Eurographics Conference on Visualization (EuroVis) 2019 M. Gleicher, H. Leitte, and I. Viola (Guest Editors)

Analysis of Long Molecular Dynamics Simulations Using Interactive Focus+Context Visualization — Supplement —

J. Byška^{†1,2}, T. Trautner^{1,3}, S. M. Marques^{4,5}, J. Damborský^{4,5}, B. Kozlíková², and M. Waldner³

¹Department of Informatics, University of Bergen, Bergen, Norway; ²Faculty of Informatics, Masaryk University, Brno, Czech Republic; ³Faculty of Informatics, TU Wien, Vienna, Austria; ⁴Loschmidt Laboratories, Masaryk University, Brno, Czech Republic;

⁵International Clinical Research Center, St. Anne's University Hospital, Brno, Czech Republic;

In this supplementary document, we present the reports of two additional case studies conducted individually by the protein engineering expert. For these two case studies, the same data sets are used as in the paper, comprising 100,000 and 50,000 simulation steps, respectively. The first case study reports insights about the MD simulation of the wild type of DhaA haloalkane dehalogenase (DhaAwt) with 1,2-dichloropropanol (DCP). The second case study was performed again on the MD simulation of the wild type of DhaA haloalkane dehalogenase (DhaAwt), which interacted with three molecules of 1,2,3-trichloropropane (TCP). The text below is the unmodified observation report. For a better understanding, we added figure captions.

1. Case Study I: DhaAwt + DCP

By observing the MD simulation of DhaAwt with the DCP and chloride ion (Cl-) products inside the active site, it was possible to find several interesting facts:

In the first part of the simulation, while still in the active site, DCP and Cl- are involved in an H-bond network for a long part of the time while in the active site, either with direct bonds (see Figure 1) or mediated by water molecules (Figure 2).

DCP is also involved in a H-bond network with the catalytic D106 (numbered as D103 in this simulation), mediated by waters (Figure 2). Such networks may be the main reason why DCP and Cl- remain in the active site for such a long time.

At least while DCP is still in the active site, the number of waters in the active site increases, and they come from the p2 (Figure 3: waters appearing from the left) and p3 tunnel (Figure 4). This seems a very important observation, because waters may be involved in the release or Cl- and DCP. This was never observed before.

By checking the derivative of waters (Figure 5), one could observe that a large increase resulted in the drifting of DCP from the tunnel. I observed that DCP was in the tunnel interacting with the Cl- ion, and then many waters surrounded it and disrupted the interactions, and after that DCP moved quickly to the tunnel mouth.

© 2019 The Author(s)

Computer Graphics Forum © 2019 The Eurographics Association and John Wiley & Sons Ltd. Published by John Wiley & Sons Ltd.

DCP rebinds to the main tunnel after roaming in the bulk solvent. While trying to find the factors leading to such rebinding, it seems that at that moment there were many water molecules in the tunnel and around the ligand. Then, promoted by the hydrophobic interactions of the protein residues with the hydrophobic part of the ligand, and by the H-bonds of DCP with the water molecules in the protein tunnel (Figure 6), DCP may have been "encouraged" to move into the tunnel. This is a potentially important mechanism to lead the ligand to rebinding into the protein tunnel.

In the second part of the simulation, after DCP was rebound and well inside the protein's main tunnel, the entering of water molecules was done mostly from the p1 tunnel. In this context, the water interactions with the ligand (Figure 7) may have contributed to increase the number of water coming inside the protein.

Globally, this tool helped us drawing some important conclusions regarding the unbinding and rebinding of the DCP and Clproducts from the protein DhaA, and the involvement of water molecules in that process. We suspected that the waters might have an influence on the release of those ligands, but this hypothesis we never tested in a robust way. Even more difficult would have been the visual analysis of the role of waters, since the currently used software tools (e.g., VMD) are extremely unfriendly when dealing with such large data. This tool brings a robust and user friendly way to visualize molecular dynamics simulations, and observe the potentially relevant events. From such visualization one can draw some hypotheses regarding causality or simultaneity (correlated events). Then, by iteratively using the embedded functions or their modified counterparts (inverse and derivative functions), one can quickly test those hypotheses.

2. Case Study II: DhaAwt + 3 TCP

By observing the MD simulation of DhaA with three molecules of the TCP substrate in the outside of the protein, it was possible to find several interesting facts:

At some point, the protein could accommodate more than one molecule of TCP inside (Figure 8). These ligands are hydrophobic,

and seem to be attracted towards the protein mostly due to the favorable hydrophobic interactions with some residues or with each other. This fact was already known before, and it was easily observed and confirmed with this tool.

The access of water molecules is very important in this molecular system, since it is one of the reagents involved in the catalytic hydrolysis of DCP. Initially we believed that the water molecules entered the active site mostly through the main tunnel (called p1), and this was confirmed here when no ligand was present inside the protein (Figure 9). However, when there were ligands within the enzyme, we observed that the water molecules could still access the active site through a "backside" tunnel (Figure 10). This was unexpected, since this tunnel is typically not among the one with highest throughput, and therefore is usually considered not relevant. These findings may change such view. The functions available in this tool that help tracking the water molecules inside the protein, together with the useful representations modes (with the spatial focus/context, plus the coloring schemes) highly facilitated our observations and conclusions.



Figure 1: DCP (balls-and-sticks) and Cl- (green) forming an H-bond with direct bonds.



Figure 2: DCP (balls-and-sticks) and Cl- (green) forming an H-bond mediated by water molecules (dark green balls-and-sticks).



Figure 3: Water molecules (balls-and-sticks colored according to the color scheme on the right) coming from the p2 tunnel to the active site around snapshot 10,000.

CAVER Analyst 2.0 BETA		Sector and	A Real Property lines	and the second second	A DECKE AND A D		
Eile View Visualization Structure	Tunnel <u>H</u> elp						
in i	: 🖙 🚡 🐔 🚳	🏠 📿 🚿	🎋 🥕 🎻 🌾 🏈	5 📌 👒 🏹 🗞	ንም 🌾 🖋 💽 🔯 🎲		
Structures Overview Structure	re Selections 🕷 🔲	😝 Visualization					Coloring % Freegidizationing
New None Atoms Residues Ch	nains SS						Global Settings
Selection3	Q 😕 🔮 👁 🚔 📓						Background
Freegidization - Focus Ele	Q 😕 🔮 🕐 🚔 🗷						Structures Selections Tunnels
Freegidization - Context El	Q 😣 🔮 🗶 🗎 🗷						Selection: Freegidization + What: All + Color
Selection7	Q 😣 🔮 👁 🗎 📓						Color: Residues v By: Watergate
B Selection8	् 😣 🔮 💋 🚔 🔳						Scheme: v Reset Save Delete
						/	Unknown
					\sum		OUTSIDE
							ACTIVE
							INSIDE
							TO ACTIVE
							FROM ACTIVE
		N ^x					
Dynamic Playback 🕷							-
Min Speed: 6 🚔 v Max Speed: 100 🐇	Aggregation NORMALIZED	Functions:	VATER AROUND LIGAND 1: 294 DCP VATER AROUND LIGAND 2: 295 CL VATER COUNT ACTIVE VATER COUNT ALL	Active Function: WATER	count all 🔹		
	0,000 15,000 20	,000 25,000	30,000 35,000 40.	45,000 50,000	55,000 60,000 65,0 TIME	00 70,000 75,000 80,00	a 86.000 00.000 05.000 100.000
Structure Dynamics (dhaawt-dcp)) 🛿 Console Structur	e Sequence					-
		Curren	t frame: 12,34; - / 100000 Anim	ation step: 1 🚔 frames	Animation speed: 6 ≑ fps 👿 Re	peat animation loop 👿 Smooth animation	
From frame: 1 / 100	000	0					To frame: 100.000 🗂 / 100000
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		u	[TO HONKE 200700 V / 200000

Figure 4: Water molecules coming from the p3 tunnel to the active site around snapshot 12,000.

CAVER Analyst 2.0 BETA	CONTRACTOR OF THE OWNER.		
File View Visualization Structure Tunnel Help			
📨 🛋 🛋 🖏 🍇 🖙 🔚 🏍 🍪 🛆	2 📿 🧐 🎋 A 🛷 🥒 P 🤌 🖅 🗲 💚 🗡 👒 💖	** 🕫 💽 🖬 🔅 🎡 🛄	
Structures Overview Structure Selections 🕷 🖃 🤗	Visualization	Coloring # I	Freegidizationing
New None Atoms Residues Chains SS		Global Settings	8
🔋 Selection3 🔍 🔁 🔮 👁 🚔 📓		Background	
🛐 Freegidization - Focus Ele 🔍 🐉 🔮 👁 🔒 📓	۹۹	Structures Se	ections Tunnele
🔄 Selection6 🔍 🐉 🔮 👁 🔒 📓			
🔋 Selection7 🔍 🔀 🕘 👁 🔒 🕱		Selection: C	ataluc_res V What: All V COOR All COOR Active
🔋 Freegidization - Context El 🔍 😫 🔮 🕢 🗃 📖		Color: Atoms	fult
🗟 catalitic res	8	Science: pe	Reset Save Deete Export Import
		H (1, Hydroge	en)
		C (6, Carbon)	
		N (7, Nitroger	0
		0 (8, Oxygen	
Dynamic Playback 🕸			
Min Speed: 4 Aggregation NORMALIZED	Functions: WATER COUNT ALL Active Function: WATER COUNT	T ALL 👻	
Max Speed: 100 🗧 🔄 Smooth	WATER COUNT INSIDE		
	WATER COUNT TO ACTIVE		
¥1			
E		E DAM I AL AMARANA LAA I AA I	2.5 8
	PPAN'N MANANA U INVINOPED A MAILENA MARKANINA MAKANA UMAKANA UMAKANA UMAKANA UMAKANA UMAKANA UMAKANA UMAKANA UM	mun mun when the south man month man	mannin lann will what on Z
20,000 10,000 15,000 20,00	3 25,000 30,000 35,000 40,000 45,000 50,000 TIME	88,000 80,000 88,000 70,000 78,000 1	80,000 85,000 90,000 95,000 100,000 >
Structure Dunamics (dhaaud dan) 10 Consola Structure S			
Structure bynamics (unaawi-ucp) # Console Structure S	Current frame: 20,35: + / 100000 Animation step: 1 + frames Anima	ion speed: 4 🜩 fps 👽 Repeat animation loop 👽 Smooth animatio	n n
From frame: 1 - 1/100000			To frame: 100,000 🚔 / 100000

Figure 5: Increase of the number of waters causes DCP (top right) to drift from the tunnel around snapshot 20,000.

CAVER Analyst 2.0 BETA	_ 0 ×
Elle View Visualization Structure Tunnel Help	
De D	
Structures Overview Structure Selections # 🖬 🛃 Visualization	Coloring - Selection8 Freegidizationing #
New None Atoms (Residues) Chains SS	
🖹 Selection3 🔍 🔅 🔮 🎱 🖩	Evaluate Temporal Spatial Focus: Water ACTIVE, INSIDE,
🕼 Freegidization - Focus Ele 🔍 🖗 🔮 🍙 🗃	Evaluate Spatial Spatial Context: Ligand Surroundings 5.0 💌
🗑 Freegidization - Context EL. 🔍 🗟 🔮 🕋 🗃	Load Turnel Data Highlight Ligand: 295 Cl -
Image: Selection 7 Q Image: Market and the selection for the s	
📝 Selection8 🔍 🕄 🖉 🖉 🗎 🔟	DHAA_WT 👻
	(V) Ghosting
	Motion Blur: 1.1 max min
Dynamic Playback 10	
Min Speed: 2 Aggregation HORMALIZED * Functions: WHER RACIAD LIGAND 1: 2940C A Active Function: ACTIVE SITE DISTANCE 1: 2 *	
Hax Speed: 100 🐨 Smooth WATER COUNT ALL V COUNT ALL	
Structure Dynamics (dhaawt-dcp) # Console Structure Sequence	
Overent frame: (66.55 🖾 / 10000) Animation stars. 💷 frame: Animation meant: 21 🖄 for: (7) Denest animation loss (7) Second animation	
From trame: 1 - /100000	To frame: 100,000 + / 100000

Figure 6: DCP is attracted by the residues (sticks) and waters (uniformly colored balls-and-sticks) to enter the tunnel around snapshot 66,500.

CAVER Analyst 2.0 BETA	_		and the second second							_ 0 ×
File View Visualization Structure	Tunnel Help									
in i	📑 🚡 🐔 🐗	🏠 🔘 🚳 🍕	e 🎤 🛷 🥒 🎸) 🔊 🧚 🖓 Y	***	💽 🖸 🝰 🎡 🛄				
Structures Overview Structure	e Selections 📽 🖃	😂 Visualization					< > v	Coloring Freeg	idizationing	
New None Atoms Residues Cha	ains SS							Global Settings		
Selection3	Q 🥵 🔮 👁 🔒 🗷							Background		
Freegidization - Focus Ele	् 😼 🔮 👁 🔒 🗏		7					Structures Selection	ns Tunnels	
Freegidization - Context El	् 🤔 🔮 🕲 🚔 📖							Selection: Freegi	lization 👻 What:	All - Color
Selection7	् 🥵 🔮 👁 🚔 📓							Color: Residues		- By: Watergate
Selection8	् 😣 🥥 💋 🗎 🗵							Scheme:	¥	Reset Save Delete
			4					Unknown		
)		OUTSIDE		
			J					ACTIVE		
								INSIDE		
								TO ACTIVE		
		D ×								
Dynamic Playback 8					,					
Min Speed: 6 🔔 / Max Speed: 100 🚖	Aggregation <u>NORMALIZE</u> Smooth	D Functions: WA WA WA	ER AROUND LIGAND 1: 294 DC ER AROUND LIGAND 2: 295 CL ER COUNT ACTIVE ER COUNT ALL	Active Function: Derivative [WATER COUNT ALL	•				
	AAA AA, AA AA AA AAAAAAAA AA AA AA AA AA	0,000 25,000	30,000 35,000 ·	40,000 45,000	50,000 55,000 TIME	60,000 65,000		000 85,000	90,000 95,000	1.0 R 0.5 COUNT 100,000
Structure Dynamics (dhaawt-dcp)	S Console Structu	ire Sequence								
		Current fr	ame: 93,374 🚔 / 100000 🗛	nimation step: 1 👘 fra	ames Animation speed:	6 🚔 fps 💟 Repeat anima	ation loop 📝 Smooth animation			^
From frame: 1 🚔 / 1000	000								To frame:	100,000 🔷 / 100000 🗉

Figure 7: Ligand interacting with water molecules inside the main tunnel around snapshot 93,000.

CAVER Analyst 2.0 BETA				_	_		_ 0 ×
File View Visualization Structure Tunnel Help							
📨 🚎 🚎 🏹 🇞 🎼 🏹 🐔) 🏠 📿 🛸 🎋 🎋 🖉 🖉	S 🗚 😳 Y 🗞 💖 🠇 🖋 🌔	🛄 🏖 🖬 🥘				
Structures Overview # Structure Selections	😣 Visualization				Coloring - Freegidization -	Context Elem	Freegidizationing 🕷 🖃
Activate All Invert Active Remove Active Remove All							
👔 dhaawt_3tcp 🛛 🔍 🐉 🥥 💥 🚔 🐼 🛞		- r v			Evaluate Temporal	Spatial Focus:	All Ligands 🗸
			X		Evaluate Spatial	Spatial Context:	Ligand Surroundings 5.0 👻
			X		Load Tunnel Data	Highlight Ligand:	296 TCP 👻
					DHAA_WT 👻		
			R		V Ghosting		
	1 designed and the second s				Motion Blur: 1.1	mi	n
Dynamic Playback 🕷							=
Min Speed: 2 & Aggregation NORMALIZE Max Speed: 200 & Smooth	D Functions: ACTIVE SITE DISTANCE 1: 294 TOP ACTIVE SITE DISTANCE 2: 295 TOP ACTIVE SITE DISTANCE 3: 296 TOP LIGAND SPEED 1: 294 TOP	Active Function: ACTIVE SITE DISTANCE 3: 2 Derivative Invert	• •				
	10,000 12,000 14,000 16,000 18,000 20,0	000 22,000 24,000 26,000 28,000 TIME	30,000 32,000 34,000 3	6,000 38,000 40,00	0 42,000 44,000 44	.000 48.000	25 50,000 52,000 22 20 50,000 52,000 22 20 50,000 52,000 22
Structure Dynamics (dhaawt_3tcp) # Console Struct	ture Sequence						
	Current frame: 38,89€ ↓ 50000 Animat	tion step: 1 🚔 frames Animation speed:	7 🕂 fps 📝 Repeat animation loop	Smooth animation			^
5 6 4 A (5000				0		T- 6	
rrom trame: 1 v / 50000				U		To frame:	50,000 - / 50000 =

Figure 8: All three TPC molecules (balls-and-sticks) inside protein at once around snapshot 39,000.

CAVED Analyst 20 BETA	
Tel View Visulization Structure Tunnel Help	
a a a a a a a a a a a a a a a a a a a	
Structures Overview Structure Selections 🕷 🕞 🦂 Wsualization	Coloring * Freegidizationing
New None Atoms Residues Chains SS	Global Settings
🗑 Selection3 🔍 🔁 🖉 🎱 👜 🖩	Background
🗑 Freegidization - Focus Ele 🔍 🔁 🖉 👁 🚔 🖩	Structures Selections Trunale
🕼 Freegidization - Context EL. 🔍 🐉 🔮 🚇 📓 🖩	
🖲 Selection7 🔍 🖗 🔮 🌰 🖩 🖩	Selection: Preeguization V What: All V Color
	Color: Residues By: Watergate
	Science: V Reset Save Delete
	Unknown
	OUTSIDE
	ACTIVE
	INSIDE
Dynamic Playback 🕷	
Min Speed: 6 Aggregation NORMALIZED V Functions: WATER ARCUND LIGAND 1: 294 DCP A Active Function: WATER COUNT ALL V	
Max Speed: 100 * Smooth WATER COURD LIGHTU 2/25 L	
WATER COUNT AL	
	10 70
	-0.5 8
	80,000 85,000 90,000 95,000 100,000
Time	Ē
Structure Dynamics (dhaavt-dcp) 🖩 Console Structure Sequence	
Current frame: 12,341 / 100000 Animation step: 11 + frames Animation speed: 61 + fos ♥ Repeat animation loop ♥ Smooth anima	tion
	To frame: 100.000 - / 100000 =

Figure 9: Water molecules are using the p1 tunnel when no ligand is inside the protein around snapshot 12,000 from the first MD simulation.

CAVER Analyst 2.0 BETA							
<u>File View Visualization Structure</u>	Tunnel <u>H</u> elp						
i 🖉 🚘 🛋 🗞 💸	🕞 🕞 🐔 🏀 🛵	😡 🧐 🎋 🖉 🥔	🤗 🗲 🧚 🐚 🦙 👒	🔊 🛠 🥐 💽 🖬 👪			
Structures Overview Structure	e Selections 🕷 🔲 属 😼 Vis	ualization			4 > -	Coloring - Freegidization - Focus	Ele 📽 Freegidizationing 🖃
New None Atoms Residues Cha	ains SS					Global Settings	
Selection1	a 🔁 🕑 🖉 🚔					Background	
Freegidization - Focus Ele	् 🔧 🔮 👁 🚔 🗷					Structures Selections Tunnels	
Freegidization - Context El	् 😣 🔮 👁 🚔 國					Selection: Freegidization +	What: All - Color.
Selection4	Q 😣 🔇 👁 🚔 📓					Color: Residues	▼ By: Watergate
Selection6	् 🔧 🔮 👁 🚔 🗷					Scheme:	v Reset Save Delete
					\mathcal{O}	Unknown	
					\prec	OUTSIDE	
						ACTIVE	
						INSIDE	
						TO ACTIVE	
						FROM ACTIVE	
	Ž	×					
Dynamic Playback 🕷							
Min Speed: 2 📮 🖌	Aggregation NORMALIZED •	Functions: WATER AROUND LIGAND 2: 295	TCP Active Function: WATER	COUNT ACTIVE			
Max Speed: 200 📩	Smooth	WATER AROUND LIGAND 3: 296 WATER COUNT ACTIVE WATER COUNT ALL	Derivative Inve	rt			
	6.000 8.000 10.000	12,000 14,000 16,000 18,000	20,000 22,000 24,000 2	6.000 28.000 30.000 32.00 TIME	0 34,000 36,000 38,000 -	0.000 42.000 44.000 48.000	48,000 50,000 52,000 ACT
Structure Dynamics (dhaawt_3tcp) 8 Console Structure Se	quence					
		Current frame: 19,13: + / 50000	Animation step: 1 🚔 frames	Animation speed: 2 🜩 fps 🔍 R	epeat animation loop 🔽 Smooth animation	1	^
From frames							To frames 50 000 1 (50000
1 1 1 5000]		TO TRAINE: 50,000 - 750000 =

Figure 10: When a ligand is present inside the protein (balls-and-sticks, top), water molecules (balls-and-sticks using the color scheme on the right) are using another tunnel around snapshot 19,000 of the second MD simulation.