Supporting Information for Molecular Dynamics Simulations Accelerated on FPGA with High-Bandwidth Memory

Jing Xiao[a,b,c], Jinfeng Chen[b,c,d], Ye Ding[b,c], You Xu[b,c], and Jing Huang*[b,c,d]

[a]Fudan University, Shanghai, 200433, China
[b]State Key Laboratory of Gene Expression, School of Life Sciences, Westlake University, Hangzhou, Zhejiang 310030, China
[c]Westlake AI Therapeutics Laboratory, Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang 310024, China
[d]Institute of Biology, Westlake Institute for Advanced Study, Hangzhou, Zhejiang 310024, China

November 2, 2025

*Corresponding author: huangjing@westlake.edu.cn

Contents

1 Illustration of basic conception in MD simulations.

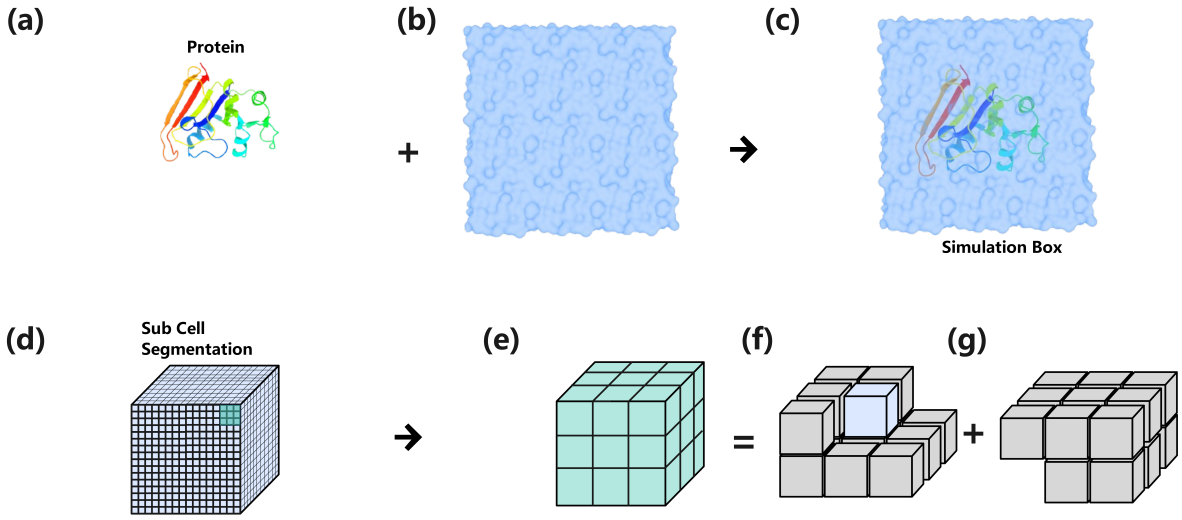

Fig. S1: (a) Three-dimensional protein structure of dihydrofolate reductase (DHFR). (b) Initial configuration of a cubic water solvation box. (c) Topology of the cubic simulation system with PBC. (d) Spatial decomposition scheme for neighbor-cell identification: Central home cell (highlighted in red). (e) 26 neighbor cells categorized by spatial orientation. Axially adjacent cells partitioned into upper and lower hemispheres. (f) Spatial distribution of 13 neighbor cells in the upper hemisphere. (g) Spatial distribution of 13 neighbor cells in the lower hemisphere.

2 Neighbour list diagram.

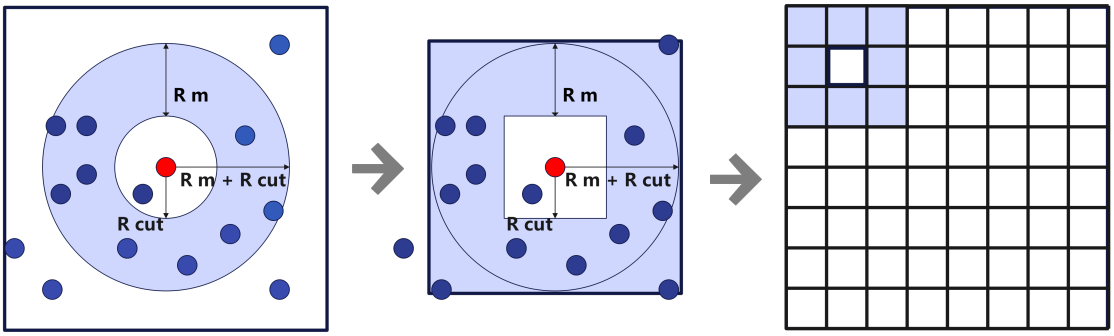

Fig. S2: During MD simulations, the center particle and the surrounding particles can be divided into equal sized squares according to the values of R_m parameter and R_{cut} cut off.

3 Performance Comparison of Accelerator Architectures.

Table S1: Performance Comparison of Accelerator Architectures

Metric	F37X (This Work)	NVIDIA T4 [a]	Xeon 8380 [b]
Compute Density (TOPS/W)	0.37	0.21	0.05
Memory Bandwidth (GB/s)	460	320	204
Programmability Level	RTL	CUDA	ISA Extensions

[a] *NVIDIA A100 Tensor Core GPU Architecture.* https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/a100/pdf/nvidia-a100-datasheet-us-nvidia-1758950-r4-web.pdf

[b] *Intel Xeon Platinum 8380 Processor Specifications.* https://ark.intel.com/content/www/us/en/ark/products/212287/intel-xeon-platinum-8380-processor-54m-cache-2-30-ghz.html

4 Dual-mode parallel data caching architecture in HBM-based MD simulations

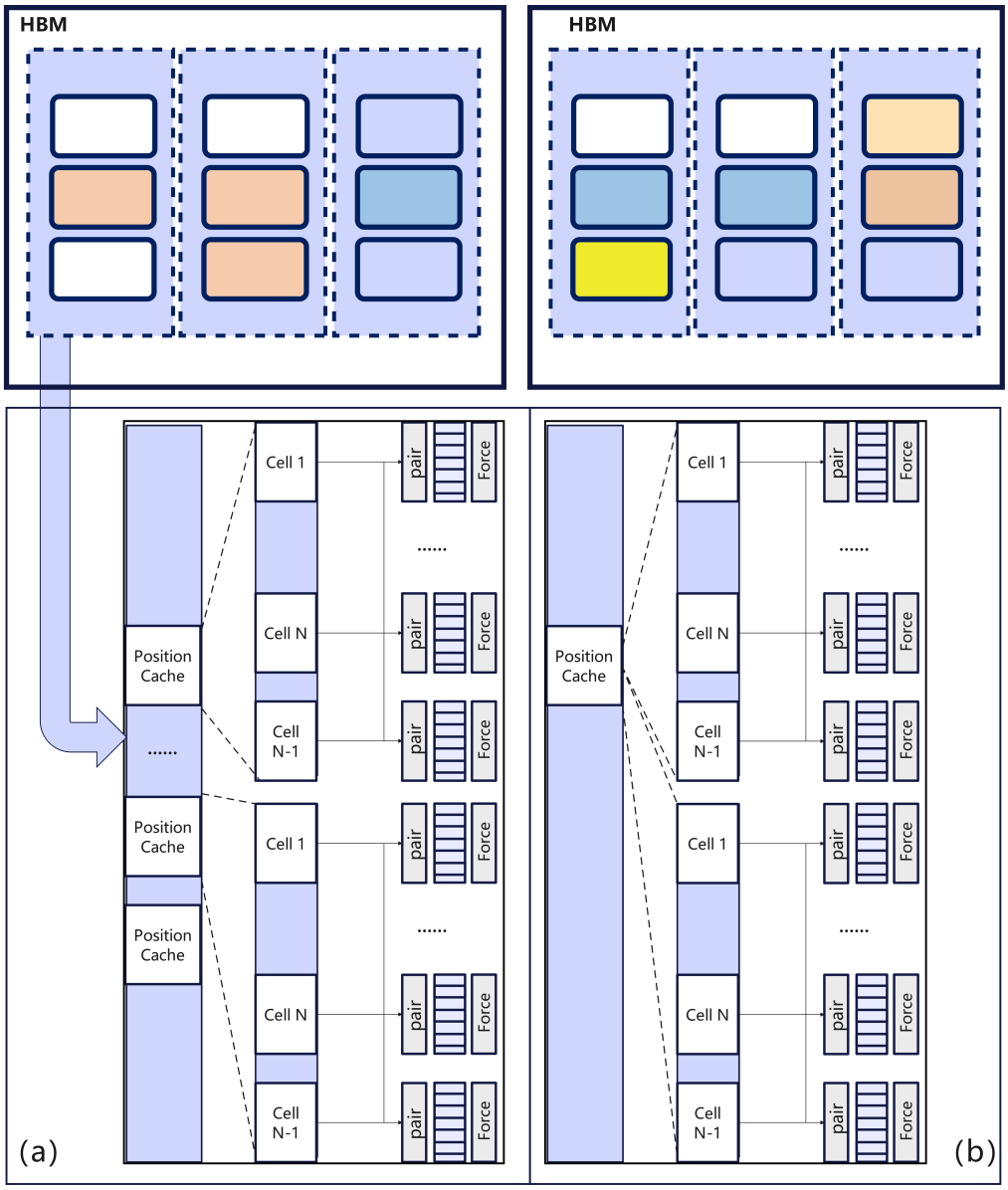

Fig. S3: Dual-mode parallel data caching architecture in HBM-based MD simulations. (a) Co-address caching: All channels access identical memory blocks (highlighted in blue). Position Cache modules synchronize particle data through shared bus (dashed lines). Force calculation units receive identical Cell data streams (Cell 1 to N). (b) Distributed-address caching: Channels access non-overlapping memory regions (color-coded orange/yellow). Independent Position Caches maintain local particle datasets. Force units process spatially partitioned Cell groups (Cell 1 to N per channel).

5 Resource utilization rates for bonded force module.

Table S2 summarizes the resource usage of the bonded force module. On the Xilinx Virtex UltraScale+ VU37P platform, the module exhibits well-balanced resource consumption: LUTs and FFs are utilized at 15% and 14%, BRAM and URAM at 24% and 13%, respectively. DSP usage is low (3%), and clock management resources (BUFG/MMCM) remain modest at 2% and 8%. These figures suggest that the bonded force module is relatively lightweight and efficiently integrated into the overall design.

Table S2: Resource utilization rates for bonded force module implemented on Xilinx Virtex UltraScale+ VU37P chip through layout routing

Resource Name	Utilization%
LUT	15
LUTRAM	4
FF	14
BRAM	24
URAM	13
DSP	3
IO	1
GT	17
BUFG	2
MMCM	8
PCIe	17

6 Validation: total LJ interaction energy calculation.

Table S3: Total LJ interaction energy calculated for a system of fifteen argon atoms on five structures

	Potential Energy (kJ/mol)		
	FPGA	CPU	GPU
Frame 1	0.923459	0.922628	0.922628
Frame 2	-2.184186	-2.182222	-2.182222
Frame 3	-3.207323	-3.204440	-3.204439
Frame 4	-1.750265	-1.748692	-1.748693
Frame 5	-1.624073	-1.622612	-1.622612