

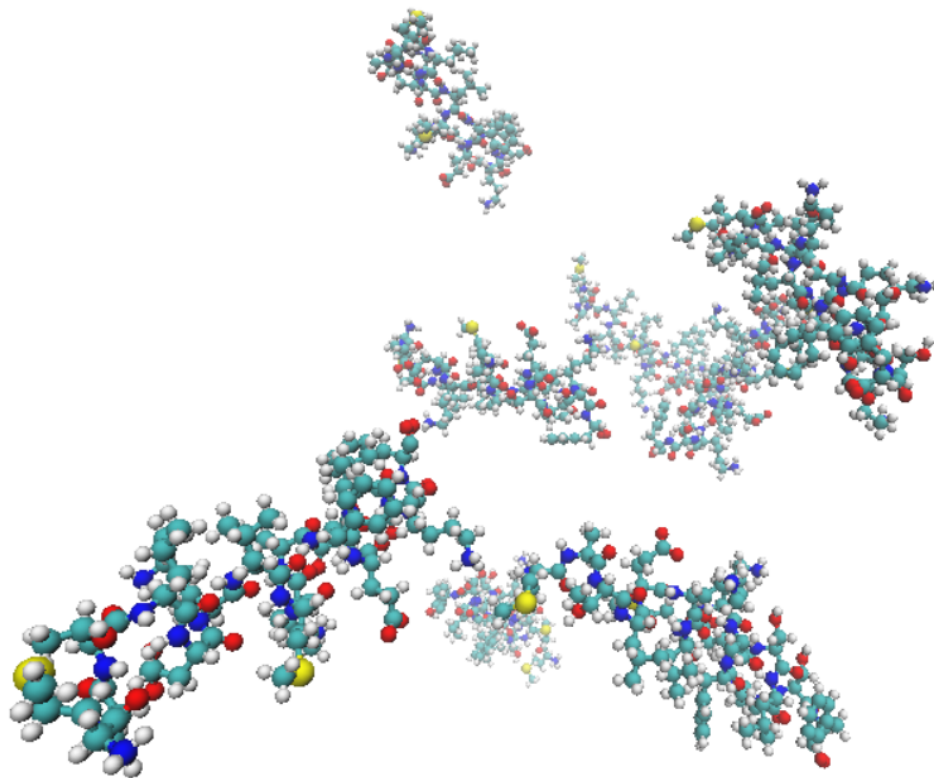


**CHALMERS**  
UNIVERSITY OF TECHNOLOGY

# CHALMERS

## University of Technology

TIFX04-21-82 - KANDIDATARBETE VID FYSIK



# Atomära molekylodynamiksimuleringar av interaktionen mellan disackarider och Huntingtinpeptider

- En undersökning av sackaros och trehalos förmåga  
att hämma proteinaggregering i olika miljöer

BORG OLIVIA  
EINARSSON ELINOR  
PODDA GRAHN ERIK  
WEBER MARIA  
ÖGREN LINNÉA

14 maj 2021

# All-atom Molecular Dynamics Simulations of the Interaction Between Disaccharides and Huntingtin Peptides

- A study of the ability of sucrose and trehalose to inhibit protein aggregation in different environments

## Abstract

Huntington's is a neurodegenerative disease for which there is currently no cure. It occurs as a consequence of a mutation in the so-called Huntingtin protein, which later leads to the protein forming toxic aggregates. This project aims to study how two major protein stabilizing saccharides - sucrose and trehalose - affect two smaller peptides - 2LD2 and 6N8C - of the larger Huntingtin protein, and whether the saccharides can reduce their aggregation. It is further investigated whether the water concentration in the environment of the peptides affects the stabilizing ability of the saccharides. To study the intermolecular interactions between the saccharides and peptides at the atomic level and with high time resolution, all-atom molecular dynamics simulations was applied. The simulations were analyzed using a number of selected analysis methods, where the results indicate both saccharides as good stabilizers, with trehalose as the foremost of the two. The effect of the water concentration on the stabilization is determined to be minimal, and a specific secondary structure interval can be linked to the maximum stabilization of the peptides in the systems studied. Lastly, several amino acids are identified as possible candidates to play a central role in the aggregation.

## Sammandrag

Huntingtons är en neurodegenerativ sjukdom vilken idag saknar botemedel. Den uppstår som konsekvens av en mutation i det så kallade Huntingtinproteinet, vilket senare leder till att proteinet bildar toxiska aggregat. Denna studie syftar till att studera hur två framstående proteinstabiliserande sackarider - sackaros och trehalos - påverkar två mindre peptider - 2LD2 och 6N8C - av det större Huntingtinproteinet, och huruvida sackariderna kan minska deras aggregering. Vidare undersöktes om vattenkoncentrationen i peptidernas omgivning påverkar sackaridernas stabiliseringsförmåga. För att studera de intermolekylära interaktionerna mellan sackariderna och peptiderna tillämpades atomära molekylodynamiksimuleringar, med vilka interaktioner hos molekylerna kan studeras på atomär nivå och med hög tidsupplösning. Simuleringarna analyserades med hjälp av ett antal utvalda analysmetoder, där resultaten indikerar båda sackariderna som goda stabilisatorer, med trehalos som den främsta av de två. Vattenmängdens påverkan på stabiliseringen fastslås som minimal, och ett specifikt sekundärstruktursintervall kan kopplas till maximal stabilisering av peptiderna i de studerade systemen. Slutligen identifieras ett antal aminosyror som möjliga kandidater centrala för aggregeringen.

# Förord

Under projektets gång har vi fått mycket stöd från våra handledare Inna Ermilova och Jan Swenson, och därmed vill vi tacka dem för deras engagemang. Utifrån de restriktioner som varit, i och med rådande pandemi, har de sett till att vara tillgängliga genom olika kommunikationsvägar. Stort tack till Inna Ermilova för goda diskussioner samt hjälp vid utförande av molekylodynamiksimuleringar och rapportskrivning. Tack till Jan Swenson för övergripande stöttning och korrekturläsning under projektets gång. Vi vill även lyfta det goda samarbetet med kandidatarbetesgruppen TIFX04-21-14 vid korrekturläsning.

Vi vill dessutom rikta ett tack till beräkningscentrum SNIC (Swedish National Infrastructure for Computing) för all CPU-tid som användes för beräkningarna i arbetet, samt klustrena; NSC (National Supercomputer Centre): SNIC2020-5-514, SNIC2020-13-38 (kluster Tetralith) och HPC2N (High Performance Computing Center North): SNIC2020-5-45, SNIC2021-5-80 (kluster Kebnekaise). Utöver dessa tackar vi också klustret HPC2N (High Performance Computing Center North): SNIC2020-6-53, SNIC2021-6-104, SNIC2020-10-22 för allt lagringsutrymme vi fick använda.

## *Handledare*

Inna Ermilova, Institutionen för Fysik  
Jan Swenson, Institutionen för Fysik

## *Examinator*

Lena Falk, Institutionen för Fysik

Kandidatarbete TIFX04-21-82  
Institutionen för Fysik  
Avdelningen för Nano- och biofysik  
Chalmers Tekniska Högskola  
412 96 Göteborg  
Telefon +46 31 772 1000

## Ordlista och förkortningar

**Affinitet** - Ämnens benägenhet att reagera med varandra

**Aggregering** - Process där partiklar anhopas

**CM** - Contact maps

**Cytotoxicitet** - Giftighet för celler

**Htt** - Huntingtinproteinet

**In silico** - Simuleringar i datorer

**In vitro** - Studie i konstgjord miljö

**In vivo** - Studie i molekylernas naturliga miljö

**Kraftfält** - Modell för potentiella energier

**MD** - Molekyldynamik

**Peptid** - Kedja innehållande upp till omkring 50 aminosyror

**RCF** - Rotationskorrelationsfunktion

**RDF** - Radiell fördelningsfunktion

**Relaxationstid** - Tiden det tar för ett stört system att återgå till jämvikt

# Innehåll

<b>1</b>	<b>Inledning</b>	<b>1</b>
<b>2</b>	<b>Teoretisk bakgrund</b>	<b>3</b>
2.1	Bakgrund till protein, aggregering och sekundärstruktur . . . . .	3
2.2	Huntingtinproteinet och dess roll i Huntingtons sjukdom . . . . .	4
2.3	Sackariders egenskaper för potentiell stabilisering av protein . . . . .	5
2.3.1	Sackaros och trehalos . . . . .	6
<b>3</b>	<b>Molekyldynamiksimuleringar</b>	<b>7</b>
3.1	Central fysik i klassiska molekyldynamiksimuleringar . . . . .	7
3.2	Kraftfält: Modeller för potentiell energi . . . . .	8
3.3	Energiminimering och jämviktning vid imitation av naturliga system	9
3.4	Metoder för att analysera simulerade system . . . . .	10
3.4.1	Radiell fördelningsfunktion . . . . .	10
3.4.2	Contact Maps . . . . .	11
3.4.3	Mellanliggande spridningsfunktion för partiklars självrörelse .	11
3.4.4	Rotationskorrelationsfunktion . . . . .	12
3.4.5	Sekundärstrukturberäkning . . . . .	12
<b>4</b>	<b>Metod</b>	<b>13</b>
4.1	Sammansättning av system . . . . .	13
4.2	Molekyldynamiksimuleringar som undersökningsmetod . . . . .	14
4.3	Utförande . . . . .	14
4.3.1	Skapande och simulering av system . . . . .	14
4.3.2	Analys av data . . . . .	15
<b>5</b>	<b>Resultat</b>	<b>17</b>
5.1	Antal vätebindningar samt molekylär RDF . . . . .	17
5.2	Atomär RDF . . . . .	19
5.3	Mellanliggande spridningsfunktion för självrörelse . . . . .	20
5.4	Rotationskorrelationsfunktion . . . . .	20
5.5	Contact maps . . . . .	22
5.6	Sekundärstruktur . . . . .	22
<b>6</b>	<b>Diskussion</b>	<b>25</b>
6.1	Diskussion om metod . . . . .	25
6.1.1	Felkällor för MD-simuleringar . . . . .	25
6.1.2	Felkällor för analysmetoder . . . . .	26
6.2	Diskussion av resultat . . . . .	27
6.2.1	Sammanfattning av resultatdiskussionen . . . . .	31
<b>7</b>	<b>Slutsats och framtidsutsikter</b>	<b>32</b>
	<b>Litteratur</b>	<b>33</b>

<b>A Bilagor</b>	<b>I</b>
A.1 Molekylär RDF mellan aminosyror och sackarider . . . . .	I
A.1.1 A-system . . . . .	I
A.1.2 B-system . . . . .	III
A.1.3 C-system . . . . .	V
A.1.4 D-system . . . . .	VII
A.1.5 E-system . . . . .	IX
A.2 Mellanliggande spridningsfunktion för partiklars självrörelse . . . . .	XIII
A.2.1 Peptider . . . . .	XIII
A.2.2 Sackarider . . . . .	XIV
A.2.3 Vatten . . . . .	XV
A.3 Contact Maps . . . . .	XVI
A.4 Sekundärstruktur . . . . .	XXXI
A.4.1 A-system . . . . .	XXXI
A.4.2 B-system . . . . .	XXXVII
A.4.3 C-system . . . . .	XLIII
A.4.4 D-system . . . . .	XLIX
A.4.5 E-system . . . . .	LV

# 1 Inledning

År 1872 publicerade den då 22 årige George Huntington en vetenskaplig rapport i den välkända medicinska veckotidningen *The Medical and Surgical Reporter*, där han beskrev en ärftlig neurodegenerativ sjukdom som för eftervärlden skulle leva kvar som Huntingtons [1]. Det skulle dröja mer än 100 år, fram till 1983, innan forskare från USA lyckades spåra orsaken till en genetisk variation på kromosompar fyra [2]. Tio år senare identifierades slutligen roten till sjukdomen; en mutation i en gen som senare gavs namnet Huntingtingenen. Idag är det känt att den muterade genen resulterar i onaturligt höga polyglutaminupprepningar hos Huntingtinproteinet (Htt) [3]. Det höga antalet polyglutaminenheter leder till att Htt bildar toxiska aggregat [4], och denna aggregering tros vara en orsak till sjukdomens uppkomst.

Symptomen, först beskrivna av George Huntington, innefattar motoriska, kognitiva och psykiska nedsättningar så som ofrivilliga rörelser, språkförsämring, demens, depression och apati [3]. Att utföra dagliga aktiviteter blir successivt mer besvärligt under sjukdomsförloppet, vars slutskede innebär en förtida död. Normalt sett inträffar detta mellan 17 och 20 år efter insjuknandet [3], som i genomsnitt sker i 35-årsåldern [4]. I världen drabbas idag ungefär 1 av 10 000 personer. [4].

Trots stora framsteg gällande förståelsen för Huntingtons sjukdom finns idag varken botemedel eller bromsmedicin, utan nuvarande läkemedel är endast symptomlindrande [3]. Genom att studera Htt, och undersöka hur proteinet interagerar med andra molekyler, kan större förståelse kring sjukdomen skapas. Tidigare studier indikerar att flera olika sockerarter, även kallade sackarider, däribland disackariderna sackaros och trehalos, är goda stabilisatorer av olika proteiner i varierande miljöer [5],[6],[7].

Utifrån de ovannämnda studierna syftar projektet till att undersöka hur sackaros och trehalos interagerar med två peptider - 2LD2 och 6N8C - från Htt, och huruvida de kan dämpa aggregering. Utöver detta syftas även att utröna hur de studerade interaktionerna påverkas av olika vattenkoncentration. Problemformuleringen analyseras utifrån flera perspektiv vilka innefattar följande frågeställningar: Påverkar sackaros och trehalos aggregeringen av Huntingtinpeptiderna 2LD2 och 6N8C? Förändrar sackaros och trehalos egenskaperna hos 2LD2 och 6N8C i vattenlösning och minskar de peptidernas rörelseförmåga? Vidare undersöks hur de molekylära interaktionerna inom, samt mellan, 2LD2 och 6N8C bildar sekundärstruktur, och huruvida sekundärstrukturen förändras vid tillägg av sackaros och trehalos. Slutligen, påverkar mängden vatten aggregeringen av 2LD2 och 6N8C, samt interaktionerna mellan peptiderna och sackariderna?

Initialt genomförs en omfattande litteraturstudie om Huntingtinproteinet, uppkomsten av Huntingtons sjukdom och hur sackarider kan påverka proteinaggregering. Utöver litteraturstudien utförs en modellbaserad studie med atomära molekylodynamiksimuleringar, MD-simuleringar. Med dessa undersöks strukturer och dynamik hos molekyler på atomär nivå och med hög tidsupplösning [8]. Detta görs med hjälp av programvaran GROMACS-2019 [9], tillsammans med lämpliga modeller för de potentiella energierna i de undersökta systemen, så kallade kraftfält.

Studien avgränsas till att studera två olika peptider, 2LD2 [10] och 6N8C [11], och två olika sockarider, sackaros och trehalos, i en miljö som efterliknar människokroppen. Valet av systemuppsättning optimeras för att erhålla kvalitativ information om hur sockerarter eventuellt påverkar aggregering. Dessutom begränsas de simulerade tidsintervallen till 300 ns. Dessa avgränsningar baseras på projektets begränsade tidsperiod.

## 2 Teoretisk bakgrund

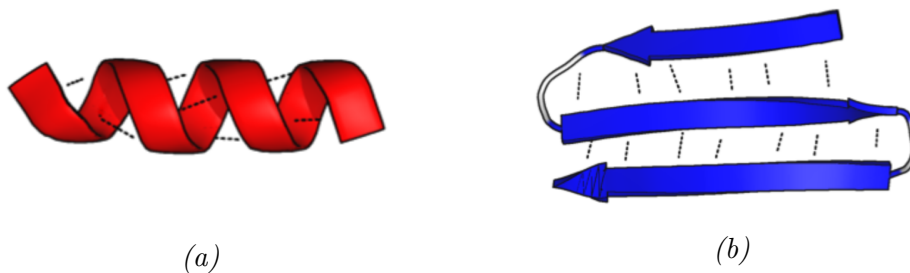
I följande kapitel behandlas bakgrundsinformation och teori nödvändig för förståelsen av innehållet i detta projekt. Inledningsvis ges en kortare genomgång av centrala molekylärbiologiska begrepp. Vidare beskrivs de strukturelaterade processer som föranleder proteinaggregering, samt dess konsekvenser. Därefter följer en förklaring av det specifika fallet Huntingtons sjukdom. Mer information ges även kring Huntingtinpeptiderna 2LD2 och 6N8C, vilka utgör de centrala undersökningsmotiven för senare molekylodynamiksimuleringar. Slutligen presenteras de dokumenterade proteininstabiliserande egenskaperna av disackariderna sackaros och trehalos.

### 2.1 Bakgrund till protein, aggregering och sekundärstruktur

För att förstå orsakerna bakom proteinaggregering, samt vilken roll den riskerar att spela i neurodegenerativa sjukdomar krävs en grundläggande beskrivning av vissa bakomliggande molekylära processer. Terminologiskt är protein den allmänna beteckningen för längre sammanlänkade kedjor av aminosyror, en klass av kemiska föreningar innehållande en gemensam ryggrad, samt en sidokedja vars struktur och egenskaper varierar kraftigt [12][13]. I biologiska system används protein som den primära funktionella enheten för en rad olika uppgifter [12]. När protein förekommer i kortare segment innehållande upp till omkring 50 aminosyror brukar de istället refereras till som peptider [13].

Korrekt funktion hos genererade proteiner sker som ett resultat av att den tidigare nämnda aminosyrakedjan sammanfogas till en tredimensionell struktur[12]. Detta sker då olika regioner av aminosyrakedjan interagerar med varandra, vilket resulterar i att den viker sig efter en termodynamiskt fördelaktig mall [14]. Ofta kan man dela in större protein i domäner, regioner av aminosyrakedjan som genomgår vikiningsprocessen oberoende av resterande kedja [12]. De strukturer som skapas genom denna process refereras till som proteinets sekundärstruktur, vilket utgör ett av fundamenten för analys av biologiska interaktioner mellan olika proteiner och deras omgivning.

De vanligaste konformationerna bland sekundärstrukturer skapas då primärt vätebindningar bildas mellan den gemensamma ryggraden hos aminosyrorna [12]. Bland dessa bör särskilt två centrala och synnerligen frekvent förekommande kategorier lyftas, så kallade  $\alpha$ -helix och  $\beta$ -flak. Samlingsnamnet  $\alpha$ -helix refererar till konformationer där vätebindningar i aminosyrakedjan framkallar en vridning vilket bildar en stabil cylinderformation likt (a) i figur 1. Konformationer där den sekundära strukturen utgörs av  $\beta$ -flak skapas då vätebindningar istället bildas på upprepade platser mellan två, antingen loopade parallella, eller antiparallella sträckor av samma aminosyrakedja vilket kan ses i (b) i figur 1.



Figur 1: Visualisering av sekundärstrukturerna a)  $\alpha$ -helix och b)  $\beta$ -flak. De prickade linjerna representerar vätebindningar. Från [15]. Omarbetad med tillstånd. CC BY-SA 4.0.

Den process som senare sammanfogar de varierande domänerna av sekundärstruktur domineras av potentialskillnader mellan hydrofoba och hydrofila interaktioner[14]. Bland aminosyrornas unika sidokedjor förekommer flera som är opolära och därmed hydrofoba[12]. Då dessa ogärna binder till de polära vattenmolekylerna i proteinets omgivning, resulterar detta i en ökad affinitet (ämnens benägenhet att reagera med varandra) till andra hydrofoba sekvenser i samma aminosyrakedja. Detta är ytterligare en av de bidragande orsakerna till proteinets tredimensionella vikning och efterföljande funktion.

Under vissa förhållanden kan dessa processer misslyckas, vilket kan resultera i ett oveckat, delvis veckat eller felaktigt veckat protein[14]. Om dessa felveckade konformationer förekommer i ett större antal kan de exponerade hydrofoba insidorna av protein binda till varandra för att reducera sin fria energi. Interaktioner av detta slag riskerar att starta en kedjereaktion, där ytterligare felveckade protein rekryteras och riskerar att forma större aggregat.

Förekomsten av onormalt veckade proteiner, samt proteinaggregat av detta slag, riskerar att orsaka cytotoxicitet (giftighet för celler) hos celler [14]. Om dessa är en del av det centrala nervsystemet kan sådan toxicitet i förlängningen medföra allvarliga neurodegenerativa konsekvenser, vilket antas vara fallet hos patienter som lider av Huntingtons, Alzheimers och Parkinsons sjukdom. Trots att dessa samband är välkända och att flera omfattande teorier lagts fram, har ännu inte den exakta anledningen till hur felveckade proteiner och dess aggregat orsakar cytotoxicitet kunnat fastställas.

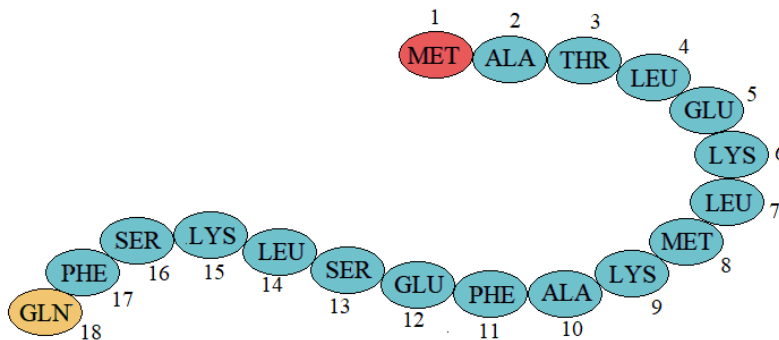
## 2.2 Huntingtinproteinet och dess roll i Huntingtons sjukdom

Den bakomliggande orsaken till Huntingtons sjukdom är en muterad variant av Huntingtinproteinet (Htt) [4]. Mutationen framträder i form av förlängda polyglutaminkanaler; sekvenser i proteinet bestående av flera enheter av aminosyran glutamin, och leder till att Htt anhopas och bildar toxiska aggregat. De molekylära mekanismerna vilka ligger bakom aggregeringen är ännu okända [16], men det är konstaterat att dessa aggregat stabiliseras med hjälp av vätebindningar [17]. Därför är just dessa sorters bindningar relevanta att studera.

Den ökade toxiciteten medför nedbrytning av nervceller främst i striatum [4], en del av storhjärnan. Utöver detta antyder experimentella studier att mutationen även genererar störningar av proteinets normala biologiska funktion. Denna normala funktion tycks innefatta förmedling av en mängd olika fysiologiska processer, med indikationer att proteinets polyglutaminkanal har en betydande neural funktion hos ryggradsdjur, däribland människor.

Htt är ett stort protein innehållande 3144 aminosyror [18], och för att minska simuleringarnas beräkningskostnad undersöks därför mindre peptider av proteinet. De peptider som behandlas i detta arbete, 2LD2 och 6N8C, består av 17 aminosyror vardera. 2LD2 innefattar de första 17 aminosyrorerna av Huntingtinproteinet (definierade enligt figur 2) [10], medan 6N8C utgörs av aminosyror 2-18 [11]. Aminosyra 18 (GLN) är den första glutaminkomponenten i polyglutaminkanalerna.

Utöver att de båda peptiderna ingår i det större Htt existerar de sannolikt även enskilt hos människor. Detta grundas i att mindre fragment av Htt återfinns i påverkade områden av hjärnan [17]. Det är också känt att peptidernas närvaro bidrar till avsevärt ökad hastighet hos polyglutaminaggregering [10],[11],[18]. De 17 första aminosyrorerna i Htt har identifierats som en kritisk region vilken spelar stor roll i bland annat aggregeringen av proteinet och dess toxicitet [18]. Utifrån det ovan nämnda är de två peptiderna av intresse att studera.



*Figur 2:* Aminosyrasekvenserna vilka utgör peptiderna 2LD2 (aminosyra 1-17) och 6N8C (aminosyra 2-18). De blåmarkerade aminosyrorerna är gemensamma för peptiderna. Aminosyra 18 (GLN) är första glutaminkomponenten i Htt:s polyglutaminkanal.

## 2.3 Sackariders egenskaper för potentiell stabilisering av protein

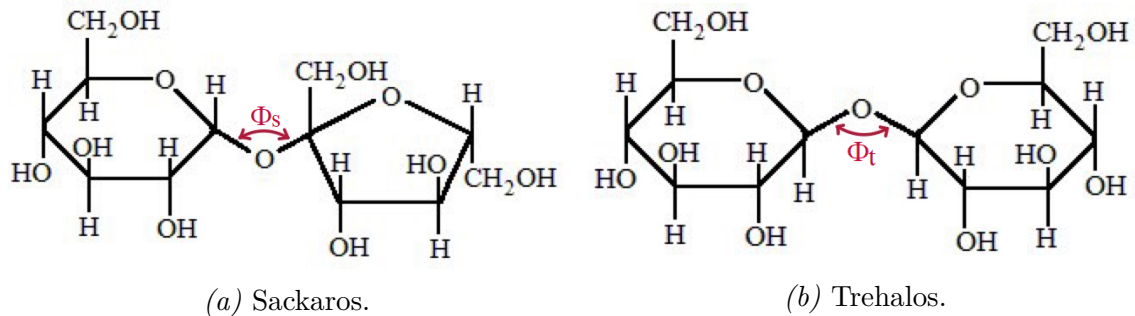
Ämnen som kan förhindra aggregering av molekyler benämns ofta som kemiska stabilisatorer. En vanligt använd sådan är sockerarter, även kallade sackarider, vilka visar egenskaper för att förhindra proteinaggregering [19]. Sockerarter kan förekomma som större och mindre molekyler, och beroende på storleken har de olika egenskaper vid interaktion med andra biomolekyler. Större sackarider kan interagera med omgivningen och minska global rörlighet i systemet mer effektivt än de mindre disackariderna, medan disackariders mindre storlek gör det möjligt att reducera lokal

rörlighet [6]. Dessa egenskaper är några av de faktorer som gör sackarider till goda stabilisatorer av proteiner i olika miljöer.

*In vitro*-studier (studier i konstgjord miljö) tyder på att olika sackarider, däribland sackaros och trehalos, kan hämma proteinaggregering som uppstår vid ökad mängd polyglutamin, så kallad polyglutaminmedierad aggregering [7],[20]. Studiernas resultat indikerar även att disackarider reducerar polyglutaminaggregat i en cellulär modell av Huntingtons sjukdom. Dessa resultat motiverar vidare undersökningar kring huruvida disackariderna sackaros och trehalos interagerar med Htt, och om de kan dämpa aggregeringen av proteinet.

### 2.3.1 Sackaros och trehalos

Disackariderna sackaros och trehalos är erkänt goda stabilisatorer av proteiner i olika miljöer [5],[20]. I figur 3 illustreras sockerarternas kemiska strukturer, där det bland annat kan noteras att strukturerna skiljer sig genom olika placering av OH-grupperna. Trehalos konfiguration gör att den föredrar att interagera med vattenmolekyler i omgivningen, genom vätebindningar, framför andra trehalosmolekyler [21]. Denna egenskap i kombination med andra kemiska egenskaper medför att trehalos kan förhindra att protein vecklar ut sig och därmed riskerar att aggregera, samt bidrar till en homogen stabiliserande miljö. Studie visar dessutom att sackariderna skulle kunna agera som steriskt hinder, vilket innebär att deras storlek och struktur vilket kan hindra peptider från att interagera med varandra [19]. Sackaros uppvisar en strukturstabiliserande effekt på system för protein med olika hög grad av vattenlöslighet, enligt en interaktionsstudie gjord mellan protein och sackaros [5].



Figur 3: Kemiska strukturer för disackariderna sackaros och trehalos. Hörnen föreställer kolatomer.

De ovannämnda resultaten för trehalos styrks ytterligare av en studie gjord *in vivo*, i de studerade molekylernas naturliga miljö, på möss som intagit sockerlösning oralt [20]. Denna visar att trehalos kan stabilisera proteiner samt hämma aggregeringen av polyglutamin.

### 3 Molekyldynamiksimuleringar

För att undersöka om sackaros och trehalos interagerar med de två Huntingtin-peptiderna tillämpades klassiska molekyldynamiksimuleringar (MD-simuleringar). Med dessa kan molekylernas beteende undersökas på atomär nivå och med mycket hög tidsupplösning [8]. I följande avsnitt behandlas de bakomliggande koncepten för MD-simuleringar, samt hur resultatet av dessa kan studeras och analyseras.

#### 3.1 Central fysik i klassiska molekyldynamiksimuleringar

Klassiska MD-simuleringar syftar till att beräkna trajektorier, det vill säga tidsvarierande positioner, hastigheter och accelerationer, för partiklar i ett system [22]. I atomära simuleringar representeras dessa partiklar av atomer och joner. Trajektorierna erhålls genom att tillämpa Newtons rörelselagar, och således kan partiklarnas rörelse med tiden tas fram. För ett system med  $N$  partiklar kan ekvationerna definieras enligt

$$m_i \ddot{\vec{r}}_i = \vec{F}_i, \quad (1)$$

där  $i = 1, \dots, N$ . Här betecknar  $m_i$  massan för partikel  $i$ ,  $\ddot{\vec{r}}_i$  är andra ordningens tidsderivata av den tredimensionella spatiella koordinaten  $\vec{r}_i$  för partikel  $i$  och  $\vec{F}_i$  är kraften som verkar på partikel  $i$ , vilken genereras av de resterande partiklarna i systemet. Då ekvationerna är differentialekvationer löses dessa numeriskt med hjälp av en integrator, men för att åstadkomma detta behöver krafterna liksom partiklarnas begynnelsekoordinater vara kända.

Om systemets totala potentiella energi,  $V(\vec{r}_1, \dots, \vec{r}_N)$ , är känd kan  $\vec{F}_i$  erhållas som

$$\vec{F}_i(\vec{r}_1, \dots, \vec{r}_N) = -\vec{\nabla}_i V(\vec{r}_1, \dots, \vec{r}_N) = -\frac{dV(\vec{r}_1, \dots, \vec{r}_N)}{d\vec{r}_i}, \quad (2)$$

där  $\nabla_i V(\vec{r}_1, \dots, \vec{r}_N)$  är gradienten av den potentiella energin med avseende på partikel  $i$  [23]. Den potentiella energin beror alltså av samtliga partiklars positionskoordinater. Genom att applicera modeller för hur energin i ett system beror av partiklarnas koordinater kan ekvation (2), och därmed (1), lösas. Sådana modeller benämns som kraftfält, och består av en kombination av analytiska uttryck och associerade parametrar för den potentiella energin  $V(\vec{r}_1, \dots, \vec{r}_N)$  [24].

För att simuleringarna ska uppnå numerisk stabilitet och så utförliga trajektorier som möjligt krävs en hög tidsupplösning, det vill säga att beräkningarna görs med mycket små tidssteg [8]. Detta eftersom de kraftfält som verkar på de enskilda partiklarna endast är giltiga för de specifika tidpunkter de tagits fram för. Typiskt är dessa tidssteg i storleksordningen omkring 1 fs ( $10^{-15}$  s). Detta tidssteg är lämpligt eftersom det är ungefär 10 gånger kortare än de bindingsvibrationer med kortast tidsperiod, vilket gäller för vätebindningar som har en tidsperiod på omkring 13 fs [25].

### 3.2 Kraftfält: Modeller för potentiell energi

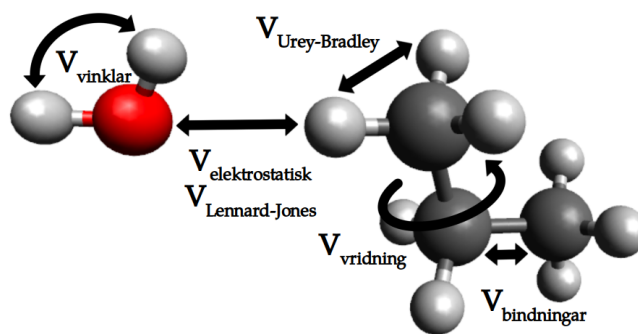
Simuleringarna baseras på lämpliga modeller, så kallade kraftfält, vilka beskriver den interatomära potentiella energin i ekvation 2. Då systemen som behandlas i detta arbete innehåller olika typer av partiklar, bland annat peptider och sackarider, behövs olika kraftfält för att representera partiklarnas bindande respektive icke-bindande interaktioner [22]. De bindande interaktionerna inkluderar rotation, tånjning och vridning av bindningar och de ickebindande tar hänsyn till elektrostatisk potential. Olika kraftfält behövs eftersom molekylerna är uppbyggda av atomer vilka innehar individuella kemiska egenskaper. I detta projekt används kraftfälten CHARMM36, [26], och General AMBER Force Field (GAFF), [27], för att representera peptidernas respektive sackaridernas energier. CHARMM36 är ett kraftfält som passar för peptider och atomerna dessa innehåller [26]. Dess bidrag från de bindande interaktionerna uttrycks som

$$\begin{aligned}
 V_{bindningar} &= \sum_{bindningar} k_r (r_{ij} - r_0)^2, \\
 V_{Urey-Bradley} &= \sum_{Urey-Bradley} k_b (b - b_0)^2, \\
 V_{vridning} &= \sum_{vridning} k_\phi (1 + \cos(n\phi - \delta)), \\
 V_{vinklar} &= \sum_{vinklar} k_\theta (\theta - \theta_0)^2,
 \end{aligned} \tag{3}$$

där  $k_r$ ,  $k_b$ ,  $k_\phi$  och  $k_\theta$  representerar kraftkonstanter;  $r_{ij}$  är bindningslängd mellan atom  $i$  och  $j$ ,  $b$  är Urey-Bradley 1,3-längd (avståndet mellan atom 1 och 3),  $\phi$  är vridningsvinkeln och  $\theta$  är bindningsvinkeln. Där de variabler med nedsänkt 0 betecknar jämviktsvärden. Bidragen från de ickebindande interaktionerna beskrivs istället som

$$\begin{aligned}
 V_{Lennard-Jones} &= \sum_{i,j \neq i} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right], \\
 V_{elektrostatisk} &= \sum_{i,j \neq i} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}},
 \end{aligned} \tag{4}$$

där  $\epsilon_{ij}$  representerar Lennard-Jones låddjup,  $\sigma_{ij}$  är Lennard-Jones minimumavstånd (avståndet då potentialen är noll),  $r_{ij}$  är avståndet mellan atomerna  $i$  och  $j$ ,  $\epsilon_0$  är den dielektriska konstanten och  $q$  är potentiell atomladdning för atomerna  $i$  och  $j$ . I figur 4 visualiseras de olika energierna, och vilka interaktioner de beskriver, inom och mellan molekylerna.



Figur 4: Visuell representation av de olika kraftfältsbidragen.

Den totala potentiella energin för CHARMM36 erhålls genom att summera de bindande och ickebindande bidragen enligt

$$V_{\text{CHARMM36}} = V_{\text{bindningar}} + V_{\text{vinklar}} + V_{\text{vridning}} + V_{\text{Urey-Bradley}} + V_{\text{Lennard-Jones}} + V_{\text{elektrostatisk}} \quad (5)$$

För att istället representera sackaridernas interatomära energier kan kraftfältet GAFF appliceras [27]. Denna bygger på liknande uttryck som för CHARMM36 för de bindande och ickebindande interaktionerna, med undantaget att Urey-Bradley-termen inte inkluderas. På samma sätt som för CHARMM36 summeras de olika bidragen till den totala potentialen för GAFF.

### 3.3 Energiminimering och jämviktning vid imitation av naturliga system

De simulerade partiklarna i systemen placeras initialt i slumpmässiga positioner, vilket kan leda till en ofördelaktigt hög energi då partiklarna kan överlappa eller befinna sig onaturligt nära varandra. Detta motiveras också av ekvationen 4, där energin närmar sig oändligheten för mycket små avstånd. För att undkomma detta krävs att förhållandena i systemet ändras, det vill säga att molekylerna placeras med längre avstånd till varandra. Genom så kallad energiminimering kan detta åstadkommas genom att justera molekylernas koordinater i systemet, vilket gör systemet mer energimässigt fördelaktigt. För att göra detta kan steepest descentmetoden tillämpas [28].

Med jämviktning kan ytterligare restriktioner på partiklarnas rörelse sättas. Dessa krävs för att erhålla trajektorier som efterliknar verkliga biologiska system, då dessa har egenskaper som Newtons rörelselagar inte tar hänsyn till. Restriktionerna löses med ensembler, vilka beskriver systemets egenskaper med hjälp av olika algoritmer, så kallade barostater och termostater. För biologiska system är ofta densiteten okänd, och det är då lämpligt att använda den så kallade NPT-ensemblen [22]. Denna kan beskrivas av Berendsen-barostaten [29] och den så kallade "velocity rescale"-termostaten [30], där antal partiklar, temperatur och tryck hålls konstant, medan volymen tillåts variera.

### 3.4 Metoder för att analysera simulerade system

För att analysera molekylernas rörelser i simuleringarna tillämpas flera olika analysmetoder. Med dessa kan olika egenskaper hos systemen studeras, så som avstånd mellan molekyler och atomer, molekylers rörelseförmåga och peptidernas sekundärstruktur. Genom att undersöka hur egenskaperna eventuellt skiljer sig mellan system med och utan sackarider, kan slutsatser dras om sackaridernas påverkan. Nedan följer förklaringar av teorin bakom några av de analysmetoder som används i detta arbete.

#### 3.4.1 Radiell fördelningsfunktion

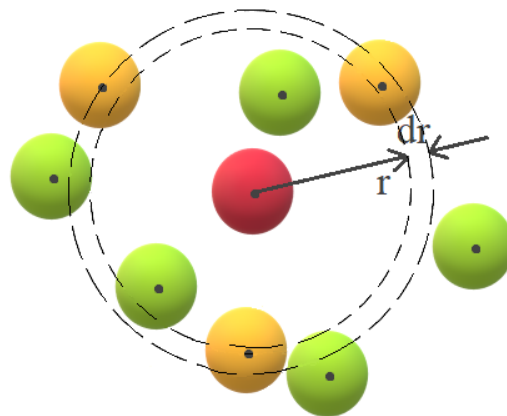
Den radiella fördelningsfunktionen (RDF),  $g(r)$ , beskriver densiteten av atomer eller molekyler i systemet som funktion av avståndet från en referenspunkt. Funktionen definieras enligt [31]

$$g(r) = \lim_{dr \rightarrow 0} \frac{p(r)}{4\pi r^2 (N_{par}/V) dr}, \quad (6)$$

där  $r$  är avståndet mellan två partiklars masscentrum,  $p(r)$  det genomsnittliga antalet atom- eller molekympar som befinner sig inom intervallet  $r$  och  $r + dr$ ,  $V$  systemets totala volym, och  $N_{par}$  antal unika atom- eller molekympar i systemet. Beräkning av  $g(r)$  sker med en uppsättning partiklar som referens, och en annan uppsättning partiklar till vilka avstånd beräknas. Dessa två partikeluppsättningar bildar atom- eller molekymparen. I figur 5 visualiseras beräkningen av  $g(r)$ , där de streckade linjerna innesluter det approximativa avståndet  $r$  från en röd referenspartikel. De gula partiklarna har masscentrum på approximativa avståndet  $r$ , medan de gröna partiklarnas masscentrum ligger utanför.

I detta projekt beräknas  $g(r)$  inledningsvis mellan sackaridmolekylerna och aminosyrorna, för att undersöka vilka aminosyror som sackariderna binder till. Därtill beräknas  $g(r)$  mellan individuella atomer i peptiderna och sackariderna. Detta för att erhålla en mer detaljerad bild över hur sackariderna binder till peptiderna via vätebindningar.

RDF beräknas och plottas för hela tidsintervall, och representerar därmed en sannolikhetsfördelning för olika interatomära eller intermolekylära avstånd som förekommer under simuleringen. Tydliga toppar i plotten där  $g(r) \geq 1$  indikerar de avstånd som majoriteten av partiklarna be-



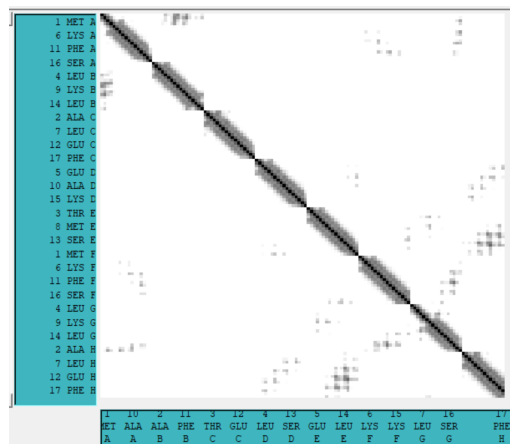
Figur 5: Visualisering av RDF-beräkning. De gula partiklarna har masscentrum på det approximativa avståndet  $r$  och de gröna utanför. De svarta prickarna markerar masscentrum.

finner sig på, och toppar med dessa värden är således de som är relevanta att studera.

### 3.4.2 Contact Maps

Contact maps (CM) kan användas för att beräkna och visualisera proteiners tredimensionella struktur [32], vilken skapas av polypeptidkedjans vikningar. CM baseras på en symmetrisk  $(n \times n)$ -matris,  $\mathbf{M}$ , där  $n$  betecknar antalet atomer i polypeptidkedjan. För att underlätta beräkningarna ytterligare förenklas varje aminosyra och antas vara unika enheter. Elementen,  $M_{i,j}$ , i avståndsmatrisen blir då avståndet mellan atomparen  $i$  och  $j$ . Vidare baseras CM på ett avbrottskriterium (det maximala tillåtna avståndet mellan atomparen), även kallat tröskel, för att exkludera irrelevanta avstånd och förenkla representationen av proteinet.

Med hjälp av till exempel COMAR-algoritmen erhålls den slutliga CM [32], som sedan kan visualiseras enligt exemplet i figur 6, där aminosyragrupperna listas på x- och y-axeln. En punkt med koordinat  $(a, b)$  implicerar att x-axelns aminosyraatom  $a$  är inom tröskelns avstånd till y-axelns aminosyraatom  $b$ . Punkternas olika färgintensitet beror på avståndet mellan aminosyrorna, där mörkare färg implicerar kortare avstånd och vice versa.



Figur 6: Exempel på Contact Map (här för peptiden 2LD2 med disackariden trehalos, tidsintervall 250-300 ns).

### 3.4.3 Mellanliggande spridningsfunktion för partiklars självrörelse

För att undersöka hur peptidernas rörelse i de olika simulerade systemen skiljer sig från varandra, samt om tillsättning av sackariderna förändrar deras rörelseförmåga, kan den mellanliggande spridningsfunktionen beräknas. Denna ger information om partiklars rörelseförmåga i ett system, och erhålls genom att beräkna den rumsliga Fouriertransformen av van Hoves fördelningsfunktion. Van Hoves funktion beskriver sannolikheten att en partikel befinner sig i en viss position vid en specifik tidpunkt, givet att en partikels position vid  $t = 0$  definierar origo [33].

Den mellanliggande spridningsfunktionen kan delas upp i två delar; en som beskriver en partikels självrörelse (SISF) och en som beskriver partiklars korrelerade rörelse [34]. Den del som i detta sammanhang är av intresse är den förstnämnda, SISF, vilken kan uttryckas utifrån partiklarnas trajektorier [35] som

$$F_s(\vec{k}, t) = \frac{1}{N} \left\langle \sum_{j=1}^N \exp\left(-i\vec{k} \cdot [\vec{r}_j(t) - \vec{r}_j(0)]\right) \right\rangle, \quad (7)$$

och därmed kan erhållas direkt från MD-simuleringar. I uttrycket representerar  $N$  antalet partiklar,  $\vec{r}_j(t)$  är positionskoordinaten för partikel  $j$  vid tiden  $t$ ,  $k = |\vec{k}| = 2\pi/r_j(t)$  är spridningsvektorn, där  $r_j(t) = |\vec{r}_j(t)|$ , och  $\langle \cdot \rangle$  beskriver genomsnittet.

$F_s(\vec{k}, t)$  beräknas för systemens olika molekyler, och plottas som funktion av logaritmerad tid. Detta görs för att undersöka hur deras rörelseförmåga skiljer sig i de olika systemen, där en kurva med längre relaxationstid implicerar långsammare rörelse och vice versa, och kan tolkas som stabilisering av partiklar. Med relaxationstid menas tiden det tar för ett stört system att återgå till jämvikt, vilket i plottarna representeras av hur snabbt en kurva går mot noll.

#### 3.4.4 Rotationskorrelationsfunktion

Rotationskorrelationsfunktionen (RCF) kan användas för att undersöka huruvida sackariderna påverkas av dess omgivning. RCF beskriver hur snabbt en molekyl roterar runt en satt rotationsvinkel,  $\theta$ , vilket är vinkeln mellan tre bestämda atomer [36]. RCF beror av tiden,  $t$ , och beskrivs enligt

$$C_2(t) = \left\langle P_2 \cos[\theta(t)] P_2 \cos[\theta(0)] \right\rangle, \quad (8)$$

där  $P_2$  är den 2:te ordningen av Legendre-polynom. Denna definieras enligt  $P_2(t) = \frac{1}{2}(3x^2 - 1)$  [37]. Funktionen plottas som funktion av logaritmerad tid, där en kurva med en kortare rotationsstid implicerar en snabbare rotation, och vice versa.

#### 3.4.5 Sekundärstrukturberäkning

För att analysera huruvida molekylära interaktioner förändrar de initiala tillståndet hos peptiderna kan sekundärstrukturen beräknas för varje peptid. De olika sekundärstrukturerna erhålls som matriser, vilka kan åskådliggöras genom ett färgkodningssystem [38], med tiden på y-axeln och aminosyranumren på x-axeln. Förändringar i färgen implicerar således ett skifte i vilken sekundärstruktur som dominerar bland den simulerade aminosyrakedjan, där  $C = \text{slumpmässig spole (random coil)}$ ,  $I = \pi\text{-helix}$ ,  $G = 3\text{-}10\text{-helix}$ ,  $H = \alpha\text{-helix}$ ,  $B = \text{isolerat } \beta\text{-ark (isolated } \beta \text{ bridge)}$ ,  $E = \text{utökad konformation (extended conformation)}$  och  $T = \beta\text{-sväng (} \beta \text{ turn)}$ .

## 4 Metod

Inledningsvis utfördes en omfattande litteraturstudie om Huntingtinproteinet och uppkomsten av Huntingtons sjukdom, samt hur sackaros och trehalos skulle kunna påverka proteinaggregering. Därefter genomfördes en modellbaserad studie *in silico*, i datorer, med atomära MD-simuleringar likt beskrivet i kapitel 3. I detta kapitel presenteras arbetssättet för den sistnämnda studien: sammansättning av system, undersökningsmetoden MD-simulering, skapande och simulering av system, samt analys av resultat.

### 4.1 Sammansättning av system

För att undersöka interaktionen mellan sackariderna och peptiderna studerades olika system med varierande uppsättningar av partiklar. Studien fokuserade på totalt 15 system. Av dessa bestod 10 stycken av 8 peptider (2LD2 och 6N8C) och 16 sackarider (sackaros och trehalos). Dessutom skapades 5 referenssystem, system utan sackarider, för att urskilja sackaridernas inverkan på aggregering och dynamik. Till alla system adderades vatten samt natrium- och klorid-joner ( $\text{Na}^+$  och  $\text{Cl}^-$ ). Två olika vattenmängder, 4800 respektive 2400 molekyler, tillsattes för att studera dess påverkan på systemen. Jonerna fungerade dels som motjoner (för att neutralisera laddningen), dels för att efterlikna kroppens egna pH-värde på 7.4, där deras antal valdes utifrån mängden peptider respektive vattenmolekyler och deras totala laddning. Kompositionen av alla system listas i tabell 1.

Tabell 1: Samtliga systemuppsättningar skapade för simuleringar.

System	2LD2	6N8C	Sackaros	Trehalos	$\text{Na}^+$	$\text{Cl}^-$	$\text{H}_2\text{O}$
A1	8	-	16	-	9	17	4800
A2	8	-	-	16	9	17	4800
A3	8	-	-	-	9	17	4800
B1	-	8	16	-	9	17	4800
B2	-	8	-	16	9	17	4800
B3	-	8	-	-	9	17	4800
C1	8	-	16	-	5	13	2400
C2	8	-	-	16	5	13	2400
C3	8	-	-	-	5	13	2400
D1	-	8	16	-	5	13	2400
D2	-	8	-	16	5	13	2400
D3	-	8	-	-	5	13	2400
E1	4	4	16	-	9	17	4800
E2	4	4	-	16	9	17	4800
E3	4	4	-	-	9	17	4800

## 4.2 Molekyldynamiksimuleringar som undersökningsmetod

Interaktionerna mellan molekylerna i de olika systemen undersöktes genom atomära MD-simuleringar, vilka implementerar de numeriska lösningarna av Newtons ekvation enligt avsnitt 3.1. MD-simuleringar på atomär nivå undersöker ett systems fysikaliska egenskaper i högre detalj i jämförelse med simuleringar som behandlar flera atomer som en enda partikel [22], och interaktioner på atomär nivå kan således studeras närmare.

Den höga noggrannheten medför däremot att simuleringarna kräver omfattande beräkningskraft, på grund av de åtskilliga interaktioner som därmed evalueras varje tidssteg i kombination med den stora mängden totala tidssteg [8]. Då arbetets tidsperiod dessutom är begränsad gjordes alla beräkningar på datorkluster, för att på så sätt minska simuleringarnas tidsåtgång. De kluster som användes var Kebnekaise och Tetralith på respektive centrum HPC2N och NSC.

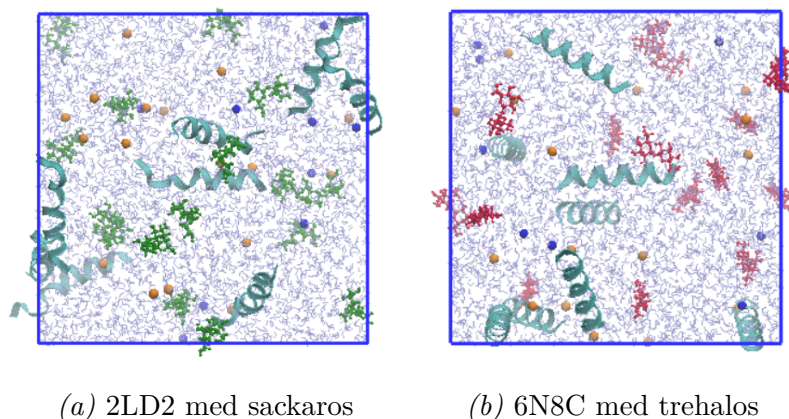
Simuleringarna utfördes med programvaran GROMACS-2019 [9], där systemen definieras utifrån deras storlek och form, antal och typ av molekyler samt positionskoordinater och hastigheter hos samtliga av systemets atomer och joner. För att representera de inter- och intramolekylära interaktionerna applicerades de tidigare nämnda kraftfälten, CHARMM36 och GAFF, vilka är kompatibla med GROMACS-2019. Utifrån de totala krafterna som verkar på samtliga atomer i systemet kan ekvation 1 lösas och trajektorier för atomerna tas fram. För att lösa Newtons ekvation numeriskt använder GROMACS-2019 leapfrog-algoritmen [39] som integrator, och för energiminimering av systemen tillämpas steepest descent metoden [28].

## 4.3 Utförande

Projektets utförande består av två delar; skapande och simulering av system samt analys av data från simulering, vilka redogörs för nedan.

### 4.3.1 Skapande och simulering av system

De uppbyggda systemen, listade i tabell 1, skapades som tidigare nämnt med hjälp av GROMACS-2019. En tredimensionell box användes för att innesluta varje system och på så sätt begränsa dess storlek. Storleken på boxarna varierades beroende på antal atomer och molekyler i samtliga system. Systemen A(1-3), B(1-3), E(1-3), gjordes i boxar med storlek 7 nm, medan C(1-3) och D(1-3), vilka innehöll mindre vatten, gjordes i boxar med storlek 5 nm. Exempel på boxar, med innehållande molekyler och joner, illustreras i figur 7, där de turkosa banden representerar peptiderna, de ljusblå strecken vattenmolekyler, de gröna respektive röda molekylerna är sackaros respektive trehalos, de blå sfärerna är  $\text{Na}^+$  och de orange sfärerna  $\text{Cl}^-$ .



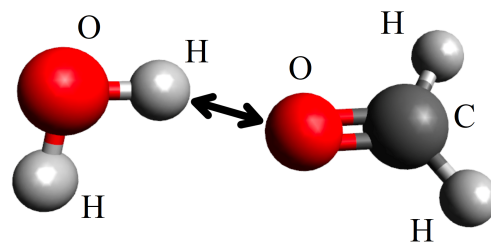
*Figur 7:* Boxar innehållande molekyler och joner i system med (a) 2LD2 och sackaros, samt (b) 6N8C med trehalos. De turkosa banden representerar peptiderna, de ljusblå strecken vattenmolekyler, de gröna respektive röda molekylerna är sackaros respektive trehalos, de blå sfärerna är  $\text{Na}^+$  och de orange sfärerna  $\text{Cl}^-$ .

Samtliga system energiminimerades därefter för att optimera avstånden mellan atomerna och molekylerna, enligt avsnitt 3.3. Därefter jämviktades systemen med NPT-ensemblen vid den konstanta temperaturen 310 K och trycket 1 atm, vilket motsvarar människans kroppstemperatur samt atmosfäriskt tryck. Systemen simulerades under ett tidsintervall på totalt 300 ns, med tidsstegen 2 fs.

### 4.3.2 Analys av data

De simulerade systemen analyserades med hjälp av GROMACS-2019 och de tidigare nämnda metoderna; RDF, CM, SISF, RCF och sekundärstruktur, tillsammans med antal vätebindningar mellan olika molekyler. Utöver GROMACS-2019 användes programvaran Visual Molecular Dynamics (VMD) för visualisering och ytterligare beräkningar.

RDF beräknades på två olika nivåer, atomär och molekylär. Atomära RDF beräknades mellan specifika syreatomer bundna till peptidens ryggrad, och väteatomer bundna till syre på sackariden, illustrerat i figur 8. Avstånden mellan atomerna som fokuserades på var upp till 0.35 nm, vilket baseras på övre gränsen för vätebindningars längd [40]. För dessa lades fokus för analys därför på toppar som påträffades inom dessa avstånd. Molekylär RDF beräknades mellan masscentrum för varje aminosyra och masscentrum för sackariderna. För dessa analyserades istället



*Figur 8:* Atomer mellan vilka atomär RDF beräknas med sackarid till vänster och peptid till höger.

toppar förekommande på avstånd mindre än 0.7 nm, för att i viss utsträckning ta hänsyn till molekylernas utbredning från masscentrum. Beräkningarna gjordes för tidsintervallet 50-300 ns, och  $g(r)$  plottades som funktion av avståndet mellan partiklarna.

Vid visualisering av den tredimensionella strukturen beräknades och plottades CM av peptiderna under tidsintervallet 50-300 ns på VMD och antal aminosyrakontakter beräknades för 50-55 ns, 175-180 ns samt 295-300 ns på GROMACS-2019. Som tröskelvärde valdes avståndet 0.7 nm, i enighet med den molekylära RDF-analysen. Graferna jämfördes sedan med varandra för att studera eventuella skillnader i kontakter hos de olika systemen.

SISF beräknades för avståndet  $r_j(t) = 0.5$  nm från origo (positionen vid  $t = 0$ ), med ett k-värde på  $2\pi/0.5 \approx 12.6$ . Beräkningarna gjordes för fyra olika kompositioner; hela systemet under tidsperioden 240-300 ns, sockerarter under 240-300 ns, vatten under 250-300 ns samt peptider under 220-300 ns. Därefter undersöktes sackaridernas rörelse med hjälp av RCF på tidsintervallet 200-300 ns. Denna räknades ut för vinklarna  $\Phi_s$  och  $\Phi_t$ , vilka definierades i figur 3 i avsnitt 2.3.1.

Respektive peptids sekundärstrukturer beräknades över tid, och genom att använda samma trajektorier som för CM-beräkningar genererades grafer över det totala tidsintervallet 50-300 ns. För de erhållna sekundärstrukturerna beräknades sedan de procentuella andelarna som de specifika sekundärstrukturerna, vilka nämns i avsnitt 3.4.5, upptar för varje system.

Slutligen beräknades antal vätebindningar mellan sockerart och peptid, vatten och peptid samt peptid och peptid under tiden 50-300 ns. Medelvärdet av antalet vätebindningar under det analyserade tidsintervallet, tillsammans med en feluppskattning, sammanställdes i en tabell.

## 5 Resultat

I detta kapitel presenteras en sammanfattning av de erhållna resultaten, i figur och tabellform, från de tillämpade analysmetoderna. En total sammanställning av resultaten återfinns i bilagor A.

### 5.1 Antal vätebindningar samt molekyllär RDF

Det erhållna medelantalet för alla vätebindningar i samtliga system listas i tabell 2, tillsammans med de beräknade feluppskattningarna. I kolumner 2-4 redovisas antal vätebindningar mellan peptider (peptid-peptid), peptider och sackarider (peptid-sackarid) respektive peptider och vattenmolekyler (peptid-vatten), avrundade till heltal.

*Tabell 2:* Uppmätta medelvärden, samt beräknade feluppskattningar, av antal vätebindningar mellan peptid-peptid, peptid-sackarid samt peptid-vatten för samtliga system.

System	# peptid-peptid	# peptid-sackarid	# peptid-vatten
A1	105 ± 0.0145	31 ± 0.0176	292 ± 0.0336
A2	97 ± 0.0124	50 ± 0.1339	289 ± 0.0297
A3	90 ± 0.0202	-	350 ± 0.0423
B1	109 ± 0.0178	26 ± 0.0142	310 ± 0.0400
B2	103 ± 0.0141	48 ± 0.0211	302 ± 0.0355
B3	122 ± 0.0162	-	313 ± 0.0341
C1	104 ± 0.0133	46 ± 0.0465	274 ± 0.0601
C2	109 ± 0.0138	45 ± 0.0161	268 ± 0.0332
C3	103 ± 0.0134	-	330 ± 0.0318
D1	111 ± 0.0188	33 ± 0.0188	297 ± 0.0348
D2	109 ± 0.0156	51 ± 0.0162	289 ± 0.0409
D3	120 ± 0.0190	-	325 ± 0.0359
E1	108 ± 0.0148	36 ± 0.0176	296 ± 0.0348
E2	100 ± 0.0164	47 ± 0.0210	295 ± 0.0369
E3	109 ± 0.0168	-	328 ± 0.0168

Inledningsvis observeras att antal vätebindningar mellan peptiderna och trehalos (A2-E2) är betydligt fler än de mellan peptiderna och sackaros (A1-E1), med undantag för system C. För bindningar mellan peptiderna och vattenmolekyler uppvisas mycket små skillnader sackariderna emellan, men de system vilka saknar sackariderna kan utläsas innehålla betydligt fler vätebindningar än systemen med sackarider. Vad gäller vattenmängdens påverkan kan inget mönster avläsas angående vätebindningar peptider emellan. Systemen med mindre vattenmängd uppvisar däremot fler vätebindningar mellan peptid och sackarid och färre bindningar mellan peptid och vatten, än motsvarande system med mer vatten.

Vidare noteras att antalet vätebindningar mellan peptiderna ökar vid tillsatts av sackariderna i systemen med 2LD2 (A och C), medan antalet minskar i systemen med 6N8C (B och D) samt för systemen med båda peptiderna (E). Både sackaros och trehalos påverkar även vätebindningar mellan peptiderna och vattenmolekylerna, där referenssystemen (A3-E3) - de system utan sackarider - genomgående observeras ha något fler vätebindningar än systemen innehållande sackariderna.

De beräknade feluppskattningarna är av storleksordning  $\approx 10\,000$  gånger mindre än antalet vätebindningar för peptid-peptid samt peptid-vatten, och av storleksordning  $\approx 1000$  gånger mindre för peptid-sackarid. Detta tyder på att totala antalet vätebindningar förändras mycket lite under det analyserade förloppet på 50-300 ns, och att felet kan antas vara försumbart.

Genom att kombinera ovan resultat med molekylära RDF, kan ytterligare information utläsas angående vilka delar av peptiderna som är delaktiga i de olika bindningarna. De beräknade  $g(r)$ , mellan masscentrum för aminosyror respektive sackarider, plottades för varje system och dessa figurer återfinns i bilagor A.1. Resultaten av kurvorna är sammanställda i tabell 3, där de aminosyror som uppvisar bindningar till sackariderna ( $g(r) \geq 1$ ) presenteras i andra kolumnen. För systemen innehållande båda peptiderna, beräknades  $g(r)$  för respektive peptid. I tabell 3 beskrivs därför dessa system två gånger.

*Tabell 3:* De aminosyror vilka uppvisar en tydlig topp där  $g(r) \geq 1$ , inom avståndet 0.7 nm till sackariderna, listas för samtliga system.

System	RDF aminosyra-sackarid
A1	MET(1), ALA(2), THR(3), MET(8), LYS(9), LYS(15)
A2	GLU(5), GLU(12)
B1	THR(3), LEU(4), GLU(5), MET(8), LYS(9), SER(13), LYS(15)
B2	GLU(5), GLU(12), GLN(18)
C1	MET(1), ALA(2), GLU(5), SER(13), SER(16)
C2	GLU(5), GLU(12)
D1	ALA(2), THR(3), GLU(5), MET(8), LYS(9), LYS(15)
D2	LEU(4), GLU(5), LYS(6), MET(8), LYS(9), GLU(12), LYS(15)
E1, 2LD2	MET(1), ALA(2), GLU(5), LYS(9), GLU(12)
E1, 6N8C	ALA(2), THR(3), LEU(4), GLU(5), LYS(9), GLU(12), LYS(15), PHE(17)
E2, 2LD2	MET(1), GLU(5), LYS(6), MET(8), LYS(9), GLU(12), SER(13), LYS(15), SER(16), PHE(17)
E2, 6N8C	ALA(2), THR(3), GLU(5), LYS(6), MET(8), LYS(9)

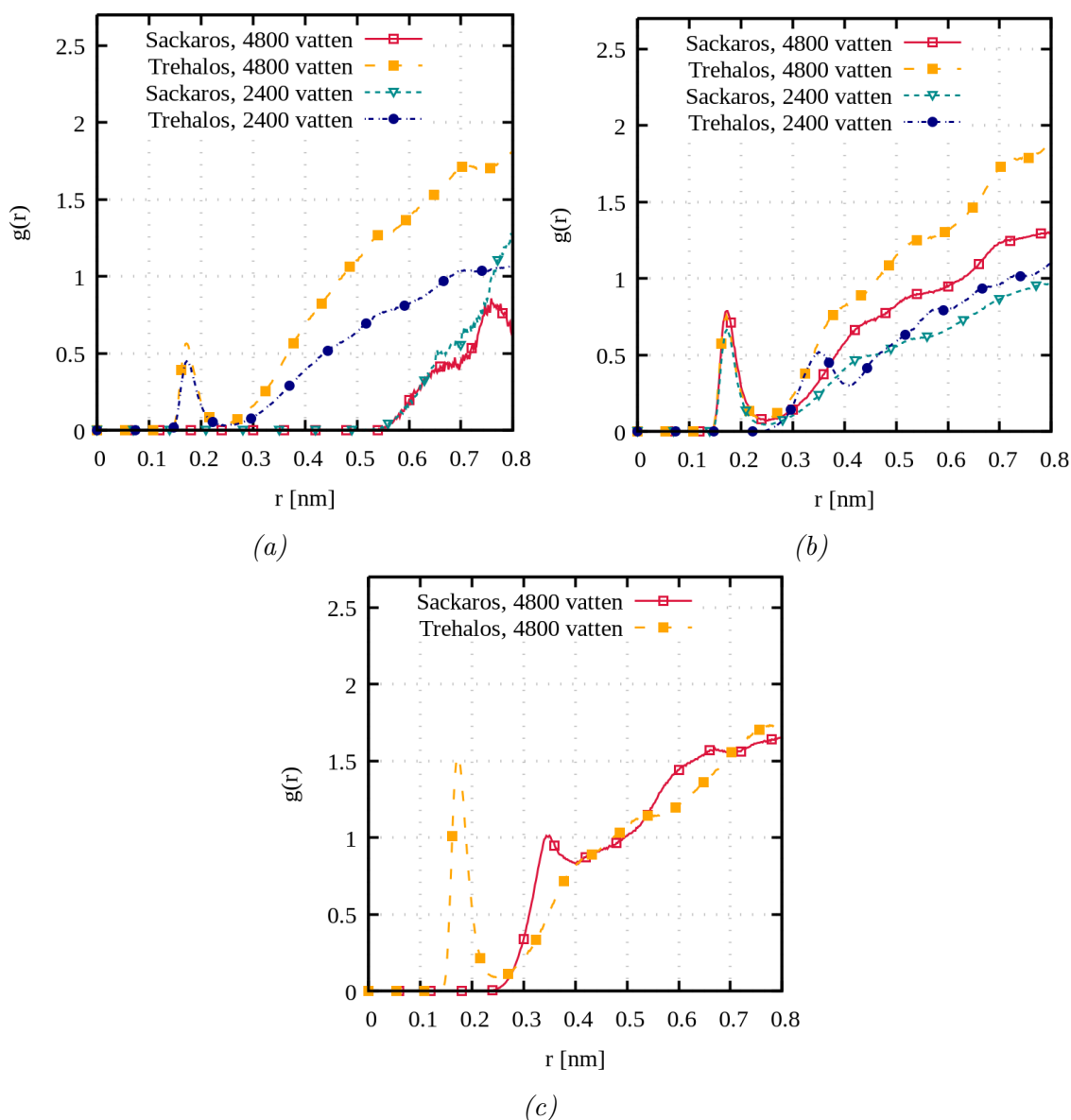
Ett mönster som kan utläsas ur tabellen är att det generellt existerar fler toppar för systemen med sackaros, till skillnad från systemen innehållande trehalos. Vidare noteras att sackaros binder till antingen MET(1) eller MET(8) samt LYS(15) i nästintill samtliga system. Samtidigt binder trehalos nästan uteslutande till GLU(5) och GLU(12). I systemen med båda peptiderna verkar fler olika aminosyror binda

till sackariderna, jämfört med system där peptiderna undersöks individuellt.

Vid förändrad vattenkoncentration kan vissa skillnader i antal och typ av aminosyror som binder till sackariderna iaktas. Ingen tydlig trend kan däremot observeras angående hur denna förändring framträder.

## 5.2 Atomär RDF

De atomära RDF för syreatomer hos peptiderna och väteatomer bundna till syreatomer hos sackariderna redovisas i figur 9.



Figur 9:  $g(r)$  mellan syreatomer hos aminosyror och väteatomer bundna till syre hos sackariderna, plottade som funktion av  $r$ . Figur a) visar system med 2LD2 (A & C), b) system innehållande 6N8C (B & D), c) system innehållande en blandning av de båda peptiderna (E).

Trenden som kan observeras för de flesta system är en topp vid avstånden  $\approx 0.18$  nm eller vid  $\approx 0.34$  nm, vilket som tidigare nämnt ligger under de relevanta avstånden som kan indikera vätebindningar. Genom att försumma de system där topparna inte sträcker sig över  $g(r) = 1$  återstår endast två system: E1 och E2, innehållandes båda peptider. Noterbart i sammanhanget är att system E2 uppvisar en svagare vätebindning än E1 då toppen är på längre avstånd. Avsaknad av tillräckligt höga och tydliga toppar i övriga system indikerar att antalet vätebindningar mellan dessa atomer är något färre.

### 5.3 Mellanliggande spridningsfunktion för självrörelse

För att undersöka hur rörelseförmågan hos partiklarna skiljer sig mellan olika system, beräknades och plottades  $F_s(\vec{k}, t)$ , med partikelavståndet från origo  $r_j(t) = 0.5$  nm ( $k = 12.6$ ). Jämförelser av relaxationstiderna är sammanställda i tabell 4, med kolumner 2-4 för peptider, och kolumner 5-6 för sackarider. Resterande resultat återfinns i bilagor A.2. Relaxationstiden redovisas som en jämförelse inom samtliga system, där till exempel A1, A2 och A3 endast jämförs med varandra. I tabellen motsvarar  $L$  långsammast relaxationstid,  $S$  snabbast relaxationstid och  $M$  relaxationstiden däremellan. Peptidernas relaxationstid jämförs med varandra, och sackaridernas med varandra.

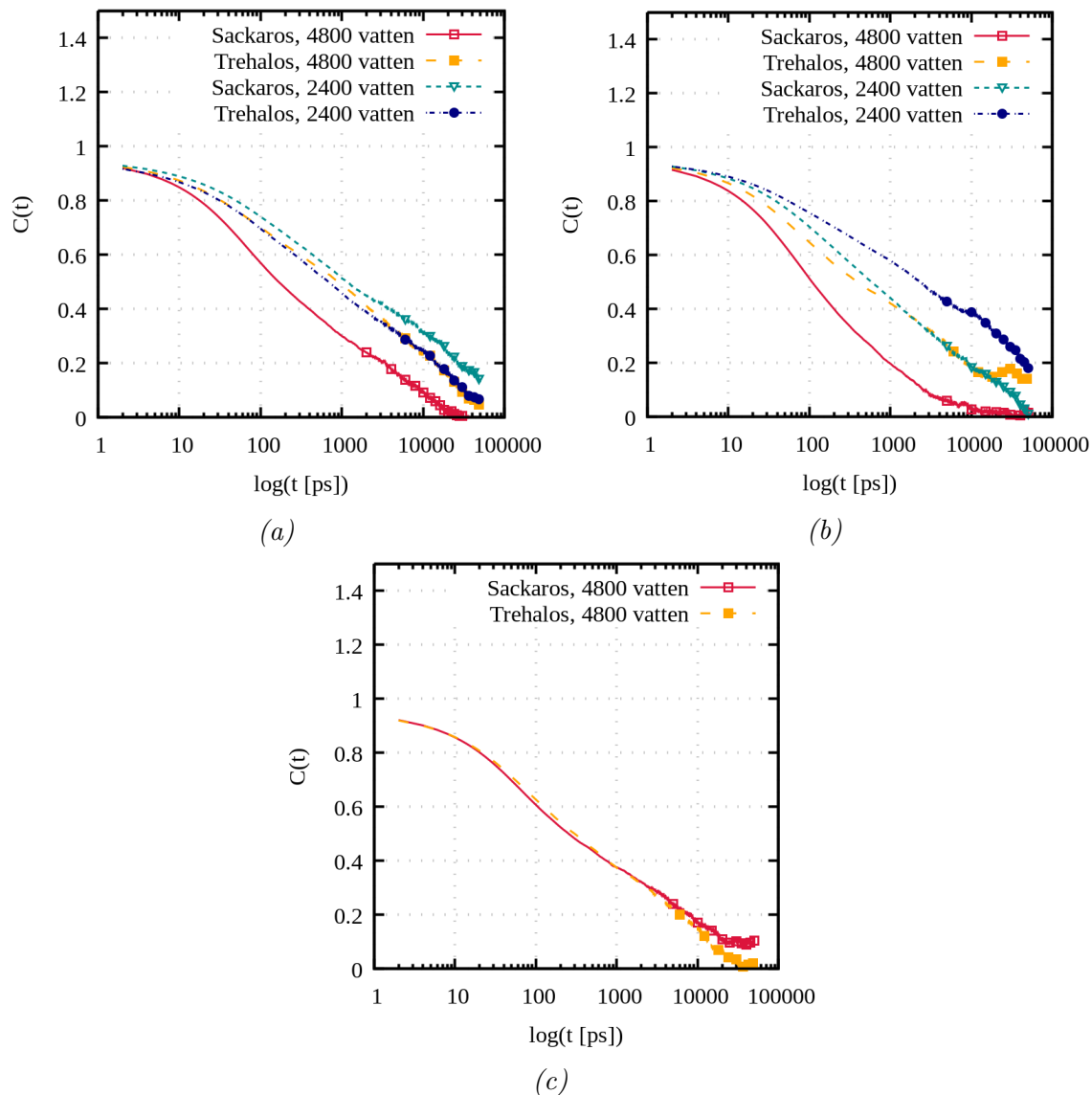
*Tabell 4:* En jämförelse av peptidernas relaxationstid inom systemen A(1-3), B(1-3), C(1-3), D(1-3) och E(1-3), under tidsintervallet 50-300 ns, för  $k = 12.6$ .  $L$  betecknar den långsammaste relaxationstiden,  $S$  den snabbaste relaxationstiden och  $M$  relaxationstiden däremellan.

System	Pep(sack)	Pep(tre)	Pep(utan sackarid)	Sackaros	Trehalos
A	$L$	$M$	$S$	$S$	$L$
B	$M$	$L$	$S$	$S$	$L$
C	$L$	$M$	$S$	$L$	$S$
D	$L$	$M$	$S$	$S$	$L$
E	$L$	$M$	$S$	$L$	$S$

Från resultatet i tabell 4 kan observeras att peptiderna i samtliga system utan sackariderna relaxerar snabbast ( $S$ ), medan relaxationstiden för peptiderna i system med sackaros är långsammast ( $L$ ) i fyra av fem fall. Vad gäller sackaridernas relaxationstid kan en mindre skillnad urskiljas, där trehalos iakttas relaxera långsammast i tre av fem fall.

### 5.4 Rotationskorrelationsfunktion

I detta avsnitt presenteras resultaten av sockerarternas benägenhet att rotera i systemen. I figur 10 visas  $C(t)$  plottade som funktion av logaritmerad tid för alla system innehållande någon sackarid.



Figur 10:  $C(t)$  för sackaros och trehalos plottade som funktion av logaritmerad tid. Figur a) visar system med 2LD2 (A & C), b) system innehållande 6N8C (B & D), c) system innehållande en blandning av de båda peptiderna (E).

I figurer 10a och 10b beskrivs de system innehållande de två peptiderna enskilt med båda sackariderna. Utifrån perspektiv av vatteninnehåll kan observeras att system med mindre vatten generellt har en långsammare rotation än sin motsvarighet med större vattenmängd. Undantag till detta är trehalos i system med peptiden 2LD2 där rotationen förblir identisk, oberoende av vatteninnehåll.

När sockerarternas skilda inverkan studeras i ovan nämnda system framgår att sackaros genomgående tycks rotera snabbare än trehalos. Återigen existerar emellertid ett undantag för system med 2LD2 och mindre vatten, där sackaros istället står för den långsammare rotationen. För systemen innehållandes båda peptiderna åskådliggörs sackaridernas påverkan på RCF i figur 10c. Där kan ingen signifikant skillnad mellan sackariderna urskiljas.

## 5.5 Contact maps

Medelantalet kontakter mellan aminosyror i peptiderna för respektive system, tillsammans med deras feluppskattningar avrundade till två decimaler, redovisas i tabell 5. Dessa listas för de satta tidsintervallen 50-55 ns, 175-180 ns, 295-300 ns. Figurer med aminosyra-aminosyrakontakter specificerade återfinns i bilagor A.3.

*Tabell 5:* Medelvärdet av antal kontakter mellan aminosyror i peptiderna under de tre tidsintervallen 50-55 ns, 175-180 ns samt 295-300 ns, med tillhörande feluppskattningar.

System	# 50-55ns	# 175-180ns	# 295-300ns
A1	336 ± 7.77	315 ± 7.12	329 ± 7.77
A2	302 ± 6.10	302 ± 6.32	304 ± 6.86
A3	355 ± 8.53	380 ± 7.40	393 ± 8.69
B1	333 ± 8.48	341 ± 7.66	341 ± 7.95
B2	320 ± 7.60	319 ± 6.96	321 ± 7.60
B3	356 ± 7.80	359 ± 7.98	364 ± 7.77
C1	358 ± 7.73	337 ± 7.35	337 ± 7.44
C2	358 ± 8.36	335 ± 7.68	342 ± 7.74
C3	378 ± 7.66	358 ± 6.83	376 ± 8.44
D1	337 ± 7.82	324 ± 6.81	340 ± 6.75
D2	322 ± 7.04	317 ± 6.70	323 ± 6.94
D3	366 ± 7.30	364 ± 6.49	363 ± 6.90
E1	309 ± 7.32	318 ± 7.28	308 ± 7.49
E2	324 ± 7.83	310 ± 7.56	322 ± 7.49
E3	349 ± 8.50	382 ± 8.44	350 ± 7.94

Utifrån tabellen kan utläsas att samtliga referenssystem har flest antal kontakter. Dessutom har peptiderna i system med trehalos generellt mindre antal kontakter än de i system med sackaros. Utmärkande undantagsfall från detta är systemen med båda peptiderna (system E) för intervallen 50-55 ns och 295-300 ns. Vidare tycks vattenmängden inte ha någon påverkan på antalet kontakter. Viktigt att ta hänsyn till är att feluppskattningarna är någorlunda stora i jämförelse med medelvärdena, och därför kan vissa avvikelser kring de erhållna resultaten existera.

## 5.6 Sekundärstruktur

Den sekundära strukturen för varje peptid i systemen visualiseras i bilagor A.4, och en sammanställning av de procentuella andelarna i respektive system, över hela tidsintervallet 50-300 ns, presenteras i tabell 6. Procentsatsen är avrundad till en decimal och har beräknats för de möjliga strukturerna: *slumpmässig spole* (*C*),  *$\pi$ -helix* (*I*), *3-10-helix* (*G*),  *$\alpha$ -helix* (*H*), *isolerat  $\beta$ -ark* (*B*), *utökad konformation* (*E*) och  *$\beta$ -sväng* (*T*).

Inledningsvis kan observeras att i samtliga system domineras peptidernas struktur

av  $\alpha$ -helix, vilken utgör ungefär 60-80 % av deras konfiguration. Likaså består en stor del (18-35 %) av peptidernas sekundära struktur av *slumpmässig spole*, och en liten andel (1-7 %) utgörs av  $\beta$ -svängar. Utöver detta består konfigurationerna delvis av  $3$ - $10$ -helix-strukturer, vilka emellertid inte förekommer i en tillräcklig mängd för att tydliga mönster ska kunna urskiljas. I tabellen återfinns även  $\pi$ -helix, isolerat  $\beta$ -ark samt *utökad konformation*, vilka upptar en minimal andel av den sekundära strukturen, och försummas därför.

Tabell 6: De procentuella andelarna av olika sekundärstrukturer för peptiderna i systemen, över hela simuleringstiden.  $C$  = *slumpmässig spole*,  $I$  =  $\pi$ -helix,  $G$  =  $3$ - $10$ -helix,  $H$  =  $\alpha$ -helix,  $B$  = isolerat  $\beta$ -ark,  $E$  = *utökad konformation* och  $T$  =  $\beta$ -sväng.

System	$C$ [%]	$I$ [%]	$G$ [%]	$H$ [%]	$B$ [%]	$E$ [%]	$T$ [%]
A1	25.4	0	0	69.7	0.2	0	4.6
A2	26.0	0	0	73.0	0	0	0.9
A3	33.6	0	0.2	59.2	0.2	0	6.8
B1	21.8	0	0.2	74.7	0	0	3.2
B2	22.3	0	0.9	71.9	0	0	4.8
B3	17.8	0	0.3	76.9	0	0	5.0
C1	20.3	0	0.1	77.4	0	0	2.2
C2	21.0	0	0.2	77.7	0	0	1.1
C3	26.1	0	0.3	66.5	0	0	7.0
D1	18.9	0	0.2	73.8	0	0	7.1
D2	21.2	0	0.1	76.1	0	0	2.5
D3	19.0	0	0.1	79.2	0	0	1.7
E1	23.9	0	0.2	73.8	0	0	2.1
E2	22.2	0	0.9	70.1	0	0	6.8
E3	23.0	0	0.4	73.0	0	0	3.6

Andelen av den mest framträdande strukturen,  $\alpha$ -helix, tycks vid tillsatts av sackariderna öka med mer än 10 procentenheter för system innehållande 2LD2 (A och C). För system med 6N8C (B och D) minskar istället andelen av strukturen vid tillsatts av sackariderna, dock inte lika kraftigt som ökningen för 2LD2. I system med båda peptiderna (E) sker enbart en mindre förändring av andelen  $\alpha$ -helix vid tillsatts av sackaros och trehalos.

Vidare observeras att för 2LD2 reduceras andelen *slumpmässig spole* med mer än 5 procentenheter för system med sackaros och trehalos. Vad beträffar peptid 6N8C ökar däremot förekomsten av *slumpmässig spole* med mer än 2 procentenheter i närvaro av sackariderna, med undantag för system med sackaros och mindre vatten. För resterande system sker ingen nämnvärd förändring av *slumpmässig spole* då sackariderna tillsätts.

Fortsatt noteras en förändring av andelen  $\beta$ -svängar i peptidernas struktur. I samtliga system med peptiderna var för sig, med undantag för system D, sker en minskning

av andelen  $\beta$ -svängar i närvaro av sackariderna. För system D sker istället en ökning av dessa strukturer. System innehållande båda peptider består av en lägre andel  $\beta$ -svängar än motsvarande system med sackaros och trehalos.

Sammanfattningsvis angående sockerarternas generella påverkan, tycks konformationen för 2LD2 öka i andelen  $\alpha$ -helix och minska i *slumpmässig spole* vid tillsatts av sackariderna. För 6N8C sker det omvända; en minskning av  $\alpha$ -helix och en ökning av *slumpmässig spole* (med undantag för system D1). Dessa förändringar minskar strukturdifferensen mellan system A-D jämfört med peptidblandningen (system E), för vilken sekundärstrukturen behålls nästintill konstant oberoende av sackaridernas närvaro.

Vad beträffar skillnaden mellan de olika sackaridernas påverkan kan några få mönster noteras. Det förefaller som att sackaros, i större utsträckning än trehalos, bidrar till en minskning av strukturen *slumpmässig spole* hos 2LD2. Omvänt påverkar trehalos *slumpmässig spole* mer än sackaros hos 6N8C. Dessutom tycks trehalos påverka strukturen  $\alpha$ -helix mer än sackaros hos båda peptiderna, oavsett om andelen tilltar eller avtar. Därutöver kan konstateras att alla observerade förändringar för *slumpmässig spole* och  $\alpha$ -helix bidrar till en utjämning av strukturskillnaderna mellan peptidblandningen och övriga system. Inget tydligt mönster kan utrönas gällande sackaridernas skiljda påverkan på  $\beta$ -sväng.

Slutligen urskiljs att vattenmängden i de olika systemen har en bidragande effekt endast gällande strukturen *slumpmässig spole*. Vid reducerad vattenkoncentration förändras denna struktur avseendevårt mindre än för den större vattenmängden. Ytterligare samband kring huruvida vattenmängden påverkar den sekundära strukturen kan inte uttydas.

## 6 Diskussion

Nedan diskuteras projektets metoder tillsammans med deras felkällor, samt de erhållna resultaten. Både MD-simuleringarna samt de tillämpade analysmetoderna, bygger på approximationer och förenklingar, men dess höga tids- och rums-upplösning bidrar ändå till att metoden är betydelsefull. De olika analysmetodernas resultat diskuteras först individuellt, för att sedan sammanställas och kopplas ihop till en slutsats angående sackaridernas påverkan på aggregeringen av Htt.

### 6.1 Diskussion om metod

Undersökningsmetoden atomära MD-simuleringar kan, som tidigare beskrivits, skildra hur system och deras komponenter beter sig med mycket hög tidsupplösning och på atomär nivå. Detta medför att system kan studeras med en detaljrikedom som inte är möjlig att uppnå med experimentella metoder. Vidare ställer MD-simuleringar inte några krav på tillgång till material eller de molekyler som studeras, och är även av denna anledning en betydelsefull undersökningsmetod. En mängd analysmetoder kan tillämpas för att studera resultaten av simuleringarna, där en handfull valdes i detta projekt. Dessa är värdefulla redskap för att kunna dra slutsatser om simuleringarnas resultat. Trots fördelarna med MD-simuleringar finns det osäkerheter och begränsningar kring dessa och de applicerade analysmetoderna, vilka redogörs för i följande två avsnitt.

#### 6.1.1 Felkällor för MD-simuleringar

De potentiella energierna approximeras som nämnt i 3.2 för de bindande och icke-bindande interaktionerna, och har enskilda parametreringskontakter vilka är framtagna empiriskt för de atomer som de designas för. Trots approximeringarna anses dock sannolikheten låg att betydande fel uppstår som konsekvens av detta, då uttrycken är delvis utvecklade för biologiska system. Mycket vikt läggs vid korrekta kraftfält, då eventuella fel i dessa kan förstärkas för varje tidssteg och därmed utgöra en central felkälla, trots ett i början litet fel. Att kraftfälten baseras på experimentell data talar dock för att de bör kunna efterlikna verkligheten relativt tillförlitligt.

En ytterligare begränsning med simuleringar *in silico* är att de hanterar diskreta tidssteg, till skillnad från verkligheten som är tidskontinuerlig. Risken existerar således att tidsstegen väljs för stora, vilket helt kan osynliggöra interaktioner på kortare tidsskala. I detta projekt är det minsta relevanta händelseförloppet vätebindningars vibration på 10-13 fs. I detta avseende anses det valda tidssteget på 2 fs som tillräckligt för att nå tillräcklig precision, även fast viss avvikelse från verkligheten inte kan uteslutas.

Vidare utgår autenticiteten av simuleringarna i detta projekt till hög grad från huruvida de skapade systemen speglar förhållandena i ett biologiskt system på ett korrekt sätt. I detta fall framgår tydligt att förenklingen kring att kroppen enbart innehåller vatten, Huntingtinpeptider samt natrium- och kloridjoner begränsar tillämpningen av resultatet. I ett verkligt scenario skulle flertalet ytterligare faktorer,

såsom interaktioner till cellulära membran, förekomst av proteiner eller peptider av annat ursprung, samt närvaro av andra salter och joner, kunna spela en bidragande och oförutsägbar roll i aggregeringsprocessen.

En ytterligare begränsning relaterad till kompositionen av biologiska system är att densiteten är okänd. Detta eftersom sammansättningen tenderar att variera kraftigt när tillräckligt små regioner betraktas. För att undgå denna problematik vid simulering används NPT-ensemblen, som bättre representerar den verkliga miljön. Dock speglas fortfarande inte verkligheten idealt, då partiklar, tryck och temperatur (vilka NPT-ensemblen sätter konstanta) varierar något i kroppen.

Projektet behandlar dessutom endast två fragment av Htt på 17 aminosyror vardera, vilket totalt sett innebär att endast 18 av Htts totalt drygt 3000 aminosyror studerades. Således kan ingen definitiv slutsats av resultaten dras om hela proteinet och dess polyglutaminkanal, eller hur sackariderna kan påverka hela proteinets aggregering. Däremot existerar sannolikt dessa peptid-fragment enskilt hos människor (se avsnitt 2.2), och har påvisats spela en kritisk roll i aggregeringen av hela proteinet och dess toxicitet. Resultat från detta projekt kan därmed, trots arbetets avgränsade omfång, ge relevanta indikationer gällande sackaridernas effekt på hela Htt.

Utöver detta studerades inte huruvida koncentrationen av sackariderna förändrar deras inverkan, då endast ett bestämt antal på 16 stycken lades till varje system. Denna koncentration valdes utifrån sackaridernas storlek i jämförelse med peptidernas. Brist på variation av sackarid-koncentration resulterar i en ovisshet angående eventuell skillnad i hur en annan sackaridmängd skulle påverka aggregeringen, även om denna skillnad kan antas vara liten.

### 6.1.2 Felkällor för analysmetoder

Beräkningarna av  $g(r)$  mellan peptiderna och sackariderna ger en tydlig bild av hur deras olika molekyler och atomer interagerar med varandra. I detta projekt beräknades atomära RDF endast mellan ett typ av atompar, och bindningar som eventuellt bildas mellan andra atomer bortses därmed från. Som komplement beräknades det totala antal vätebindningar, vilka istället begränsas av att det inte specificeras vilka atompar som bildat de noterade vätebindningarna. Det bör även nämnas att antalet vätebindningar inte helt säkert indikerar vilken sackarid som främst stabiliserar peptiderna, då stabilisering även kan ske på andra sätt. Exempelvis skulle sackariderna kunna agera som steriskt hinder, likt beskrivet i 2.3.1.

De molekylära RDF begränsas av det faktum att funktionen utgår från molekylernas masscentrum, och därmed inte beaktar deras utbredning. För att till viss utsträckning undgå detta användes en högre toleranslängd på 0.7 nm mellan två masscentrum, i jämförelse med de atomära RDF, vilka använder omkring 0.3 nm för att klassificera en vätebindning. Trots att denna anpassning anses lämplig, medför detta osäkerhet kring vilka avstånd som faktiskt representerar en bindning och inte. Dessutom bortses fortfarande avstånd större än 0.7 nm, som också skulle kunna vara bindningar, även om sannolikheten avtar för ökande avstånd. Att denna sannolikhet har en betydelse bidrar till en viss osäkerhet kring beräkningarna. På

grund av funktionens beroende av molekylernas storlek kan heller inte topparnas storlek jämföras med varandra på ett meningsfullt sätt.

De två metoderna SISF och RCF kan indikera hur snabbt molekylerna rör sig, respektive hur snabbt sackariderna roterar, under simuleringstiden. Det bör dock tas i beaktande att  $F_s(\vec{k}, t)$  (SISF) ej tar hänsyn till bindningar eller molekylernas närhet till varandra, och att  $C(t)$  (RCF) inte förklarar orsaken bakom ökad eller minskad rotationshastighet hos sackariderna i de olika systemen.

CM är ett användbart verktyg för att visualisera peptid-peptidinteraktioner, vilket utgör en av de tydligaste indikationerna till aggregering i detta projekt. En väsentlig begränsning med denna metod härrör dock från det faktum att resultatet ej särskiljer mellan interaktioner två peptider emellan, och interaktioner inom samma aminosyrakedja, då en peptid bildar sekundärstruktur. Vidare tas i resultatet endast hänsyn till det totala antalet kontakter, och särskiljer därmed inte på antalet kontakter mellan specifika aminosyror. Dessutom beräknas CM endast för tre tidsintervall på 5 ns vardera, och antalet kontakter under resterande delar av simuleringarna bortses från. Denna avgränsning görs för att i tabellform enklare kunna särskilja hur antalet kontakter förändras över intervallen.

För att tydligare konkretisera hur interaktioner mellan Huntingtinpeptider och molekyler i dess omgivning påverkar peptidernas konformation, är gestaltning av peptidernas sekundärstruktur ett mycket användbart redskap. Vidare är identifikation av eventuella kopplingar mellan sekundärstruktur och aggregering av största vikt för framtida studier. Viktigt att notera i relation till detta är dock de tidigare diskuterade begränsningarna av den simulerade miljön, då ett biologiskt system innehåller flertalet ytterligare faktorer vilka kan ha en effekt på sekundärstrukturen, som här förbises. Dessutom måste beaktas att möjlig korrelation mellan en specifik sekundärstruktur och aggregering inte nödvändigtvis betyder att ett orsakssamband existerar, även om indikationer på ett sådant är intressant inför framtida undersökningar.

Sammanfattningsvis kan påpekas att analysmetoderna i kombinerad form kan ge kvalitativ information om de ingående molekylernas beteenden. De individuella metodernas resultat bör däremot kopplas till aggregering med varsamhet, på grund av de ovan nämnda individuella begränsningarna i respektive metod.

## 6.2 Diskussion av resultat

I detta avsnitt diskuteras de erhållna resultaten som beskrevs i kapitel 5. Dessa behandlas individuellt och i kombination, för att undersöka vilka slutsatser som kan dras angående sackaros och trehalos påverkan på aggregering i olika miljöer.

Det erhållna medelantalet kontakter mellan aminosyror i peptiderna, beräknade med CM enligt tabell 5, är med marginal störst för alla referenssystemen (system utan sackarider). Detta indikerar att sackariderna tydligt reducerar antalet kontakter mellan aminosyrorna. Trehalos uppvisar generellt sett en större förmåga att minska kontakterna mellan aminosyrorna jämfört med sackaros. Feluppskattningarna för metoden är däremot relativt stora i relation till detta, men resultatet anses ändå

övertygande nog för att fastslå att dessa enbart kan ansvara för enstaka avvikelser, och inte för att bryta den större trenden.

Ovannämnda trend, att sackariderna minskar aminosyrakontakterna, skulle kunna styrkas av det faktum att sackariderna kan ses binda till peptiderna genom vätebindningar i tabell 2. Trehalos vätebinder emellertid betydligt mer till peptiderna än sackaros, vilket överensstämmer med att vätebindningar peptiderna emellan kan ses minska i närvaro av trehalos. Detta kan tolkas som att vätebindningarna mellan trehalos och peptiderna konkurrerar ut dem mellan peptiderna. För sackaros ökar däremot peptid-peptidbindningarna, och ovan nämnda teori kan därför inte appliceras som förklaring till varför sackaros minskar aminosyrakontakterna mellan peptider. Det bör däremot återigen nämnas, likt i diskuterat i avsnitt 6.1.2, att aminosyrakontakter inte behöver minskas på grund av att sackariderna binder till dessa med vätebindningar, utan även kan ske till följd av andra orsaker som exempelvis steriska effekter. Detta skulle kunna förklara sackaros förmåga att minska aminosyrakontakter.

Trehalos spekuleras som tidigare nämnt förhindra aggregering genom att bland annat ersätta peptid-peptidbindningar med peptid-sackaridbindningar. Följaktligen kan trehalos även observeras binda till aminosyrorna LEU(4), GLU(5), LYS(6), MET(8), LYS(9), GLU(12), LYS(15), samt GLN(18), vilket markerar dessa som potentiella kandidater för interaktioner mellan peptider, och därmed aggregering. Av dessa utmärker sig glutaminyheten GLN(18) särskilt, då 6N8C, vilken innehåller GLN(18), i samtliga system har 5-30 fler bindningar peptider emellan än 2LD2. Detta kan indikera att glutaminyheten, som inte finns i 2LD2, bidrar till fler peptid-peptidbindningar och därmed eventuellt även till aggregering. Detta styrks även av fynd från litteraturstudien, i avsnitt 2.2, där Huntingtinproteinets aggregering kartlagts till ett ökat antal polyglutaminyheter.

Det totala antalet vätebindningar är, som tidigare nämnt, fler mellan trehalos och peptiderna än sackaros och peptiderna. Intressant att notera är att resultaten för molekylär RDF indikerar att sackaros regelbundet binder till fler olika sorters aminosyror än trehalos, trots att trehalos har större antal vätebindningar totalt. Dessa resultat kan tolkas som att sackaros interagerar med peptiderna genom interaktioner av annan karaktär än vätebindningar, i större utsträckning än vad trehalos gör. En potentiell förklaring till dessa samband kan relateras till de strukturella skillnaderna hos trehalos och sackaros, där de olika placeringarna av sackaridernas OH-grupper skulle kunna främja fler vätebindningar för trehalos.

Vattenmängden tycks inte ha någon signifikant påverkan på antalet vätebindningar, med undantag för peptid-vattenbindningar, där ett minskat antal kan observeras för system med mindre vatten. Att ett reducerat antal vattenmolekyler inte genererar lika många bindningar till peptiderna attribueras dock inte till en reducerad förmåga för enskilda vattenmolekyler att binda, utan härleds enbart till det minskade antalet. Däremot bör poängteras att antalet bindningar mellan vattenmolekylerna och peptiderna inte halveras då vattenmängden halveras, utan att endast en mindre förändring kan iakttas. Detta kan tyda på att det maximala antalet möjliga bindningar till peptiderna uppnås, och överskrids, i systemen innehållandes mycket vat-

ten, vilket är av vikt vid design av framtida system. En ytterligare trend angående vattenmängden, som kan iakttas i figur 10, är att ett reducerat antal vattenmolekyler genomgående kopplas till långsammare rotation av sackariderna samtidigt som bindningarna mellan sackarid och vatten även de observeras minska. Att fler bindningar skulle medföra en snabbare rotation är emellertid inte vad som kan förväntas mekaniskt, och orsaken bakom denna korrelation förblir okänd. Vattenkoncentrationens påverkan på sekundärstruktur begränsas till strukturen *slumpmässig spole*, vilken ses förändras något mindre för system med färre vattenmolekyler. När projektets frågeställningar, relaterade till aggregering, beaktas verkar dessa samband däremot inte spela någon avgörande roll.

Antalet vätebindningar mellan peptider och vatten kan också observeras avta när sackariderna närvarar i systemet, med den största minskningen då trehalos närvarar. Detta indikerar att sackaridernas affinitet för att binda till peptiderna är överlägsen i systemet, och därmed konkurrerar ut vattenmolekylernas bindningar till peptider.

Vad beträffar sekundärstrukturerna kan en möjlig korrelation till stabiliseringsgrad göras. Inledningsvis observeras att i E-systemen, innehållande båda peptiderna, bevaras sekundärstrukturen förhållandevis konstant även efter tillägg av sackaros och trehalos. Vidare uppvisar referenssystem E3 minst antal kontakter mellan aminosyror i peptiderna i tabell 5 jämfört med de andra referenssystemen, vilket tyder på att aggregeringsgraden är lägre i dessa peptidblandningar. Utan tillägg av sackariderna tycks 6N8C strukturera sig som  $\alpha$ -*helix* mer frekvent än 2LD2, som istället har en större benägenhet att bilda *slumpmässig spole*. Vid tillägg av sackariderna tycks de tidigare nämnda skillnaderna i sekundärstruktur dock jämnas ut. Anmärkningsvärt kan även noteras att dessa system, efter de iaktagna strukturutjämnarna, skiljer sig mycket lite från E-systemen. Detta tyder starkt på att en sekundärstruktur bestående av mellan 19-25% *slumpmässig spole* och 70-77%  $\alpha$ -*helix* skulle kunna utgöra den mest termodynamiskt fördelaktiga konfigurationen, minst benägen att aggregera, för peptiderna 6N8C och 2LD2, i de studerade systemen.

Med ovannämnda resonemang i åtanke tycks det även som att tillägg av sackaros och trehalos, dock i olika utsträckning, kan underlätta den vikningsprocess vilken är nödvändig för att peptiderna ska uppnå den eventuellt mest stabila sekundärstrukturen, vilken diskuterats ovan. Trehalos observeras nästintill genomgående utjämna skillnaderna mellan strukturerna hos system A-D jämfört med E, mer än sackaros. Det enda tydliga undantag till detta är andelen *slumpmässig spole* hos 2LD2, där sackaros utjämnar mer. Under antagandet att sekundärstrukturen för E-systemens peptider hämmar aggregering, skulle detta indikera att trehalos är en bättre stabilisator än sackaros. Anledningen till detta skulle återigen kunna härledas till de strukturella skillnaderna hos trehalos och sackaros. Vätebindningar kan i sådant fall antas vara av stor vikt för bildning av sekundärstrukturen  $\alpha$ -*helix* för båda peptider samt *slumpmässig spole* hos 6N8C. Att sackaros tycks utjämna skillnaden i strukturen *slumpmässig spole* mer för 2LD2, skulle möjligtvis kunna implicera att steriska effekter är av större vikt i detta fall. Detta samband är emellertid för diffust för att kunna avgöras säkert.

Peptidblandningens utmärkande egenskaper kan även urskiljas i resultaten för atomära RDF. Där kan avläsas att vätebindningar, mellan de studerade atomerna, endast bildas i relevant andel för system E1 och E2. Detta skulle kunna förklaras av att den eventuellt termodynamiskt fördelaktiga konformationen, vilken förekommer till störst del i dessa system, tycks främja vätebindningar mellan de specifika atomer som undersöktes hos peptiderna och sackariderna.

I nästintill samtliga system innehållande en typ av peptid kan trehalos kopplas till den största minskningen av aminosyrakontakter mellan peptiderna. I system med båda peptiderna står istället sackaros för den största minskningen av kontakter. En potentiell förklaring till detta kan härledas till de skiljda interaktionsvägarna hos sackaros och trehalos som det tidigare spekulerats kring. Då det tidigare indikerades att trehalos skulle kunna kopplas till stabilisering genom vätebindningar med peptiderna, i större utsträckning än sackaros, kan detta implicera att interaktioner av denna typ skulle kunna spela en central roll i stabiliseringen av de enskilda peptiderna. Istället skulle sackaros, och interaktioner genom steriska hinder, potentiellt kunna associeras med en högre grad av stabilisering av peptidblandningen. Den sistnämnda förklaringsmodellen är emellertid diffus, då starkare underlag skulle behövas för att styrka att peptidblandningen stabiliseras ytterligare av sackaros.

För att undersöka orsaken bakom varför sekundärstrukturen i peptidblandningen skiljer sig från resterande system behöver konformationerna jämföras. Detta kan åstadkommas genom jämförelse av referenssystem E3 med både A3 och B3, vilka innehåller peptiderna enskilt men i övrigt är identiska. Då kan noteras att E3:s andelar av de olika strukturerna inte motsvarar det som kan förväntas teoretiskt; ett medelvärde av andelarna för A3:s och B3:s sekundärstrukturer. Detta stämmer inte överens med vad som erhålls av simuleringarna, vilka kan ses i tabell 6, och tyder på att interaktionerna mellan de två olika peptiderna förändrar sekundärstrukturerna. Stöd för detta återfinns även i resultat för molekylär RDF, som tidigare nämnts uppvisa en större variation mellan vilka aminosyror sackariderna binder till i peptidblandningen.

Från resultaten av SISF för peptider, sammanfattade i tabell 4, kan utläsas att peptiderna i system utan sackariderna relaxerar snabbast, vilket tyder på att peptiderna stabiliseras vid närvaro av sackariderna. Av sackariderna iaktas sackaros vara den främsta stabilisatorn i denna aspekt, då de i betydligt större utsträckning kopplas till en långsammare relaxation. Detta skulle återigen kunna förklaras av strukturella skillnader sackariderna emellan. Viktig särskiljning i relation till detta är att peptidrelaxation inte i sig själv bör kopplas till aggregering. Likt tidigare diskuterat kan emellertid tillägg av sackarider tydligt kopplas till en utjämning mot vad tydligt pekar på en energetiskt fördelaktig sekundärstrukturskonformation. Dessa två samband i kombination kan utgöra en tillförlitlig indikator över aggregering.

Fortsättningsvis impliceras av RCF, i figur 10, att trehalos generellt roterar långsammare än sackaros. När SISF för sackariderna, i tabell 4, betraktas syns även att trehalos uppvisar en långsammare relaxation jämfört med sackaros. Detta är i enighet med det faktum att trehalos typiskt vätebinder till peptiderna mest frekvent, vilket skulle kunna påverka dess rörelse mer. Sackaros snabbare rotation och

relaxation skulle då kunna härstamma från att de färre bindningarna tillåter molekylerna att röra sig snabbare, eller att rörelseförmågan är en inneboende egenskap relaterad till sackaridernas strukturskillnader. En alternativ förklaring är att sackaros snabbare rörelse reducerar dess bindningsförmåga. Intressant vidare observation är att för system med båda peptiderna tycks dessa samband upphöra, samtidigt som nya bindningsmönster kan observeras i molekylär RDF. Detta styrker att rotationsförmågan skulle kunna kopplas till bindningarnas karaktär, eftersom sackaridernas strukturer förblir oförändrade.

### 6.2.1 Sammanfattning av resultatdiskussionen

Enligt diskussionen ovan kan konstateras att systemen utan sackarider, referenssystemen, har en snabbare peptidrelaxation kombinerat med en instabil sekundärstruktur, samt fler interna kontakter mellan aminosyror jämfört med systemen innehållandes sackariderna. Detta kan tyda på en högre grad av aggregering i dessa referenssystem. Fortsättningsvis visar resultaten att kontakter mellan aminosyror i peptiderna reduceras, att relaxationstiden för peptiderna ökar, samt att skillnaderna i sekundärstrukturer mellan systemen utjämnas, då sackariderna läggs till. Sammantaget indikerar detta att aggregeringen motverkats vid tillsatts av sackariderna.

Kring sackaridernas individuella förmåga att stabilisera peptiderna iakttas ett antal skillnader. Trehalos visas vätebinda till peptiderna i betydligt högre grad än sackaros, samtidigt som vätebindningarna mellan peptiderna minskar. Trehalos minskar även antalet kontakter mellan peptiderna mer än sackaros, samt bidrar till större utjämning av sekundärstrukturerna. Sackaros ger emellertid en genomgående långsammare relaxation av peptiderna, samt även reducerar kontakterna mellan aminosyror mer än trehalos, då båda peptider förekommer i samma system.

Sammantaget kan båda sackariderna sägas bidra till stabilisering, dock tycks trehalos vara en något effektivare stabilisator. Tillvägagångssättet som denna stabilisering sker på verkar dock skilja sig mellan sackaros och trehalos.

Vattenmängden i systemen visas påverka antalet bindningar mellan peptid-vatten, rotationsförmågan för peptiderna, samt i minimal grad även hur sackariderna förändrar sekundärstrukturen för peptiderna. Sammanfattningsvis kan dock sägas att oberoende av vad som studeras tycks vattenmängden inte påverka systemens stabilisering väsentligt.

Vidare identifieras även ett specifikt sekundärstruktursintervall med 19-25 % *slumpmässig spole* och 70-77 %  *$\alpha$ -helix* som förekommer i de fall där peptiderna kan sägas vara mindre aggregerade. Det bör emellertid påpekas att detta intervall uppnås oberoende av sackariderna då båda peptider närvarar samtidigt, ett samband som i detta projekt saknar förklaring. Slutligen påvisas att aminosyrakandidaterna LEU(4), GLU(5), LYS(6), MET(8), LYS(9), GLU(12), LYS(15), samt GLN(18) möjligtvis spelar en central roll i aggregeringen av peptiderna 6N8C och 2LD2.

## 7 Slutsats och framtidsutsikter

Projektets centrala syfte var att undersöka hur sackaros och trehalos kan påverka aggregering av de två Huntingtinpeptiderna 2LD2 och 6N8C, samt vattenkoncentrationens inverkan. Resultaten tyder sammanfattningsvis på att främst trehalos, men även sackaros, reducerar aggregeringen av peptiderna. Sambandet som iakttas, för nästintill samtliga applicerade analysmetoder, är att sackariderna både reducerar de interna kontakterna mellan peptider, och minskar peptidernas rörelseförmåga i systemet. Skillnaderna mellan sackaros och trehalos skulle kunna härledas till deras individuella strukturer.

I de system där sekundärstrukturerna är instabila förefaller sackariderna stabilisera dessa, och även här ses trehalos vara den främsta stabilisatorn. Ett specifikt sekundärstrukturintervall kopplades till maximal stabilisering av peptiderna i de studerade systemen. Det presenteras också en rad möjliga kandidater bland peptidernas aminosyror som potentiellt spelar en avgörande roll i peptidaggregeringen. Även vattenmängden kan observeras ha viss påverkan på peptid- och sackarid-interaktionerna, men detta bedöms inte omfattande nog för att kunna bidra med avgörande skillnader relaterade till aggregering.

För att säkerställa resultaten påvisade i detta projekt, och vidare undersöka de mekanismer vilka ligger bakom både aggregeringen och sackaridernas påverkan på denna, skulle de molekylära interaktionerna kunna analyseras i högre detalj. Dessutom skulle fler delar av Huntingtinproteinet, samt undersökning av fler peptidblandningar, vara av intresse att studera då peptidernas interna interaktioner tycks påverka deras stabilitet. Avslutningsvis lägger detta projekt en grund för vidareutveckling av de erhållna sambanden och framtida forskning kring läkemedelsutveckling för Huntingtons sjukdom.

## Referenser

- [1] A. Wexler, E. J. Wild och S. J. Tabrizi, "George Huntington: a legacy of inquiry, empathy and hope," *Brain: A Journal of Neurology*, årg. 139, nr 8, s. 2326–2333, Jul. 2016. DOI: 10.1093/brain/aww165.
- [2] H. Chial, "Huntington's Disease: The Discovery of the Huntingtin Gene," *Nature Education*, årg. 1, nr 1, 2008. Tillgänglig: <https://www.nature.com/scitable/topicpage/huntington-s-disease-the-discovery-of-the-851/>.
- [3] R. A. C. Roos, "Huntington's disease: a clinical review," *Orphanet Journal of Rare Diseases*, årg. 5, nr 40, Dec. 2010. DOI: 10.1186/1750-1172-5-40.
- [4] J. Schulte och J. T. Littleton, "The biological function of the Huntingtin protein and its relevance to Huntington's Disease pathology," *Current Trends in Neurology*, årg. 5, s. 65–78, Jan. 2011.
- [5] G. Graziano, "How does sucrose stabilize the native state of globular proteins?" *International Journal of Biological Macromolecules*, årg. 50, nr 1, s. 230–235, Jan. 2012. DOI: 10.1016/j.ijbiomac.2011.10.025.
- [6] M. A. Mensink, H. W. Frijlink, K. van der Voort Maarschalk och W. L. J. Hinrichs, "How sugars protect proteins in the solid state and during drying (review): Mechanisms of stabilization in relation to stress conditions," *European Journal of Pharmaceutics and Biopharmaceutics*, årg. 114, s. 288–295, Maj. 2017. DOI: 10.1016/j.ejpb.2017.01.024.
- [7] C. Olsson och J. Swenson, "The role of disaccharides for protein–protein interactions – a SANS study," *Molecular Physics*, årg. 117, nr 22, s. 3408–3416, Jul. 2019. DOI: 10.1080/00268976.2019.1640400.
- [8] D. Vlachakis, E. Bencurova, N. Papangelopoulos och S. Kossida, "Current State-of-the-Art Molecular Dynamics Methods and Applications," i *Advances in Protein Chemistry and Structural Biology*, R. Donev, utg., vol. 94, Burlington: Academic Press, 2014, kap. 3, s. 274–276.
- [9] D. V. D. Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark och H. J. C. Berendsen, "GROMACS: fast, flexible, and free," *Journal of computational chemistry*, årg. 26, nr 16, s. 1701–1718, Dec. 2005. DOI: 10.1002/jcc.20291.
- [10] M. Michalek, E. S. Salnikov och B. Bechinger, "Structure and Topology of the Huntingtin 1–17 Membrane Anchor by a Combined Solution and Solid-State NMR Approach," *Biophysical Journal*, årg. 105, nr 3, s. 699–710, Aug. 2013. DOI: 10.1016/j.bpj.2013.06.030.
- [11] S. A. Kotler et al., "Probing initial transient oligomerization events facilitating Huntingtin fibril nucleation at atomic resolution by relaxation-based NMR," *PNAS*, årg. 116, nr 9, s. 3562–3571, Feb. 2019. DOI: 10.1073/pnas.1821216116.
- [12] Bruce Alberts et al., *Molecular biology of the cell*. Garland Science, 2015. Tillgänglig: <https://search.ebscohost.com/login.aspx?direct=true&db=catt07470a&AN=c1c.e2c39cfe.4f74.4255.ae8d.2994b7365104&site=eds-live&scope=site&authtype=guest&custid=s3911979&groupid=main&profile=eds>.
- [13] J. M. Manning, "Peptide," *AccessScience*, August 2019. DOI: 10.1036/1097-8542.497210.

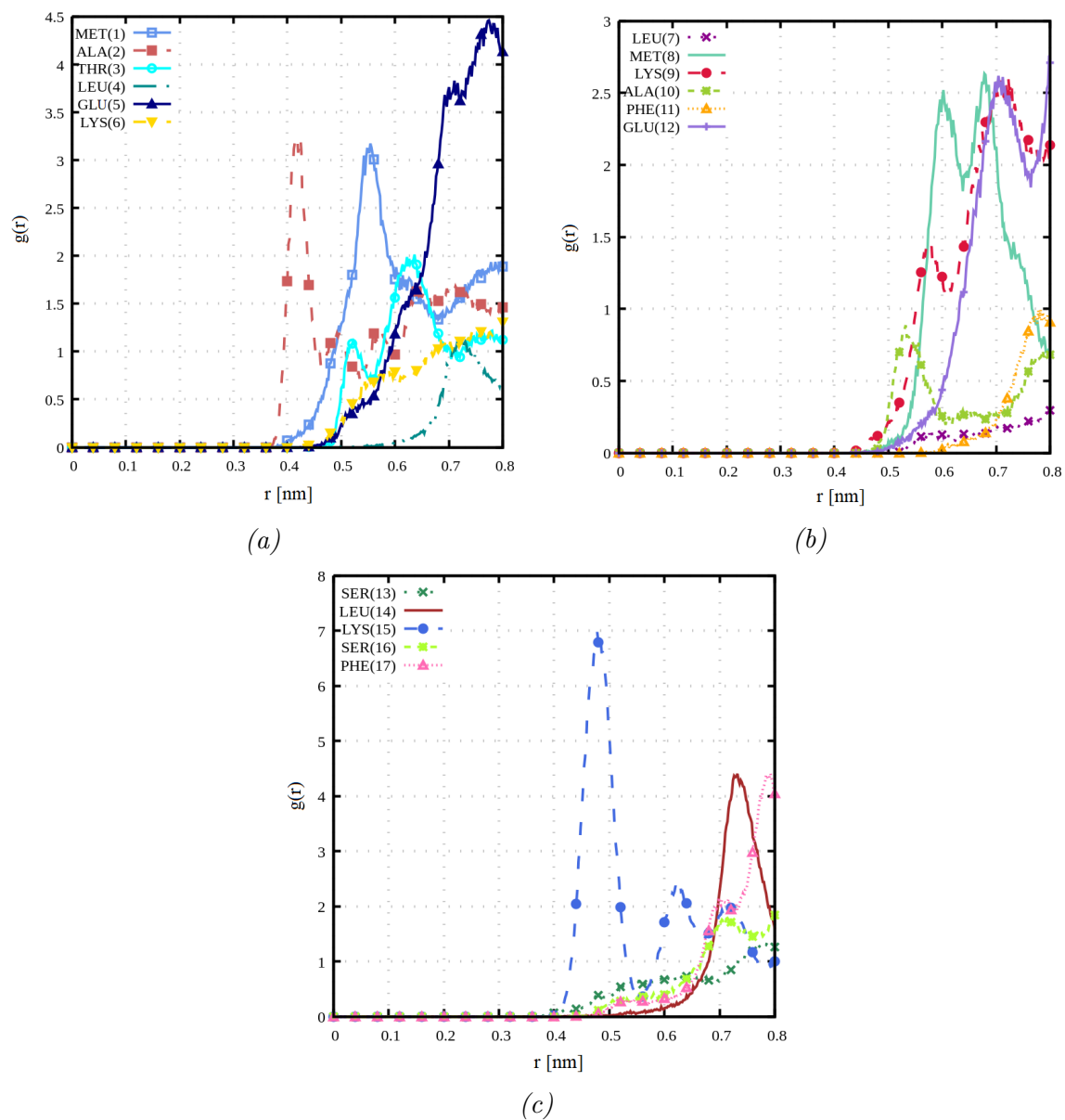
- [14] D. A. Stein, "Protein Aggregation," i *Microbiology Research Advances*, Nova Science Publishers, Inc, 2011. Tillgänglig: <https://search.ebscohost.com/login.aspx?direct=true&db=edsebk&AN=539227&site=eds-live&scope=site&authtype=guest&custid=s3911979&groupid=main&profile=eds>, s. 223–235.
- [15] T. Shafee, *Hydrogen bonds in protein secondary structure. Cartoon above, atoms below with nitrogen in blue, oxygen in red*, Feb. 2017.
- [16] T. Yushchenko, E. Deuerling och K. Hauser, "Insights into the Aggregation Mechanism of PolyQ Proteins with Different Glutamine Repeat Lengths," *Biophysical Journal*, årg. 114, nr 8, s. 1847–1857, Apr. 2018. DOI: 10.1016/j.bpj.2018.02.037.
- [17] G. Hoffner, S. Souès och P. Djian, "Aggregation of Expanded Huntingtin in the Brains of Patients With Huntington Disease," *Prion*, årg. 1, nr 1, s. 26–31, Jan. 2007. DOI: 10.4161/pri.1.1.4056.
- [18] M. P. Parsons och L. A. Raymond, *Huntington Disease*. Academic Press, 2015, kap. 20, s. 303–320. DOI: 10.1016/B978-0-12-398270-4.00020-3.
- [19] C. Olsson och J. Swenson, "The role of disaccharides for protein–protein interactions – a SANS study," *An International Journal at the Interface Between Chemistry and Physics*, årg. 117, nr 22, s. 3408–3416, Jul. 2019. DOI: 10.1080/00268976.2019.1640400.
- [20] M. Tanaka et al., "Trehalose alleviates polyglutamine-mediated pathology in a mouse model of Huntington disease," *Nature Medicine*, årg. 10, nr 2, s. 148–154, Jan. 2004. DOI: 10.1038/nm985.
- [21] C. Olsson, "Interaction of trehalose with water and protein for the understanding of biological stabilization," licentiatavhandling, Fysikinstitutionen, Chalmers Tekniska Högskola, Göteborg, Sverige, 2016. [Online]. Tillgänglig: <https://research.chalmers.se/publication/244770>.
- [22] I. Ermilova, "Modeling of biomembranes: from computational toxicology to simulations of neurodegenerative diseases," doktorsavhandling, Institutionen för Material- och Miljö kemi, Stockholms Universitet, Stockholm, Sverige, Mar. 2019. DOI: 10.13140/RG.2.2.24409.57447.
- [23] M. Cederwall och C. Forssén, "Rotationsfria vektorfält, skalära potentialer," i *En första kurs i matematisk fysik*, Institutionen för fysik, Chalmers tekniska högskola, Göteborg, Sverige, 2018, kap. 8.1, s. 113. [Online]. Tillgänglig: [http://fy.chalmers.se/subatom/cforssen/vektorfalt/download/book\\_4ed.pdf](http://fy.chalmers.se/subatom/cforssen/vektorfalt/download/book_4ed.pdf). Hämtad 2021-03-04.
- [24] O. Guvench och A. D. M. Jr, *Comparison of Protein Force Fields for Molecular Dynamics Simulations*, ser. Methods in Molecular Biology. Department of Pharmaceutical Sciences, University of Maryland, Baltimore, USA: Humana Press, 2008, vol. 443, s. 63–88. DOI: 10.1007/978-1-59745-177-2\_4.
- [25] K. A. Feenstra, B. Hess och H. J. C. Berendsen, "Improving Efficiency of Large Time-Scale Molecular Dynamics Simulations of Hydrogen-Rich Systems," *Journal of Computational Chemistry*, årg. 20, nr 8, s. 786–798, Maj. 1999. DOI: 10.1002/(SICI)1096-987X(199906)20:8<786::AID-JCC5>3.0.CO;2-B.
- [26] J. Huang och A. D. M. Jr, "CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data," *Journal of Computa-*

- tional Chemistry*, årg. 34, nr 25, s. 2135–2145, Sep. 2013. DOI: 10.1002/jcc.23354.
- [27] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman och D. A. Case, "Development and testing of a general amber force field," *Journal of Computational Chemistry*, årg. 25, nr 9, s. 1157–1174, Jul. 2004. DOI: 110.1002/jcc.20035.
- [28] J. M. Lewis och S. L. S. Dhall, "10 - Optimization: steepest descent method," *Dynamic Data Assimilation*, s. 169–189, Dec. 2009. DOI: <https://doi.org/10.1017/CB09780511526480.011>.
- [29] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola och J. R. Haak, "Molecular dynamics with coupling to an external bath," *The Journal of Chemical Physics*, Aug. 1998. DOI: <https://doi.org/10.1063/1.448118>.
- [30] G. Bussi, D. Donadio och M. Parrinello, "Canonical sampling through velocity rescaling," *The Journal of Chemical Physics*, årg. 126, nr 1, Jan. 2007. DOI: 10.1063/1.2408420.
- [31] K. Reuter och J. Köfinger, "CADISHI: Fast parallel calculation of particle-pair distance histograms on CPUs and GPUs," *Computer Physics Communications*, årg. 236, s. 274–284, Mar. 2019. DOI: 10.1016/j.cpc.2018.10.018.
- [32] Marco Vassura et al., "Reconstruction of 3D Structures from Protein Contact Maps," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, årg. 5, nr 3, s. 357–367, Mar. 2008. DOI: 10.1109/TCBB.2008.27.
- [33] P. Hopkins, A. Fortini, A. J. Archer och M. Schmidt, "The van Hove distribution function for Brownian hard spheres: Dynamical test particle theory and computer simulations for bulk dynamics," *The Journal of Chemical Physics*, årg. 133, nr 22, Dec. 2010. DOI: 10.1063/1.3511719.
- [34] A. L. Thorneywork, D. G. A. L. Aarts, J. Horbach och R. P. A. Dullens, "On the Gaussian approximation in colloidal hard sphere fluids," *Soft Matter*, årg. 12, nr 18, s. 4129–4134, Apr. 2016. DOI: 10.1039/c5sm03049h.
- [35] H. Voigt och S. Hess, "Comparison of the intensity correlation function and the intermediate scattering function of fluids: a molecular dynamics study of the Siegert relation," *Physica A: Statistical Mechanics and its Applications*, årg. 202, nr 1-2, s. 145–164, Jan. 1994. DOI: 10.1016/0378-4371(94)90171-6.
- [36] G. D. G. Hinze och T. Basché, "Rotational Correlation Functions of Single Molecules," *Physical Review Letters*, årg. 93, nr 20, Nov. 2004. DOI: 10.4161/pri.1.1.4056.
- [37] G. B. Folland, "Legendre Polynomials," i *Fourier Analysis and Its Applications*, American Mathematical Society, 2009, kap. 6.2, s. 166.
- [38] B. Isralewitz, *Timeline: a VMD plugin for trajectory analysis*, Mars 2011, Tillgänglig: [https://www.ks.uiuc.edu/Training/Tutorials/science/timeline/tutorial\\_timeline.pdf](https://www.ks.uiuc.edu/Training/Tutorials/science/timeline/tutorial_timeline.pdf).
- [39] M. A. Cuendet och W. F. van Gunsteren, "On the calculation of velocity-dependent properties in molecular dynamics simulations using the leapfrog integration algorithm," *The Journal of Chemical Physics*, årg. 127, nr 18, Dec. 2007. DOI: 10.1063/1.2779878.
- [40] S.-Y. Sheu, D.-Y. Yang, H. L. Selzle och E. W. Schlag, "Energetics of hydrogen bonds in peptides," *PNAS*, årg. 100, nr 22, s. 12 683–12 687, Okt. 2003. DOI: 10.1073/pnas.2133366100.

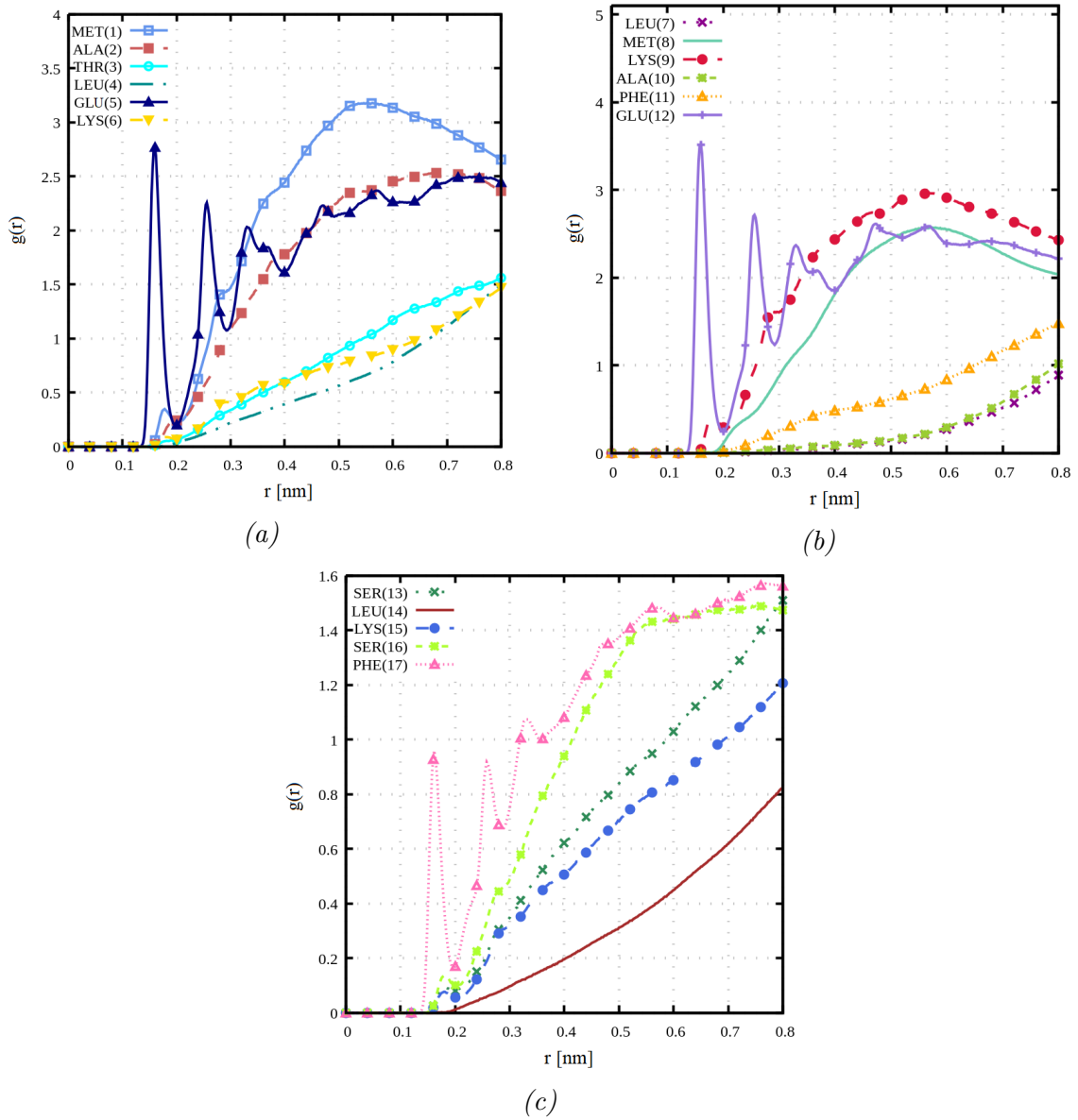
# A Bilagor

## A.1 Molekylär RDF mellan aminosyror och sackarider

### A.1.1 A-system

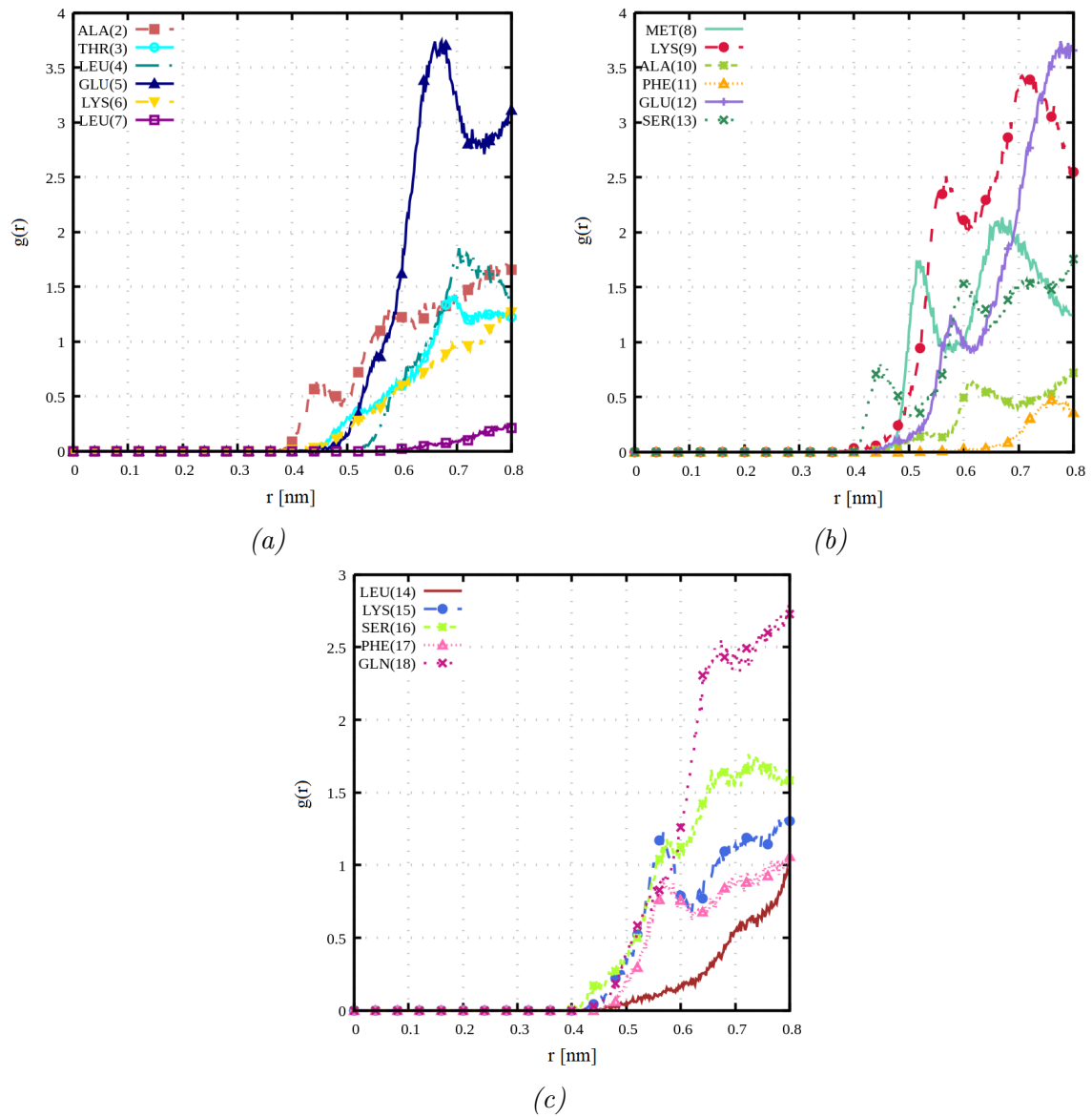


Figur 11:  $g(r)$  mellan masscentrum hos aminosyror i 2LD2 och sackaros, som funktion av avståndet  $r$ , för system A1 (4800 vatten). Figur a) visar aminosyror 1-6, b) aminosyror 7-12 och c) aminosyror 13-17.

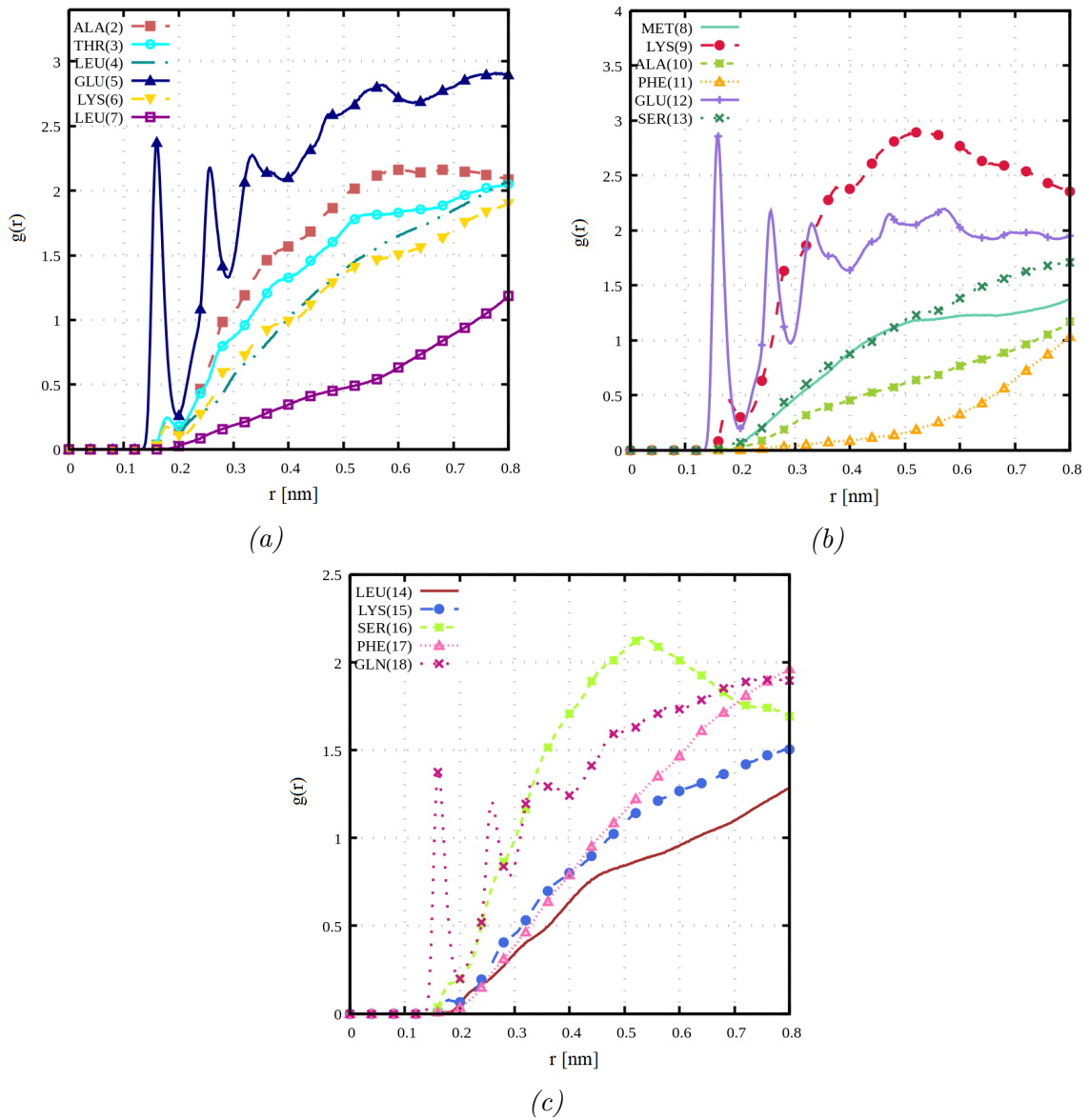


Figur 12:  $g(r)$  mellan masscentrum hos aminosyror i 2LD2 och trehalos, som funktion av avståndet  $r$ , för system A2 (4800 vatten). Figur a) visar aminosyror 1-6, b) aminosyror 7-12 och c) aminosyror 13-17.

### A.1.2 B-system

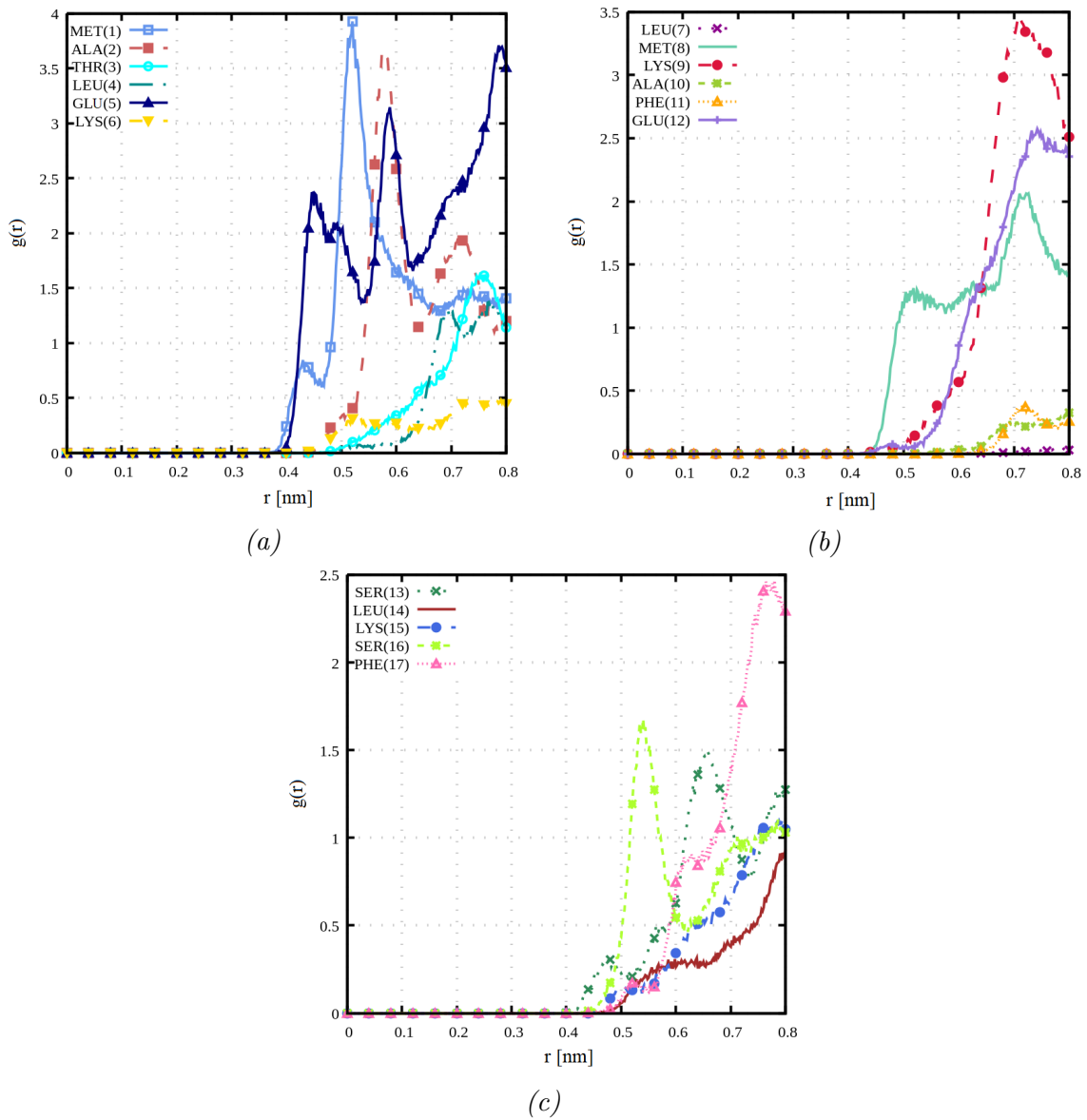


Figur 13:  $g(r)$  mellan masscentrum hos aminosyror i 6N8C och sackaros, som funktion av avståndet  $r$ , för system B1 (4800 vatten). Figur a) visar aminosyror 2-7, b) aminosyror 8-13 och c) aminosyror 14-18.

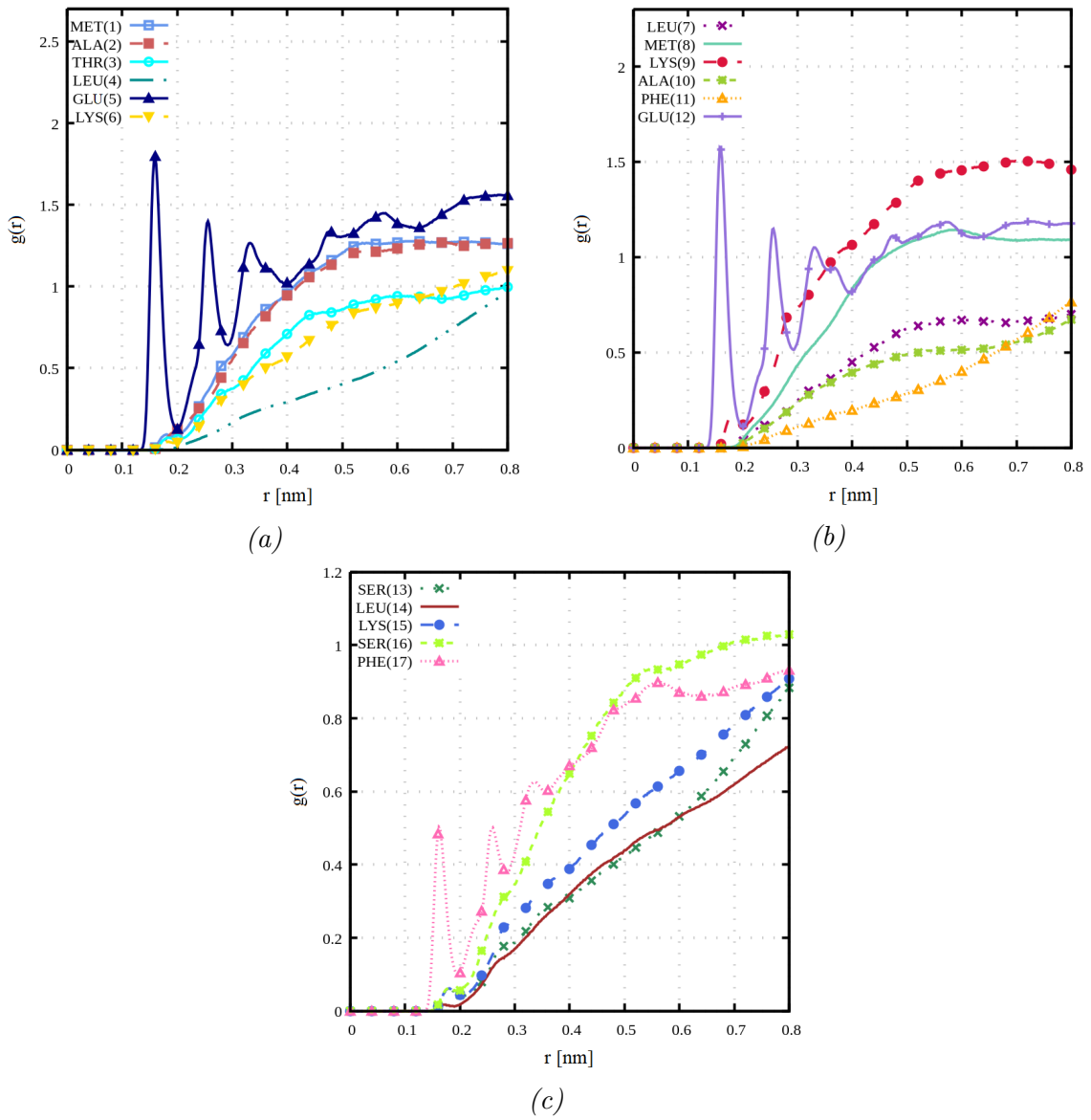


Figur 14:  $g(r)$  mellan masscentrum hos aminosyror i 6N8C och trehalos, som funktion av avståndet  $r$ , för system B2 (4800 vatten). Figur a) visar aminosyror 2-7, b) aminosyror 8-13 och c) aminosyror 14-18.

### A.1.3 C-system

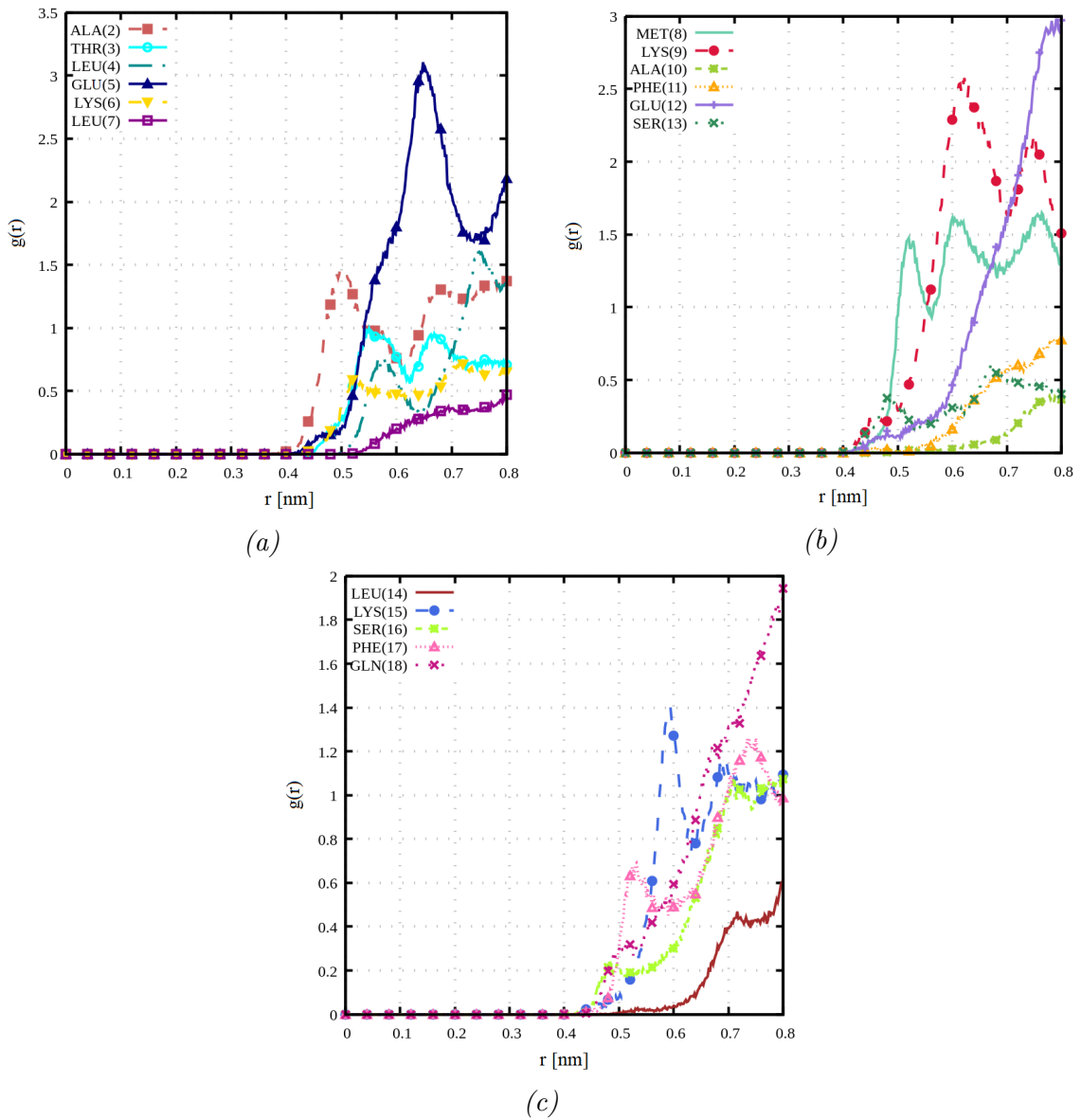


Figur 15:  $g(r)$  mellan masscentrum hos aminosyror i 2LD2 och sackaros, som funktion av avståndet  $r$ , för system C1 (2400 vatten). Figur a) visar aminosyror 1-6, b) aminosyror 7-12 och c) aminosyror 13-17.

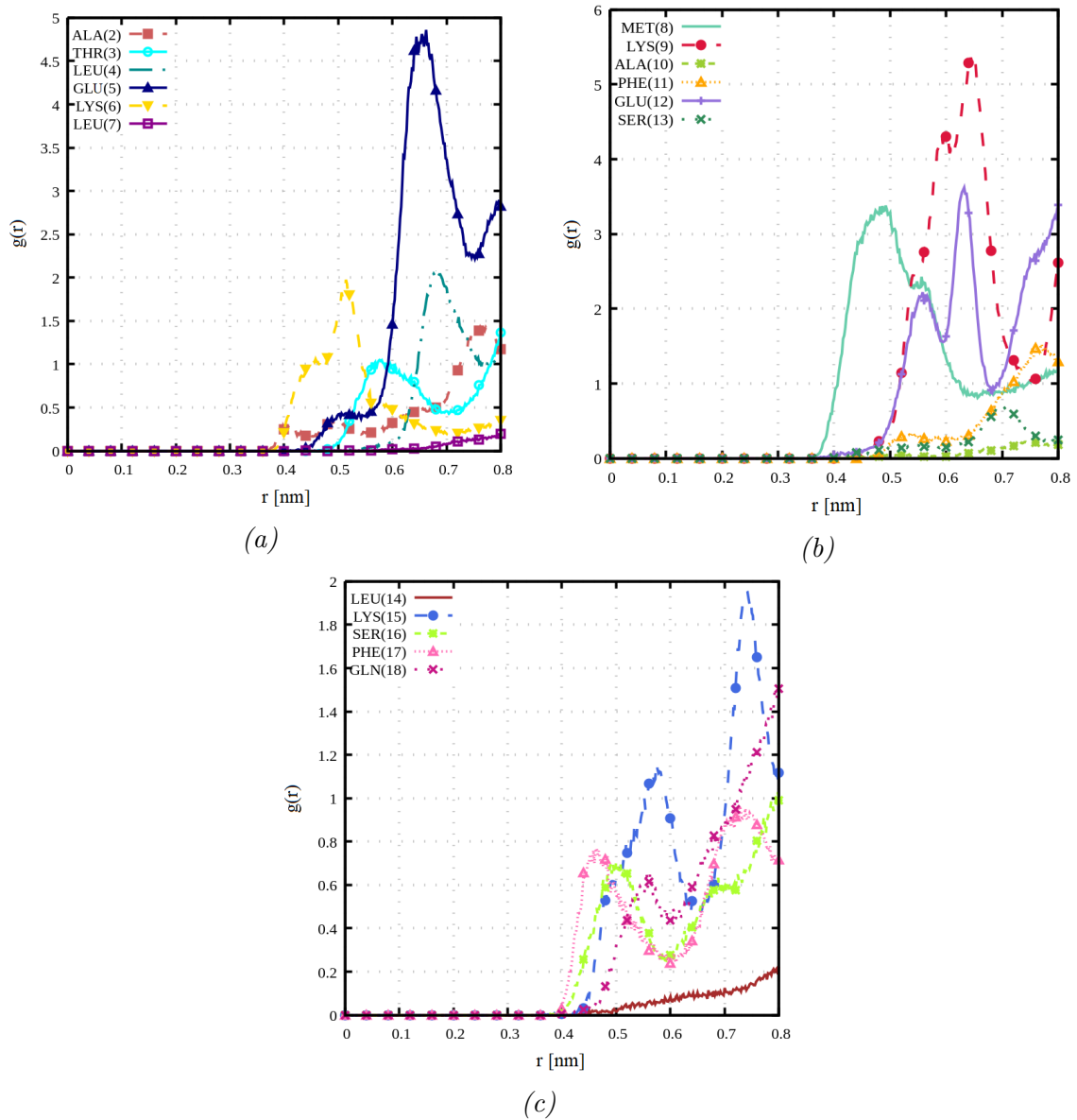


Figur 16:  $g(r)$  mellan masscentrum hos aminosyror i 2LD2 och trehalos, som funktion av avståndet  $r$ , för system C2 (2400 vatten). Figur a) visar aminosyror 1-6, b) aminosyror 7-12 och c) aminosyror 13-17.

### A.1.4 D-system

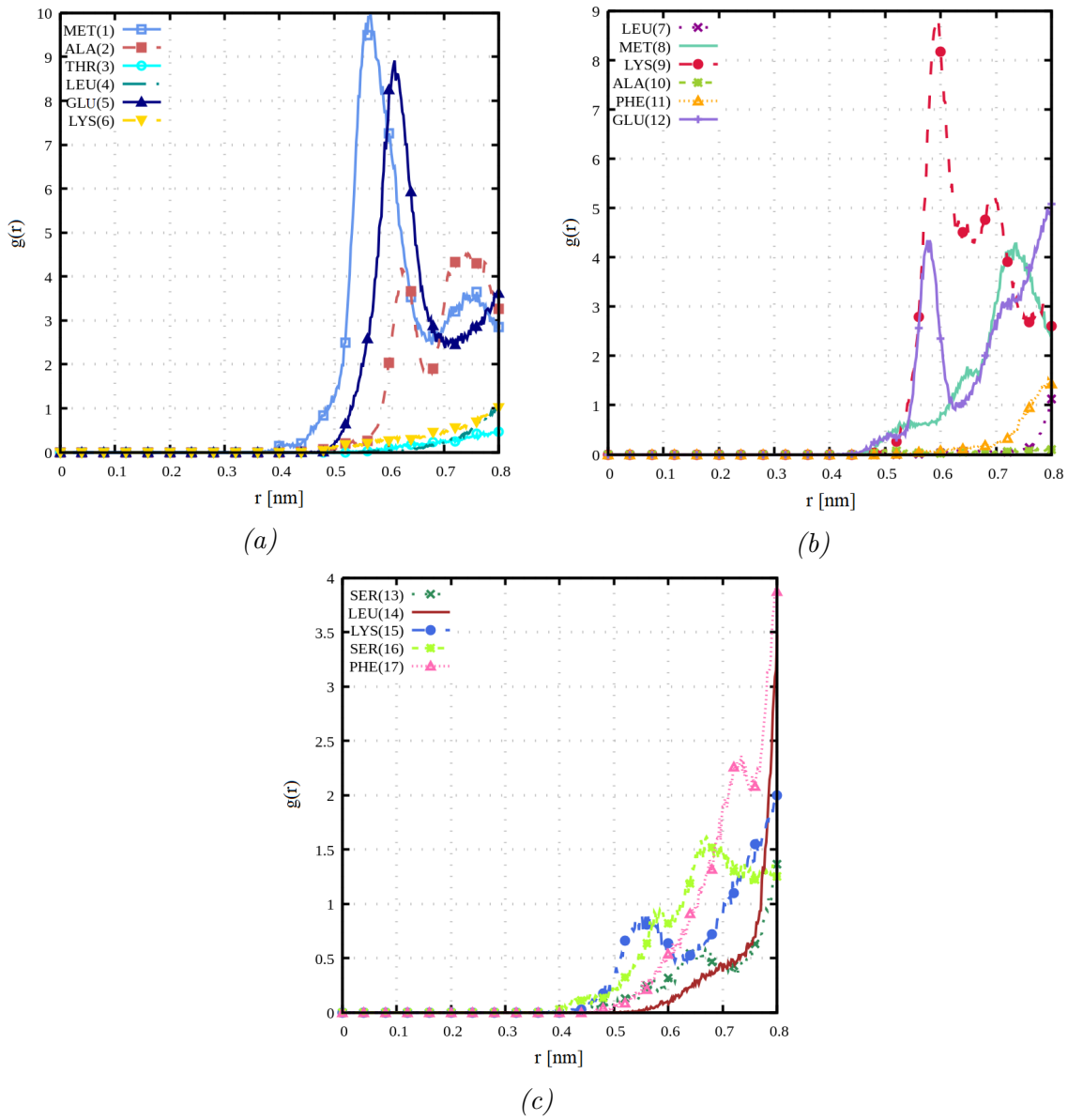


Figur 17:  $g(r)$  mellan masscentrum hos aminosyror i 6N8C och sackaros, som funktion av avståndet  $r$ , för system D1 (2400 vatten). Figur a) visar aminosyror 2-7, b) aminosyror 8-13 och c) aminosyror 14-18.

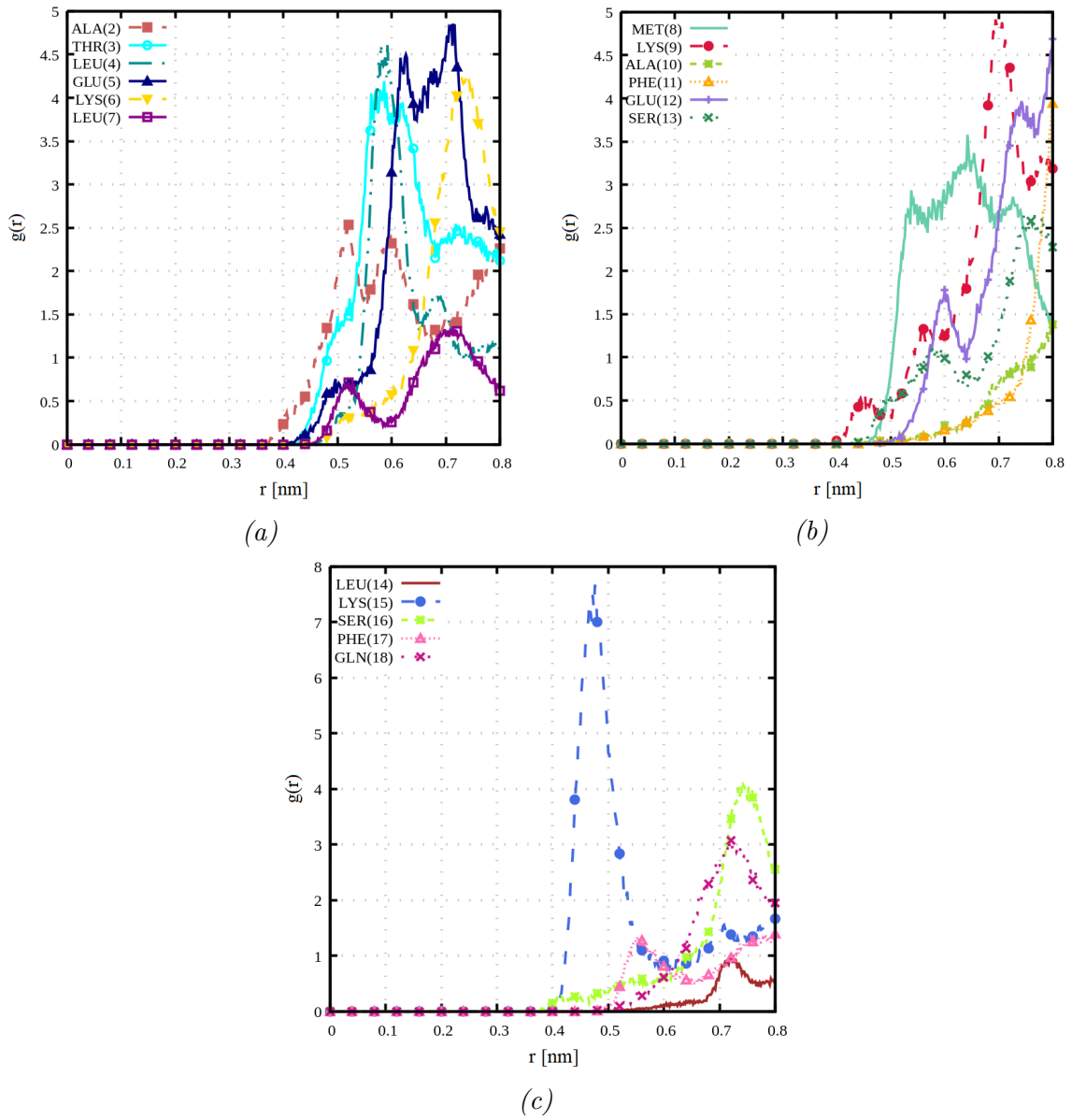


Figur 18: RDF för aminosyror i 6N8C med trehalos, system D2 (2400 vatten), där figur a) visar aminosyror 1-6, b) aminosyror 7-12 och c) aminosyror 13-17.  $g(r)$  mellan masscentrum hos aminosyror i 6N8C och trehalos, som funktion av avståndet  $r$ , för system D2 (2400 vatten). Figur a) visar aminosyror 2-7, b) aminosyror 8-13 och c) aminosyror 14-18.

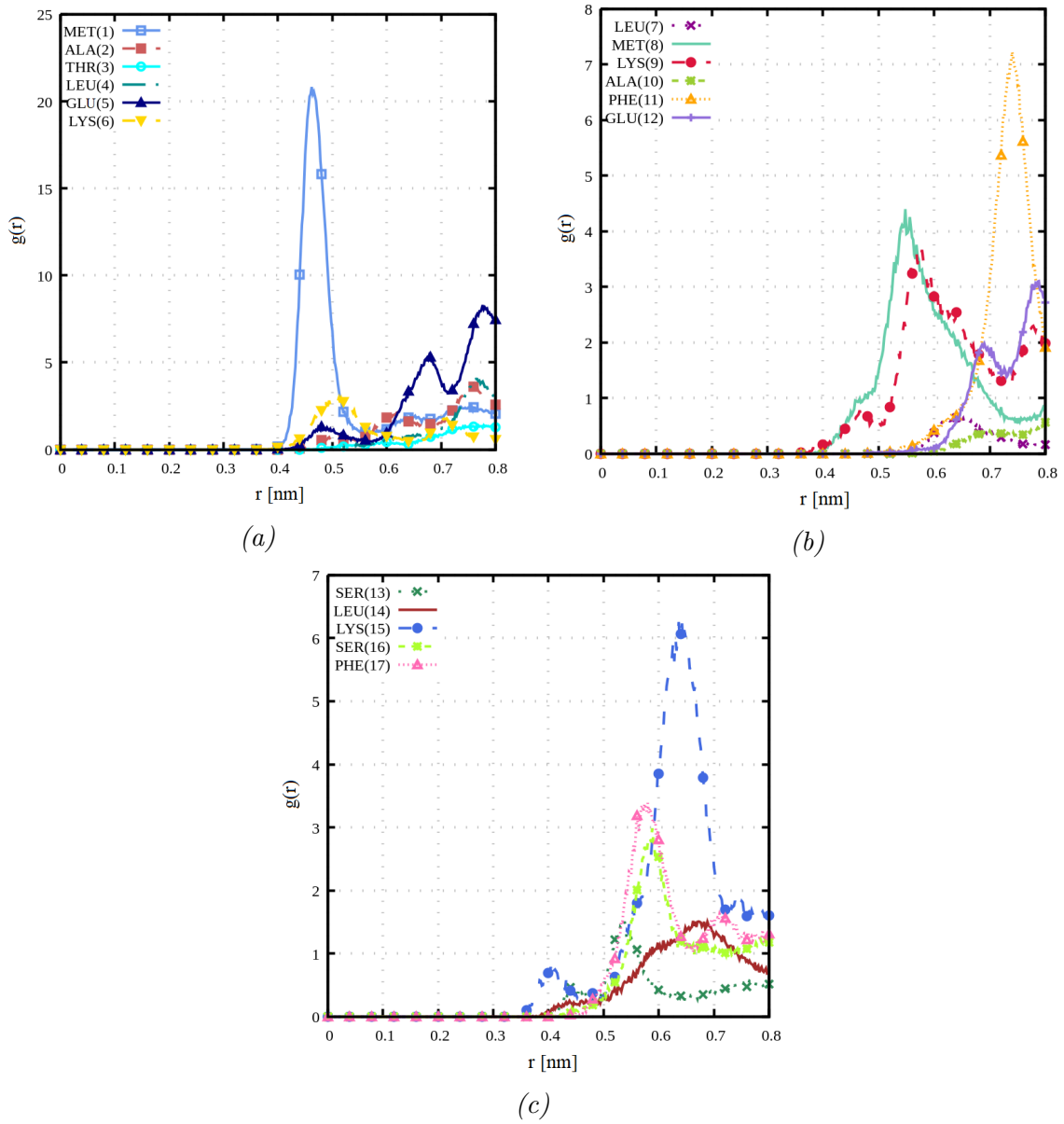
### A.1.5 E-system



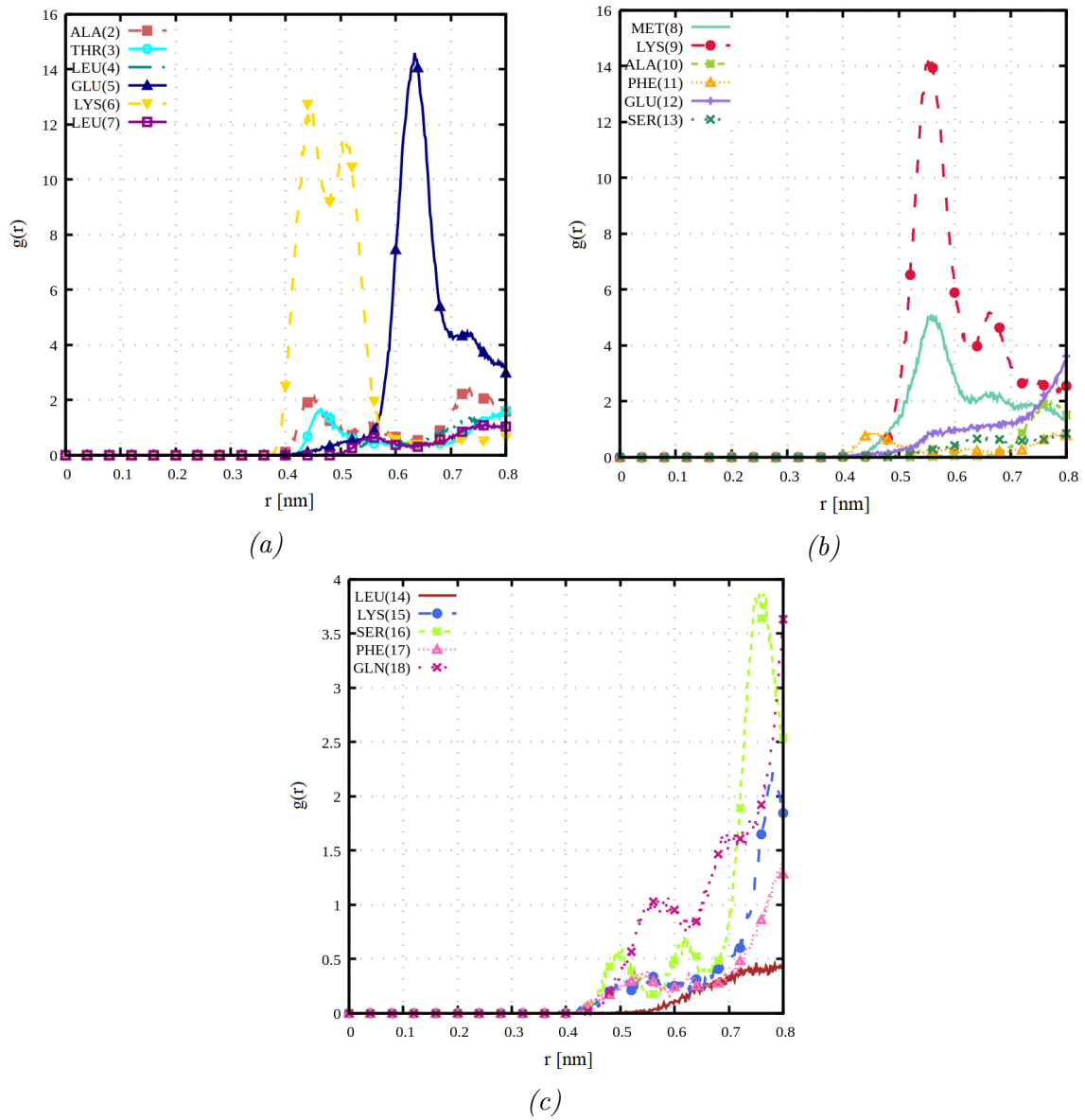
Figur 19:  $g(r)$  mellan masscentrum hos aminosyror i 2LD2 och sackaros, som funktion av avståndet  $r$ , för system E1 (båda peptider, 4800 vatten). Figur a) visar aminosyror 1-6, b) aminosyror 7-12 och c) aminosyror 13-17.



Figur 20:  $g(r)$  mellan masscentrum hos aminosyror i 2LD2 och sackaros, som funktion av avståndet  $r$ , för system E1 (båda peptider, 4800 vatten). Figur a) visar aminosyror 1-6, b) aminosyror 7-12 och c) aminosyror 13-17.



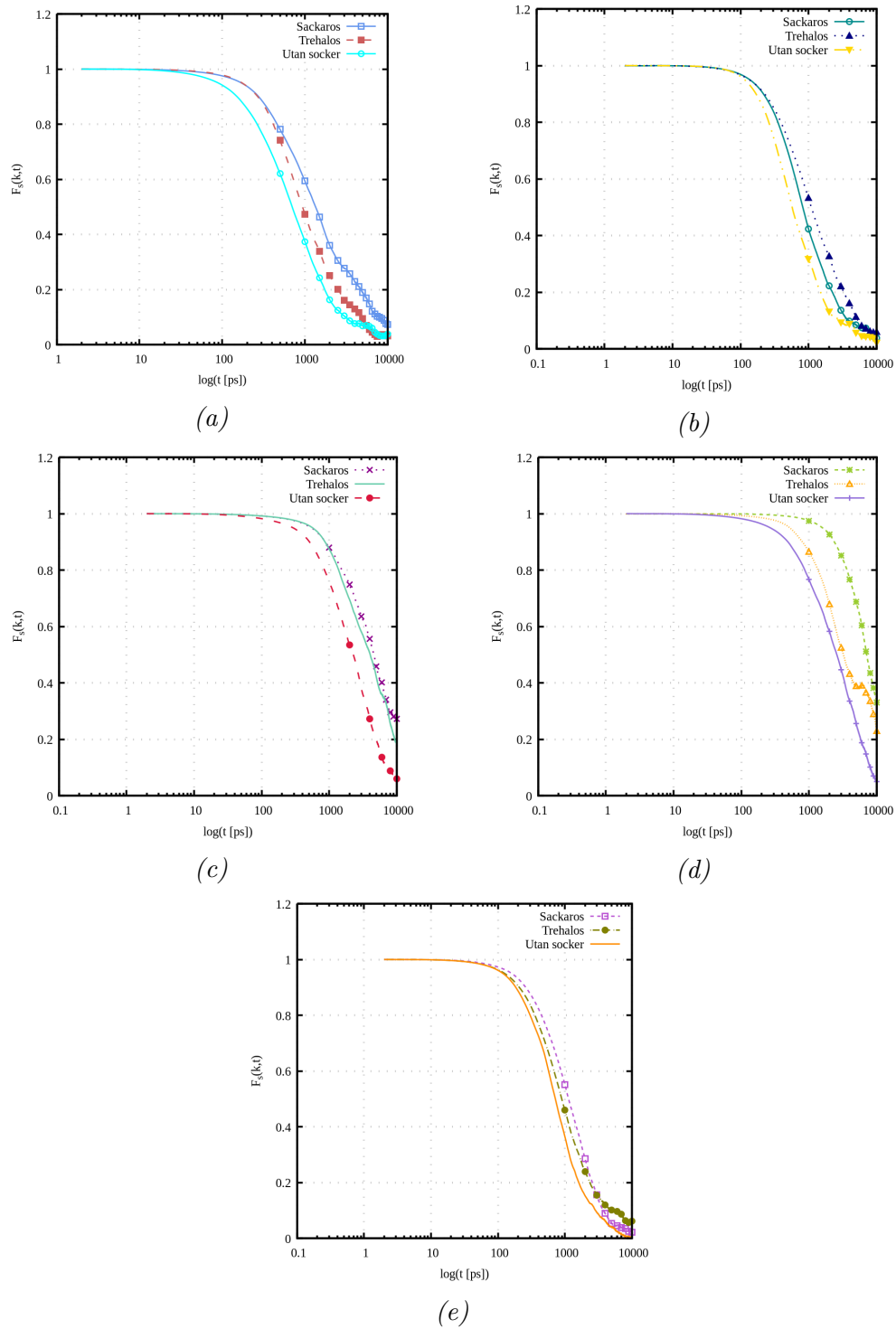
Figur 21:  $g(r)$  mellan masscentrum hos aminosyror i 6N8C och trehalos, som funktion av avståndet  $r$ , för system E2 (båda peptider, 4800 vatten). Figur a) visar aminosyror 2-7, b) aminosyror 8-13 och c) aminosyror 14-18.



Figur 22:  $g(r)$  mellan masscentrum hos aminosyror i 6N8C och trehalos, som funktion av avståndet  $r$ , för system E2 (båda peptider, 4800 vatten). Figur a) visar aminosyror 2-7, b) aminosyror 8-13 och c) aminosyror 14-18.

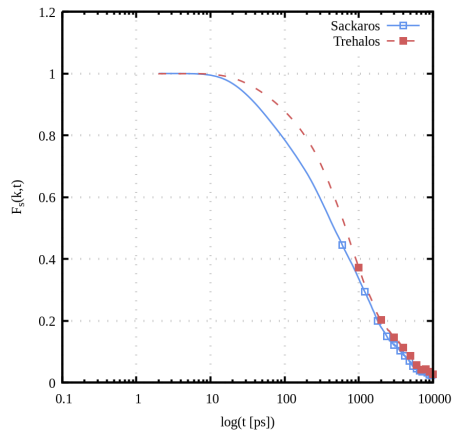
## A.2 Mellanliggande spridningsfunktion för partiklars självrörelse

### A.2.1 Peptider

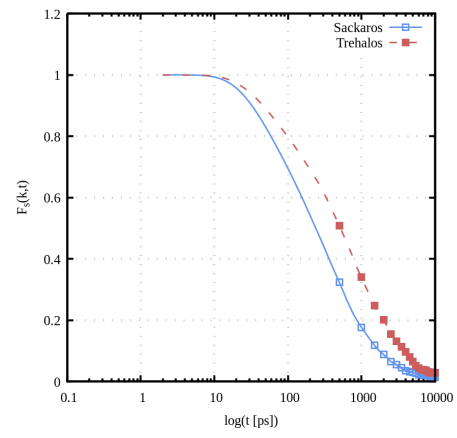


Figur 23: SISF för peptider, system A-E, där figur a) visar system A, b) system B, c) system C, d) system D och e) system E.

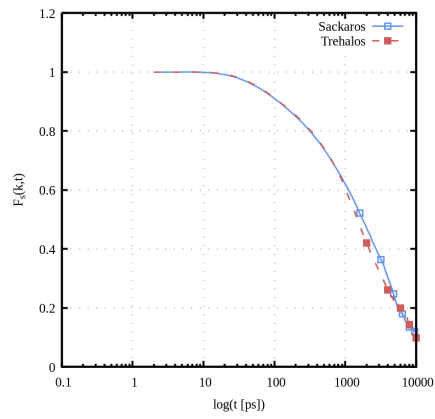
## A.2.2 Sackarider



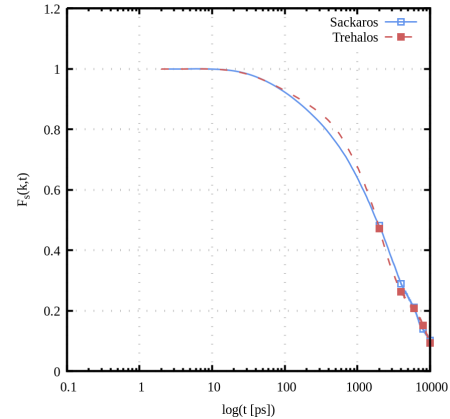
(a)



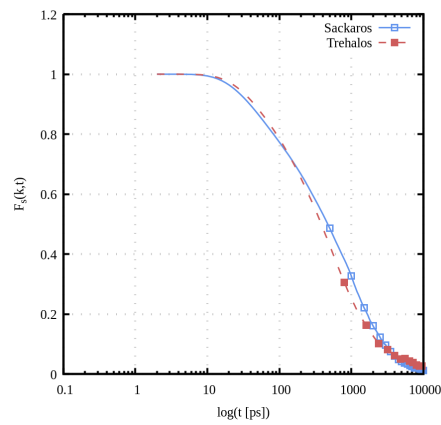
(b)



(c)



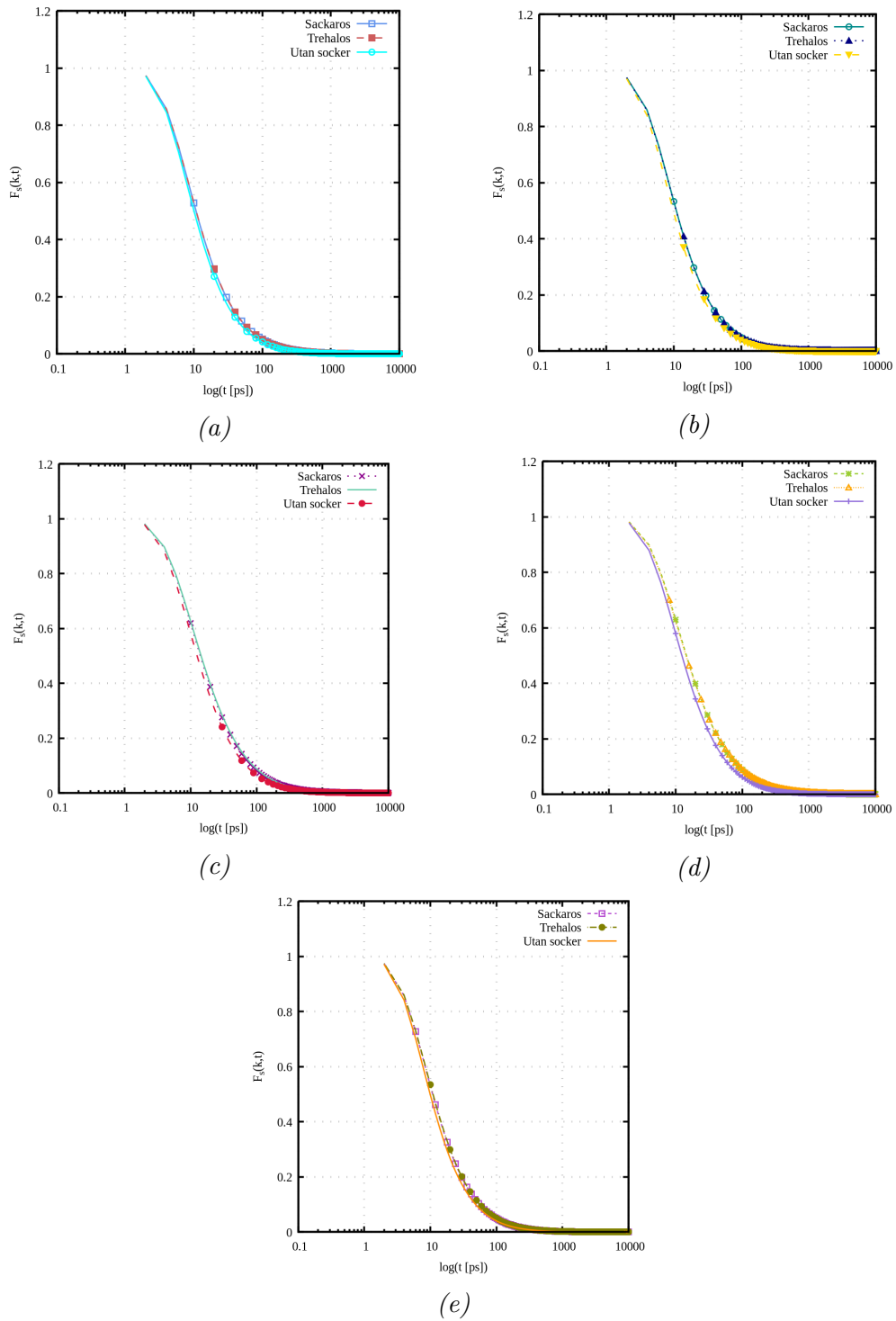
(d)



(e)

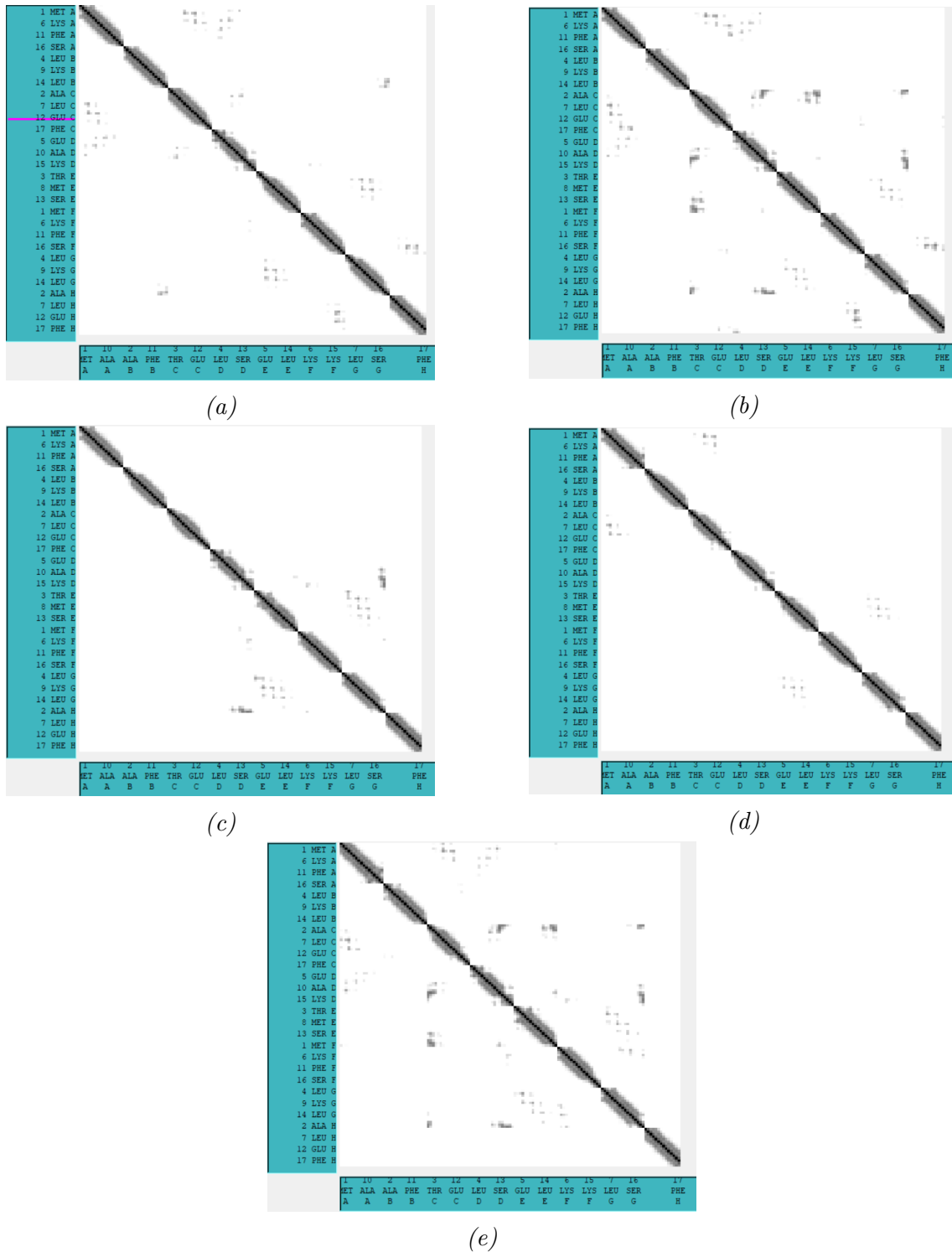
Figur 24: SISF för sackarider, system A-E, där figur a) visar system A, b) system B, c) system C, d) system D och e) system E.

### A.2.3 Vatten

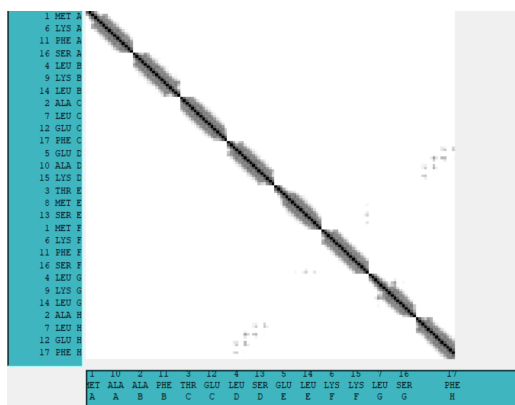


Figur 25: SISF för vatten, system A-E, där figur a) visar system A, b) system B, c) system C, d) system D och e) system E.

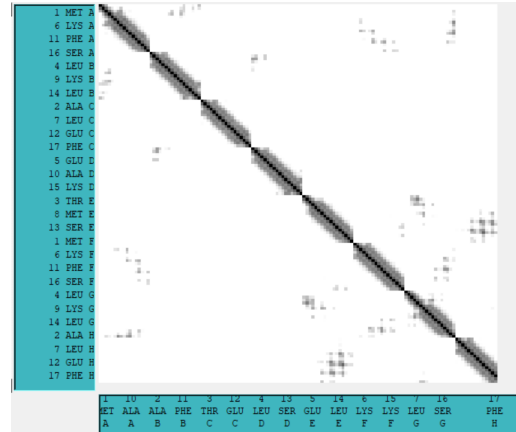
### A.3 Contact Maps



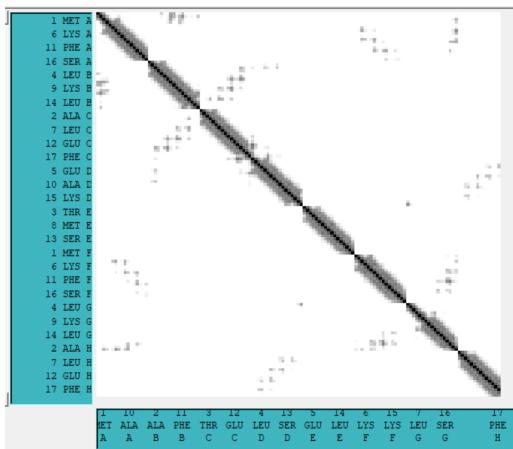
Figur 26: Contact maps för 2LD2 med sackaros och 4800 vatten (A1). Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



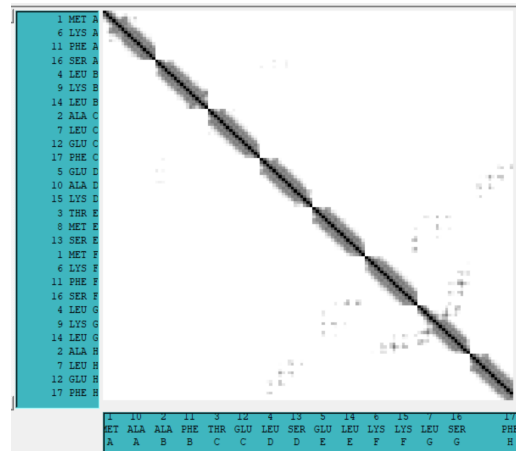
(a)



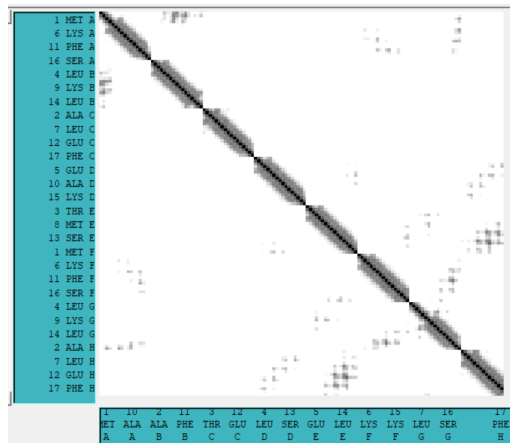
(b)



(c)

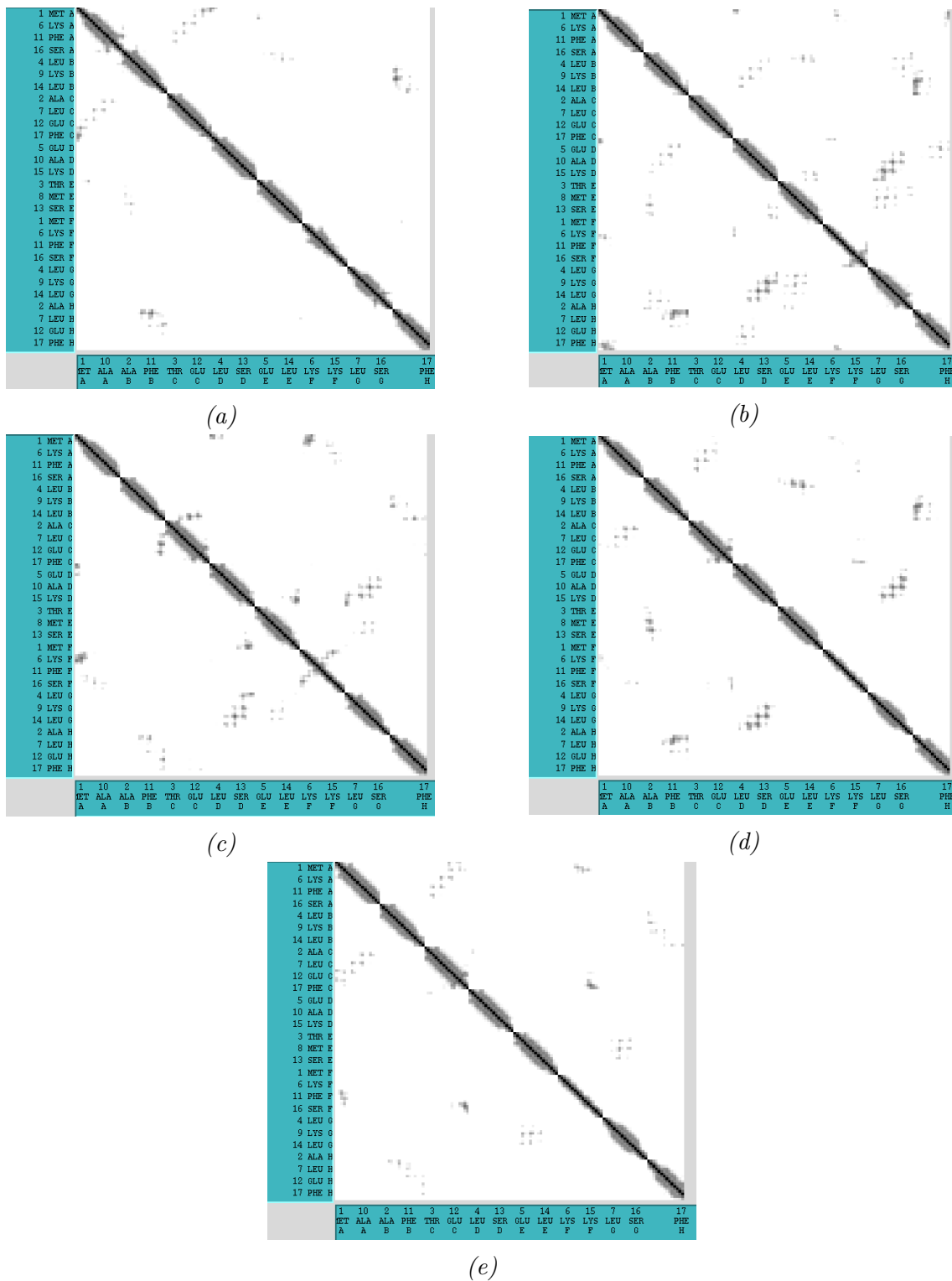


(d)

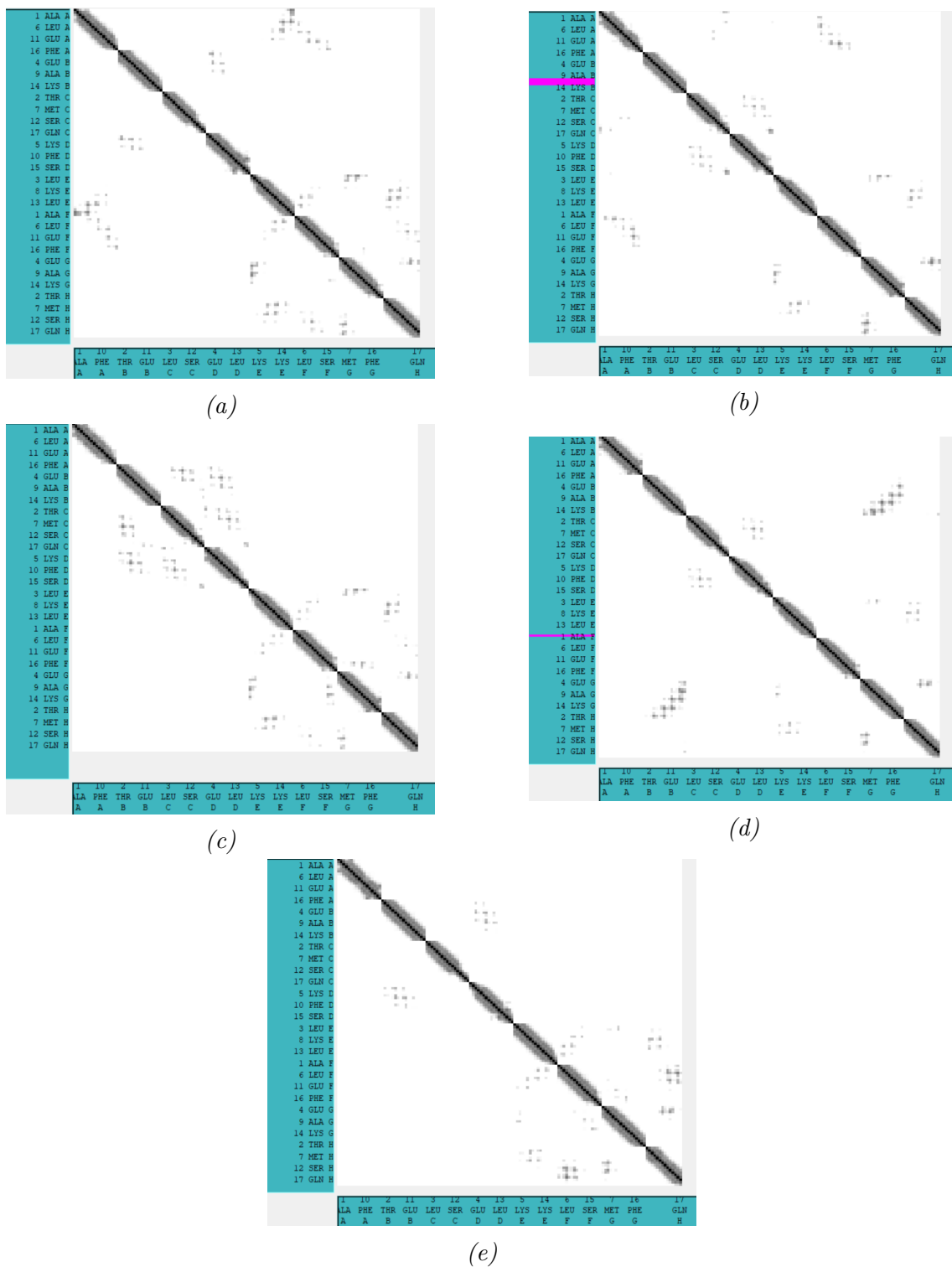


(e)

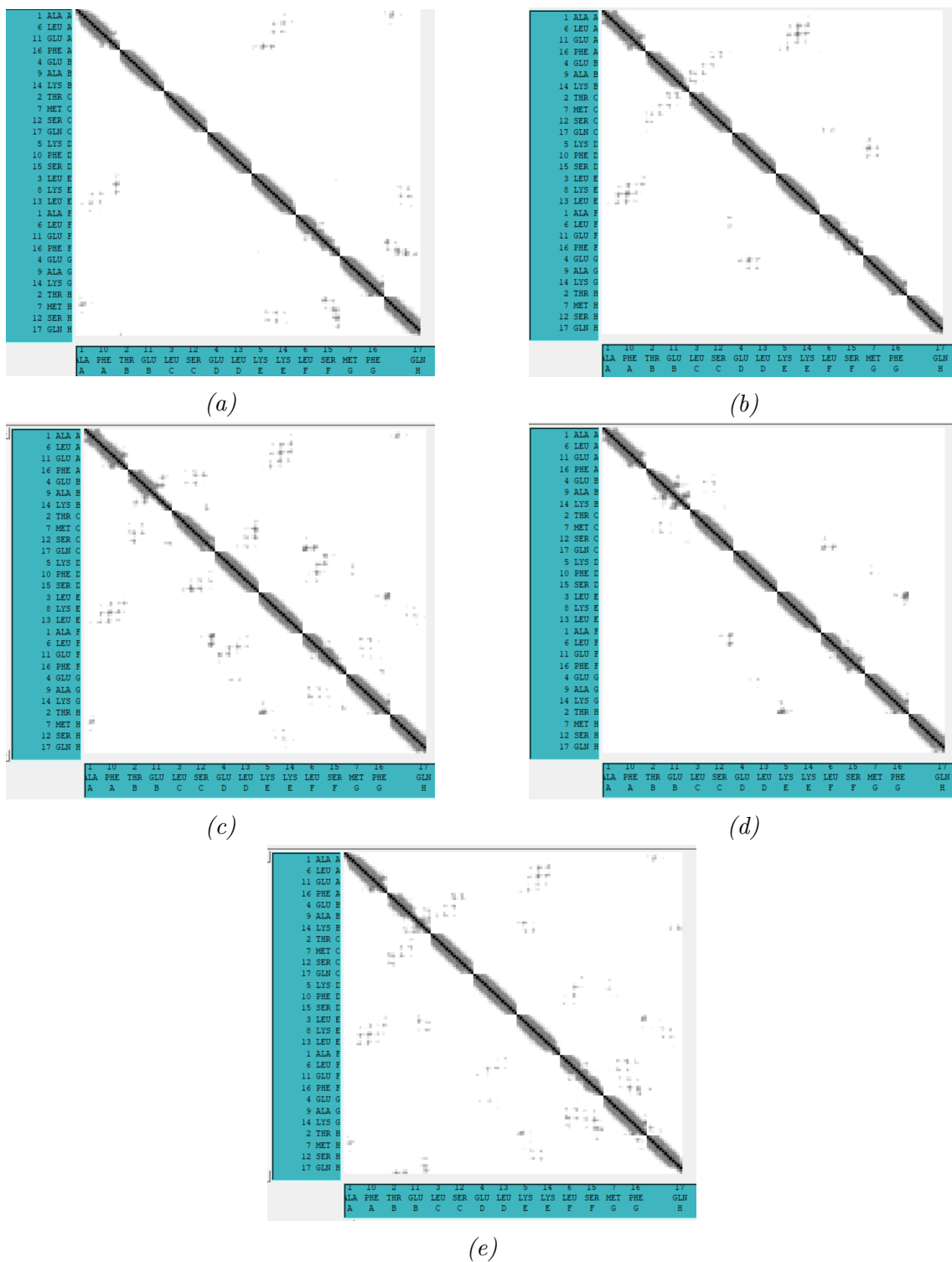
Figur 27: Contact maps för 2LD2 med trehalos och 4800 vatten (A2). Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



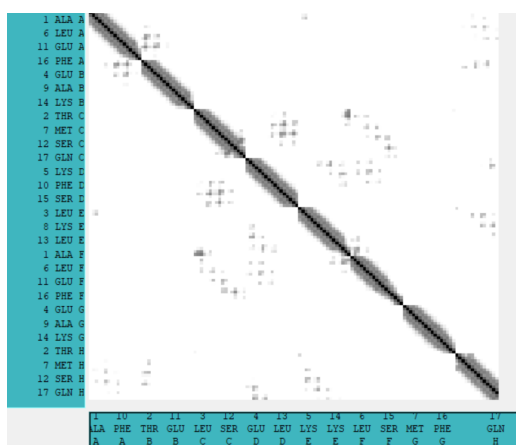
Figur 28: Contact maps för 2LD2 utan sockarider, med 4800 vatten (A3). Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



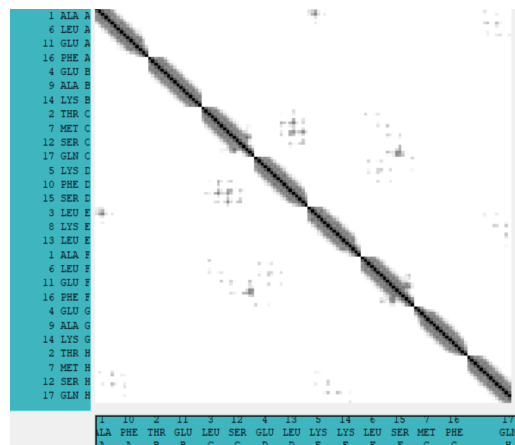
Figur 29: Contact maps för 6N8C med sackaros och 4800 vatten (B1). Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



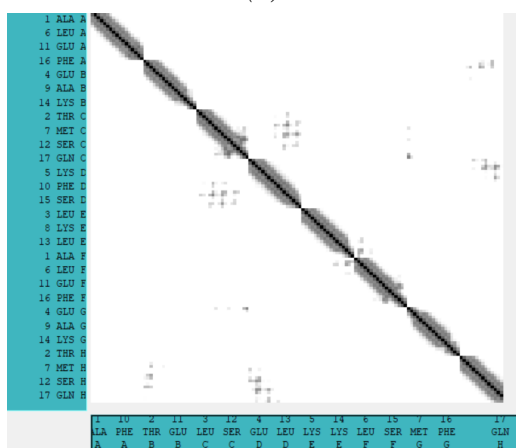
Figur 30: Contact maps för 6N8C med trehalos och 4800 vatten (B2). Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



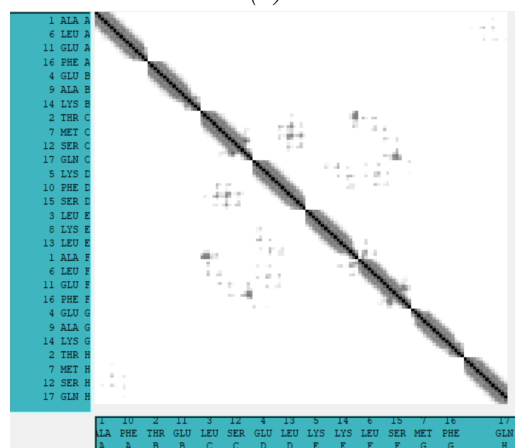
(a)



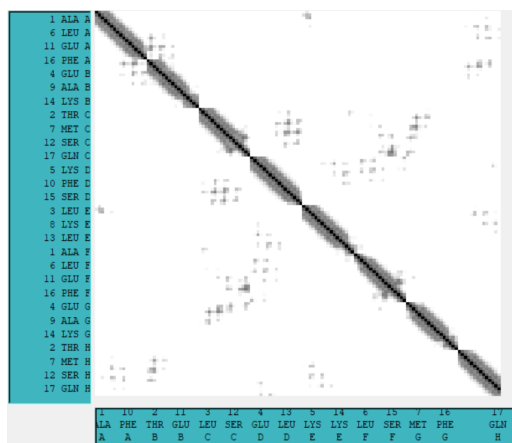
(b)



(c)

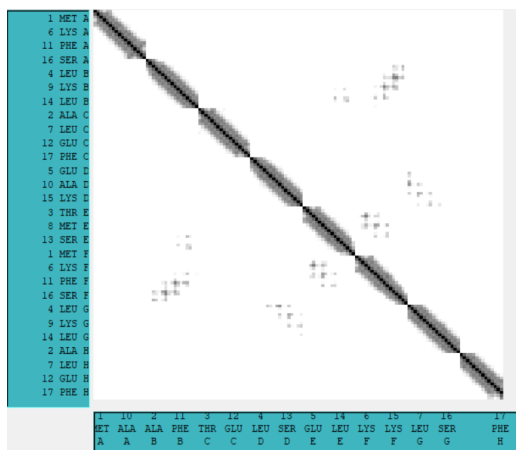


(d)

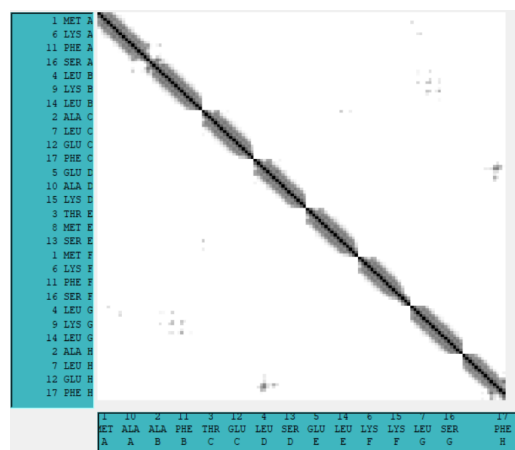


(e)

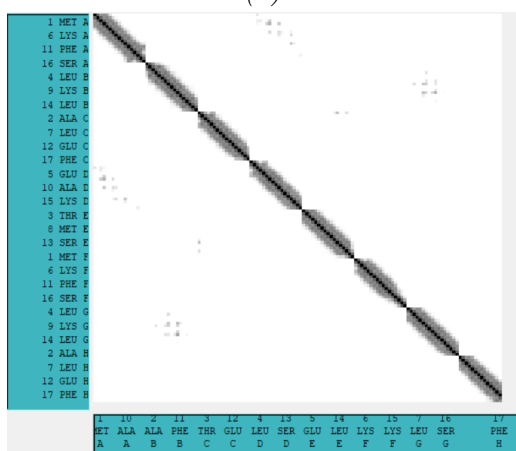
Figur 31: Contact maps för 6N8C utan socker, med 4800 vatten (B3). Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



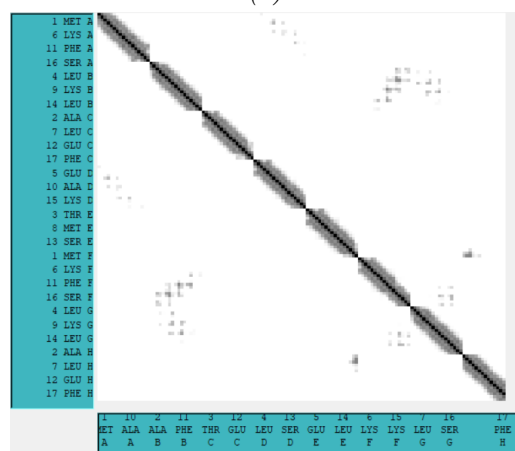
(a)



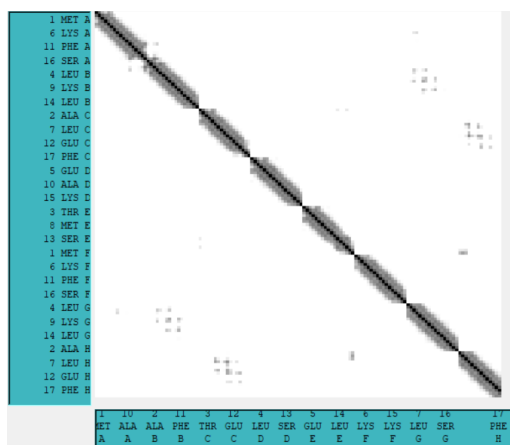
(b)



(c)

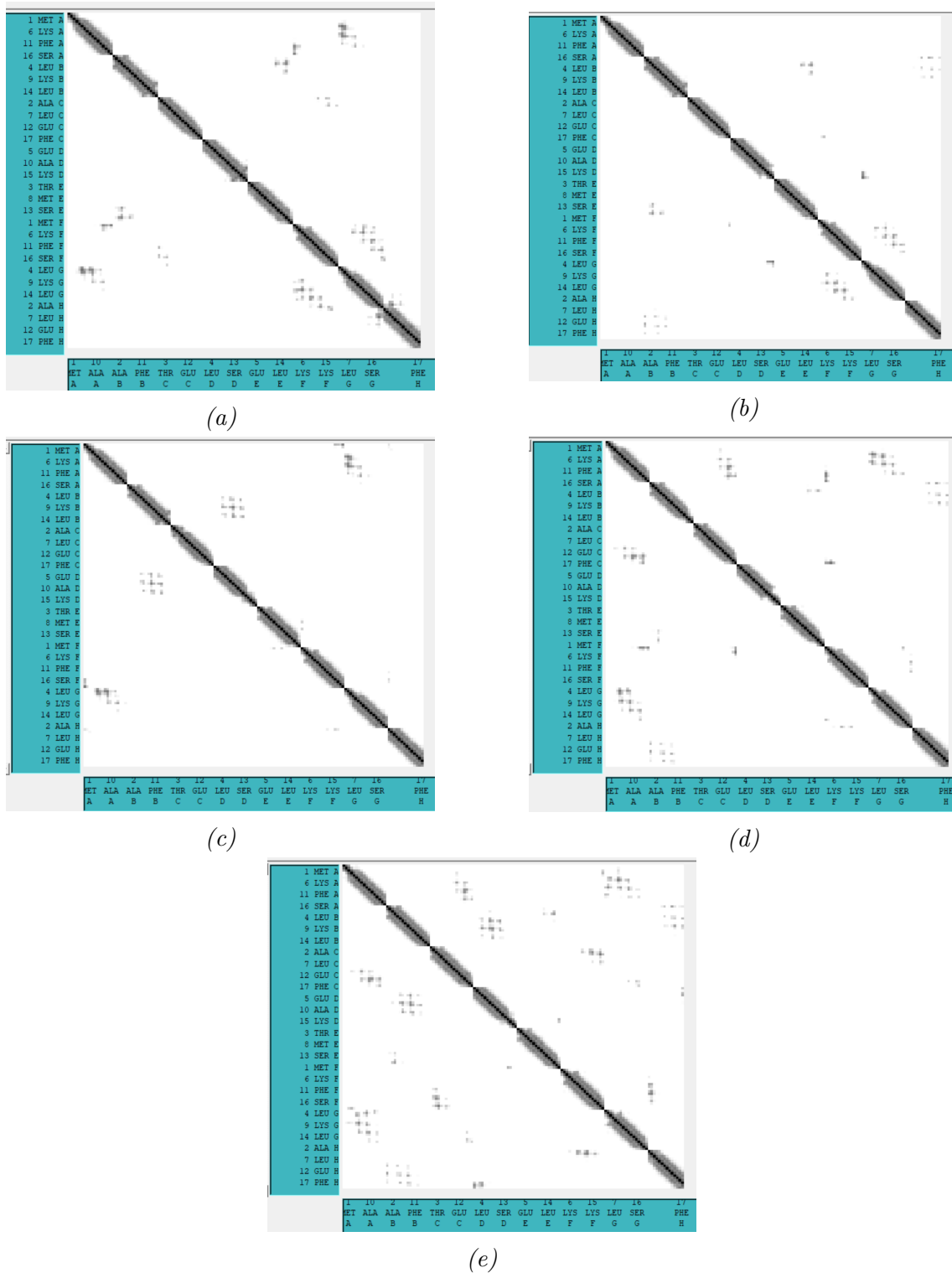


(d)

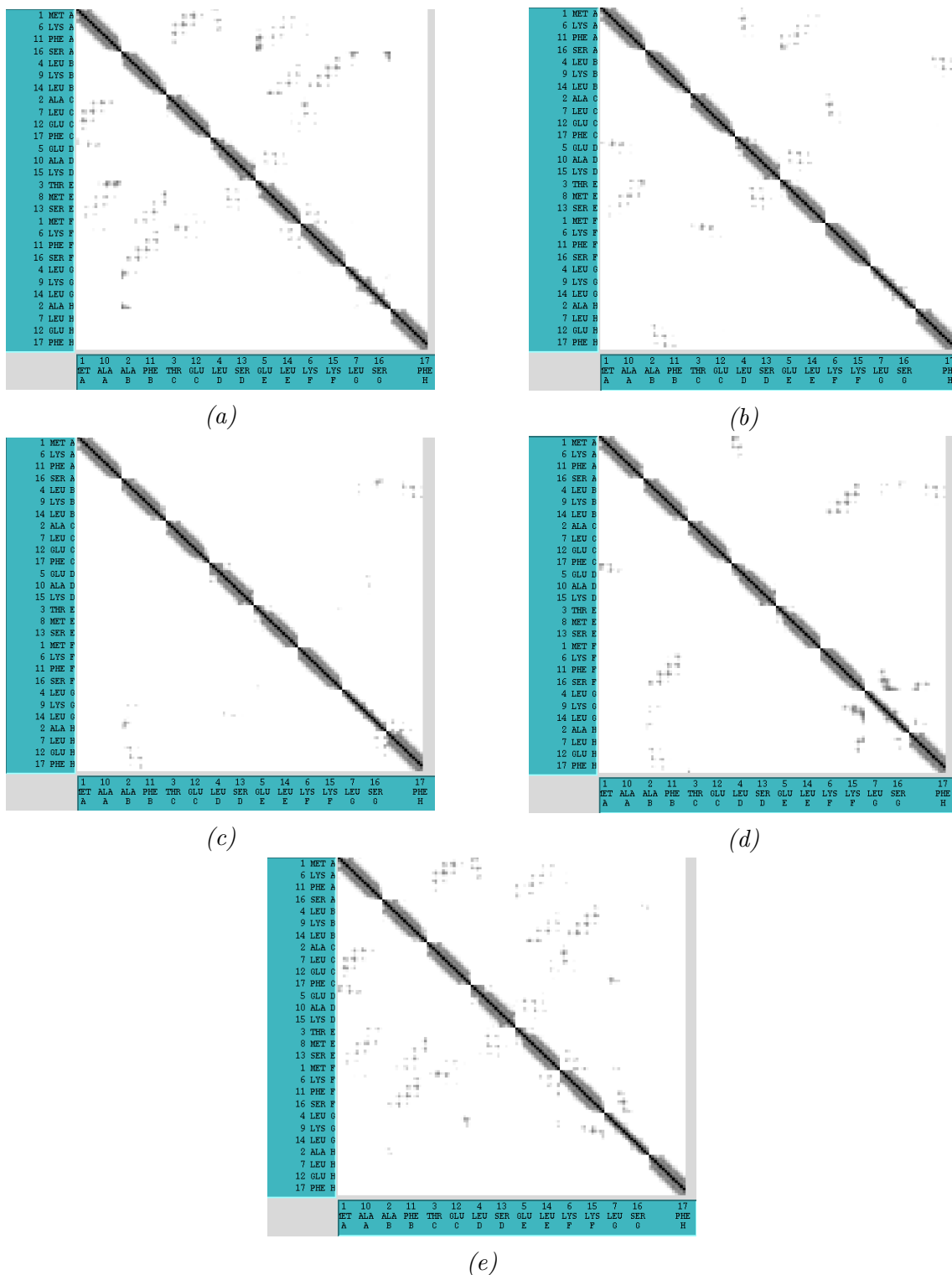


(e)

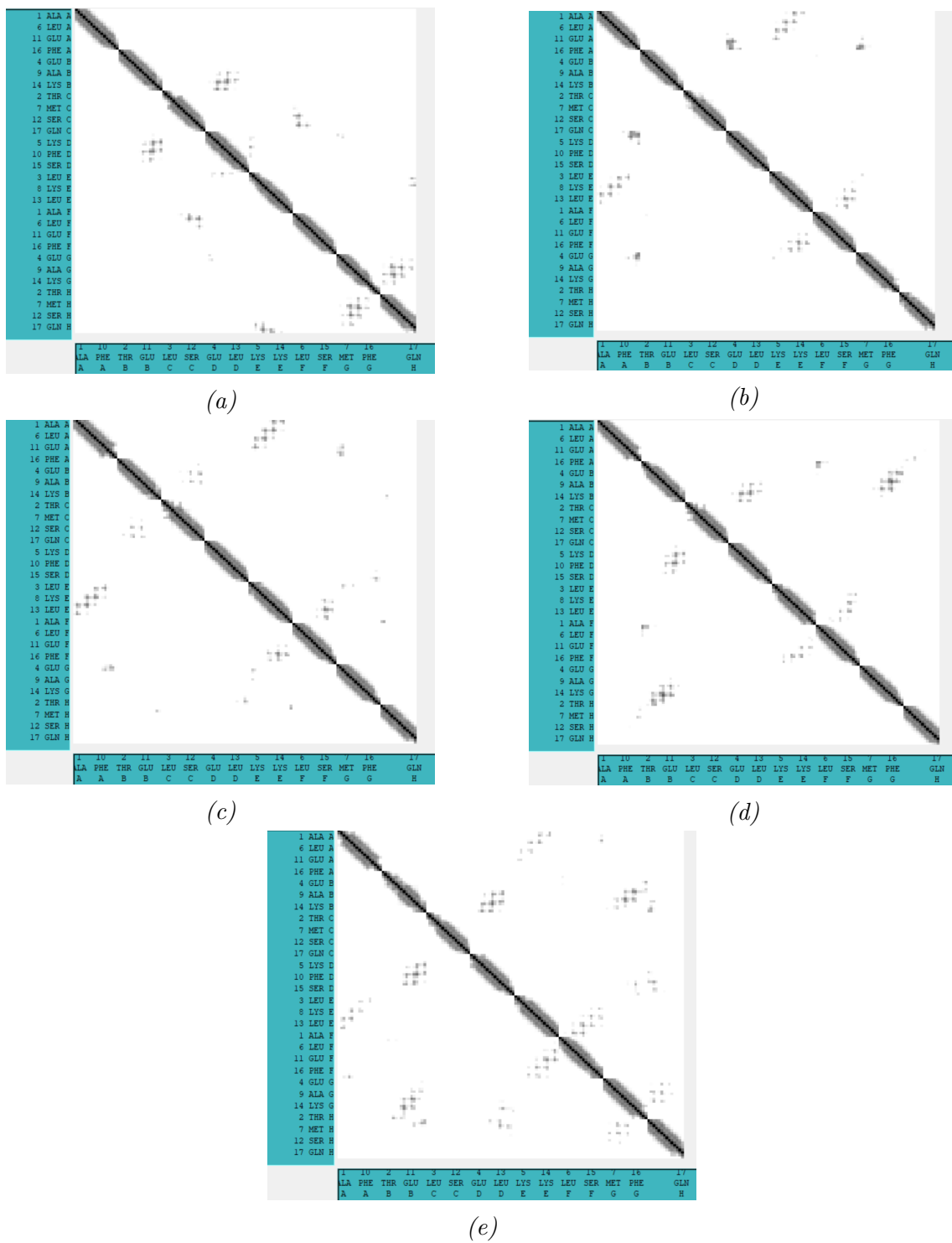
Figur 32: Contact maps för 2LD2 med sackaros och 2400 vatten (C1). Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



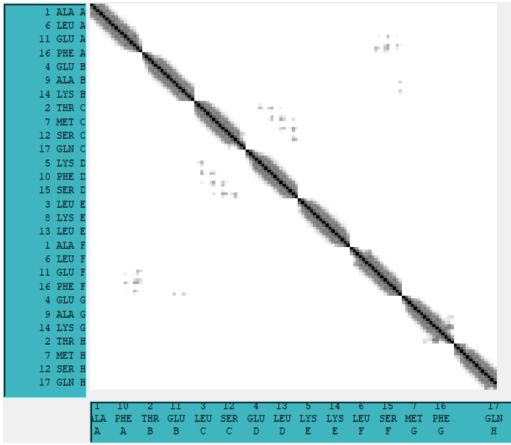
Figur 33: Contact maps för 2LD2 med trehalos och 2400 vatten (C2). Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



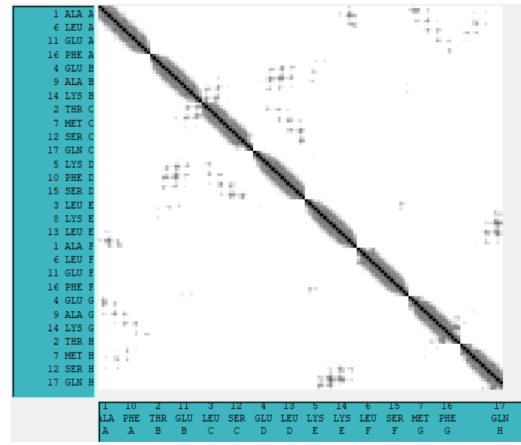
Figur 34: Contact maps för 2LD2, utan sackarider med 2400 vatten (C3). Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



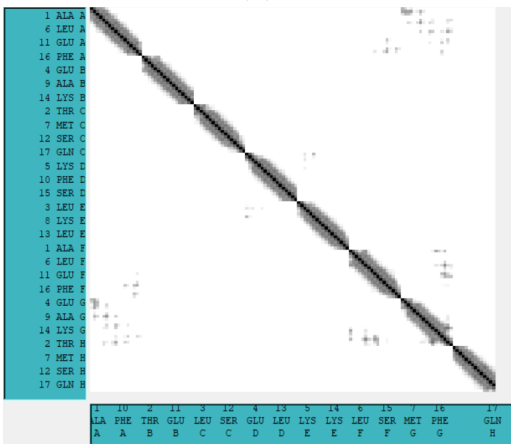
Figur 35: Contact maps för 6N8C med sackaros och 2400 vatten (D1). Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



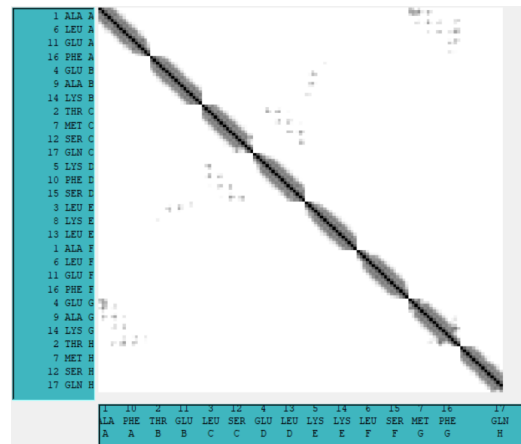
(a)



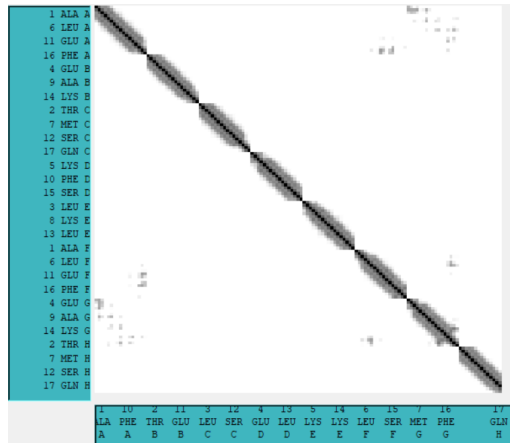
(b)



(c)

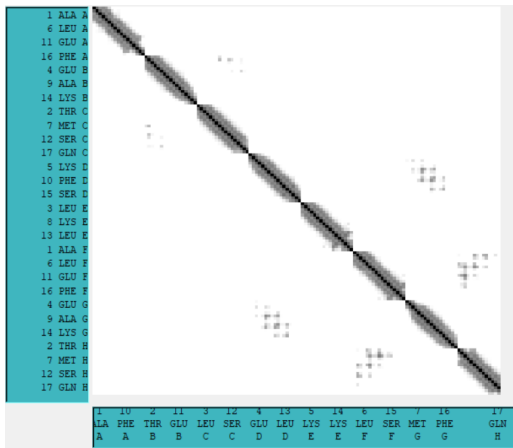


(d)

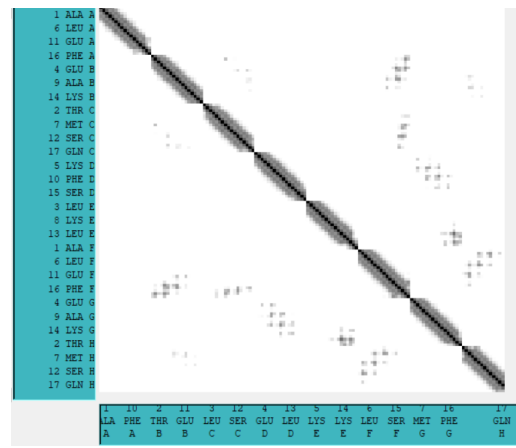


(e)

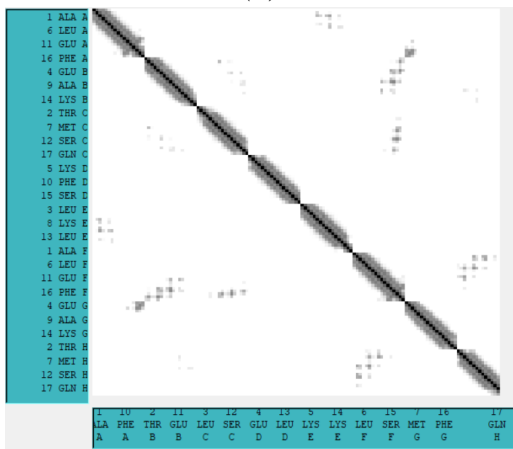
Figur 36: Contact maps för 6N8C med trehalos och 2400 vatten (D2). Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



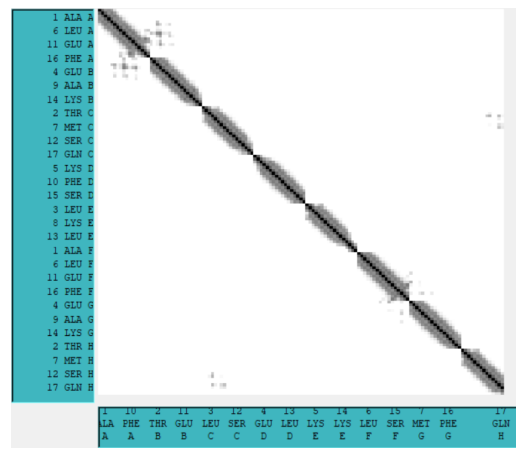
(a)



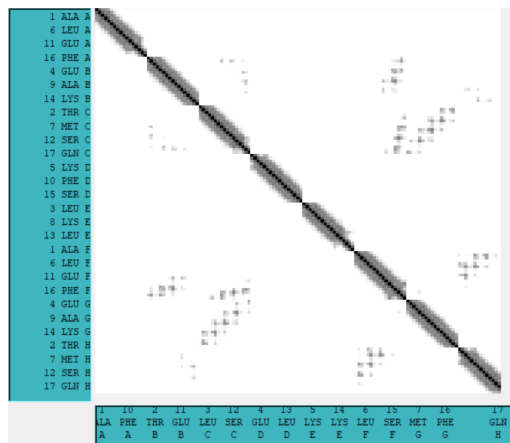
(b)



(c)

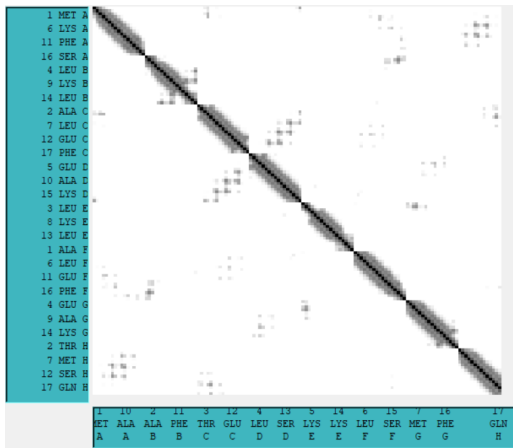


(d)

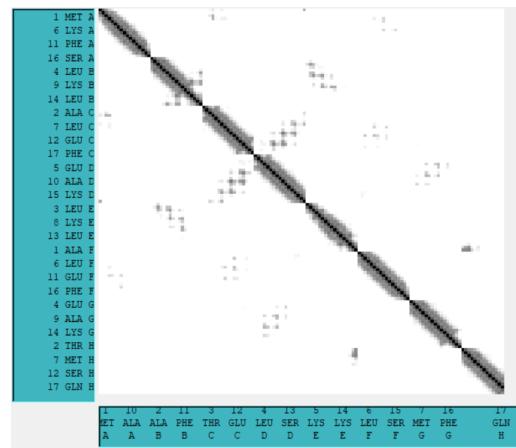


(e)

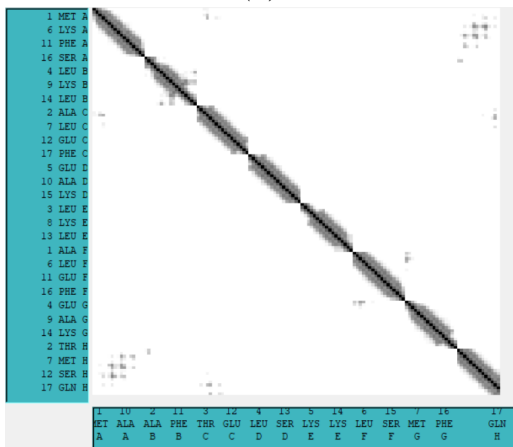
Figur 37: Contact maps för 6N8C utan socker, med 2400 vatten (D3). Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



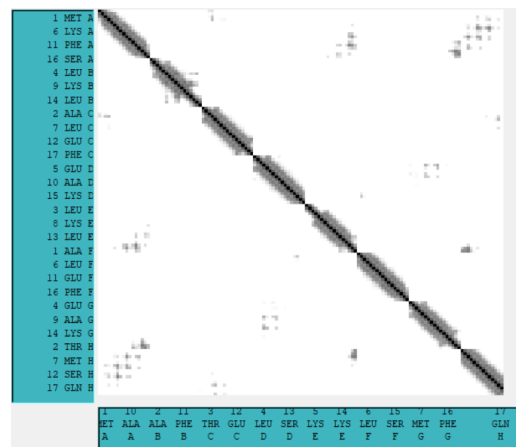
(a)



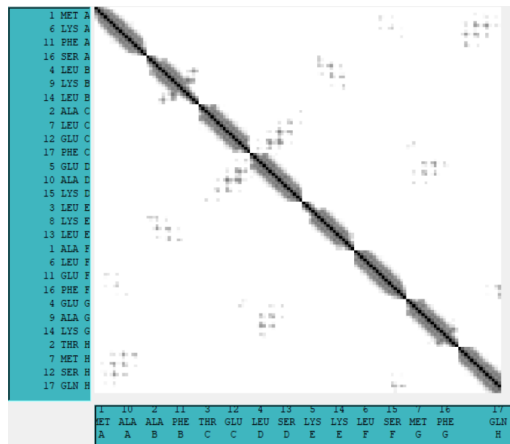
(b)



(c)

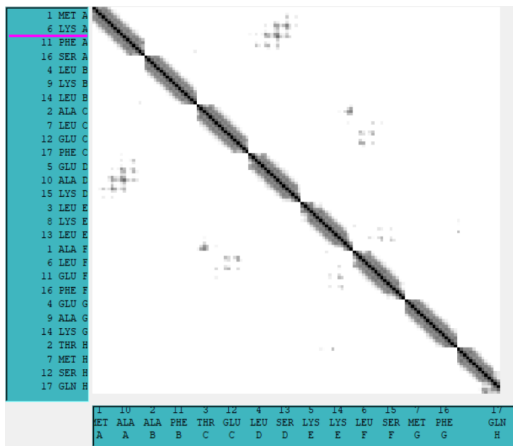


(d)

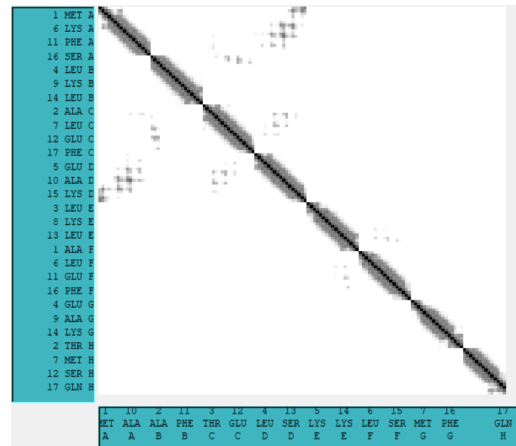


(e)

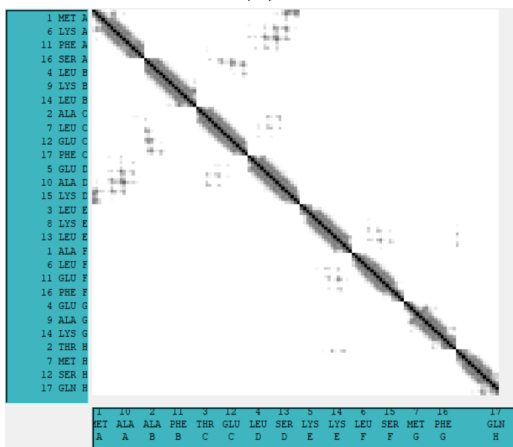
Figur 38: Contact maps för 2LD2 och 6N8C med sackaros och 2400 vatten (E1).  
 Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för  
 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



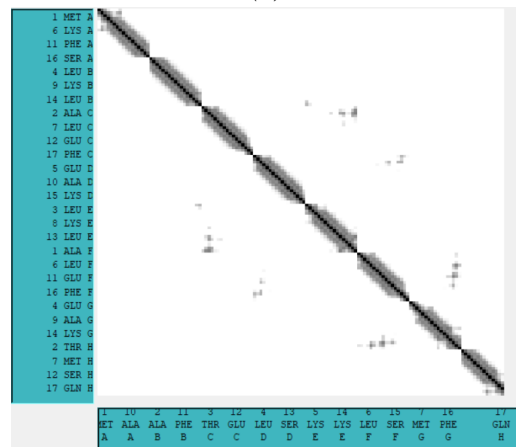
(a)



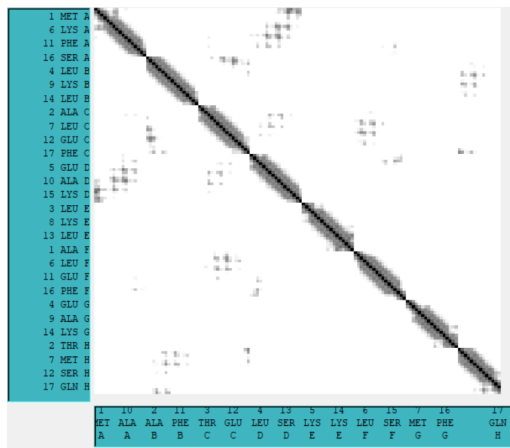
(b)



(c)

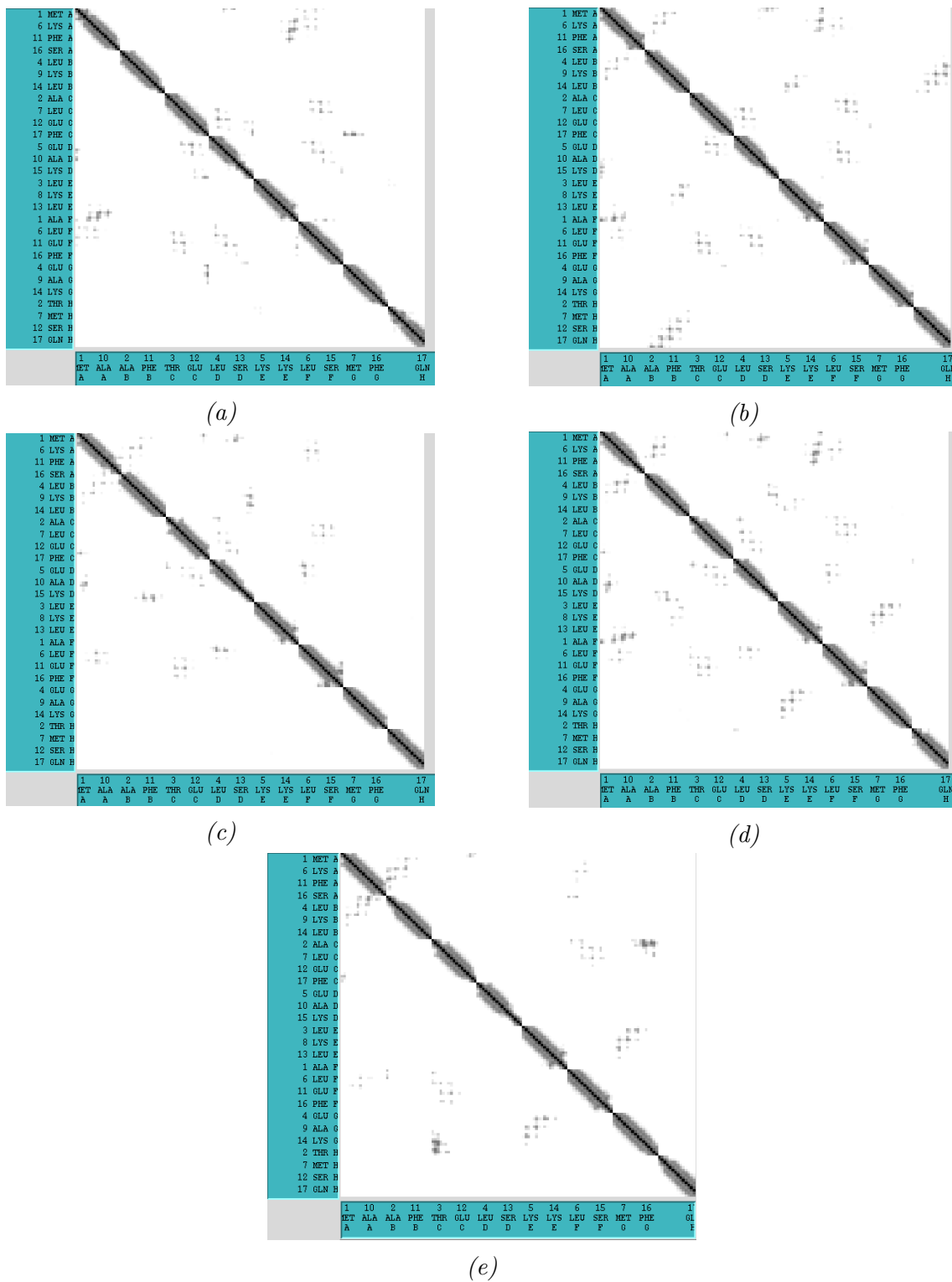


(d)



(e)

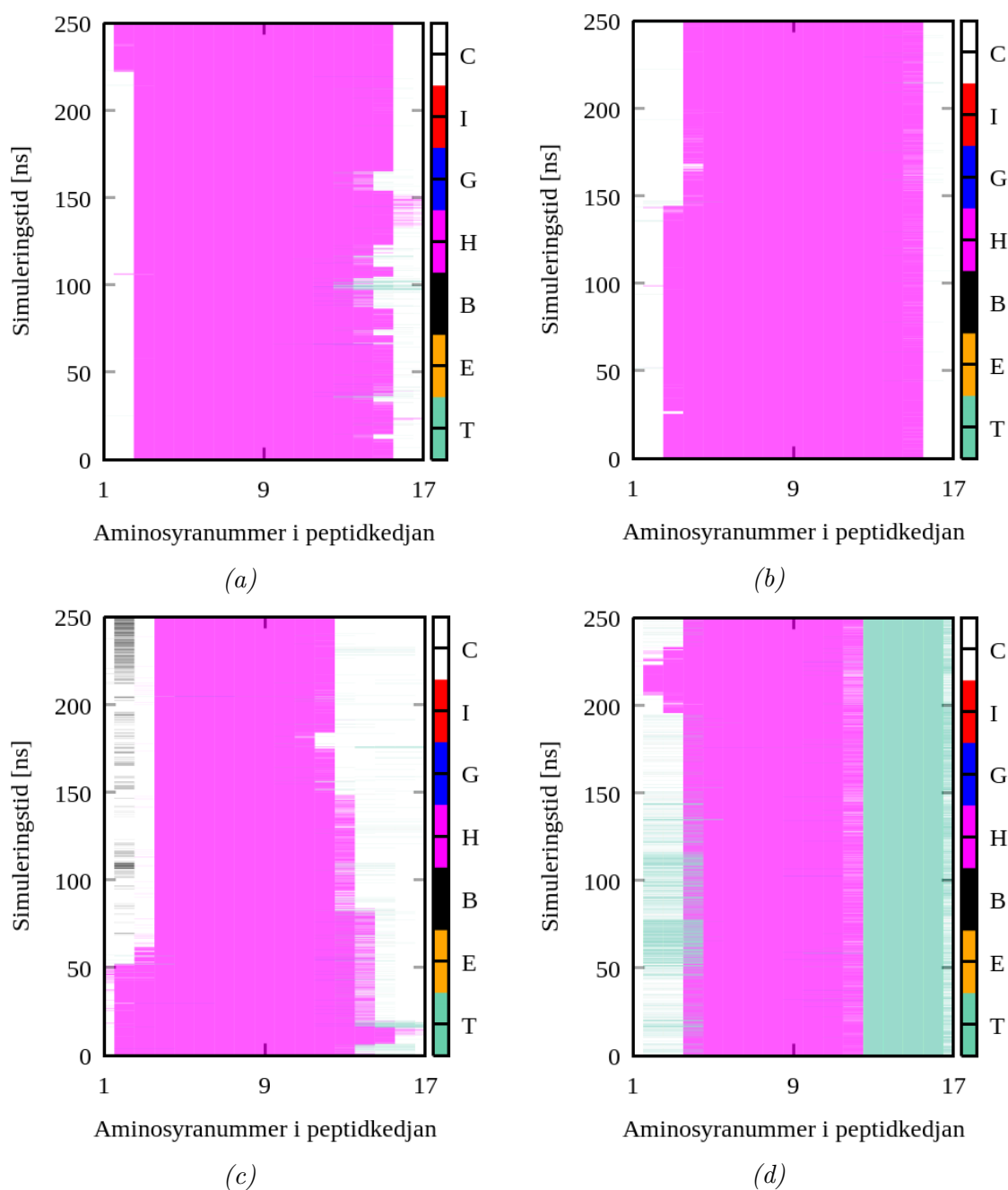
Figur 39: Contact maps för 2LD2 och 6N8C med trehalos och 2400 vatten (E2).  
 Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.



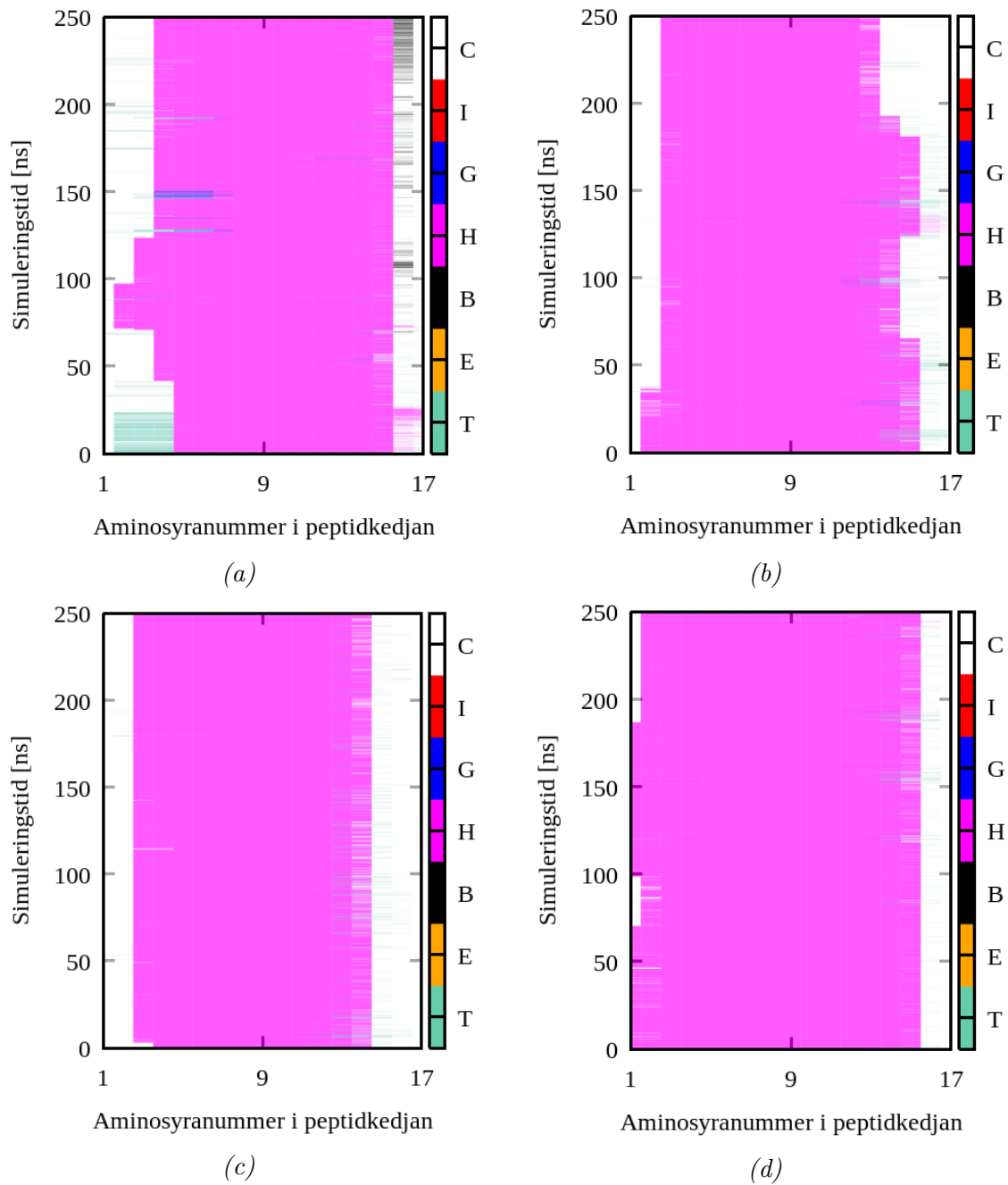
Figur 40: Contact maps för 2LD2 och 6N8C utan socker, med 2400 vatten (E3).  
 Figur a) visar CM för tidsintervallet 50-100 ns, b) för 100-150 ns, c) för 150-200 ns, d) för 200-250 ns och e) för 250-300 ns.

## A.4 Sekundärstruktur

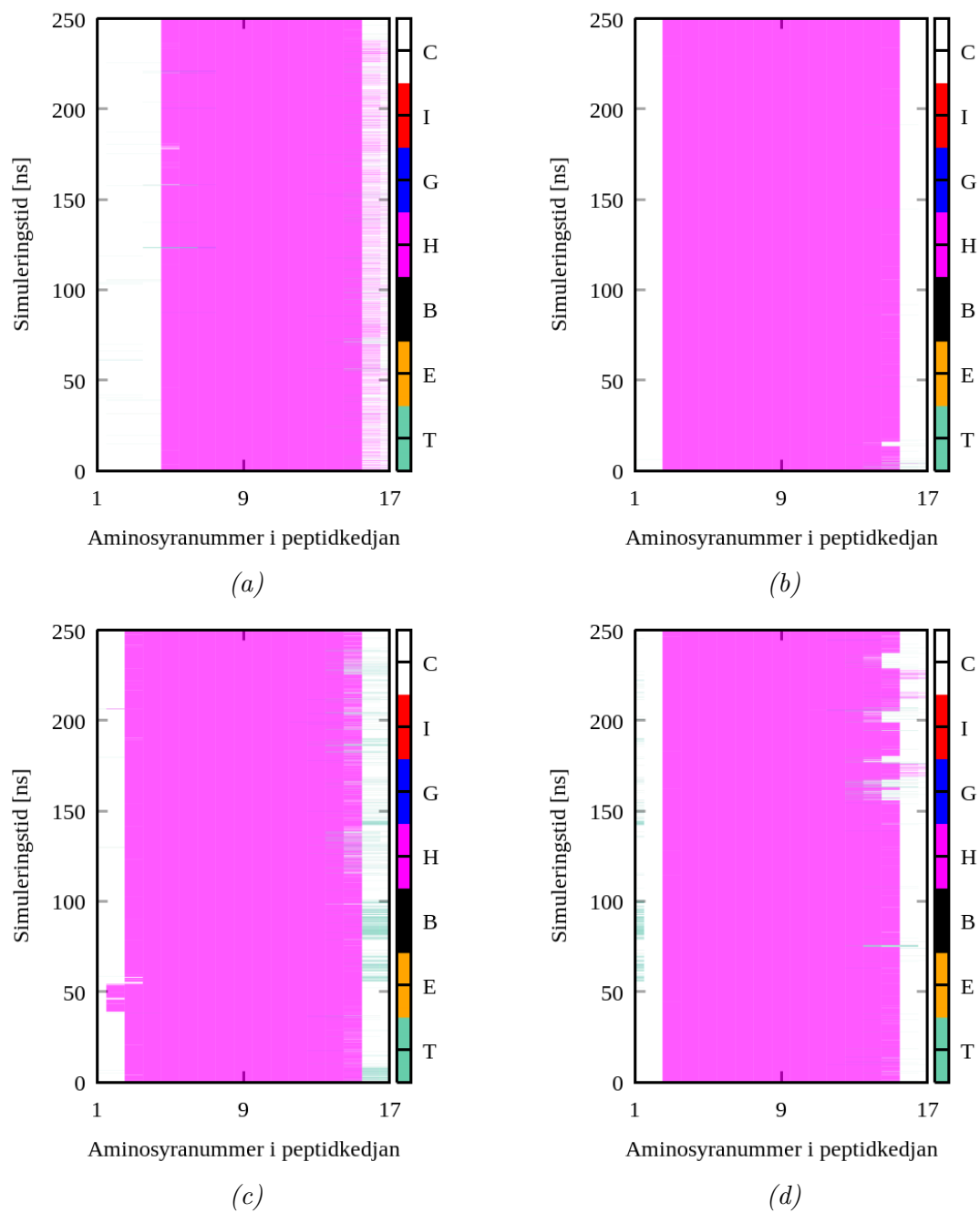
### A.4.1 A-system



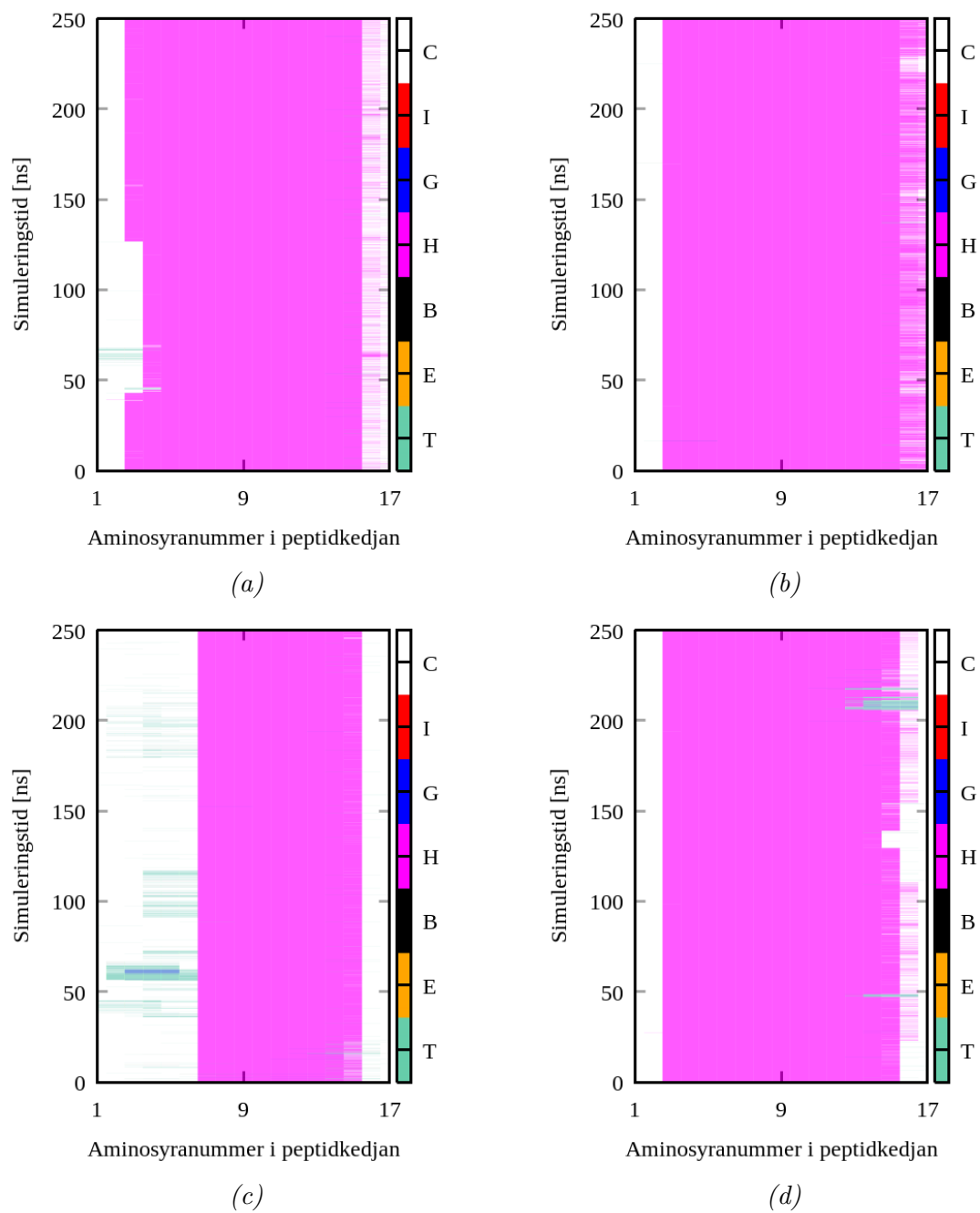
Figur 41: Sekundärstruktur för 2LD2-peptider A-D, i system med sackaros och 4800 vatten (A1). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.



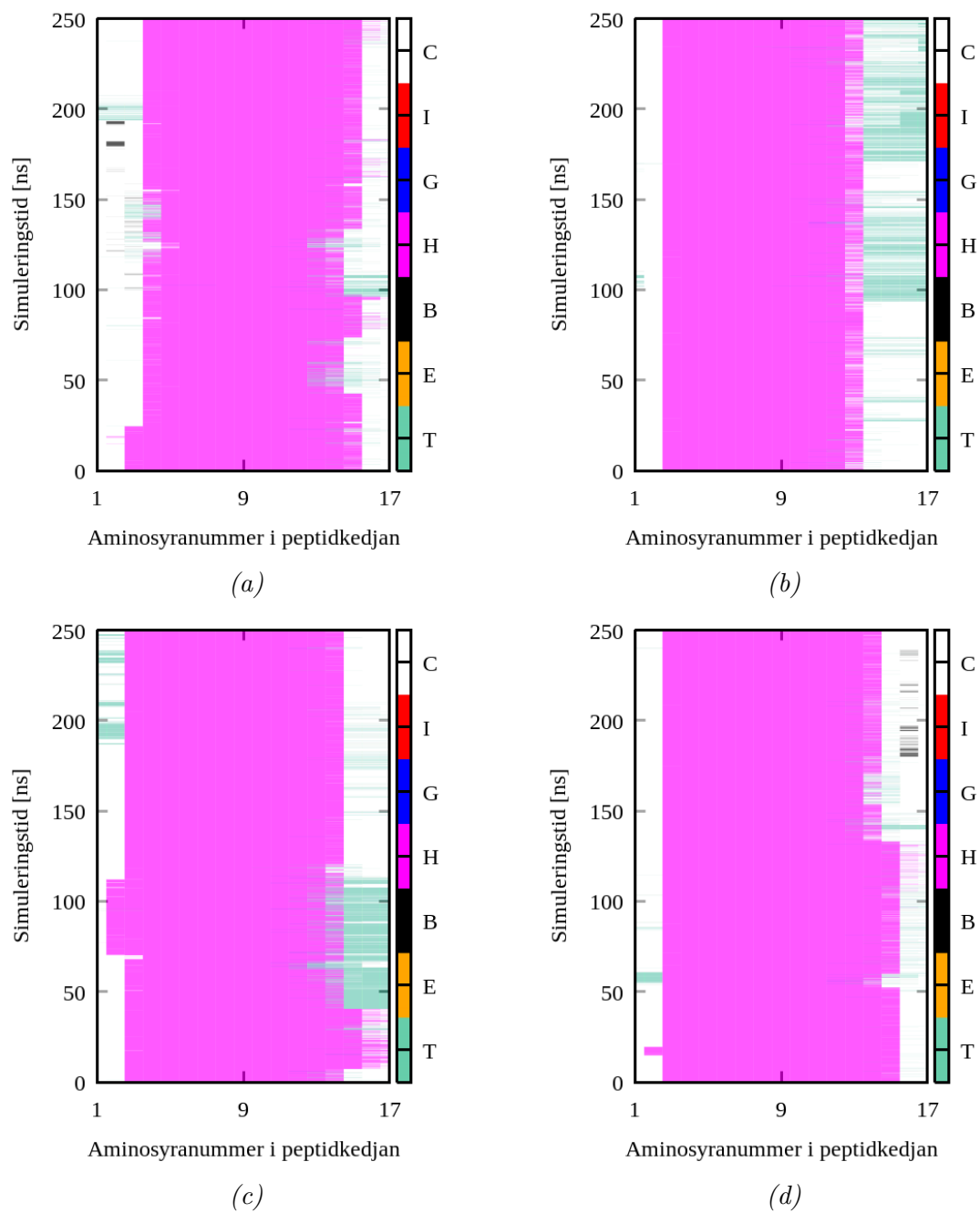
Figur 42: Sekundärstruktur för 2LD2-peptider E-H, i system med sackaros och 4800 vatten (A1). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.



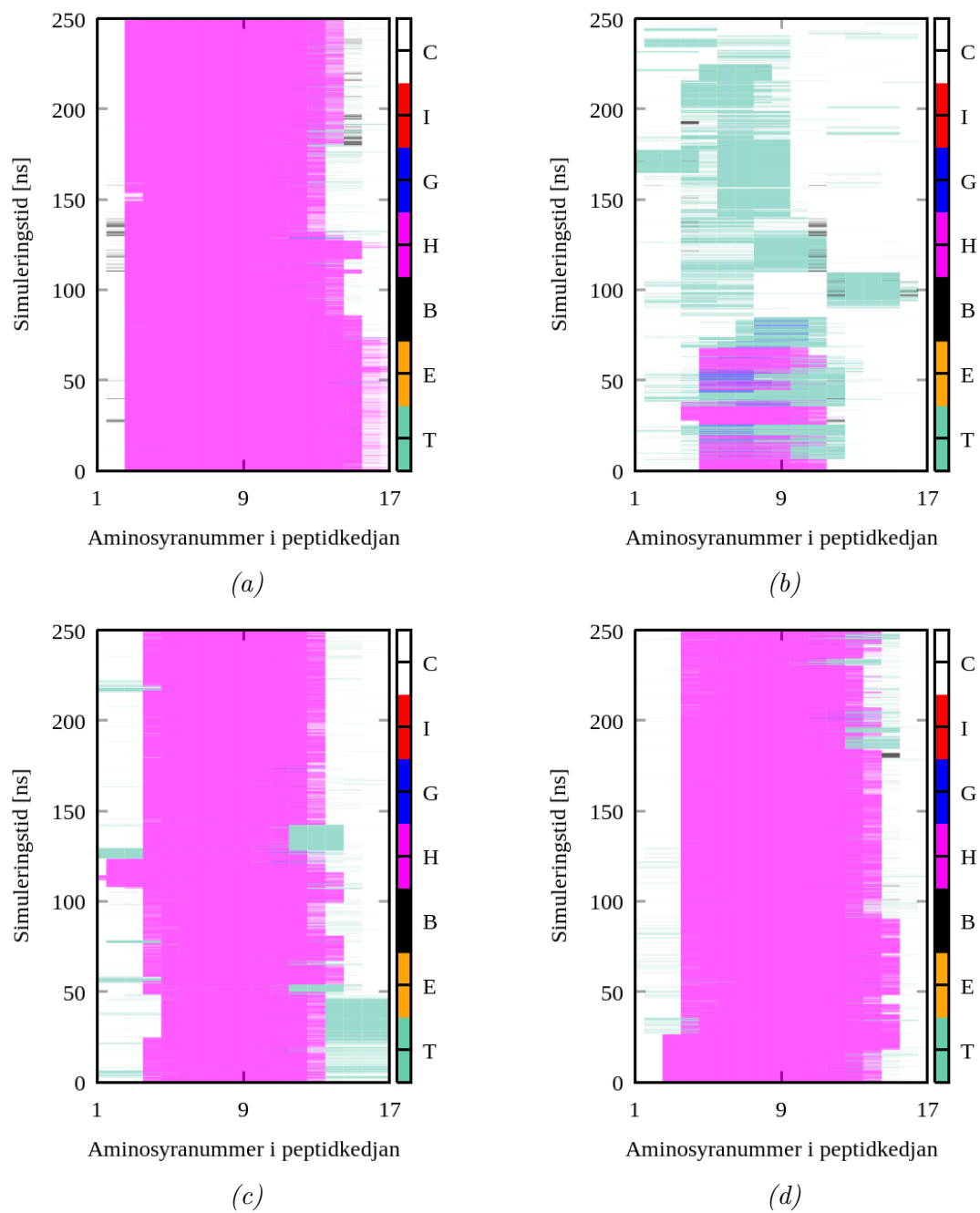
Figur 43: Sekundärstruktur för 2LD2-peptider A-D, i system med trehalos och 4800 vatten (A2). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.



Figur 44: Sekundärstruktur för 2LD2-peptider E-H, i system med trehalos och 4800 vatten (A2). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.

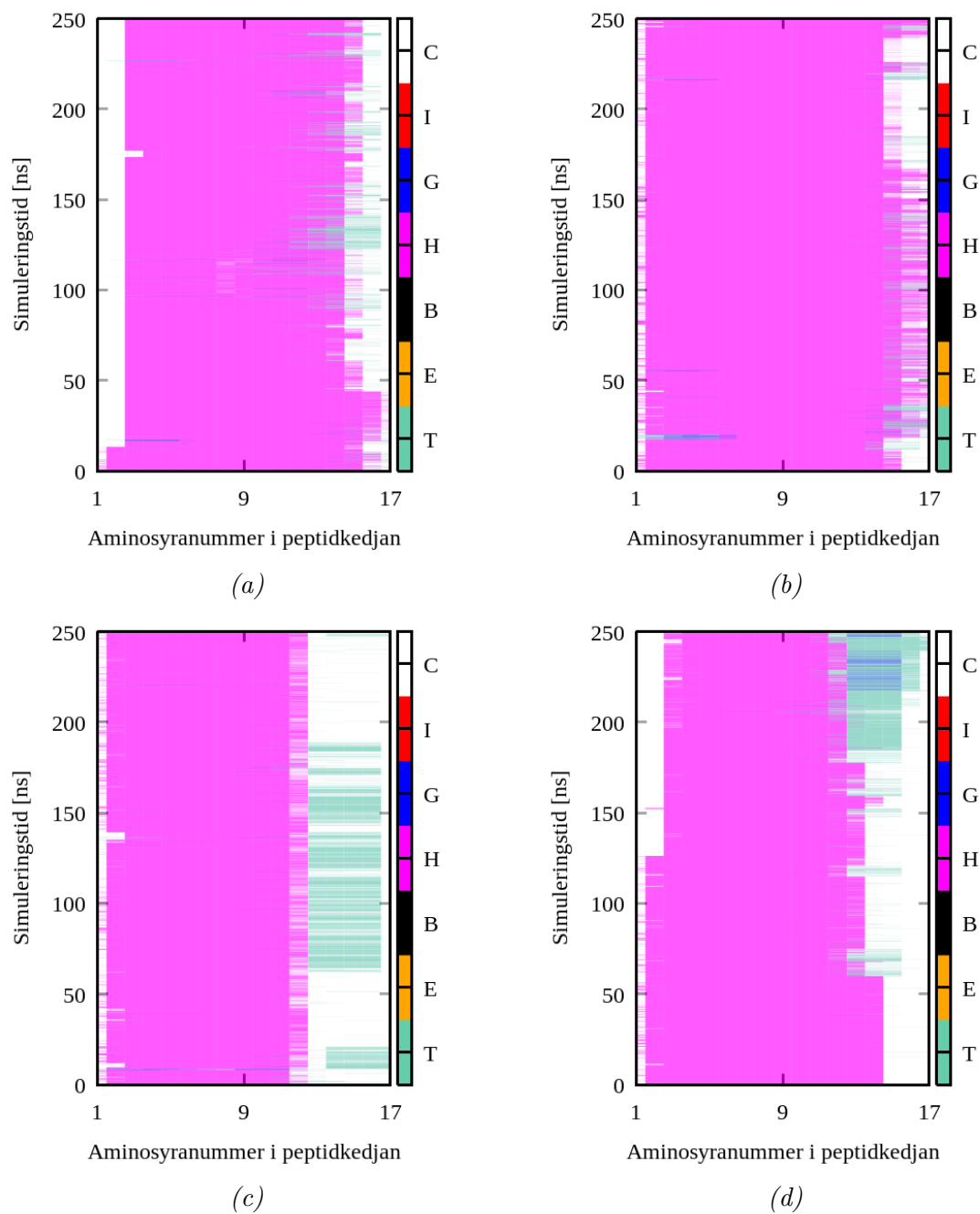


Figur 45: Sekundärstruktur för 2LD2-peptider A-D och 4800 vatten (A3). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.

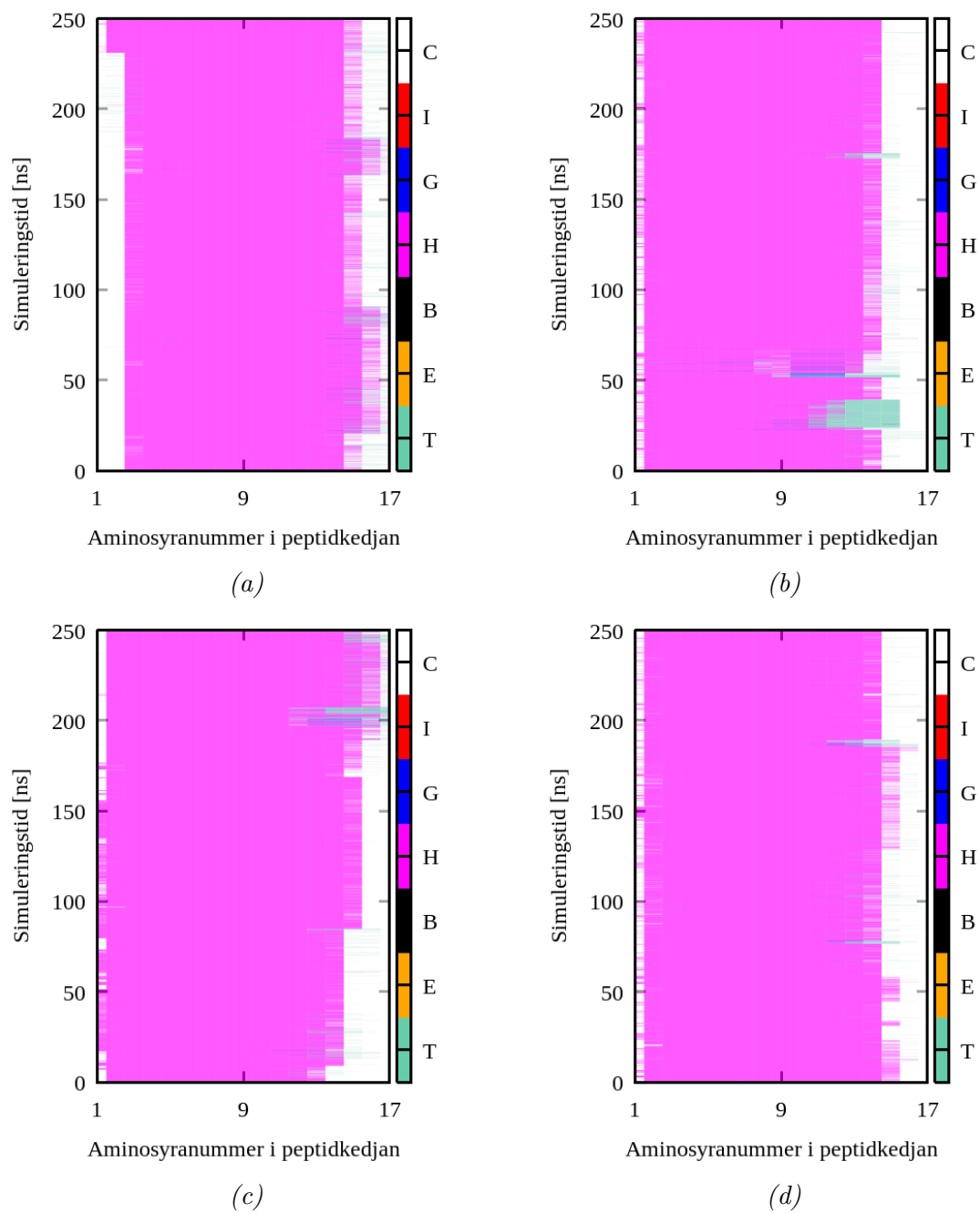


Figur 46: Sekundärstruktur för 2LD2-peptider E-H och 4800 vatten (A3). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.

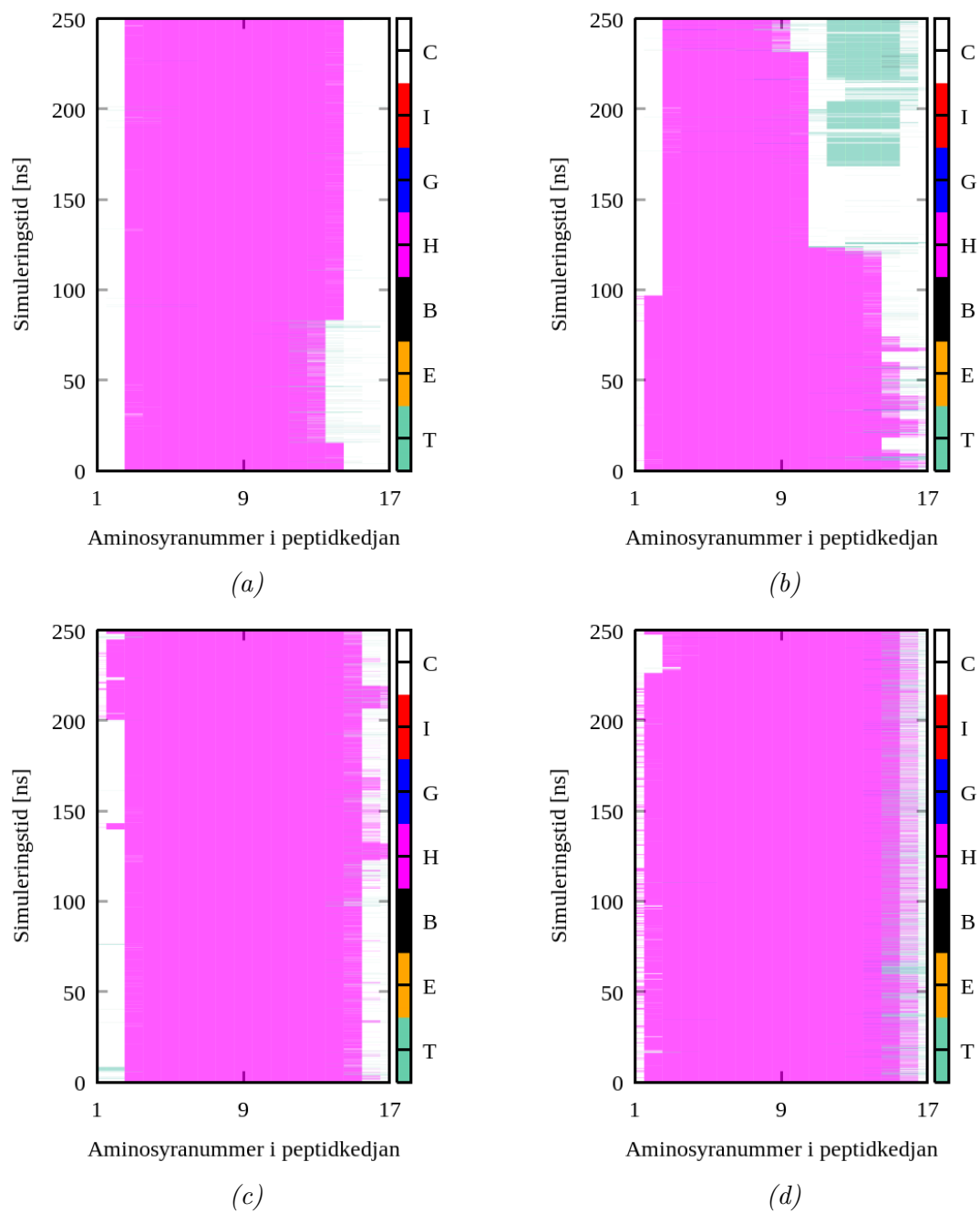
#### A.4.2 B-system



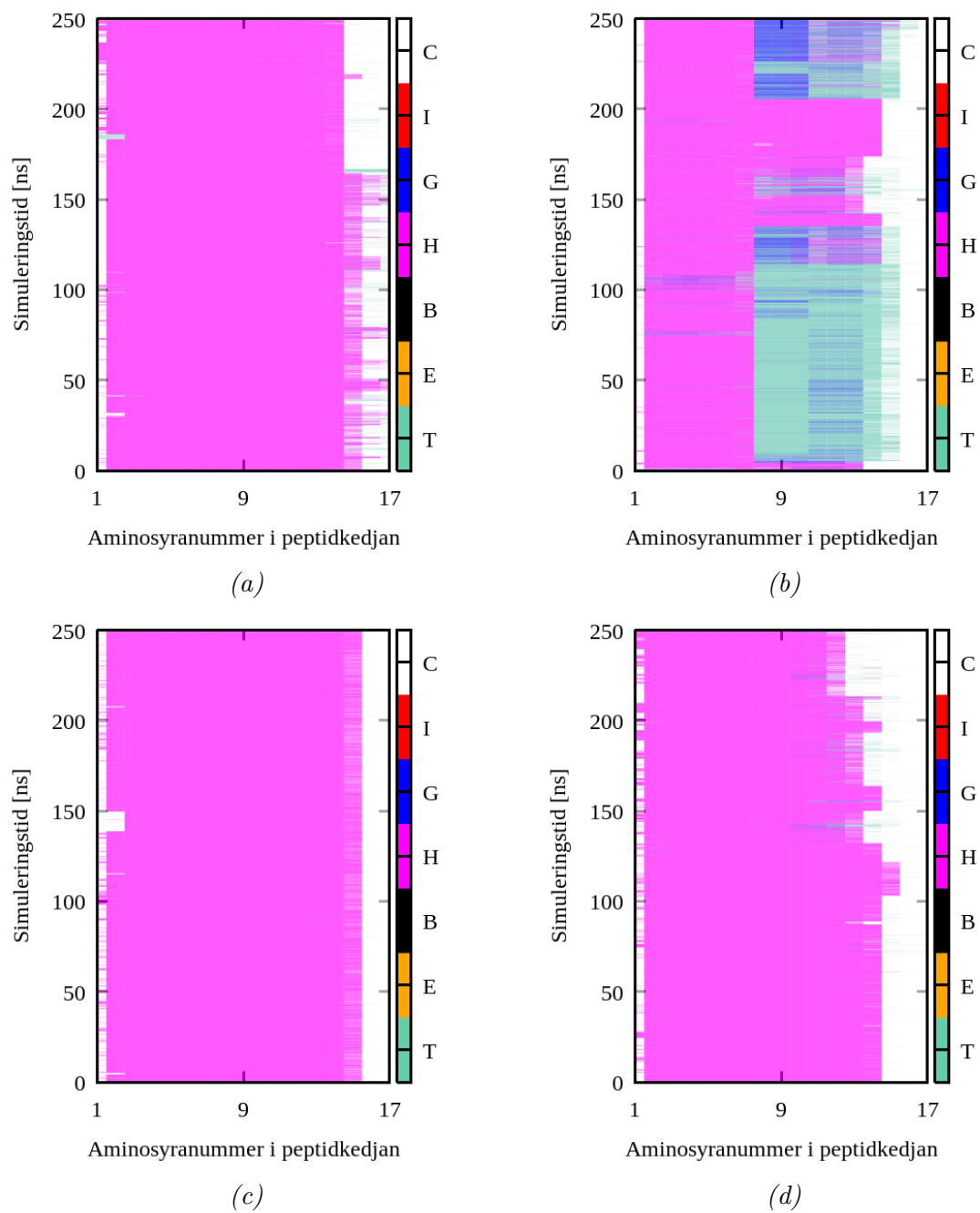
Figur 47: Sekundärstruktur för 6N8C-peptider A-D, i system med sackaros och 4800 vatten (B1). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.



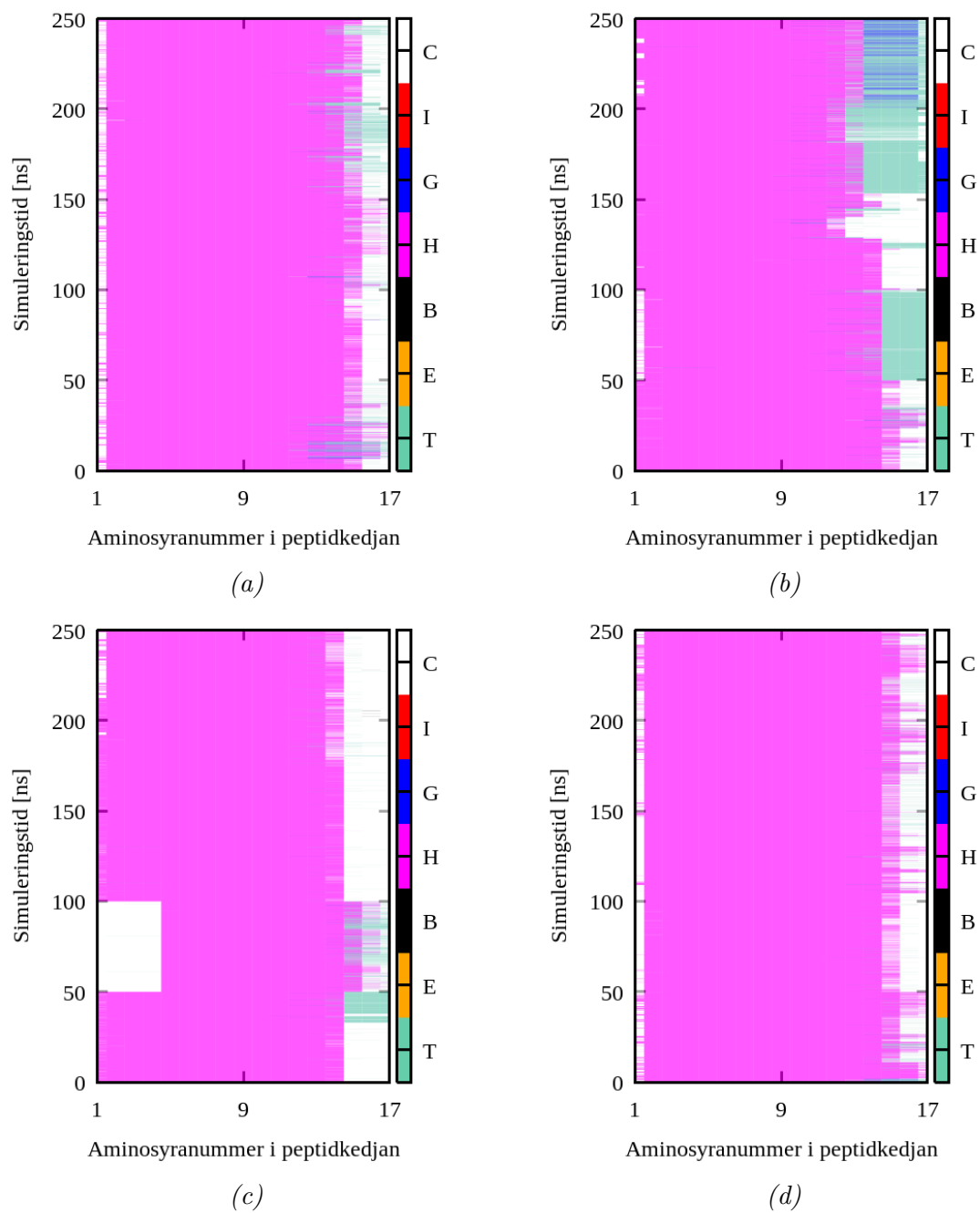
Figur 48: Sekundärstruktur för 6N8C-peptider E-H, i system med sackaros och 4800 vatten (B1). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.



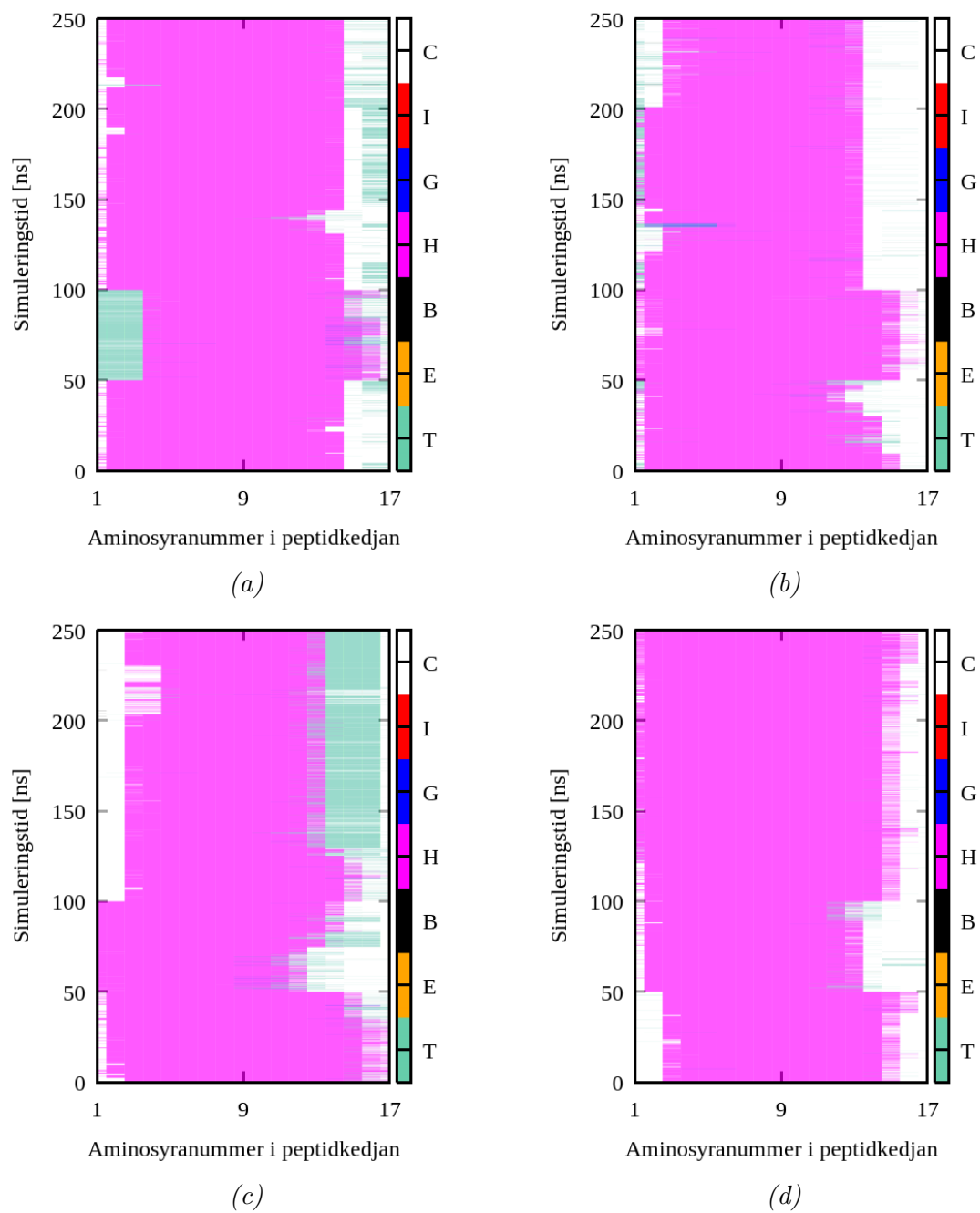
Figur 49: Sekundärstruktur för 6N8C-peptider A-D, i system med trehalos och 4800 vatten (B2). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.



Figur 50: Sekundärstruktur för 6N8C-peptider E-H, i system med trehalos och 4800 vatten (B2). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.

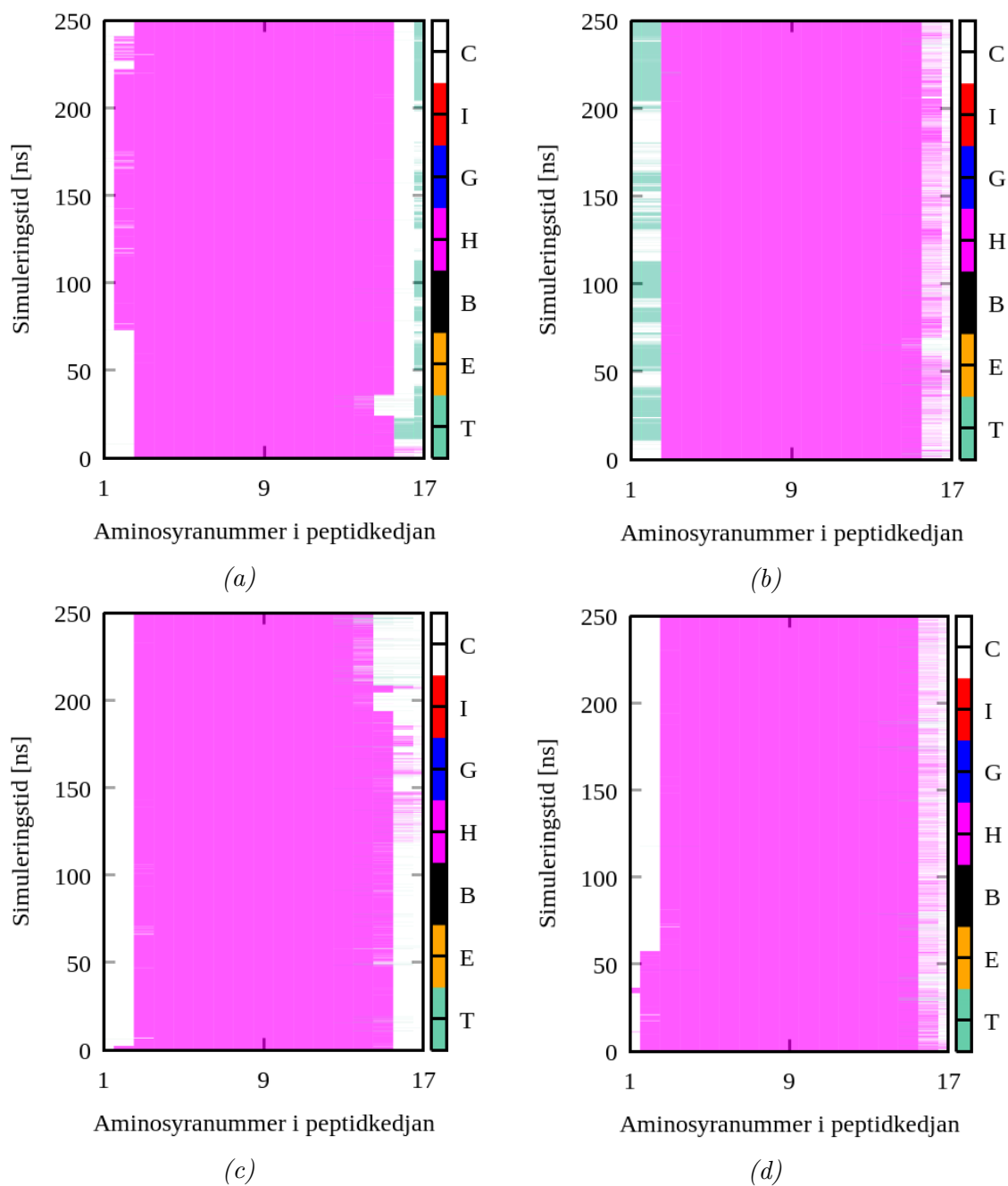


Figur 51: Sekundärstruktur för 6N8C-peptider A-D, i system med 4800 vatten (B3).  
 Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.

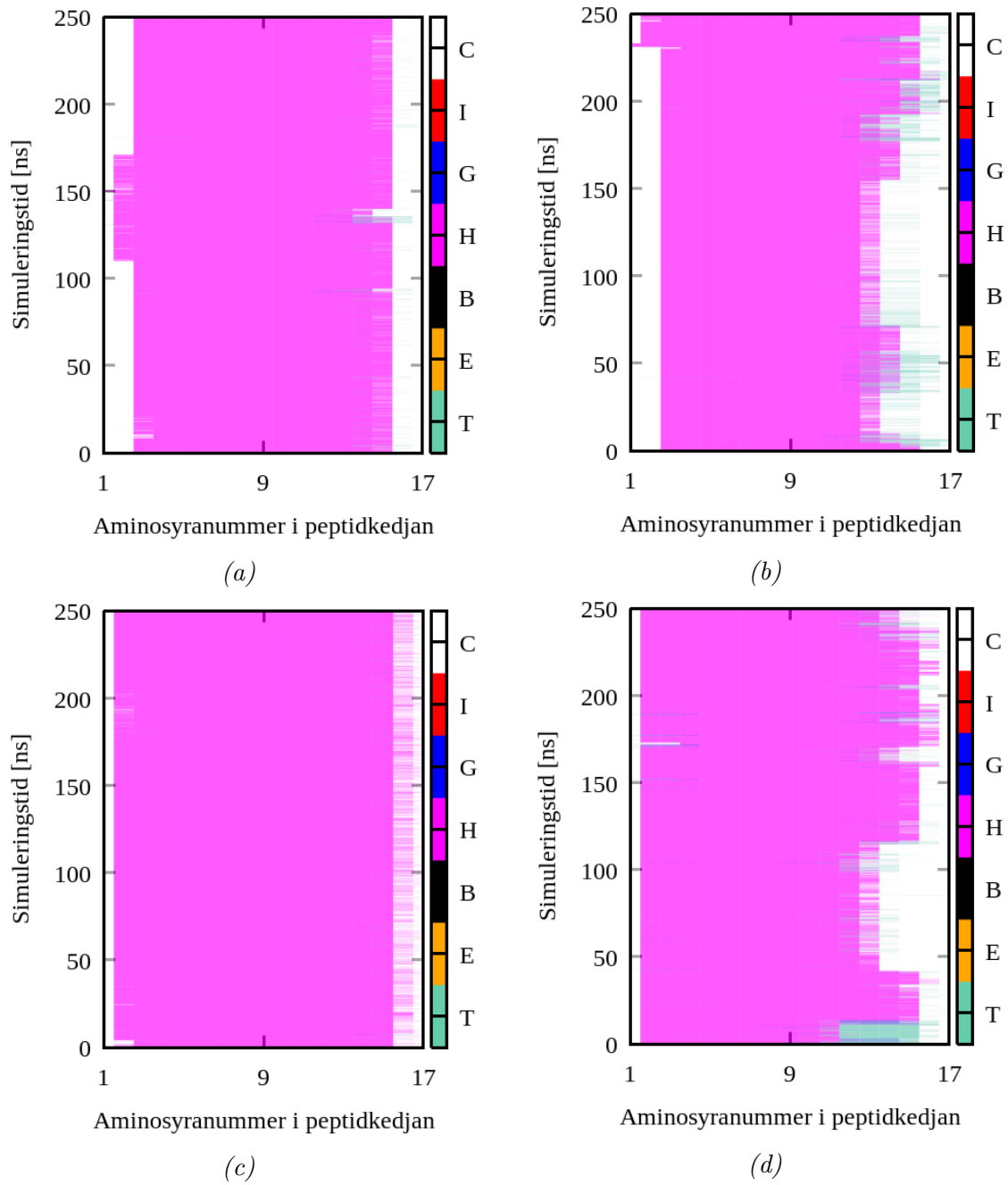


Figur 52: Sekundärstruktur för 6N8C-peptider E-H, i system med 4800 vatten (B3).  
 Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.

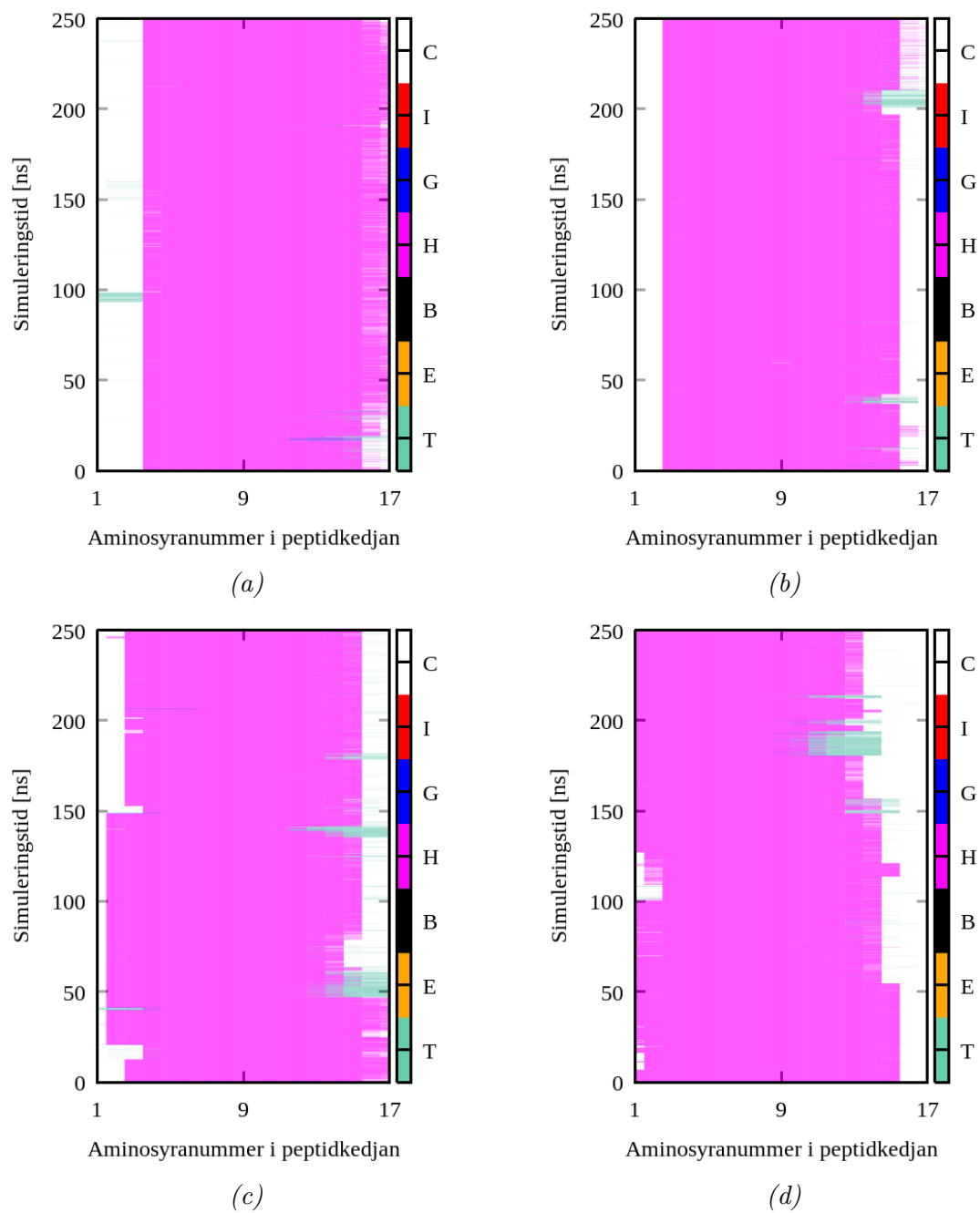
### A.4.3 C-system



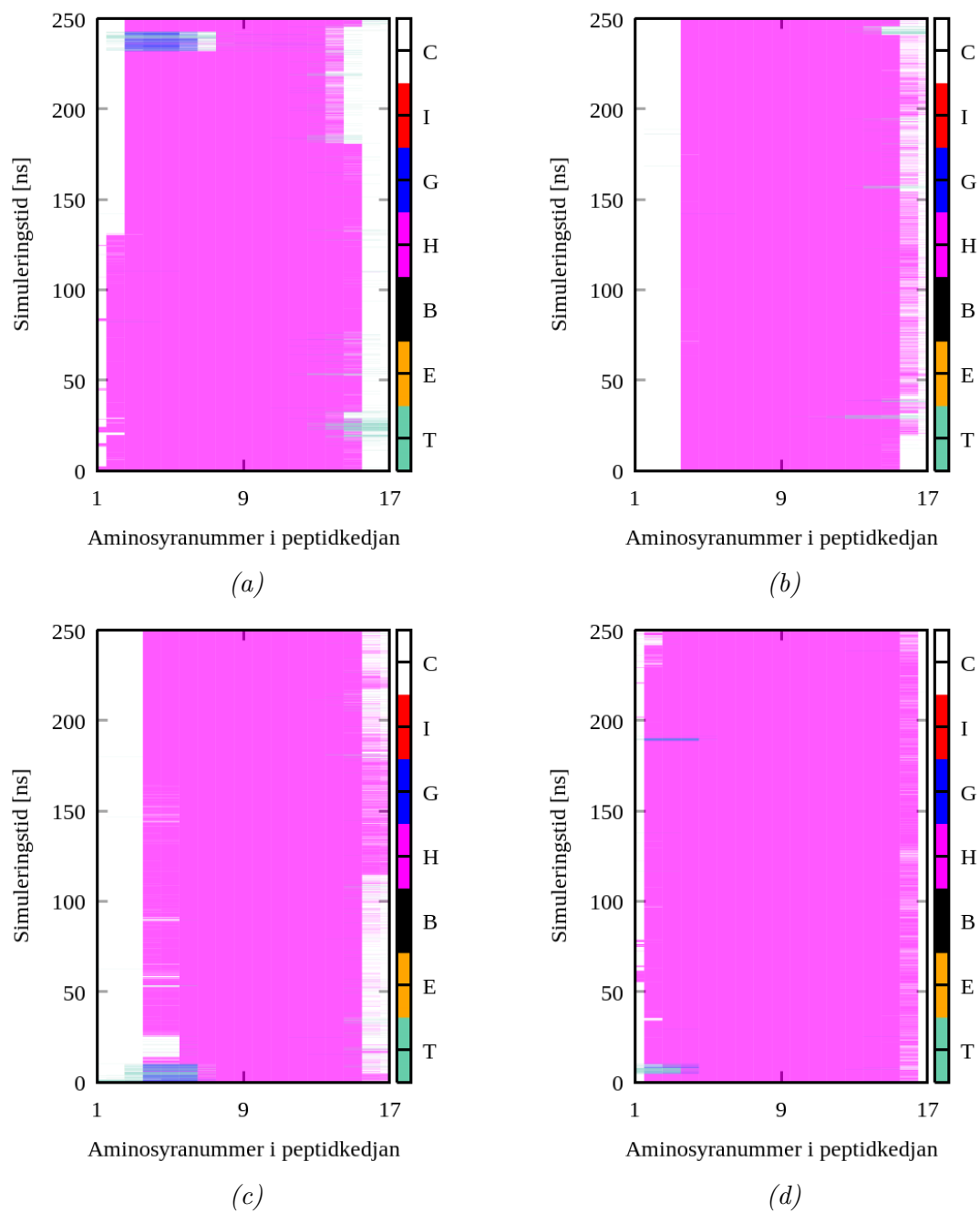
Figur 53: Sekundärstruktur för 2LD2-peptider A-D, i system med sackaros och 2400 vatten (C1). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.



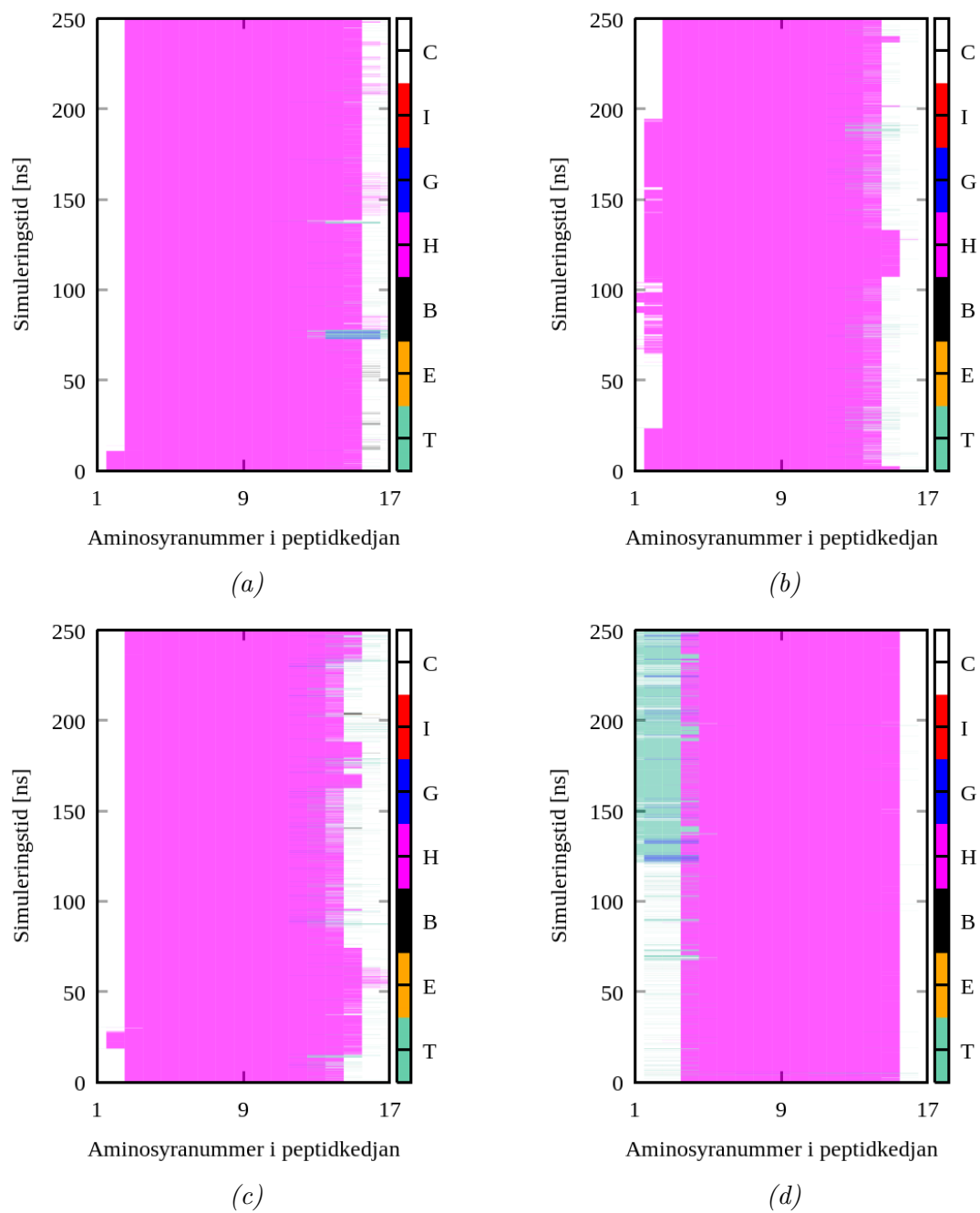
Figur 54: Sekundärstruktur för 2LD2-peptider E-H, i system med sackaros och 2400 vatten (C1). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.



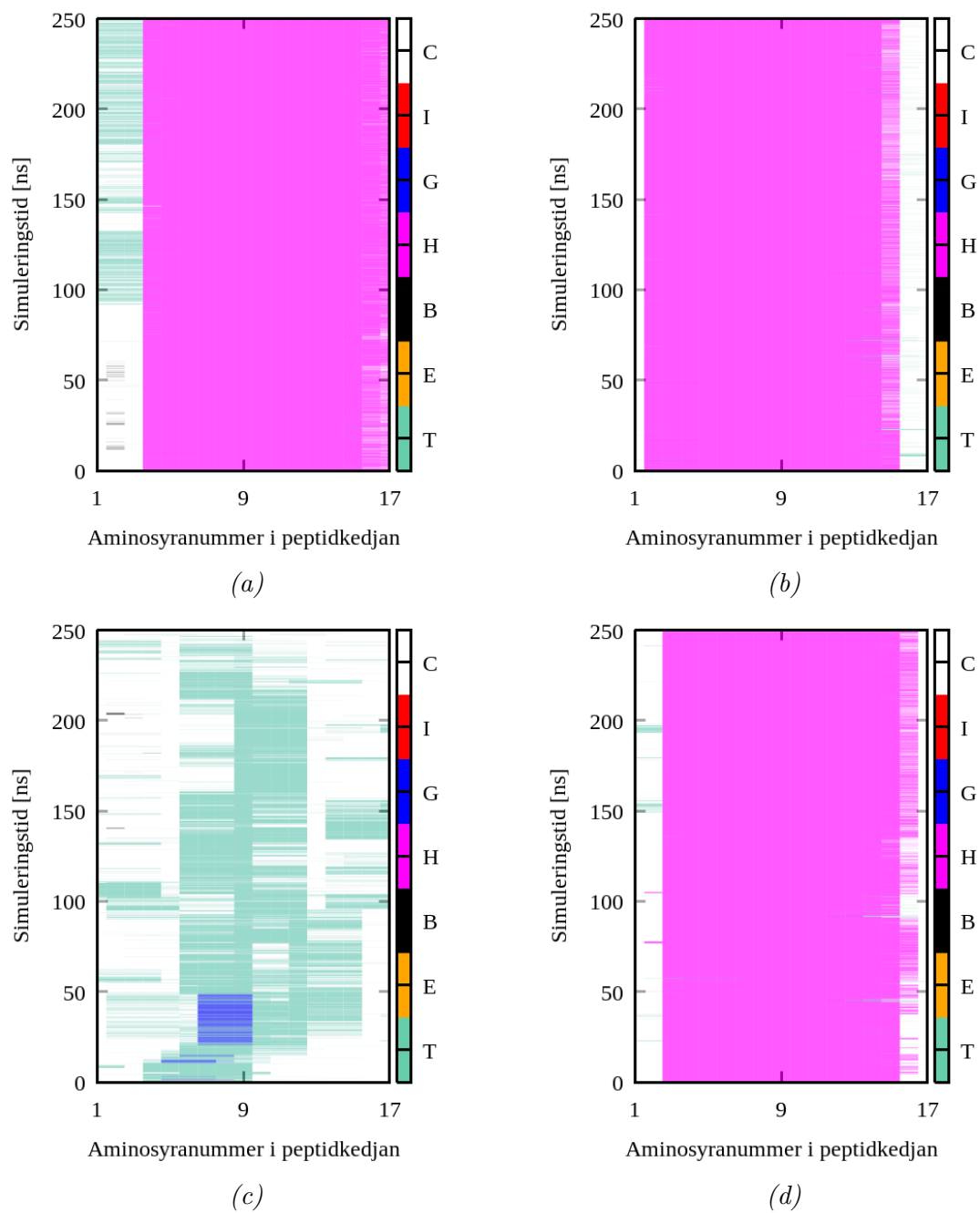
Figur 55: Sekundärstruktur för 2LD2-peptider A-D, i system med trehalos och 2400 vatten (C2). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.



Figur 56: Sekundärstruktur för 2LD2-peptider E-H, i system med trehalos och 2400 vatten (C2). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.

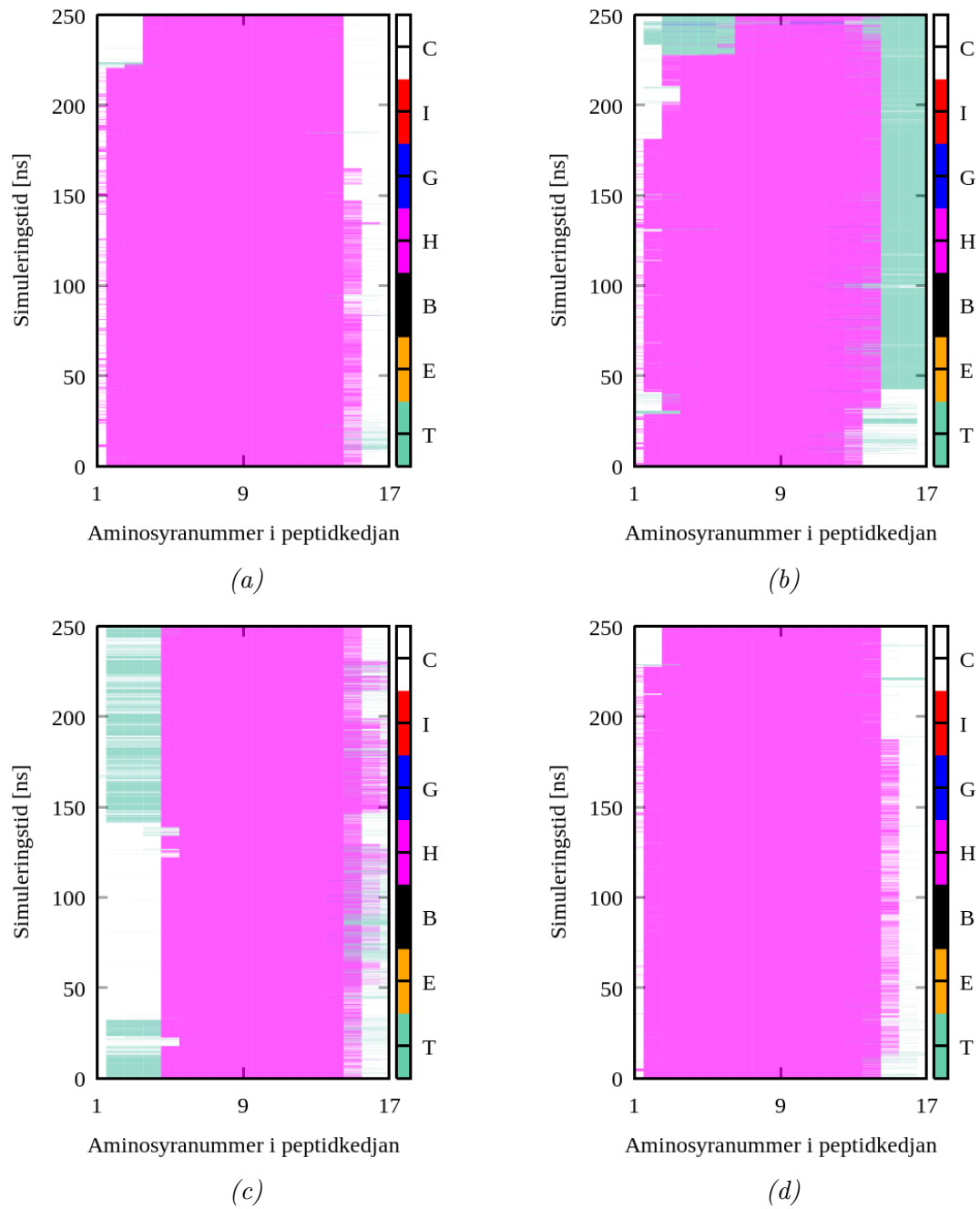


Figur 57: Sekundärstruktur för 2LD2-peptider A-D och 2400 vatten (C3). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.

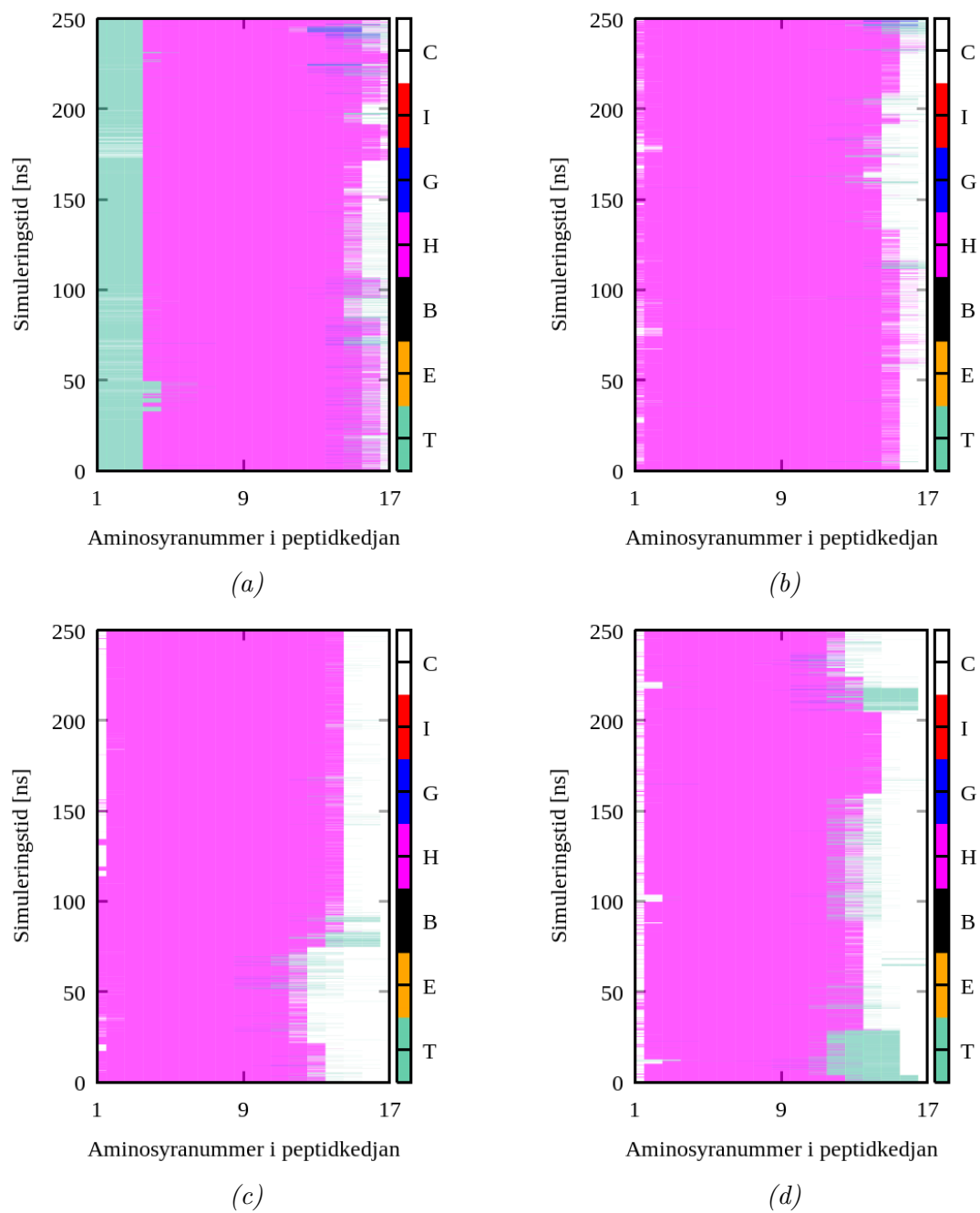


Figur 58: Sekundärstruktur för 2LD2-peptider E-H och 2400 vatten (C3). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.

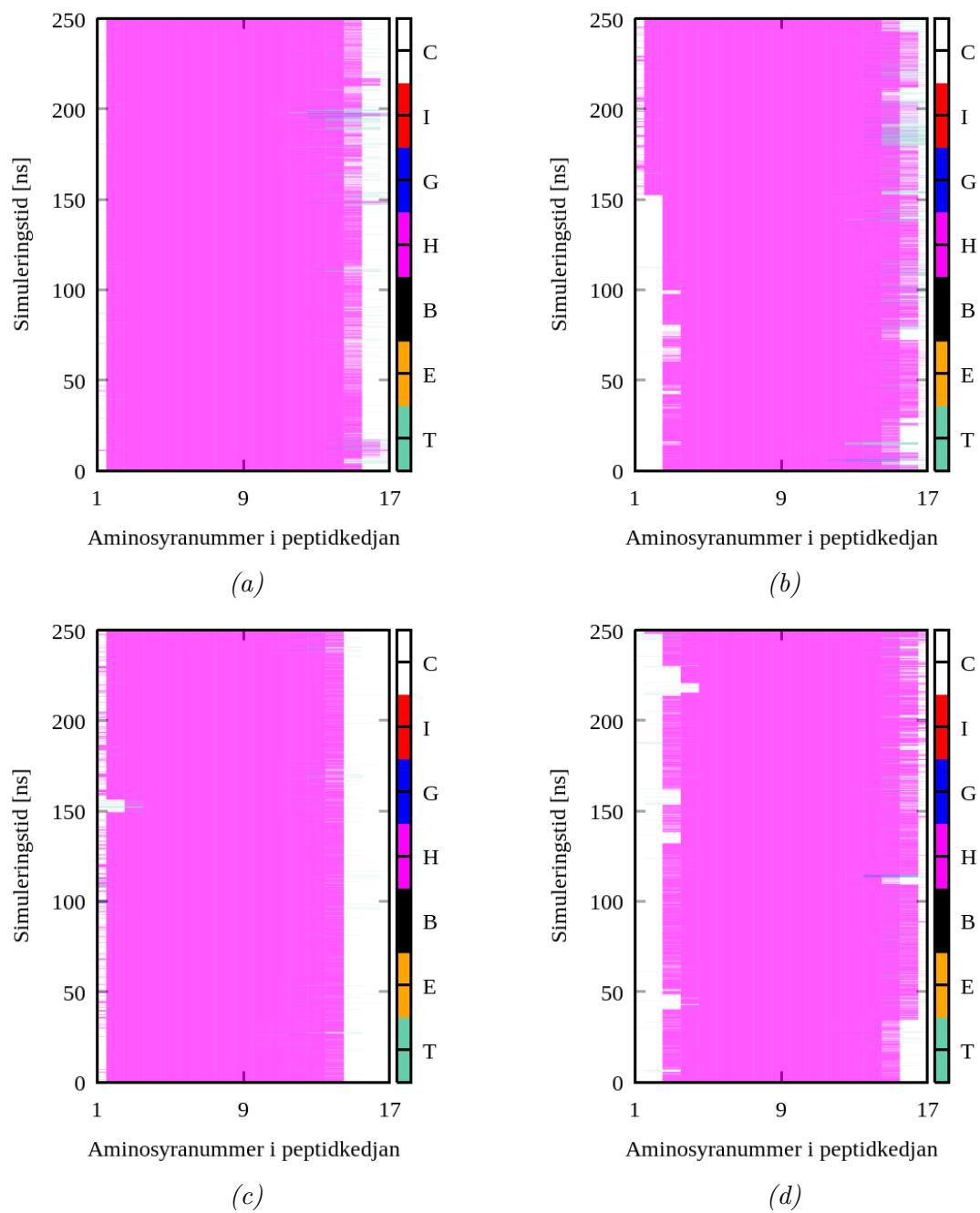
#### A.4.4 D-system



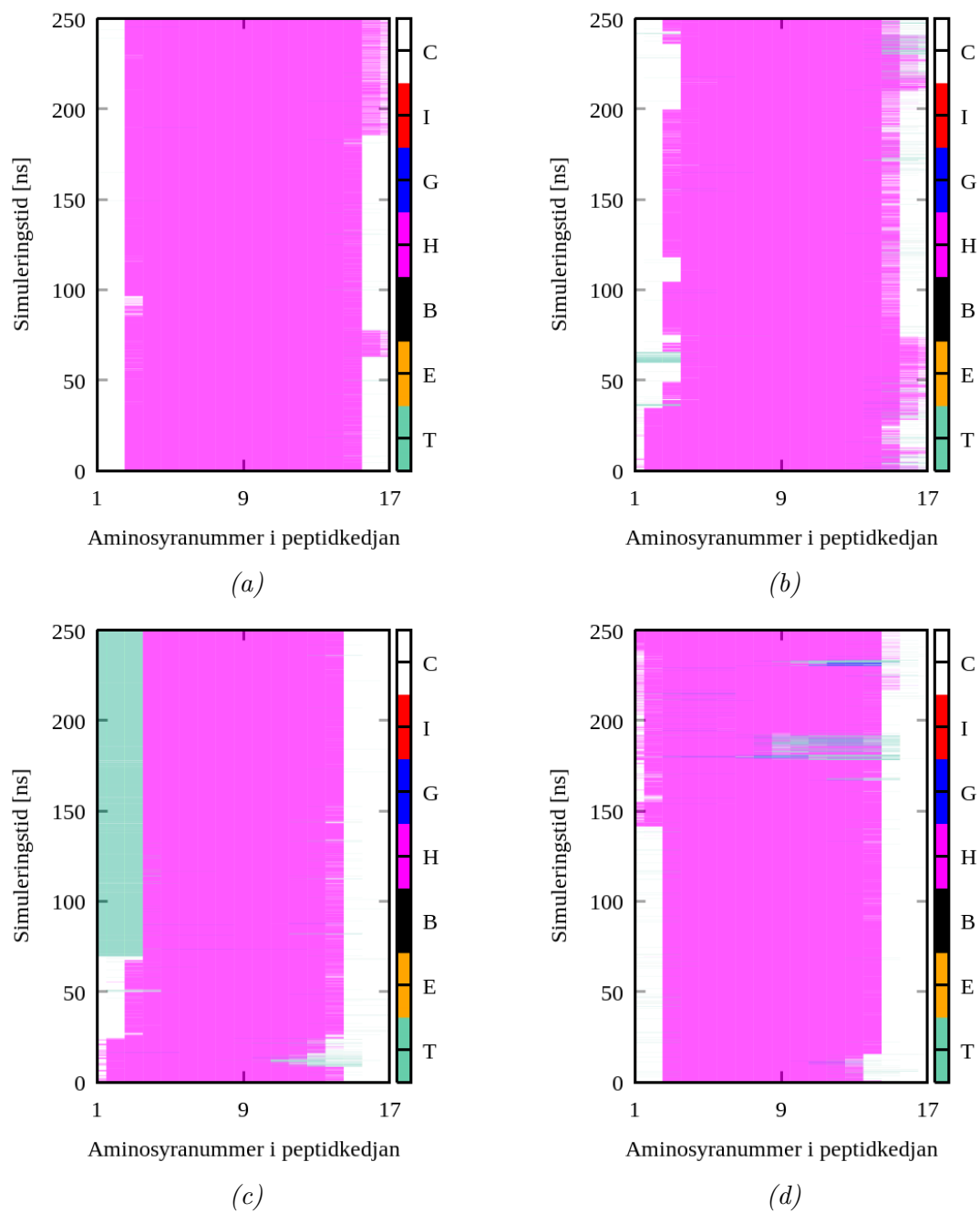
Figur 59: Sekundärstruktur för 6N8C-peptider A-D, med sackaros och med 2400 vatten (D1). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.



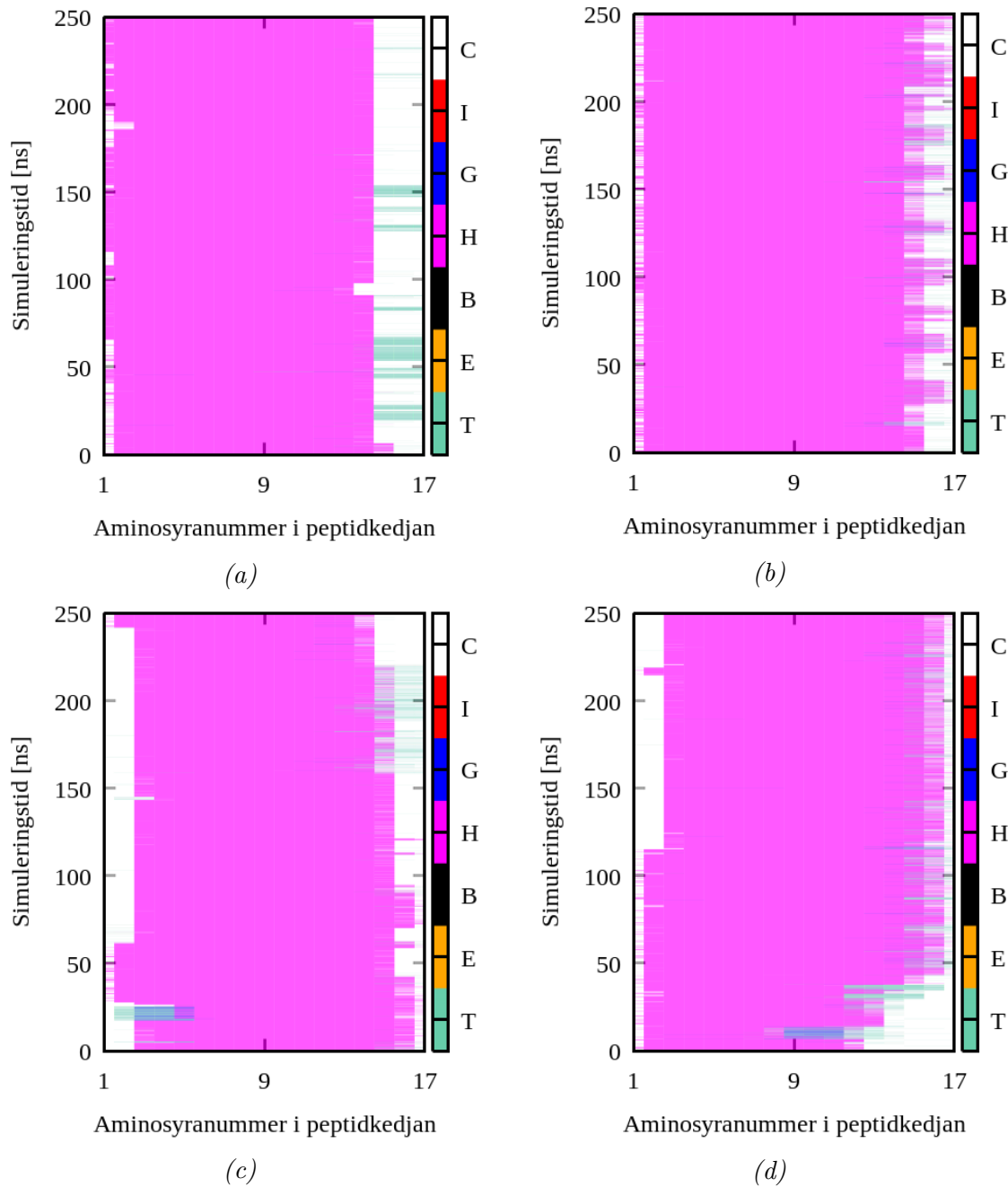
Figur 60: Sekundärstruktur för 6N8C-peptider E-H, i system med sackaros och 2400 vatten (D1). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.



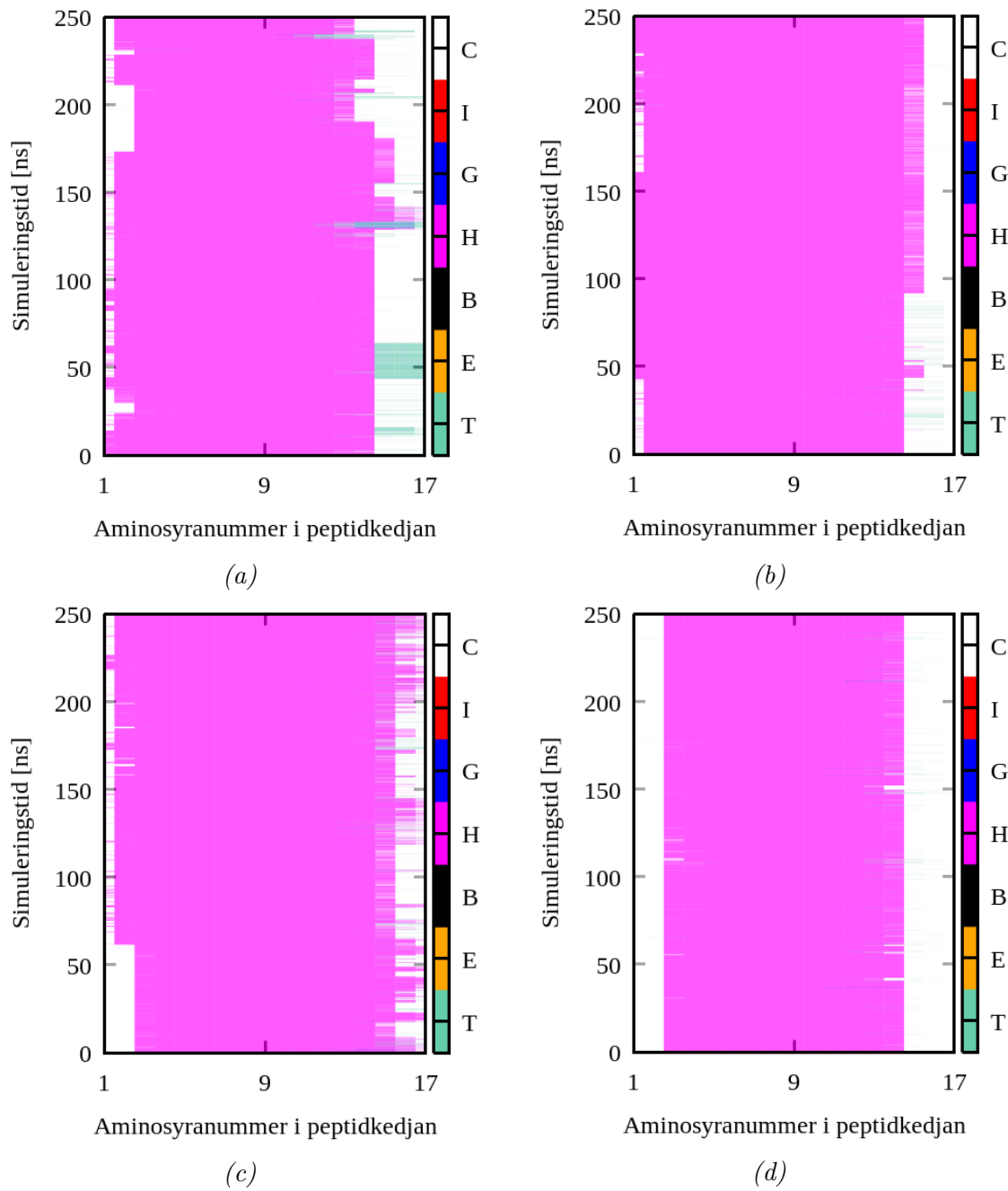
Figur 61: Sekundärstruktur för 6N8C-peptider A-D, i system med trehalos och 2400 vatten (D2). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.



Figur 62: Sekundärstruktur för 6N8C-peptider E-H, i system med trehalos och 2400 vatten (D2). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.

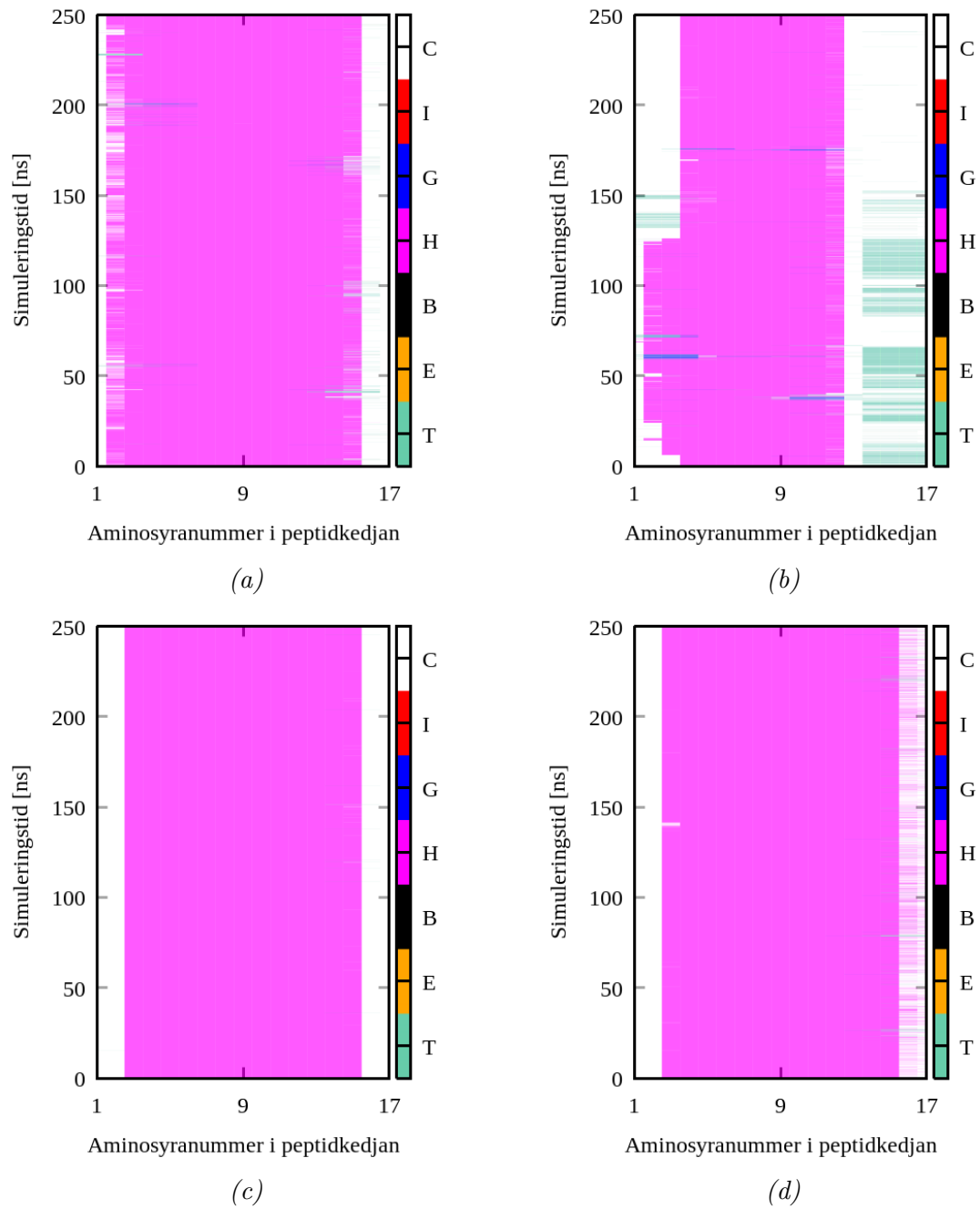


Figur 63: Sekundärstruktur för 6N8C-peptider A-D och med 2400 vatten (D3). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.

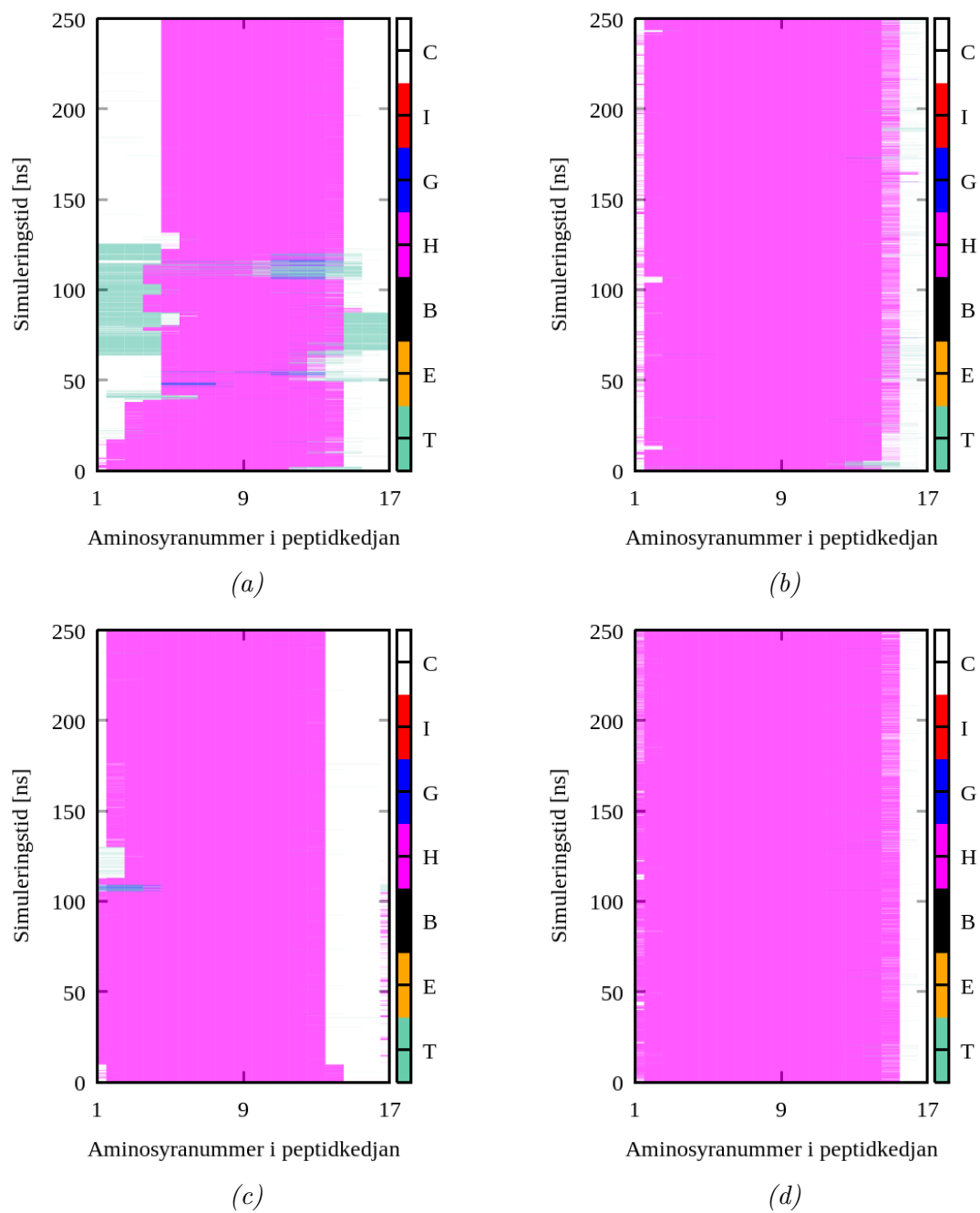


Figur 64: Sekundärstruktur för 6N8C-peptider E-H och med 2400 vatten (D3). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.

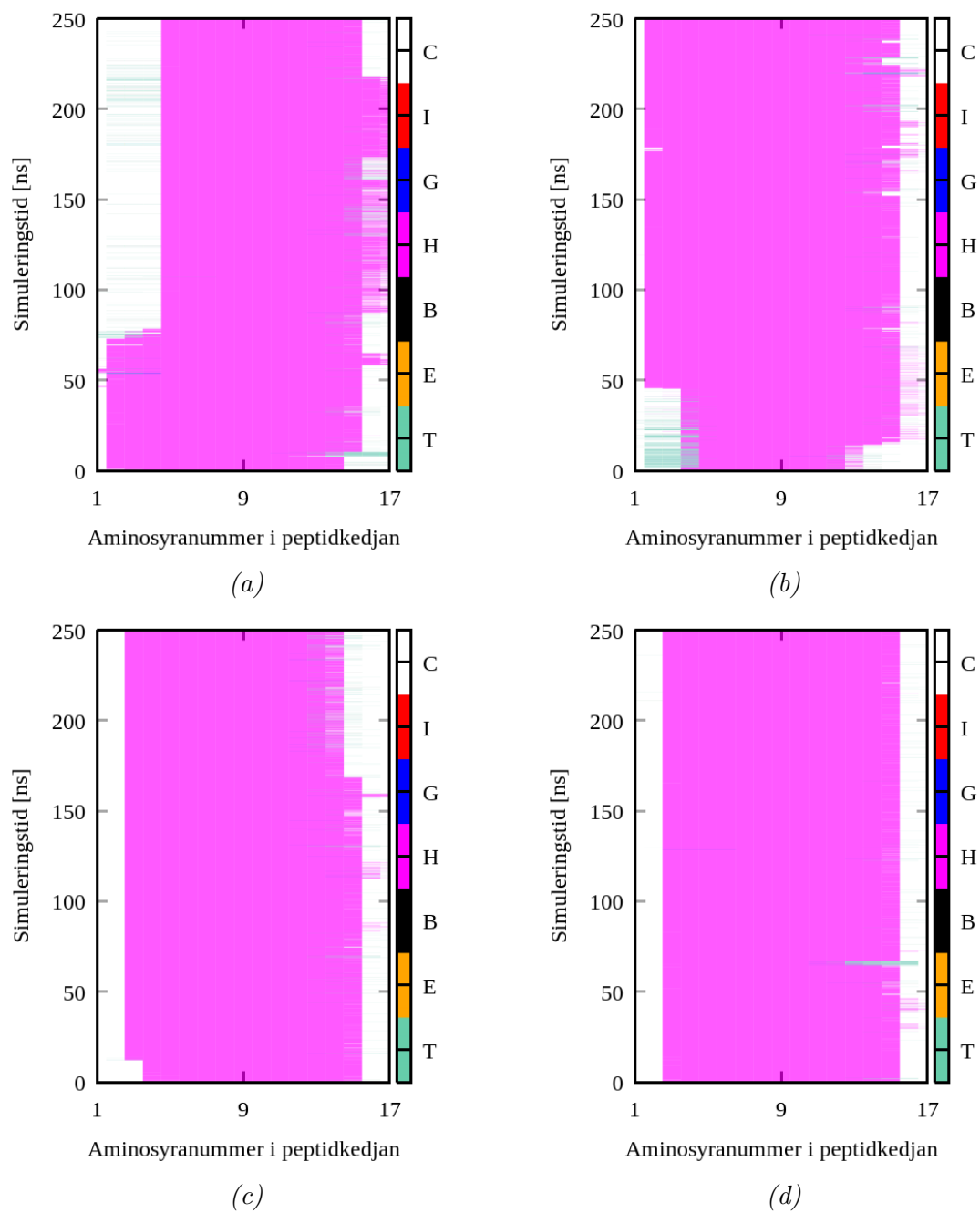
### A.4.5 E-system



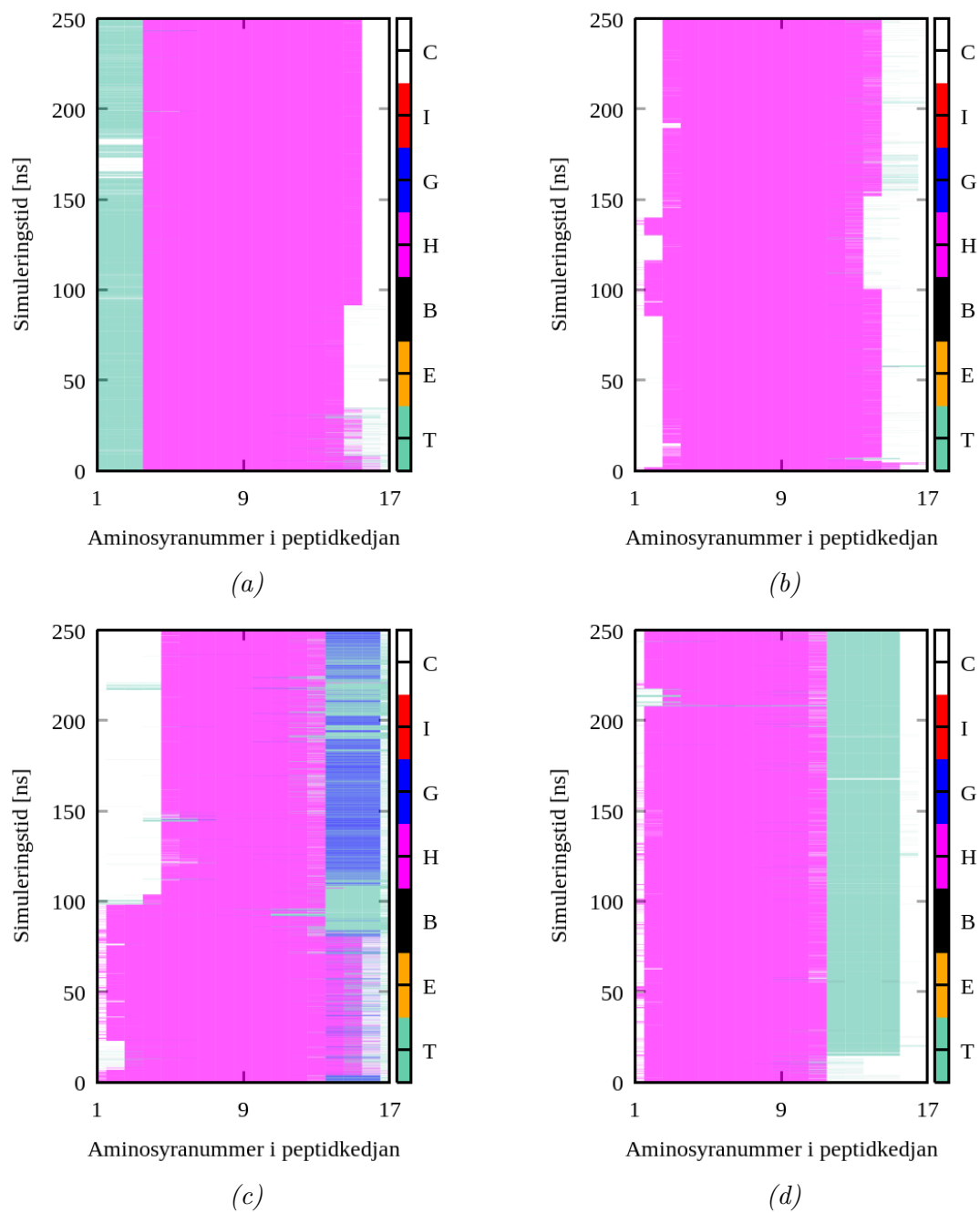
Figur 65: Sekundärstruktur för 2LD2-peptider med 6N8C-peptider A-D, sackaros och 2400 vatten (E1). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.



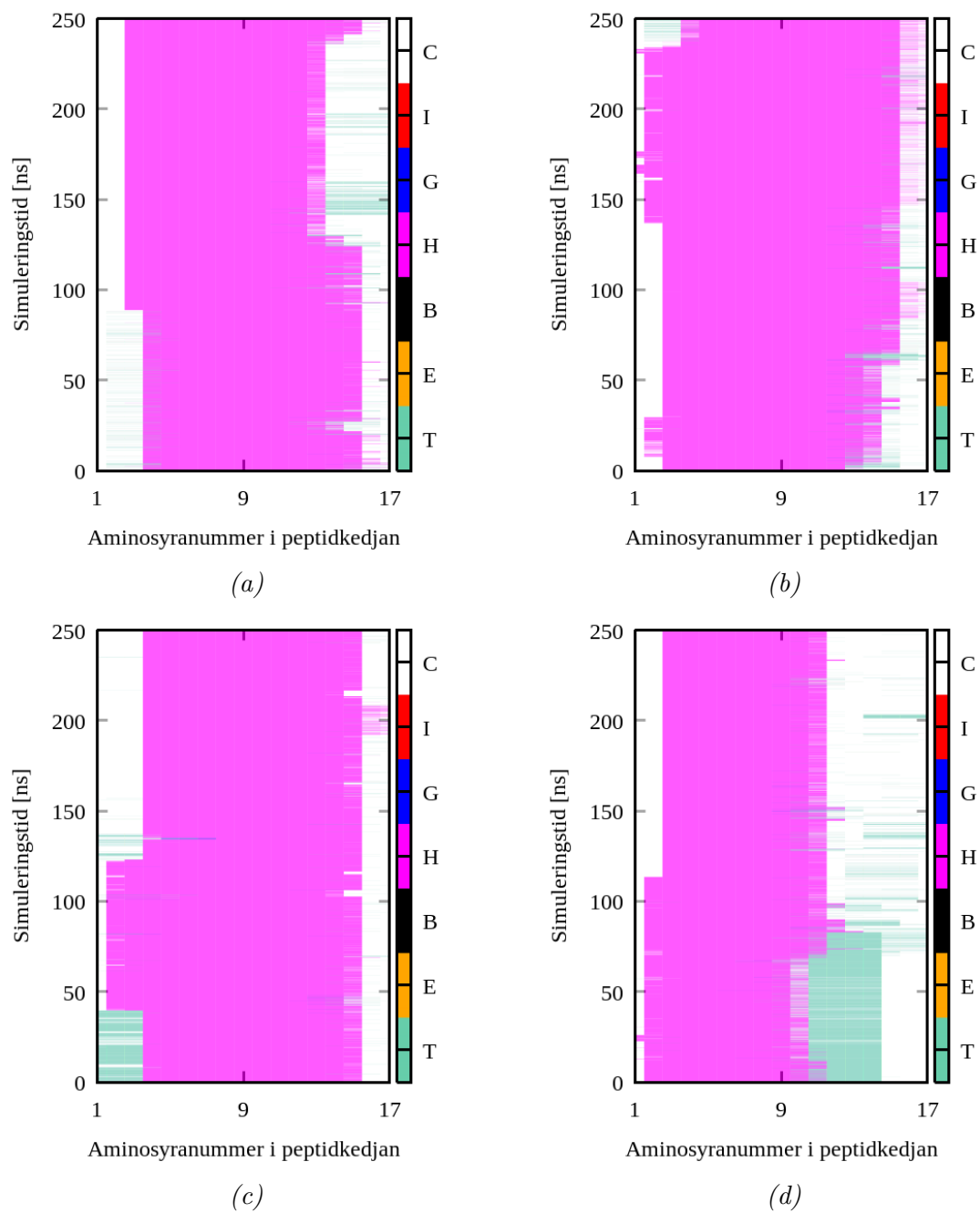
Figur 66: Sekundärstruktur för 2LD2-peptider med 6N8C-peptider E-H, sackaros och 2400 vatten (E1). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.



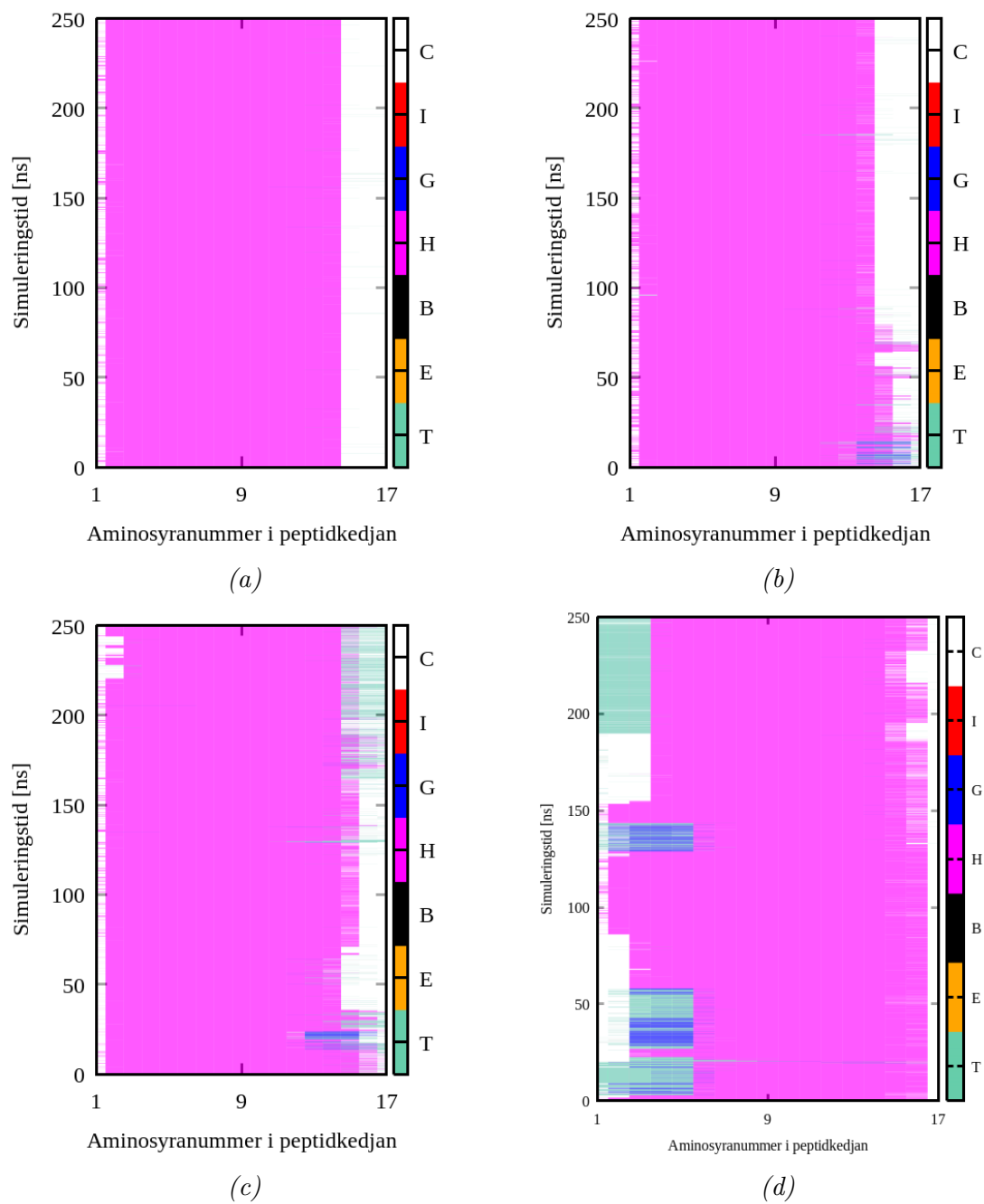
Figur 67: Sekundärstruktur för 2LD2-peptider med 6N8C-peptider A-D, trehalos och 2400 vatten (E2). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.



Figur 68: Sekundärstruktur för 2LD2-peptider med 6N8C-peptider E-H, trehalos och 2400 vatten (E2). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.



Figur 69: Sekundärstruktur för 2LD2-peptider med 6N8C-peptider A-D och 4800 vatten (E3). Figur a) visar för peptid A, b) peptid B, c) peptid C, d) peptid D.



Figur 70: Sekundärstruktur för 2LD2-peptider med 6N8C-peptider E-H och 4800 vatten (E3). Figur a) visar för peptid E, b) peptid F, c) peptid G, d) peptid H.