detta komplex demonstrerar den relativt stora skillnad i energi (5,0 kcal/mol) som kan finnas mellan olika stereoisomerer.



Figur 5: Reaktionsvägen för N1E2 med stabilt reaktantkomplex, förreaktionskomplex, övergångstillstånd och produktkomplex.

Reaktion N2E1 har den lägsta barriären (11,3 kcal/mol) vilket gör den intressant då den reaktionen kommer att ske snabbast. Även om det är den enda reaktionen av sitt slag, en nukleofil attack på en imin, finns det en likhet med N2E2 och N2E3 då alla tre attackerna sker på ett kol bundet med mer än en enkelbindning till ett kväve. Skillnaderna i barriärernas storlekar kan förklaras i att cyanogruppen i N2E2 och N2E3 är mindre reaktiv än iminen i N2E1. Detta då trippelbindningen i de tidigare två fallen är stabilare än dubbelbindningen i det senare.

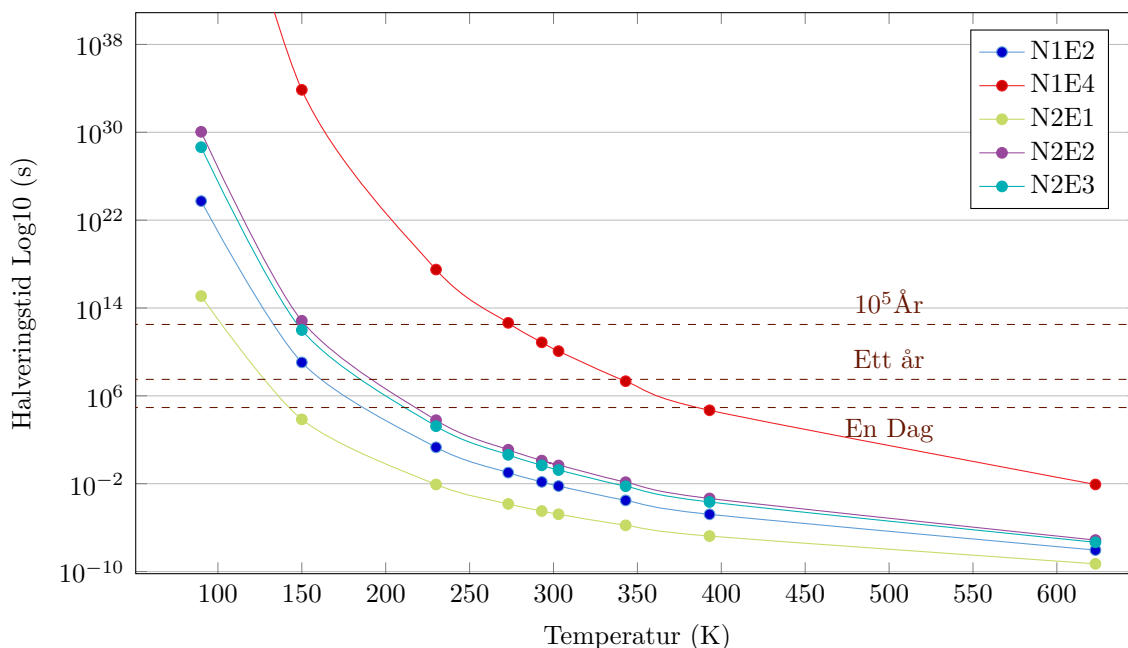
Det är även intressant att jämföra produktenergin för N2E1 (-6,5 kcal/mol) med den för N1E2 (-15,9 kcal/mol). Att N2E1 har den lägre barriären innebär att N2E1

kommer att ske fortare än N1E2, men då produkten inte är lika stabil kommer även bakåtreaktionen att ske snabbare. N2E1 är därför kinetiskt gynnad, och i ett slutet system där produkterna inte lätt kan avlägsnas, till exempel på ytan av en komet, lär dess produkt dominera vid låga temperaturer. Vid höga temperaturer lär däremot produkten från N1E2 dominera då bakåtreaktionen inte är lika snabb.

3.1.1 Halveringstider för M4-reaktionerna

I Figur 6 visas halveringstider för de olika M4-reaktionerna beräknade med ekvationerna (6) och (7). Vid Titans yttemperatur har den snabbaste reaktionen, N2E1, en halveringstid på flera miljoner år. Därmed kommer ingen av dessa reaktioner ske i någon betydande utsträckning vid så låga temperaturer. Det är vid 150 K som de flesta av reaktionerna börjar kunna ske på en tidsskala av hundratusentals år. Sådana reaktionshastigheter kan vara rimliga om man ser på, till exempel, ytan av en komet som kan gå mycket långa perioder utan att reaktionerna skulle störas av utomstående påverkan. Vid 230 K har alla reaktioner förutom N1E4 en halveringstid som är under en dag.

Figur 6 demonstrerar även att N2E1 är kinetiskt gynnad gentemot N1E2. Vid 150 K har N2E1 en halveringstid under en dag jämfört med den för N1E2 som är över ett år.



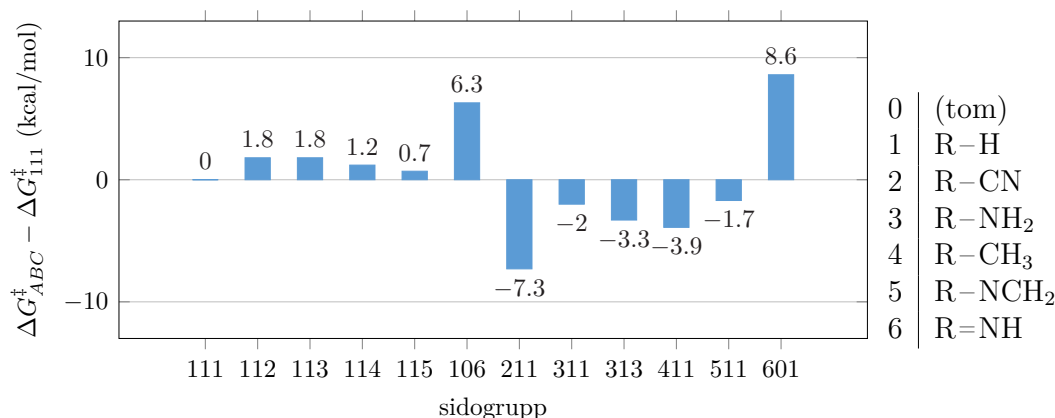
Figur 6: Halveringstid i sekunder för de fem olika M4-reaktionerna i logaritmisk skala mot temperaturen i Kelvin. Avskuren punkt för N1E4 ligger vid 10^6 s.

3.2 Vidare undersökning av cykliseringsreaktionen

För cykliseringsreaktionen N1E2 visas här hur reaktionsbarriärer och produktenergier påverkas av olika sidogrupeer. Som referens beräknades först

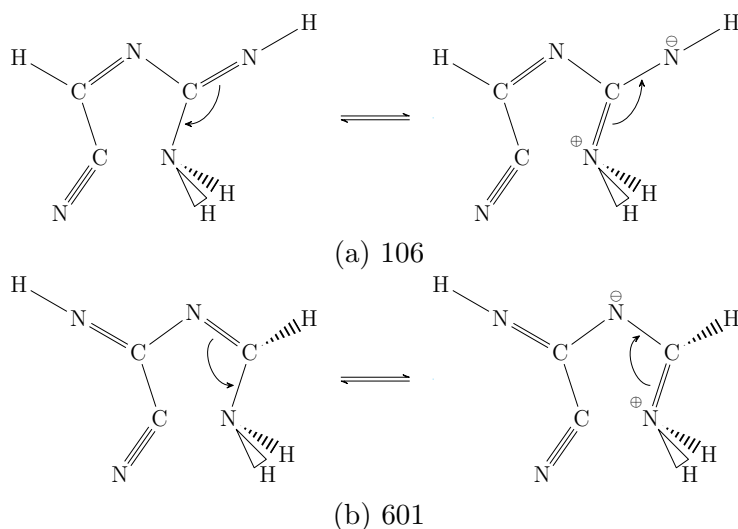
energin för respektive molekyls reaktionskomplex.

Alla produktkomplexens energier blev lägre än för reaktionskomplexen, dessa värden finns tabellerade i Bilaga III. Reaktionsbarriärerna är därför mer intressanta att studera närmare och redovisas i Figur 7. Till skillnad från tidigare anges istället värdena relativt barriären för 111-reaktionen. Denna valdes som referens då den har samma grupp (en väteatom) i alla tre positioner och kan ses som ett enklaste fall. De faktiska energibarriärerna syns i Bilaga III.



Figur 7: Beräknade energibarriärer för alla sidogrupper relativt de för 111. Till höger en påminnelse om vilka sidogrupperna är.

Den tydligaste trenden är att positionen för sidogruppen är avgörande. Ta som exempel grupp 4 (metylgrupp) vilken höjer barriären när den placeras i position C (114) men sänker den till näst lägst när den placeras i position A (411). Undantaget är grupp 6 som höjer barriären markant i båda positionerna. Att den avviker på detta sätt kan förklaras med att reaktanterna med denna grupp har resonansstrukturer som stabiliserar molekylen och gör den mindre reaktiv, dessa resonansstrukturer visas i Figur 8.

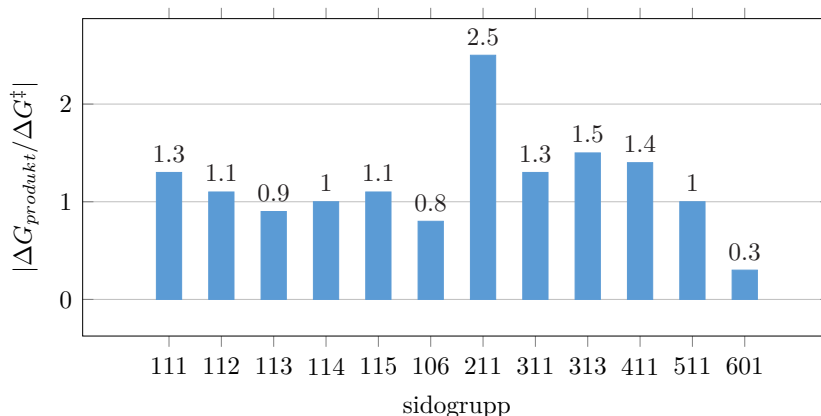


Figur 8: Resonansstrukturer för 601 och 106.

Av de undersökta substitutionerna är det bara 112 och 211 som är isomerer av M4. 211 var även den molekyl vars reaktion hade överlägset lägst barriär. Detta är särskilt intressant då det bidrar med ytterligare argument för att HCN-baserade reaktioner är en viktig del av prebiotisk kemi.

För att studera vilken av positionerna A,B och C som var mest inflytelserik beräknades även energierna för 313, en molekyl med aminogrupeer i A och C. Att lägga till ytterligare en grupp sänker barriären relativt både 113 och 311 till ett värde som ligger närmare det för 311 än för 113. Detta antyder att sidogruppen närmast det elektrofila kolet (position A) verkar ha en större vikt i reaktionen än vad den närmast det nukleofila kvävet (position C) har.

Likt N1E2 så har en del av cykliseringarna en produkt vars energi är i belopp större än energibarriären för reaktionen. I Figur 9 visas skillnaden i bråkform, där ett värde över 1 innebär att reaktionen kan teoretiskt vara självdrivande energimässigt. 211 har ett högt bråk på 2.5 vilket betyder att en sådan reaktionen skulle teoretiskt kunna ge energi till två till.

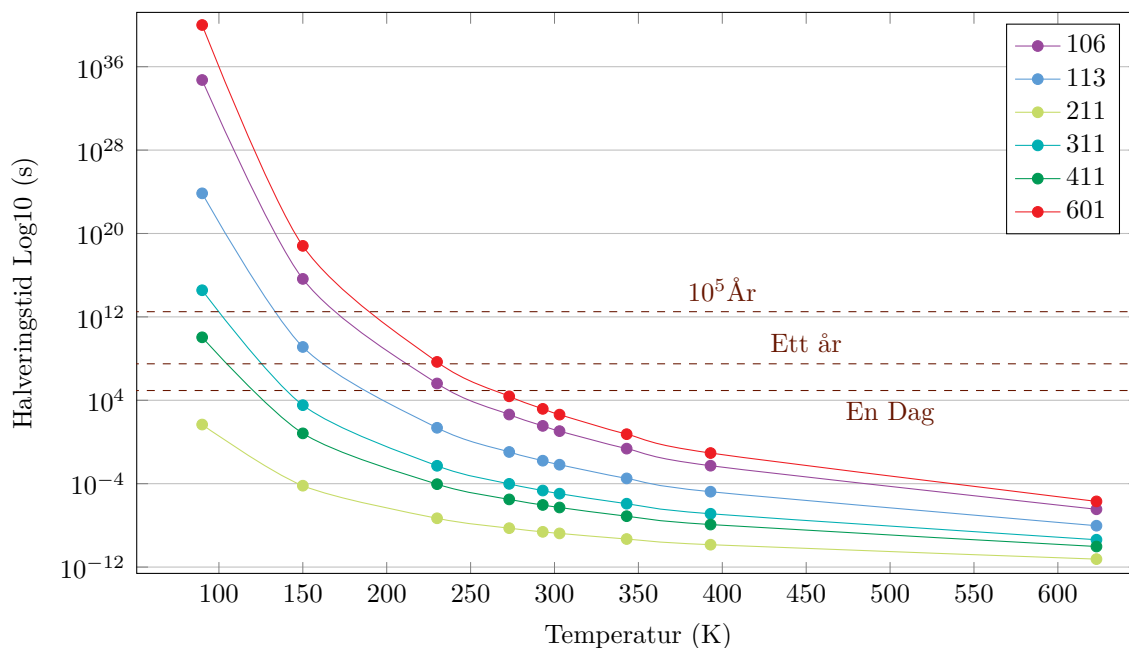


Figur 9: Cykliseringsreaktionernas produktenergi genom barriärsenergi.

3.2.1 Halveringstider för cyklisering

I Figur 10 visas halveringstiderna för ett urval av de undersökta självcykliseringarna. Resterande reaktioner faller mellan 411 och 113, se Bilaga IV för figur med samtliga halveringstider.

211 har väldigt kort halveringstid och redan vid 90 K är den på några tiotal sekunder, vilket är 10^8 gånger snabbare än den närmaste reaktionen, 411.



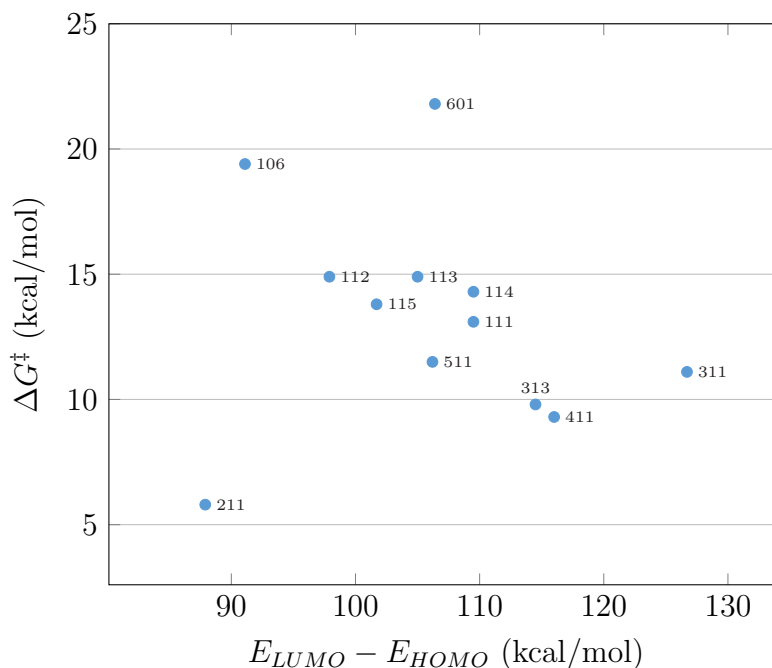
Figur 10: Halveringstid för ett urval undersökta självcykliseringar. Halveringstider för samtliga cykliseringar hittas i Bilaga III.

3.2.2 Molekylorbitalanalys

För att försöka förklara skillnaderna mellan de olika sidogrupperna analyseras HOMO och LUMO. I Figur 11 visas energibarriärerna för varje reaktion mot energin i LUMO för respektive molekyls reaktionskomplex minus energin i HOMO för en CN^- -jon. Värdena för flera olika reaktionskomplex analyserades och alla gav liknande resultat som Figur 11. Resterande figurer kan hittas i Bilaga V.

Med undantag för 211 och 601 verkar skillnaden i energi mellan LUMO och HOMO vara omvänt proportionell mot energibarriären, vilket är det motsatta förhållandet som teorin förutsäger. Detta skulle kunna bero på att KS-orbitalerna som används i DFT inte exakt representerar molekylorbitaler, eller att teorin som kopplar ihop HOMO- och LUMO-energierna med reaktivitet inte lämpar sig väl för att beskriva intramolekylära reaktioner. Det skulle också kunna bero på att alla molekylers övergångstillstånd är mycket stabila på grund av dess ringstruktur, vilket då skulle höja barriären.

Energiskillnaden mellan HOMO och LUMO för 211 och 106 är relativt lika, men de har mycket olika barriärer. Det antyder att andra faktorer än energiskillnaden i HOMO och LUMO påverkar hur höga barriärerna blir.

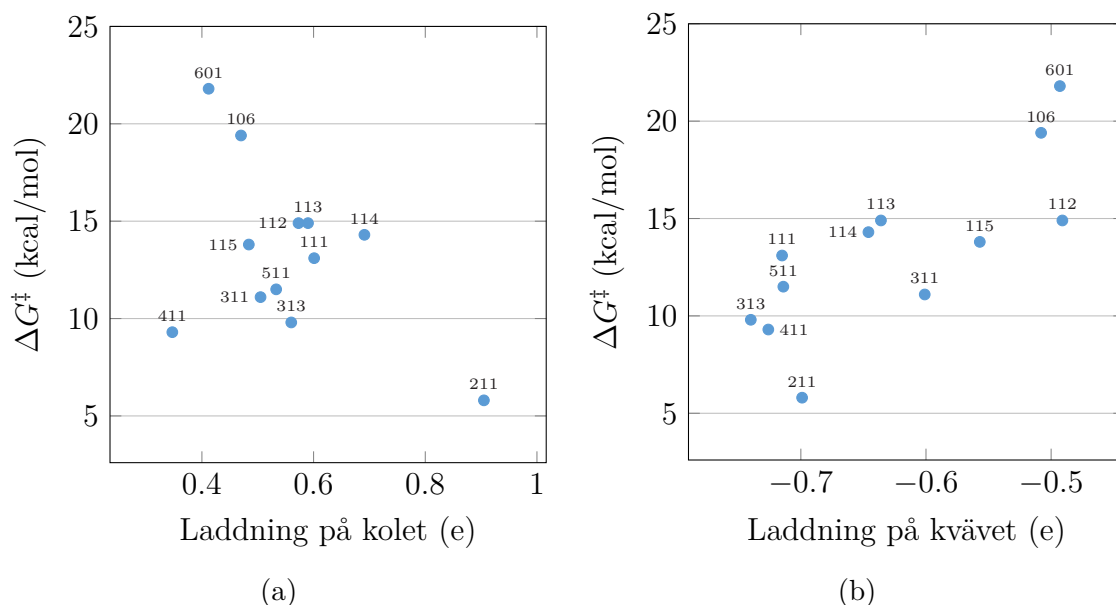


Figur 11: Energibarriären för varje molekyl plottade mot skillnaden i energi mellan LUMO för reaktionskomplex utan CN^- och HOMO för en CN^- -jon.

3.2.3 Analys av laddningar

För att hitta ytterligare förklaringar till varför sidogrupperna påverkar barriärerna studerades även molekylernas partiella laddningar. En jämförelse mellan barriärerna och laddningarna hos det elektrofila kolet respektive det nukleofila kvävet syns i Figur 12.

I laddningarna på kolet i Figur 12a syns en möjlig trend där en större positiv laddning sänker barriären. 411 avviker från denna trend, vilket möjligen kan bero på att det reagerande kolet är mer elektronegativt än det i metylgruppen. En motsvarande trend, att barriärerna sjunker med en mer negativ laddning på kvävet, syns i Figur 12b där barriärerna istället jämförs med kvävet's laddning.



Figur 12: Energibarriären för molekylerna mot laddningen på det elektrofila kolet a) och det nukleofila kvävet (b). Notera att laddningarna i (b) är negativa.

Båda fallen i Figur 12 passar in i teorin om nukleofila attacker. Reaktiviteten gynnas av att elektrofilen är mer positiv och av att nukleofilen är mer negativ. De här resultaten ger en möjlig förklaring till varför de olika sidogrupperna ger olika reaktionsbarriärer. Frågan som kvarstår är varför sidogrupperna påverkar de reagerande atomernas laddningar på detta sätt. Cyanogruppen är mycket elektrondragande vilket kan ses i Figur 12 på det mycket positiva kolet i 211 och den låga negativa laddningen på kvävet i 112. En imingrupp gör kvävet mer positiv oavsett på vilken sida den sitter, som ses av 106 och 601 i Figur 12b. Detta kan vara på grund av resonansen i de två molekylerna som diskuterats innan och visas i Figur 8. Alla grupper verkar agera elektrondragande när de sitter på kvävet sida, och alla grupper förutom cyanogruppen verkar agera elektrondonerande när de sitter på kolets sida. Detta är underligt då kväve är mer elektronegativt än kol, åtminstone ensamt, men det skulle kunna förklaras av att kolet ingår i en cyanogrupp.

4 Slutsats

Vi ser att det finns konkurrerande reaktioner som är mer kinetiskt gynnade, men att cykliseringsreaktionen är mer termodynamiskt gynnad och kan bli den mest dominanta reaktionen under högre temperaturer.

Vi har även funnit ett flertal rimliga cykliseringsreaktioner som kan bilda fematomiga ringar. Två av reaktionerna, 112 och 211, inkluderar reaktanter som är tetramerer av HCN, vilket vidare styrker tanken att HCN kan ha bidragit till uppkomsten av biologiskt intressanta molekyler. Särskilt intressant är att 211 har överlägset lägst barriär av samtliga studerade reaktioner.

Vårt arbete antyder även att information om energinivåer för HOMO och LUMO är av begränsat värde för att studera cykliseringsreaktioner i HCN-polymerer. Arbetet visar dock på värdet av att undersöka effekterna av reaktanternas partiella laddningar på reaktiviteten.

4.1 Framtidsutsikter

Det här arbetet har förhoppningsvis bidragit med ett litet steg på den prebiotiska kemins väg mot exempelvis adenin. Framtida arbeten kan fokusera på att hitta fler steg på den reaktionsvägen eller möjligen studera reaktioner med H_2O för att hitta steg på reaktionsvägarna till andra nukleobaser.

Då nukleobaser är baserade på heterocycliska ringar skulle framtida arbeten också kunna vidare studera andra möjliga cykliseringsreaktioner, liknande de som studerades här. Det skulle också vara intressant att undersöka hur våra cykliserade produkter kan reagera vidare. Även för cykliseringsreaktioner kan reaktioner med H_2O studeras, men även andra sidogrupper kan vara intressanta.

Molekylen M4 som vi har utgått ifrån valdes främst för att den innehöll flera funktionella grupper samt att den var en tetramer av HCN. Utifrån vår kännedom så saknas det kunskap om hur denna molekylen bildas, vilket är ytterligare ett område som framtida arbeten kan fokusera på.

Erkännanden

Beräkningarna och simuleringarna var möjliggjorda genom resurser tillhandahållna av "The Swedish National Infrastructure for Computing" (*SNIC*) vid "Chalmers Center for Computational Science and Engineering" (*C3SE*) som delvis är finansierat av vetenskapsrådet genom bidragsavtal nr. 2018-05973.

Ett stort tack RahmLab för deras stöd och handledning.

Litteraturförteckning

- [1] C.P. McKay och W.L. Davis: *Chapter 10 - Astrobiology*. Elsevier, Boston, tredje utgåva, 2014, ISBN 978-0-12-415845-0. <https://www.sciencedirect.com/science/article/pii/B9780124158450000104>.
- [2] S.A. Sandford, M. Nuevo, P.P. Bera och T.J. Lee: *Prebiotic Astrochemistry and the Formation of Molecules of Astrobiological Interest in Interstellar Clouds and Protostellar Disks*. Chemical Reviews, 120(11):4616–4659, 2020. <https://doi.org/10.1021/acs.chemrev.9b00560>, PMID: 32227889.
- [3] J. Orò: *Mechanism of Synthesis of Adenine from Hydrogen Cyanide under Possible Primitive Earth Conditions*. Nature, 191:1193–1194, 1961.
- [4] M. Levy, S.L. Miller, K. Brinton och J.L. Bada: *Prebiotic Synthesis of Adenine and Amino Acids Under Europa-like Conditions*. Icarus, 145(2):609–613, 2000, ISSN 0019-1035. <https://www.sciencedirect.com/science/article/pii/S0019103500963656>.
- [5] Y. Oba et al.: *Identifying the wide diversity of extraterrestrial purine and pyrimidine nucleobases in carbonaceous meteorites*. Nat Commun, 13(2008), 2022.
- [6] T. Das, S. Ghule och K. Vanka: *Insights Into the Origin of Life: Did It Begin from HCN and H₂O?* ACS Central Science, 5(9):1532–1540, 2019.
- [7] J.E. Šponer, A. Mládek, J. Šponer och M. Fuentes-Cabrera: *Formamide-Based Prebiotic Synthesis of Nucleobases: A Kinetically Accessible Reaction Route*. The Journal of Physical Chemistry A, 116(1):720–726, 2012. <https://doi.org/10.1021/jp209886b>, PMID: 22129168.
- [8] R.A. Sanchez, J.P. Ferbis och L.E. Orgel: *Studies in Prebiotic Synthesis: II. Synthesis of purine precursors and amino acids from aqueous hydrogen cyanide*. Journal of Molecular Biology, 30(2):223–253, 1967, ISSN 0022-2836. <https://www.sciencedirect.com/science/article/pii/S0022283667800378>.
- [9] J.P. Ferris och L.E. Orgel: *Aminomalononitrile and 4-Amino-5-cyanoimidazole in Hydrogen Cyanide Polymerization and Adenine Synthesis*. Journal of the American Chemical Society, 87(21):4976–4977, 1965.

- [10] R. Glaser, B. Hodgen, D. Farrelly och E. McKee: *Adenine Synthesis in Interstellar Space: Mechanisms of Prebiotic Pyrimidine-Ring Formation of Monocyclic HCN-Pentamers*. *Astrobiology*, 7(3):455–470, 2007. <https://doi.org/10.1089/ast.2006.0112>, PMID: 17630841.
- [11] S. Slavova och V. Enchev: *Self-catalytic mechanism of prebiotic reactions: From formamide to purine bases*. *International Journal of Quantum Chemistry*, 120(19):e26362, 2020.
- [12] C.N. Matthews och R.D. Minard: *Hydrogen cyanide polymers, comets and the origin of life*. *Faraday Discuss.*, 133:393–401, 2006. <http://dx.doi.org/10.1039/B516791D>.
- [13] M. Ruiz-Bermejo, J. L. de la Fuente, C Pérez-Fernández och E. Mateo-Martí: *A Comprehensive Review of HCN-Derived Polymers*. *Processes*, 9(4), 2021, ISSN 2227-9717. <https://www.mdpi.com/2227-9717/9/4/597>.
- [14] G.A. Zhurko: *Chemcraft - graphical program for visualization of quantum chemistry computations*, 2005. <https://chemcraftprog.com>.
- [15] M.J. Frisch et al.: *Gaussian~16 Revision C.01*, 2016. Gaussian Inc. Wallingford CT.
- [16] H. Sandström och M. Rahm: *The Beginning of HCN Polymerization: Iminoacetonitrile Formation and Its Implications in Astrochemical Environments*. *ACS Earth and Space Chemistry*, 5(8):2152–2159, 2021. <https://doi.org/10.1021/acsearthspacechem.1c00195>.
- [17] P. Atkins, J. de Paula och R. Friedman: *Physical Chemistry Quanta, Matter and Change*. Oxford University Press, 2013, ISBN 978-0-190-960981-9.
- [18] K. Burke och L.O. Wagner: *DFT in a nutshell*. *International Journal of Quantum Chemistry*, 113(2):96–101, 2013.
- [19] E. Wimmer: *Computational materials design and processing: perspectives for atomistic approaches. I C,H,N and O in Si and Characterization and Simulation of Materials and Processes*, European Materials Research Society Symposia Proceedings, sidor 72–82. Elsevier, Oxford, 1996. <https://www.sciencedirect.com/science/article/pii/B9780444824134500809>.
- [20] S.F. Sousa, P.A. Fernandes och M.J. Ramos: *General Performance of Density Functionals*. *The Journal of Physical Chemistry A*, 111(42):10439–10452, 2007. <https://doi.org/10.1021/jp0734474>, PMID: 17718548.
- [21] J. Tirado-Rives och W.L. Jorgensen: *Performance of B3LYP Density Functional Methods for a Large Set of Organic Molecules*. *Journal of Chemical Theory and Computation*, 4(2):297–306, 2008. <https://doi.org/10.1021/ct700248k>, PMID: 26620661.
- [22] J. Clayden, N. Greeves och S. Warren: *Organic Chemistry*. OUP Oxford, 2012, ISBN 9780199270293.

- [23] D.E. Jennings et al.: *Surface temperatures on Titan during northern winter and spring*. The Astrophysical Journal, 816(1):L17, jan 2016.
<https://doi.org/10.3847/2041-8205/816/1/L17>.
- [24] F. Tosi et al.: *The changing temperature of the nucleus of comet 67P induced by morphological and seasonal effects*. Nature Astronomy, 3(7):649–658, Jul 2019, ISSN 2397-3366. <https://doi.org/10.1038/s41550-019-0740-0>.
- [25] A.D. Rogers et al.: *The discovery of new deep-sea hydrothermal vent communities in the southern ocean and implications for biogeography*. PLoS biology, 10(1):e1001234–e1001234, Jan 2012, ISSN 1545-7885.
<https://pubmed.ncbi.nlm.nih.gov/22235194>, 22235194[pmid].
- [26] A.M. Hessler: *Earth's Earliest Climate*. Nature Education Knowledge, 3(10):24, 2011.
- [27] K. Kashefi och D.R. Lovley: *Extending the Upper Temperature Limit for Life*. Science, 301(5635):934–934, 2003.
<https://www.science.org/doi/abs/10.1126/science.1086823>.

I Härledning av uttryck för halveringstiden

Härledning av uttrycket för halveringstiden $t_{1/2}$ nedan från reaktionens hastighetsekvation genom variabelseparation och integrering över intervallet $([C_1]_0, \frac{[C_1]_0}{2})$ och $(0, t_{1/2})$ där $[C_1]_0$ är en godtycklig startkoncentration av ämne C_1 och $t_{1/2}$ är halveringstiden. Halveringstiden blir en funktion av reaktionens hastighetskonstant.

$$\frac{-1}{[C_1]} d[C_1] = k dt \quad (8a)$$

$$\int_{[C_1]_0}^{\frac{[C_1]_0}{2}} \frac{-1}{[C_1]} d[C_1] = \int_0^{t_{1/2}} k dt \quad (8b)$$

$$- \left(\ln\left(\frac{[C_1]_0}{2}\right) - \ln([C_1]_0) \right) = k(t_{1/2} - 0) \quad (8c)$$

$$- \ln\left(\frac{[C_1]_0}{2}\right) = -\ln\left(\frac{1}{2}\right) = kt_{1/2} \quad (8d)$$

$$t_{1/2} = \frac{-\ln\left(\frac{1}{2}\right)}{k} \quad (8e)$$

II Energier för undersökta M4-stereoisomerer

Fria energier för de fyra undersökta stereoisomerna till molekyl M4 återfinns i Tabell B1, värdet för 4 HCN avser fallet då HCN-molekyler är placerade på rad efter varandra, där vätebindningar har uppstått mellan väten och kväven. Struktur S2 är mest stabil, och används i rapportens fortsatta beräkningar.

Tabell B1: Energier för de fyra undersökta M4-isomererna i kcal/mol, värdet ΔG är relativt energin för fyra HCN på rad.

Struktur	G (kcal/mol)	ΔG (kcal/mol)
4 HCN	-234529,4	0.0
S1	-234542,3	-13.0
S2	-234543,5	-14.2
S3	-234541,1	-11.7
S4	-234539,7	-10.4

III Cykliseringsenergier för reaktant- och produktkomplex

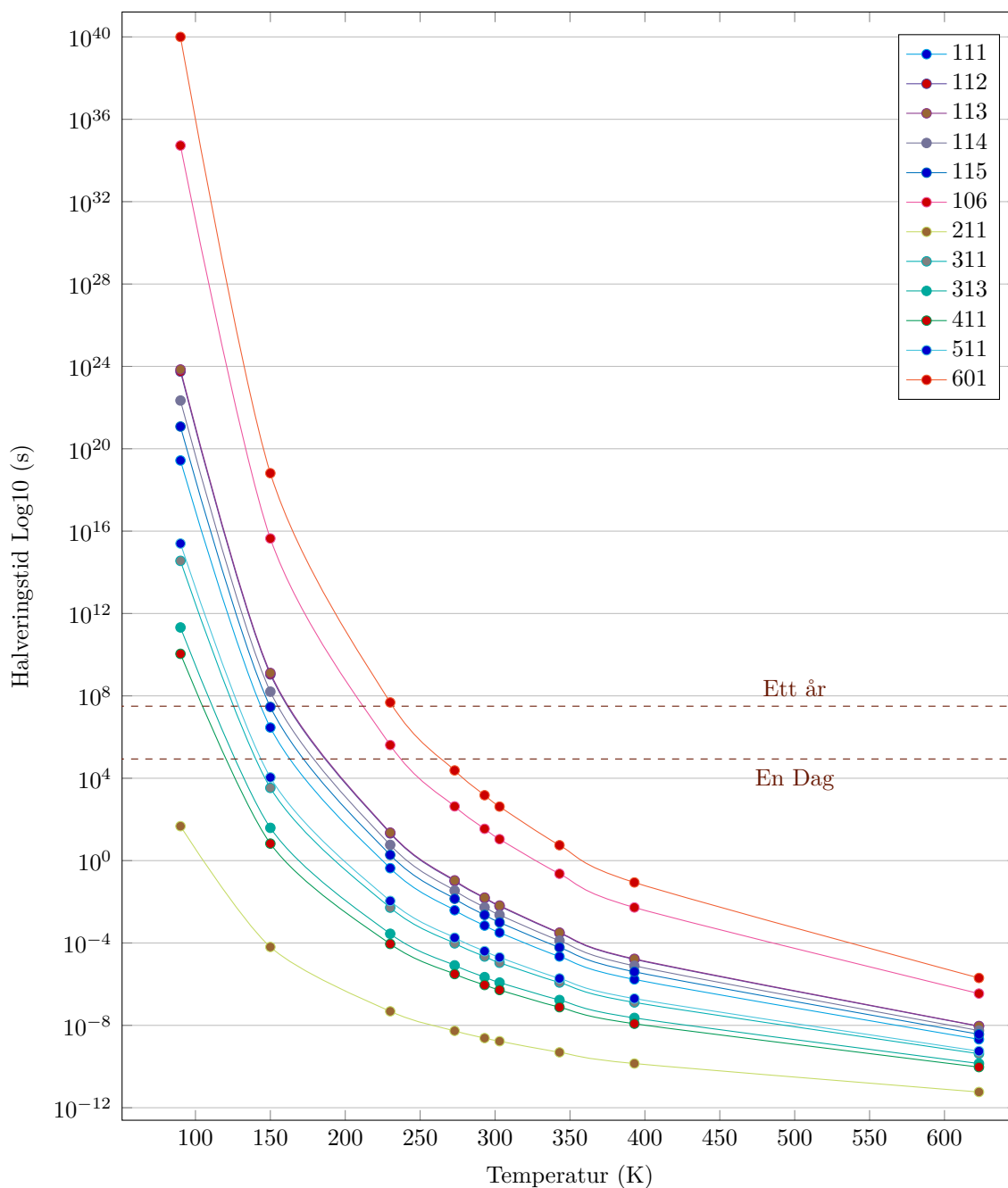
För de undersökta cykliseringsreaktionerna optimerades både reaktions- och produktkomplex, där samtliga produktkomplex var stabilare än respektive reaktantkomplex. Det kan utläsas från Tabell B2, där ΔG (kcal/mol), differensen mellan produkt och reaktant, är negativ i samtliga fall. Även reaktionernas beräknade energibarriärer (notera ej relativt 111) visas i Tabell B2.

Tabell B2: Energier relativt respektive reaktionskomplex för övergångstillstånd, ΔG_{TS} , och produktenergier, ΔG_{prod} , för cykliseringsreaktionen.

Struktur	$\Delta G^\ddagger$ (kcal/mol)	ΔG_{prod} (kcal/mol)
111	13,1	-16,6
112	14,9	-14,3
113	14,9	-13,9
114	14,3	-15,0
115	13,8	-14,6
106	19,4	-15,7
211	5,8	-14,3
311	11,1	-14,5
313	9,8	-14,4
411	9,3	-12,6
511	11,5	-11,3
601	21,8	-6,40

IV Halveringstider för samtliga undersökta cykliseringsreaktioner

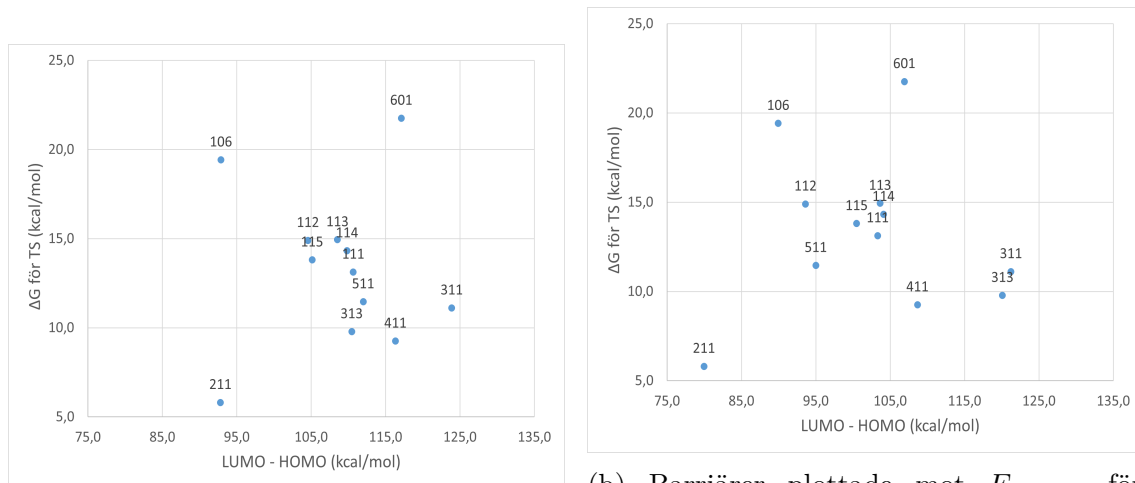
Halveringstider för samtliga tolv undersökta självcykliseringsreaktioner presenteras i sin helhet i Figur B1.



Figur B1: sammanställning av alla halveringstider för cykliserings reaktionen

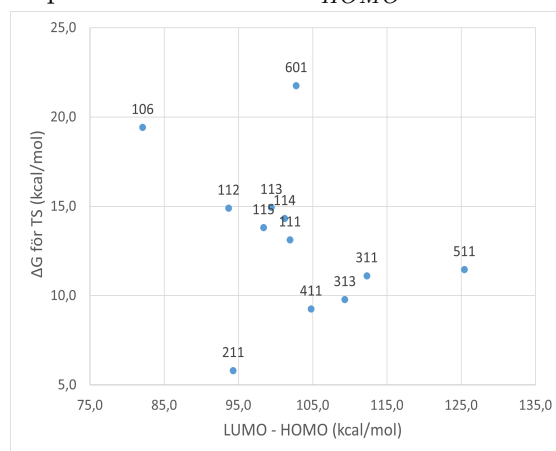
V Molekylorbitalanalys

I Figur B2 visas fler diagram med energibarriärerna jämfört med skillnaden av HOMO- och LUMO-energierna för olika komplex och isomerer. Figurerna liknar varandra och har en möjlig generell trend av lägre barriärer vid större skillnad på HOMO och LUMO, förutom för 211 och 601.



(a) Barriärer plottade mot E_{LUMO} minus E_{HOMO} för reaktionskomplex med CN^-

(b) Barriärer plottade mot E_{LUMO} för stabilt reaktionskomplex utan CN^- minus E_{HOMO} för CN^-



(c) Barriärer plottade mot E_{LUMO} minus E_{HOMO} för övergångstillstånden

Figur B2: Energibarriärer för cykliseringsreaktionerna i förhållande till HOMO- och LUMO-energierna.

INSTITUTIONEN FÖR KEMI OCH KEMITEKNIK
AVDELNINGEN FÖR KEMI OCH BIOKEMI
CHALMERS TEKNISKA HÖGSKOLA
Göteborg, Sverige 2022
www.chalmers.se



CHALMERS
UNIVERSITY OF TECHNOLOGY