



Grant agreement no. 675451

CompBioMed

Research and Innovation Action

H2020-EINFRA-2015-1

Topic: Centres of Excellence for Computing Applications

D2.1 Report on Application Software Readiness and Fast Track Exploitation

Work Package: WP2 - Biomedical Research Activities

Due date of deliverable: Month 12

Actual submission date: 29 / Sep / 2017

Start date of project: October 01, 2016 Duration: 36 months

Lead beneficiary for this deliverable: *BSC*
Contributors: *UOXF, UNIGE, USFD, UPF, ACE, UCL, UvA*

Project co-funded by the European Commission within the H2020 Programme (2014-2020)		
Dissemination Level		
PU	Public	YES
CO	Confidential, only for members of the consortium (including the Commission Services)	
CI	Classified, as referred to in Commission Decision 2001/844/EC	

Disclaimer

The content of this deliverable does not reflect the official opinion of the European Union. Responsibility for the information and views expressed herein lies entirely with the author(s).

Table of Contents

1	Version Log	3
2	Contributors	3
3	Definition and Acronyms	4
4	Introduction.....	5
5	Activities Carried Out.....	7
5.1	General overview: A proposal for a multidimensional taxonomy	7
5.1.1	Description	7
5.1.2	Research Exemplars.....	10
5.2	Tasks' outcomes.....	15
6	Bibliography / References	21
7	Conclusions.....	23
8	Appendix.....	24
8.1	Appendix A. CompBioMed software stack.....	24
8.2	Appendix B. Selected publications.....	26

List of Tables and Figures

Figure 1.	The multiple hierarchical levels of modelling and simulation that arise in computational biomedicine. The axes denotes the spatial scale (“Scale levels”) , the physiological level of organisation (“Organisational levels”), and the computational domain of modelling and simulation (“Action levels”).	6
Figure 2.	CompBioMed layered software structure. The layers are different stages on the computational workflow, for intance: Data Ingestion, Pre-process, Simulation, Data analysis, Post-process, Visualization, etc. The interfaces are the linking pieces of software between one layer and the next one.	9

1 Version Log

Version	Date	Released by	Nature of Change
V1.0	25/AUG/2017	Mariano Vázquez	First Draft
V1.1	08/SEP/2017	Mariano Vázquez	Internally reviewed draft
V1.2	15/SEP/2017	Mariano Vázquez	internally reviewed document

2 Contributors

Name	Institution	Role
Mariano Vázquez	BSC	Author
Robin Richardson	UCL	Author
Alfons Hoekstra	UvA	Author
Peter Coveney	UCL	Author
Francesc Levrero	UOXF	Author
Blanca Rodríguez	UOXF	Author
Bastien Chopard	UNIGE	Author
Jonas Latt	UNIGE	Author
Xinshan Li	USFD	Author
David Wright	UCL	Reviewer
Bastien Chopard	UNIGE	Reviewer
Peter Coveney	UCL	Reviewer
Emily Lumley	UCL	Reviewer

3 Definition and Acronyms

Acronyms	Definitions
HPC	High Performance Computing
FT	Fast Track
DT	Deep Track
WP	Work Package
LBM	Lattice-Boltzmann Methods
FD	Finite differences

4 Introduction

CompBioMed's WP2 deals with the Biomedical Research Activities of the project and is coordinated by Barcelona Supercomputing Centre (BSC). This work package performs computational and translational biomedical research, focused in three biomedical exemplar domains: cardiovascular, molecularly-based and neuro-musculoskeletal medicine.

The main WP2 objective is developing and advancing the start-of-the-art in biomedical modelling and simulation from the desktop to the most powerful HPC (Tier-0) systems, with an emphasis on the use of multi-petaflop HPC environments. WP2 is developing strategies, including new algorithms, to ensure that current and future high fidelity biomedical codes will perform optimally on emerging exascale and other novel architecture. Codes installed and maintained on supercomputing centres, profusely documented and tested will assure an extensive reach for biomedical researchers. The idea is to demonstrate how verified and validated (multiscale) modelling and simulation can impact industrial and clinical research and decision-making, leading to innovation in the healthcare sector.

WP2 strongly interacts with WP5 Resource and Infrastructure Support, a WP that is concerned with porting, deployment, and optimisation of the CompBioMed codes to and on supercomputing resources, and **with WP6 Empowering Biomedical Applications**, which facilitates the development and execution of complex workflows.

WP2 Work Scheme

The work package goal is to establish a layered Applications Pipeline, consisting of a set of simulation and data analysis codes, to perform simulations within the three aforementioned research exemplar domains. WP2 work is structured around two main concepts: the fast track (FT) and the deep track (DT), which apply to all three selected exemplars and are defined as follows:

- **Fast Track:**

The Fast Track (FT) builds on the existing capabilities of the partners in the use of HPC for modelling and simulation. The FT is focused on HPC-based methods that already handle multi-physics and multi-scale features to produce integrated high fidelity personalised human models. While the efficient use of HPC resources for the individual codes will be a major priority, we wish to emphasise that the FT also focuses on providing seamless integration interfaces and strategies to enable component codes to be deployed within workflows (which are addressed by WP6). In this respect, by linking together codes developed by different partners, the CompBioMed Centre of Excellence will have a much wider range of activity than merely supporting its individual partners' own particular research areas.

For the FT, the work package end-users have identified concrete problems that can be attacked with relatively minor adaptations to the simulation codes of the work package developers. This allows CompBioMed to address key biomedical challenges from the start of the project using the existing e-infrastructure.

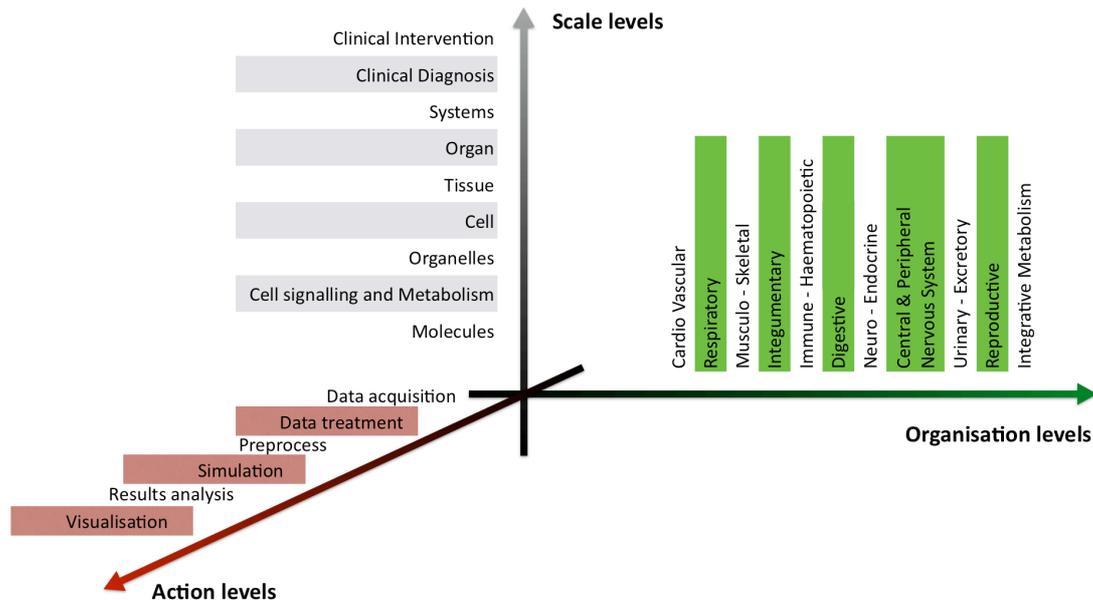


Figure 1. The multiple hierarchical levels of modelling and simulation that arise in computational biomedicine. The axes denotes the spatial scale (“Scale levels”), the physiological level of organisation (“Organisational levels”), and the computational domain of modelling and simulation (“Action levels”).

- **Deep Track:**

The Deep Track (DT) addresses additional medium-term challenges. DT extends the capabilities available from the FT to other parts of the “level matrix” shown in Figure 1, providing new capabilities where needed, addressing more complex requirements that require a deeper and longer level of development. What is studied thanks to the Fast Track will be recycled and projected into new areas through the Deep Track. The standardised interfaces and integrative tools designed in the FT will be extended to meet the needs of different groups, including clinicians, researchers and industrial users, helping to compose pipelines which can then be executed on HPC and other computing resources.

In the DT, we will upscale the software towards the exploitation of future (exascale) HPC environments, through extreme scaling of monolithic codes, where relevant, and through the design and development of workflows and coupling interfaces for multiscale and multiphysics applications that can be deployed on such machines. These activities will require close attention to be paid to the detailed, heterogeneous nature of the HPC resources.

5 Activities Carried Out

In this section, we will describe the main results obtained so far within the Work Package 2. We will start by a general technical description of the results, which inform the CompBioMed software infrastructure design and implementation. Next, we go over each of the WP Tasks and its scientific and technical outcome.

5.1 General overview: A proposal for a multidimensional taxonomy

5.1.1 Description

During the first 12 months, an initial prototype of the CompBioMed software infrastructure has been designed, profiled and installed. The **CompBioMed software infrastructure** is the centre's software stack which will be deployed in the supercomputing centres partnering CompBioMed, including workflow control software, dataset examples and documentation. This software infrastructure will be maintained by CompBioMed and free of access for the user's community.

To begin with, the applications supported by CompBioMed can be described in many ways. We propose three different points of view, relevant for the scope of the CoE, a user-driven centre of excellence. These different points of view address the following aspects:

- *Motivation for using HPC*: what are the reasons, if any, for using HPC resources?
- The *user groups* for CompBioMed applications: what classes of users do the applications target?
- *The application structure in terms of components*: what is the current development status of the application, piece by piece? Can all these components be installed freely (or not) in all supercomputing centres?

Let us analyze individually each point of view.

Motivation for using HPC

According to our long experience, there are a several drivers related to an increased complexity requirement, that prompt biomedical researchers, who are already users of computational tools, to become users of HPC:

- Multi-physics simulations: coupled models, where several physical, chemical or physiological models are tightly linked together. Each of the models can be computationally very expensive, with the cost of the coupled solution at least the sum of these constituents.
- Multi-scale simulations: as above, but now coupled models, where each model solves a different time or space scale (molecules, cells, tissue, organs, etc.).
- Coupled simulations workflows: model coupling can be very complex, involving complicated branches and conditionally executed steps.
- Pre-process and model setup: the simulation scenario preparation can require a strong computational effort.

- Post-process and results analysis: to explore very large result datasets needs a highly efficient computational effort. This analysis can include visualization tools deployed in large-scale parallel computers.

User groups

Users from the biomedical realm in need of HPC resources come from the following groups:

- Academic research: these users generally require large-scale runs to test their models beyond their lab capabilities. Eventually, these runs can be at the upper-end of the current computational capabilities of Tier-0 facilities.
- Industry: these users predominantly require many smaller size runs to improve their products, rather than massive runs.
- Clinical trials: are similar to industry users in their needs.
- HPC for *in silico* medicine services: this fourth group refers to startups that offer all-around simulation services to industries, and are either too small to own their own computational environment or are happy to sell Software as a Service deployed on third-parties' hardware.

The application structure in terms of components

This third point of view has two main objectives. Firstly, it allows us to see if there is scope to reuse software components across applications. Secondly, from an operational standpoint, it allows CompBioMed to analyze what to deploy and how to do it.

The CompBioMed software pipeline (Figure 2) goes from medical observation and clinical data to numerical simulation and analysis, aiming to make it an aid to decision making. The pipeline can be imagined as a layered structure:

- Layer 1 - Data processing: signal processing, image processing, etc.
- Layer 2 - Data Digestion: big data analytics, Kalman filtering, etc.
- Layer 3 - Model pre-processing: model generation, meshing, etc.
- Layer 4 - Simulation (Solvers): model solution
- Layer 5 - Model Post-Processing: scientific visualisation, results data analytics, reporting, Kalman filtering, etc.
- Layer 6 - Workflows: orchestrations of multiple models to capture a multi-step, multiscale, or multi-physics structure of the application

Moreover, and according to the CompBioMed tracks concept, each layer will eventually have Fast and Deep Track components. It is worth remarking that although the aforementioned six layers can be identified in all domains, it is possible that some of the layers will not be present in all exemplars within the scope of CompBioMed. Therefore, for each application, the CompBioMed users and developers define the main HPC motivation, the main User Group, and the layered structure. As the project evolves, these three scopes should evolve towards new needs.

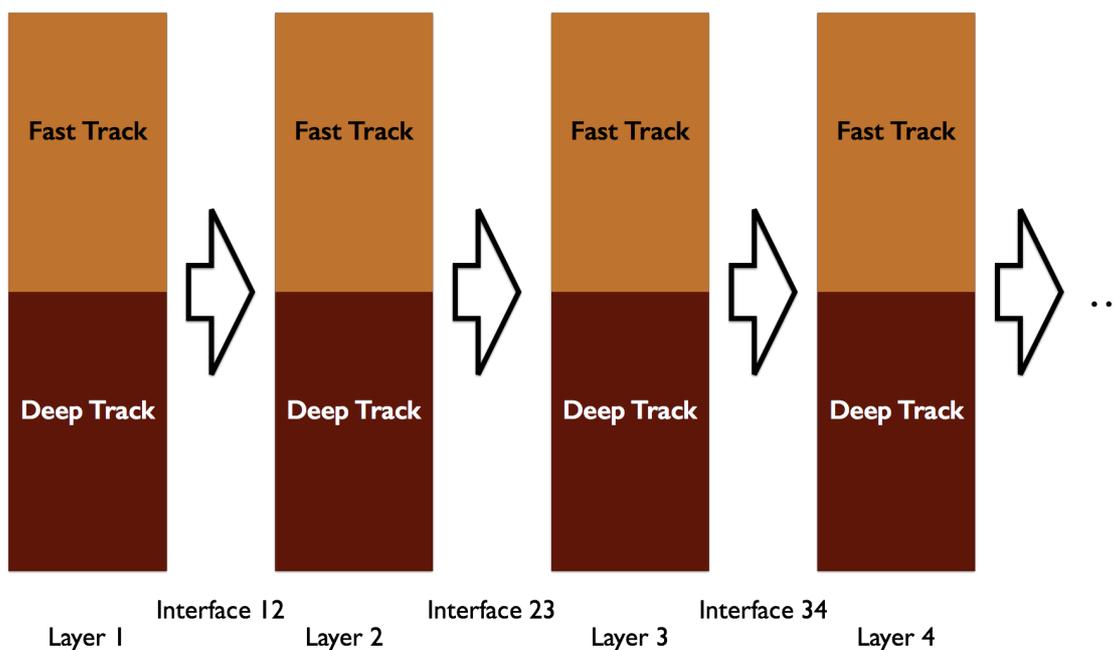


Figure 2. CompBioMed layered software structure. The layers are different stages on the computational workflow, for instance: Data Ingestion, Pre-process, Simulation, Data analysis, Post-process, Visualization, etc. The interfaces are the linking pieces of software between one layer and the next one.

The pipeline design and action depends on the research exemplar, individually described in the following sections. However, some common features can be identified:

- The pipeline is initially deployed at CompBioMed supercomputing centres: EPCC, SURFsara and BSC. In the future, we will explore the possibility of adding more sites, including other HPC providers (especially those that allow High Throughput Computing, requiring "urgent computing", account sharing, etc.) such as Amazon Web Services, Microsoft Azure, etc. It is important to remark that while initially the CompBioMed pipeline targets HPC resources available on academic supercomputing centres, on future stages it will probably be deployed "in the Cloud". As a matter of fact, some cloud providers already offer HPC capabilities (cf Microsoft's Azure HPC).
- CompBioMed includes in its web site a "Hub" (<http://www.compbioMed.eu/software-hub/>) that is a centralized directory to facilitate access to these resources, gathering links to the software developers sites, documentation on codes, datasets, etc.
- Layers and Interfaces will be composed of software owned by (or licensed to) CompBioMed either as a consortium or as individual partners. In the case of commercial software, the pipeline will be fully operational in those systems with the appropriate licences.
- Each Research Exemplar will define its own pipeline, whilst always trying to extract the common features to allow as much recycling as possible.
- Each layer could have a "FT" and a "DT".
- The FT is ready to go, with a relatively simple definition of the connecting pipeline and all the required software already deployed, being compiled with the appropriate performance optimization flags for each supercomputer.
- The DT requires longer-term research, which will be done within CompBioMed.

5.1.2 Research Exemplars

Cardiovascular Biomechanics Simulation

HPC-motivation:

Run large scale cardiac (coupled fluid-electro-mechanical) and vascular simulations (coupled fluid-mechanical)

Cardiac simulations are gaining traction in the study of the heart's pumping function, both healthy and under failure. This problem presents many difficulties, especially the extreme complexity of tight coupling between models from three different fields. Vascular simulations are much more mature, and are now capable of exploring fluid-structure interactions. In both cases, HPC-based simulations are a must.

User groups:

HPC for Researchers

Cardiovascular biomedical researchers usually develop computational tools that, from a larger perspective, can be seen as parts of a much larger model: a material model for arterial tissue, a cell model for electrophysiology, a coupling electromechanical model on a piece of cardiac tissue, etc. CompBioMed will allow them to test their models in a comprehensive simulation tool, capable of running large-scale cardiovascular simulations. For instance, a researcher who develops a cell model for cardiac myocytes can test his/her model on CompBioMed software, including highly parallel solvers to run cardiac electrophysiology on complex geometries with well-defined interfaces. This researcher will link his/her model to cardiac simulation and thus test it within a much larger environment.

HPC for clinical trials

CompBioMed tools will allow the simulation of clinical trial scenarios, for diagnosis, treatment planning and drug or medical device testing.

HPC for Industry

It is expected that cardiovascular biomedical related industries will become users of the CompBioMed simulation suite. Particularly Pharmaceuticals, CRO and Medical Devices Manufacturers.

Software structure:

Layer name	Within CompBioMed?
Data Processing: signal processing, image processing, etc.	YES
Data Digestion: big data analytics, Kalman filtering, etc.	NO
Pre-processing: model generation, meshing, etc.	YES
Simulation (solvers): model solution	YES
Post-Processing: scientific visualisation, results data analytics, reporting	YES
Workflows: orchestrations of multiple models to capture a multi-step, multiscale, or multi-physics structure of the application	YES

Data processing Layer

FT/DT: Segmentation, Co-registration, etc.

Software:

- VMTK (<http://www.vmtk.org/>): libraries and tools
- Amira (<https://www.fei.com/software/amira-3d-for-life-sciences/>): commercial software for image processing
- Matlab (<https://www.mathworks.com/products/matlab.html>): commercial software

Pre-processing Layer

FT/DT: Mesh generators for structured (LBM, FD,...) or unstructured (the rest) meshes.

Software:

- ICEM (<http://www.ansys.com/>): mesh generation commercial software
- ANSA (<https://www.beta-cae.com/ansa.htm>): mesh generation commercial software
- GiD (<https://www.gidhome.com/>): mesh generation commercial software
- IRIS (<http://www.bsc.es>): mesh generation (PI of BSC), open access

Simulation Layers:

FT: Single Physics - Single Scale

a) Vascular: CFD

Software:

- Palabos (<http://www.palabos.org/>): LBM simulation software (PI of UNIGE), open source
- Ficsion - HemoCell (<https://www.hemocell.eu/>): LBM-IB cell simulation software (the LBM part is Palabos, cells simulation part is from UvA), limited distribution
- HemeLB (<https://github.com/UCL/hemelb>): LBM simulation software (PI of UCL)
- Alya (<http://www.bsc.es>): Multi-physics simulation software (PI of BSC), open access

Simulation scenario examples:

Different vascular simulation scenarios are set up and run: coronary, carotids, aneurisms, thrombus formation, etc. LBM is specifically designed for smaller vessels, complicated geometry, and is being developed to deal with the interaction between blood cells and plasma.

b) Cardiac: Electrophysiology (EP)

Software:

- CHASTE (<http://www.cs.ox.ac.uk/chaste/>): EP simulation software (PI of Oxford), open source
- Alya (<http://www.bsc.es>): Multi-physics simulation software (PI of BSC), open access

Simulation scenario examples:

Cardiac electrophysiology simulations of different kinds: small pieces of cardiac tissue to test drug action and electrophysiology models, biventricular geometries, atria geometries, complete hearts or torso to compute ECG.

DT: Multi Physics - Multi Scale

a) Vascular: FSI and Cell Based Models

Software:

- Same as for the Fast Track

Simulation scenario examples:

Fluid-structure interaction simulations are done to compute stent deployment and action, flow-diverters, valves, etc.

b) Cardiac: Fluid-Electro-Mechanics

Software:

- Same as for the FT

Simulation scenario examples:

These scenarios are very similar to those of the FT but now including fluid-electro-mechanical coupling.

Post-processing layer

FT: HPC-based visualization

Software:

- ParaView (<https://www.paraview.org/>): visualization, open source
- VisIt (<https://wci.llnl.gov/simulation/computer-codes/visit/>): visualization, open source

DT: HPC + HP Data Analytics (HPDA)

Software:

- Apache Cassandra (<http://cassandra.apache.org/>): distributed database, open source
- Apache Spark (<https://spark.apache.org/>): database engine, open source
- Alya (<http://www.bsc.es>): Multi-physics simulation software (PI of BSC), open access

Molecular medicine

HPC motivation

In order to provide molecular insight on the timescales required for industrial use cases (such as drug discovery and design) and clinical applications simulation must be executed as fast as or faster than alternative chemical approaches. In practice this means turn around times in the

order of one or two days and at most a week. In order to meet these demands the use of HPC is essential.

User groups

HPC for Researchers

One of the major barriers preventing biomedical researchers from engaging with molecular modelling has traditionally been the difficulty in model building and the arcane interfaces of simulation software. CompBioMed will provide tools to both facilitate the 'on ramp' by building initial structures and providing appropriate forcefield representations but also flexible workflows that abstract the complexities of running simulations. Such tools can have a transformative impact not just in the field of protein-ligand interactions (which are the primary focus internally to the project) but in wider investigations of structural and dynamic properties of proteins and other biomacromolecules.

HPC for Industry

The pharmaceutical industry already makes use of both BAC and Acellera tools to a limited extent, so far primarily in *post hoc* studies looking to explain trends found in previous drug design programmes and evaluate the power of computational approaches. Most of these studies have been accomplished through direct collaborations with the developers. It is expected that by refining both workflows and calculations within CompBioMed greater use will be made of these tools, including their deployment in active drug design.

HPC for clinical trials and in silico medical services

One of the deep track aims of the molecular modelling exemplar will be to provide validated workflows for clinically relevant scenarios with interfaces simple enough for clinicians to execute.

Structure

Layer name	Within CompBioMed?
Data Processing: signal processing, image processing, etc.	NO
Data Digestion: big data analytics, Kalman filtering, etc.	NO
Pre-processing: model generation, meshing, etc.	YES
Simulation (solvers): model solution	YES
Post-Processing: scientific visualisation, results data analytics, reporting	YES
Workflows: orchestrations of multiple models to capture a multi-step, multiscale, or multi-physics structure of the application	YES

Data processing Layer

Software

- BAC - Binding Affinity Calculator (<https://ccs.chem.ucl.ac.uk>): Computational molecular dynamics simulations
- HTMD (<https://www.acellera.com/>): High-throughput molecular dynamics
- AmberTools (<https://www.acellera.com/>): Computational molecular dynamics simulations
- VMD (<http://www.ks.uiuc.edu/Research/vmd/>): Molecular visualization program.

FT/DT Simulation Layer

- NAMD (<http://www.ks.uiuc.edu/Research/namd/>): Computational molecular dynamics simulations
- AmberTools (<https://www.acellera.com/>): Computational molecular dynamics simulations
- AceMD (<https://www.acellera.com/>): Computational molecular dynamics simulations

Post Processing Layer

Bespoke scripts

Simulation Scenario example:

A set of candidate drug/target models are built, an ensemble of replica simulations are executed and then analyzed using BAC or HTMD.

Neuro-musculoskeletal**HPC motivation**

Run full order model to solve unreducible model

Bone tissue is incredibly complex from a morphological point of view, but not sufficiently regular to use safely periodic homogenisation or other similar order-reduction techniques. Thus, any finite element model of bone tissue ranges between 100 millions to some billions unconstrained degrees of freedom.

User groups

HPC for Researchers

Bone biomechanics researchers worldwide will be interested in accessing this collection of research tools to process clinical images from bone experiment, develop bone material models, and test them.

HPC for Industry

Biomedical companies that develop products interacting with bone tissue, such as penduncular screws, fracture fixation devices, bone regeneration products, pharmaceutical applications, etc. could have interest in using some of CompBioMed software stack with the support of CompBioMed consultants.

Structure

Layer name	Within CompBioMed?
Data Processing: signal processing, image processing, etc.	YES
Data Digestion: big data analytics, Kalman filtering, etc.	NO
Pre-processing: model generation, meshing, etc.	YES
Simulation (solvers): model solution	YES
Post-Processing: scientific visualisation, results data analytics, reporting	YES
Workflows: orchestrations of multiple models to capture a multi-step, multiscale, or multi-physics structure of the application	YES

Data processing layer:

Software:

- DVC-Upload (<https://www.sheffield.ac.uk/>): Digital volume correlation software.
- DVC-preproc (<https://www.sheffield.ac.uk/>): Digital volume correlation software.

Pre-processing Layer

Software

- DVC-BC (<https://www.sheffield.ac.uk/>): Digital volume correlation software.
- Sheffield Image Registration Toolkit (SHIRT): Image registration

Simulation Layer

Software

- DVC/Bone (<https://www.sheffield.ac.uk/>): Workflow for estimating displacement/strains from images.
- Ansys/Mechanical MPI (<http://www.ansys.com/>): general purpose simulation commercial software
- Matlab (<https://www.mathworks.com/products/matlab.html>): commercial software

Post Processing Layer

Software

- DVC-download (<https://www.sheffield.ac.uk/>): Digital volume correlation software.
- DVC-Strain (<https://www.sheffield.ac.uk/>): Digital volume correlation software.

Simulation Scenarios examples:

Load multiple microCT scans of a bone tissue specimen under staged compression, and download resulting displacement and strain field, as estimated by digital volume correlation (DVC). A simple Taverna workflow orchestrates DVC-preproc, DVC/Bone, and DVC-Strain modules. In the fast track activity, a homogenizer repeatedly invokes ParaFEM/Bone to identify continuum scale constitutive equations. Palabos/VP can then provide a simulation of the PMMA injection in bone porous tissue.

5.2 Tasks' outcomes

Let us now briefly analyse each of the Work Package tasks.

Task 2.1: Deploy and Utilise Application Ready Environments (M1-M12) [Fast Track]

Leader: BSC (4 PM); Partners: UCL (2 PM), SARA (2), UEDIN (2), UvA (2), USFD (2)

So far, 45 simulation and data analysis codes have been deployed on HPC system at the CompBioMed supercomputing centres (BSC, EPCC and SURFsara) (see Appendix): 18 codes for the first biomedical exemplar domain, 9 for the second one and 18 for the third one. The codes are disparate in nature, being Open Source (both compiled binary and sources available), Open Access (compiled binary available, free licensing for CompBioMed users) and Commercial (compiled binary available, license fees paid by the supercomputing centre). We have documented part of this effort in a paper [2].

Task 2.2: Cardiovascular Exemplar Research (M1-M36) [Fast-Track]

Leader: UOXF (20 PM), Partners: LTG (12), UvA (10), UNIGE (8), UCL (4 PM), BSC (4), SARA (2)

Fast Track:

Within the Cardiovascular Exemplar Research domain, there are two main strands of work. On the one hand, UCL, UNIGE, UvA and USFD are working on different aspects of blood flow models, from vessels up to the arterial network. On the other hand, BSC and UOXF are working on cardiac modelling.

With regard to vascular blood flow models, three different lattice-Boltzmann-based solvers are used to attack different problems. Additionally a 1D arterial network model is developed.

At UCL, the lattice-Boltzmann solver **HemeLB** has been applied to the study of blood flow in patient specific cerebrovascular geometries (obtained through collaborating clinicians based at Hamad Medical Corporation (HMC) in Qatar, and National Hospital for Neurology and Neurosurgery (NHNN) in London). Stent designs have been obtained from collaborators at the Royal Free Hospital, London. Most of the work to date has concentrated on the validation of inflow conditions, such as obtaining *in vivo* Transcranial Doppler (TCD) measurements of peak flow velocity through various sections of the Middle Cerebral Artery, and then comparing these with the results of our HemeLB simulations on the same patient geometry [8]. Similarly, UCL has investigated the effects of changes in inflow velocity and mesh resolution on the resultant velocity and Wall Shear Stress profiles in the circle of Willis, using weak coupling with a 1-D flow solver (PyNS) to obtain the inlet velocities.

Also using the lattice-Boltzmann scheme, UvA develops **Hemocell** which is a parallel computing framework which implements validated mechanical models for red blood cells and is capable of reproducing the emergent transport characteristics of such a complex cellular system. It is computationally capable of handling large domain sizes, thus it is able to bridge the cell-based micro-scale and macroscopic domains. Hemocell introduces a new material model for resolving the mechanical responses of red blood cell membranes under various flow conditions and compare it with a well established model. Although this new constitutive model has similar accuracy under relaxed flow conditions, it performs better for shear rates over $1500s^{-1}$. We also introduce a new method to generate randomized initial conditions for dense mixtures of different cell types free of initial positioning artifacts [1].

Using a third lattice-Boltzmann scheme-based software named **PALABOS**, UNIGE focuses on the numerical simulation of flow diverters like stents. This study contributes to the development and improvement of endovascular stenting procedures, leading ultimately to an improved treatment of intracranial aneurysms. Due to the scale difference between the struts of flow diverters and the full artery, it is common to avoid fully resolved simulations at the level of the stent porosity. Instead, the effect of stents on the flow is represented by a heuristic continuum model. However, the commonly used porous media models describe the properties of flow diverters only partially, because they do not explicitly account for the deflection of the flow direction by the stent. UNIGE shows that this deficiency can be circumvented by adopting the theoretical framework of screen models. After doing a wide review of existing screen models, UNIGE proposes an explicit formula for the drag and the deflection coefficient, as predicted by each model, for both perpendicular and inclined angles. The results of 2D numerical simulations are used to formulate a generalization of these

formulas, to achieve best results in the case of stent modeling. The obtained model is then validated, again through 2D numerical simulation [13].

OpenBF is a numerical library aimed at the solution of one-dimensional blood flow equations in a network of elastic vessels. At USFD, this is being used to perform the calculation of pulse and volumetric flow rate waves of a complete model of the arterial systemic circulation comprising the major large vessels and the cerebral circulation, for a total of 78 arteries. By varying the model parameters within their respective physiological and pathological ranges, openBF will be used to create a virtual population whose waveforms will be used by other higher dimensional models as boundary conditions in a loose coupling approach (one-way coupling). All the parameters and waveforms will be employed to populate a database that will be made available through the CompBioMed consortium. OpenBF is currently running and being tested on SURFsara private cloud environment and on iceberg (USFD tier-3).

With regard to cardiac modelling, BSC and Oxford are working on setting up large scale electrophysiology simulations including the torso, to directly compute the electrocardiogram (ECG). We have implemented several new features in **Alya**. The electrophysiology module has been heavily re-engineered, obtaining a strong performance improvement, including a semi-implicit version and new cell models (Ten Tusscher - Panfilov and O'Hara). We have also included a new electromechanical coupling model and a Windkessel model to take into consideration the pressure. We have created two new complete cardiac geometries, which are being tested [9,10,11].

Deep Track:

At UCL, Magnetic Drug Targeting (via magnetic steering of injected colloidal particles) has been successfully introduced to HemeLB, and tested for physical correctness. Large scale runs of thousands of magnetic particles travelling through the circle of Willis are currently being carried out. The aim is to study the effects of applied magnetic fields on the statistics of particle motion and dispersion.

There has also been substantial work on improving the performance (particularly as regards memory usage) of HemeLB when scaling to large core counts. So far, we have very good scaling performance at 96,000 cores on the ARCHER supercomputer (EPCC), and we intend to move on to higher core counts shortly.

Partners UvA, SURFsara, USFD and BSC plan to explore strongly-coupled approaches for the Deep Track applications. The 1D circulatory model will be updated with the solutions from the organ scale models tested at USFD.

At BSC, with regard to cardiac applications, we are working on code performance analysis and porting to new architectures. We have performed a thorough analysis of parallel performance for the coupled multi-physics FSI using a Compact Interface Quasi-Newton scheme (CIQN) [12]. For the new architectures, we are testing the Alya parallel linear solvers (CG and GMRES) on GPUs. The code has also been ported to Intel's KNL to focus specifically on vectorisation and in both cases, we have very good preliminary results.

Task 2.3 Molecularly-based Medicine Exemplar Research (M1-M36) [Fast Track]

Leader: UCL (20 PM), Partners: UPF (12 PM), EVOUK (6), ACE (6), JAN (6)

Fast Track:

Within UCL, the Binding Affinity Calculator (BAC) has been deployed to study a range of systems, many in collaboration with industrial collaborators (Core Partners Janssen and Evotec, and Associate Partner GSK). This work has resulted in two papers [4,5] and to the development of new cloud deployments of the BAC workflow (WP6). During the first 9 months of the project we have published a new protocol for rapidly, accurately and precisely computing relative binding affinities (known as TIES) [6]. Current activity involves studying the influence of mutations in a variety of cancer target systems and investigations into the range of applicability of the analysis methods currently available in BAC.

Additionally, we have published work [3] on coarse grained modelling of the type of GPCR systems upon which this exemplar is focused. This provides a platform for future work using molecular simulations to understand dimerisation processes in these systems. Work is underway to facilitate simulation and analysis of GPCR systems in BAC as a deep track activity.

Janssen, in a close collaboration with UCL, is interested in evaluating the performance of several free energy of binding calculation methods on challenging internal and public data sets. They are also interested in defining best practices and settings for running the calculations, based on the results that are obtained. Janssen is identifying data sets where a physics based approach is expected to be required, this includes: 1) affinity cliffs, how well can FEP predict these, 2) entropy-enthalpy compensation, a problematic effect to predict, and 3) how well can FEP predict changes in water network structures that appear when compounds are modified. Inspired by the multi-armed bandit problem, UPF and Acellera have formalised, as a machine learning process, the sampling of the ligand binding problem. This could be used to speed-up GPCR ligand-binding simulations by adaptive sampling, and to use gradients based on physical properties (contacts, RMSD) to guide MD simulations to the binding pose, which can be reached in a few microseconds.

Drug-receptor residence time (RT) is a key kinetic property for drug efficacy [7]. A retrospective analysis of existing drugs revealed that their efficacy may be attributed to their long drug-target residence times. Recently, several notable reviews suggest that RT should be optimised in the hit-to-lead and lead optimisation phases of drug discovery to prevent “fail late, fail expensive” scenarios. Evotec and UCL have taken the initiative in developing integrated experimental-computational technology to address the residence time of GPCR ligands.

Deep Track:

The development of an updated version of BAC that is capable of making use of a wider range of simulation engines and preparing more complicated simulation environments (such as biological membranes) has been initiated. Work is underway that is aimed at deploying more advanced versions of BAC on very large (“Tier-0”) HPC systems on the path to the exascale. This involves the use of innovative job submission, management and farming systems, including those provided by the ComPat European Project (<http://www.compat-project.eu/>) stack and RADICAL-Cybertools (<https://radical-cybertools.github.io/>). These approaches allow the simple and transparent automation of large numbers of jobs, ensuring optimal turnaround time on even the largest super computers currently available. The development of these tools help to facilitate the use of ensemble simulations to improve the accuracy and reproducibility of

simulation results and to provide reliable uncertainty quantification. The potential of this work was demonstrated when a prototype of our task farming approach was used to employ the entirety of the SuperMUC machine (*ca* 250k cores) at Associate Partner LRZ for a recent study of protein-drug binding. Our focus is currently on targeting BlueWaters in the US which provides 4k GPU accelerated nodes in addition to approx. 400k CPU cores.

Early work on BAC within CompBioMed has already produced an automated tool for creating the hybrid topologies required for various binding affinity calculations. This work has reduced a task that took many hours into one which can be completed in a matter of seconds. The current iteration of the TIES protocol within BAC can only be applied to very similar drugs; we are now exploring options to extend its range of applicability and improve the statistical uncertainties of the calculation (using dummy atom models and Multistate Bennett Acceptance Ratio (MBAR) analyses respectively).

As part of efforts to bring new compute engines into BAC, we have held collaborative discussions with GROMACS (<http://www.gromacs.org/>) developers. This happened as part of wider activities centred on the recent Free Energy Workshop organised jointly by CompBioMed and BioExcel (<http://bioexcel.eu/>) (see section covering WP3).

Task 2.4 Neuro-musculoskeletal Exemplar Research (M1-M36) [Fast Track]

Leader: USFD (20 PM); Partners: UNIGE (8 PM), BSC (2)

Fast Track:

The BoneDVC service has been successfully deployed on both ShARC (USFD Tier-3) and ARCHER. This workflow represents the latest and most accurate method to estimate full displacement/strain field from two sets of input images. The workflow is currently being used to validate additional experimental work using micro-CT images of animal and human bones.

USFD has conducted substantial work on profiling the performance of BoneDVC on HPC systems, with a manuscript in preparation. This exercise showed some limited scalability of the code when ported to ARCHER, and identified the bottleneck as being the Sheffield Image Registration Toolkit (SHIRT). The current implementation of SHIRT cannot be parallelised. Therefore, in the deep track activity, efforts will be allocated to rewrite this part of the code in order to improve its performance on HPC platforms.

Deep Track:

SHIRT will be reimplemented as part of the BoneDVC service, to make it easier to parallelise on HPC systems. Once the code has been rewritten, USFD will work with UEDIN, SURFsara, and BSC to deploy and optimise the new code on Tier-1 HPC systems.

In addition, we also plan to provide a full scale simulation environment for the analysis and modelling of bone tissues on HPC system. This will be achieved by integrating USFD's BoneDVC and microFE meshing, together with Levrero's ParaFEM solver (UOXF), and UNIGE's PALABOS cement injection simulation. This environment will enable the automatic creation of microFE models from micro-CT images, with PALABOS to simulate PMMA flow into a volume of cancellous bone, and predicting the resulting cement augmentation pattern, a procedure known as vertebroplasty. The BoneDVC will be used to calculate the experimental strain field, and compared against simulated values using ParaFEM, which also takes into account the post-elastic mechanical properties of bone tissues. We aim to port all these codes to ARCHER, which

will provide a complete workflow for microFE validation studies, including the potential for a vertebroplasty planning tool.

Task 2.5 (Post-processing and Error Analysis)

This task starts in M12.

Task 2.6 (Develop Plans for and Implement the Upscaling of CompBioMed Production Applications for Future HPC Platforms, Including those Heading Toward the Exascale)

This task starts in M12.

Task 2.7: Develop and Implement Multiscale and Workflow Interfaces (M6-M30 [Deep Track] Leader: BSC (10 PM); Partners: UvA (8 PM), UNIGE (6))

Multiscale and workflow tasks have started, in close connection with WP5 and WP6. WP2 duties are mostly related to the deployment and testing of codes in the context of a larger infrastructure.

Notably and in a joint effort, CompBioMed partners have produced a paper on porting real-world simulation biomedical research applications to an HPC environment. In this paper, we address computational workflows, which aimed at clinical decision support, consist of complex orchestration of multiple codes, often under commercial license, whose results should be obtained in the shortest time possible. These requirements differ from those of traditional HPC applications, therefore, the use of HPC systems in computational medicine is not as widespread as in other scientific research domains. In this paper, the porting of one musculoskeletal exemplar application from a Tier-3 to a Tier-1 system is described. The workflow is tested for a representative range of use cases while monitoring execution time and memory consumption. The results indicate that there is not an immediate benefit in porting the application as it is on the Tier-1 system. The paper also includes recommendations on optimisation and possible improvements of computational medicine workflows are proposed.

Additionally, and related to multi-physics coupling, BSC has developed a new multi-code fluid-structure-interaction (FSI) scheme, known as the Compact Interface Quasi-Newton scheme (CIQN) [12]. This FSI scheme is used to handle the blood-mechanical coupling with very good results on robustness, accuracy and parallel performance. In the CIQN, a MPI multi-code partitioned scheme, designed for computationally intensive fluid and solid problems is used. Thus, special care is taken to efficiently use the parallel computational resources, based in two main precepts. On one hand, memory allocation and computational operations for intermediate variables is avoided when possible. On the other hand, an efficient MPI communication scheme between parallel code must be set. Through 3D examples of increasing complexity, we show that the CIQN algorithm can efficiently deal with added mass instability and two-field coupled problems. The results compare very well against the Aitken and Broyden relaxation schemes, with at least, equal performance and accuracy and always with much improved robustness. Additionally, reduced simulation time, median and interquartile range for the distribution in coupling iterations, is obtained for the CIQN scheme. As a final test, we published a parallel performance analysis, performed using up to 4096 processors.

6 Bibliography / References

CompBioMed partners are using the Application Pipeline intensively for collaborative research. After 9 months, our published research is mostly related to the Fast Track. Appendix 2 includes some of the selected papers and pre-prints that correspond to research done in the first year of CompBioMed's WP2.

This is a list of the most significant papers published or on the publishing process so far:

- [1]. G. Závodszy, B. van Rooij, V. Azizi, and A. Hoekstra, Cellular level *in silico* modelling of blood rheology with an improved material model for red blood cells., *Frontiers in Physiology*, 2017.
- [2]. A. Melis, E. Dall'Ara, A. Marzo, M. Vazquez, G. Pringle, M. Viceconti, and CompBioMed Consortium - HPC for the rest of us: porting real-world research applications to and HPC environment. *Journal of Computational Science*, submitted 2017.
- [3]. N. Altwaijry, M. Baron, D. Wright, P. V. Coveney, A. Townsend-Nicholson. An Ensemble-Based Protocol for the Computational Prediction of Helix-Helix Interactions in G Protein-Coupled Receptors using Coarse-Grained Molecular Dynamics. *J. Chem. Theory Comput.*, 13 (5), 2254-2270, 2017
- [4]. S. Wan, A. Bhati, S. Skerratt, K. Omoto, V. Shanmugasundaram, S. Bagal, P. V. Coveney, Evaluation and Characterization of Trk Kinase Inhibitors for the Treatment of Pain: Reliable Binding Affinity Predictions from Theory and Computation. *J. Chem. Inf. Model.*, 57 (4), 897-909, 2017
- [5]. S. Wan, A. P. Bhati, S. J. Zasada, I. Wall, D. Green, P. Bamborough, and P. V. Coveney. Rapid and Reliable Binding Affinity Prediction of Bromodomain Inhibitors: a Computational Study. *J. Chem. Theory Comput.*, 13 (2), 784–795 (2017)
- [6]. A. Bhati, S. Wan, D. Wright and P. V. Coveney; Rapid, accurate, precise and reliable relative free energy prediction using ensemble based thermodynamic integration. *J. Chem. Theory Comput.*, 13 (1), 210–222 (2017)
- [7]. R. C. Eccleston, S. Wan, N. Dalchau, and P. V. Coveney, The role of multiscale protein dynamics in antigen presentation and T lymphocyte recognition. *Front Immunol.*, 10 (8), 797 (2017)
- [8]. D. Groen, R. A. Richardson, R. Coy, U. D. Schiller, H. Chandrashekar, F. Robertson, and P. V. Coveney, Validation of patient-specific cerebral blood flow simulation using transcranial Doppler measurements, submitted, 2017.
- [9]. J. Aguado-Sierra, R. Arís, E. Casoni, A. Santiago, M. López, J.C. Cajas, S. Gómez, G. Houzeaux, M. Rivero, F. Sacco, C. Butakoff, M. Vázquez, Coupled fluid-electro-mechanical cardiac computational model for device testing applications. *Frontiers in Physiology*, abstract submitted, 2017.
- [10]. F. Levrero-Florencio, E. Zacur, A. Mincholé, M. Vázquez, J. Aguado, G. Houzeaux, V. Grau, B. Rodríguez, HPC electromechanical simulations of the human heart: an MRI-based validation. *Frontiers in Physiology*, abstract submitted, 2017.
- [11]. R. Arís, J. Aguado, A. Santiago, E. Casoni, M. Rivero, G. Houzeaux, D. Gil, A. Borrás, R. Sebastian, M. Vázquez, Analysis of fiber orientation and activation sequence response on cardiac computational electromechanics. *Frontiers in Physiology*, abstract submitted, 2017.

- [12]. A. Santiago, M. Zavala-Aké, M. Vázquez, J.C. Cajas and G. Houzeaux. HPC-based compact interface quasi-Newton algorithm for fluid-structure interaction. *SIAM Journal of Scientific Computing*, submitted, 2017.
- [13]. S. Li, J. Latt and B. Chopard. Model for pressure drop and flow deflection in the numerical simulation of stents in aneurysms. *International Journal for Numerical Methods in Biomedical Engineering*, in revision, 2017.

7 Conclusions

CompBioMed's Work Package 2 is progressing at full speed. After a short initial discussion, taxonomy and design stage, the bulk of the simulation and analysis software stack to be used during the project's time span is already installed and in production at the three supercomputing partners' facilities as well as others. This porting and optimization work has been completed in coordination with WP5. From the Fast Track, a large amount of results are being produced as shown by the publication record, for the three exemplars, also manifesting integrated work among the partners. Additionally and in all three cases, Deep Track work is being prepared, analysed and, in many cases, now started. Finally, we have begun to address code-coupling schemes for multi-physics, multi-scale and complex workflows.

8 Appendix

8.1 Appendix A. CompBioMed software stack

So far (M12) 45 pieces of software are installed and being used by CompBioMed partners. It follows the list, where:

- **Code** is software name.
- **Usage** addresses the kind of usage and licensing, being
 - Open Source (O): code can be freely distributed and used, under any kind of Open Source License.
 - Limited Distribution (L): there exist some kind of limitation to source code distribution but not on the binary usage. These are typically "Open Access" codes, where the installed binary is used with no limitations.
 - Commercial (C): there exist some kind of limitation to source code distribution and binary usage. Licensing policies depend on the commercial agreement between the code vendor and the partner.
- **HPC** labels whether software is (Yes, Y) or is not (No, N) ported to run efficiently in a parallel computer and up to a certain (and arbitrary) number of computing units. We set this number to 100, which can be cores for memory distributed computers or parallel threads for GPGPUs or similar. We assign a "Y" when linear scalability is recorded up to this limit.
- **Research Exemplar** is one of CompBioMed exemplars: Cardiovascular, Musculo-skeletal or Molecular Medicine.

Code / CompBioMed developer (if applicable)	Usage	HPC	Research Exemplar
Alya / BSC	L	Y	Cardio
CHASTE / Oxford	O	Y	Cardio
FICSION - HemoCell / UvA	L	Y	Cardio
HemeLB / UCL	L	Y	Cardio
openBF	L	N	Cardio
OpenFOAM	O	Y	Cardio
PALABOS / UNIGE	O	Y	Cardio
VMTK	O	N	Cardio
Amira	C	N	Cardio
Matlab	C	N	Cardio
ICEM	C	N	Cardio
Ansa	C	N	Cardio
GiD	C	N	Cardio
BSC's IRIS MESH / BSC	L	N	Cardio
ParaView	O	N	Cardio
VisIt	O	N	Cardio
Apache Spark	O	Y	Cardio

Apache Cassandra	O	Y	Cardio
ACEMD	C	Y	Mol
AMBER	C	Y	Mol
AmberTools	O	Y	Mol
Charmm	O	Y	Mol
Desmond	C	Y	Mol
GAMESS-US	O	Y	Mol
Gromacs	O	Y	Mol
HTMD (acemd)	C	Y	Mol
NAMD	O	Y	Mol
Ansys (Mechanical APDL, ICEM)	C	Y	Musc
BONEMAT / USFD	O	N	Musc
BuilderM20 / USFD	O	N	Musc
ShIRT executable Matlab	L	N	Musc
Taverna	O	Y	Musc
ITK-Snap	O	N	Musc
DVC-Upload / USFD	L	N	Musc
DVC-preproc / USFD	L	N	Musc
DVC-download / USFD	L	N	Musc
DVC-Strain / USFD	L	N	Musc
MicroFE_Mesher	O	N	Musc
Simpleware +FE	C	N	Musc
MicroFE_DVC-BC / USFD	L	N	Musc
DVC/Bone / USFD	L	N	Musc
ParaFEM/Bone	O	Y	Musc
Homogenizer	O	N	Musc
Bone Tissue DVC workflow / USFD	O	N	Musc
Bone tissue homogeneisation / USFD	O	N	Musc

8.2 Appendix B. Selected publications

This Appendix contains the full text of a selected set of papers, representative of CompBioMed research for the first year. Research papers are a deeper description of research activities. They are included as a technical appendix in order to avoid the technical burden to the non-expert reader. In any case and for the sake of completion, this technical section is the key record of CompBioMed action. It is worth to remark that some of the papers are already published and some others are pre-prints already submitted to the journals.

Summarizing, the documents included are:

- G. Závodszy, B. van Rooij, V. Azizi, and A. Hoekstra, Cellular level *in silico* modelling of blood rheology with an improved material model for red blood cells., *Frontiers in Physiology*, 2017.
- A. Melis, E. Dall'Ara, A. Marzo, M. Vazquez, G. Pringle, M. Viceconti, and CompBioMed Consortium - HPC for the rest of us: porting real-world research applications to and HPC environment. *Journal of Computational Science*, submitted 2017.
- N. Altwaijry, M. Baron, D. Wright, P. V. Coveney, A. Townsend-Nicholson. An Ensemble-Based Protocol for the Computational Prediction of Helix-Helix Interactions in G Protein-Coupled Receptors using Coarse-Grained Molecular Dynamics. *J. Chem. Theory Comput.*, 13 (5), 2254-2270, 2017
- S. Wan, A. Bhati, S. Skerratt, K. Omoto, V. Shanmugasundaram, S. Bagal, P. V. Coveney, Evaluation and Characterization of Trk Kinase Inhibitors for the Treatment of Pain: Reliable Binding Affinity Predictions from Theory and Computation. *J. Chem. Inf. Model.*, 57 (4), 897-909, 2017
- S. Wan, A. P. Bhati, S. J. Zasada, I. Wall, D. Green, P. Bamborough, and P. V. Coveney. Rapid and Reliable Binding Affinity Prediction of Bromodomain Inhibitors: a Computational Study. *J. Chem. Theory Comput.*, 13 (2), 784–795 (2017)
- A. Santiago, M. Zavala-Aké, M. Vázquez, J.C. Cajas and G. Houzeaux. HPC-based compact interface quasi-Newton algorithm for fluid-structure interaction. *SIAM Journal of Scientific Computing*, submitted, 2017.
- S. Li, J. Latt and B. Chopard. Model for pressure drop and flow deflection in the numerical simulation of stents in aneurysms. *International Journal for Numerical Methods in Biomedical Engineering*, submitted, 2017.



Cellular Level *In-silico* Modeling of Blood Rheology with An Improved Material Model for Red Blood Cells

Gábor Závodszy^{1,2*}, Britt van Rooij¹, Victor Azizi¹ and Alfons Hoekstra^{1,3}

¹ Computational Science Lab, Faculty of Science, Institute for Informatics, University of Amsterdam, Amsterdam, Netherlands, ² Department of Hydrodynamic Systems, Budapest University of Technology and Economics, Budapest, Hungary, ³ ITMO University, Saint Petersburg, Russia

Many of the intriguing properties of blood originate from its cellular nature. Therefore, accurate modeling of blood flow related phenomena requires a description of the dynamics at the level of individual cells. This, however, presents several computational challenges that can only be addressed by high performance computing. We present Hemocell, a parallel computing framework which implements validated mechanical models for red blood cells and is capable of reproducing the emergent transport characteristics of such a complex cellular system. It is computationally capable of handling large domain sizes, thus it is able to bridge the cell-based micro-scale and macroscopic domains. We introduce a new material model for resolving the mechanical responses of red blood cell membranes under various flow conditions and compare it with a well established model. Our new constitutive model has similar accuracy under relaxed flow conditions, however, it performs better for shear rates over $1,500 \text{ s}^{-1}$. We also introduce a new method to generate randomized initial conditions for dense mixtures of different cell types free of initial positioning artifacts.

Keywords: blood rheology, RBC material model, cellular flow, high-performance computing, dense cell initialization

OPEN ACCESS

Edited by:

Timothy W. Secomb,
University of Arizona, United States

Reviewed by:

Panagiotis Dimitrakopoulos,
University of Maryland, College Park,
United States

Dmitry A. Fedosov,
Forschungszentrum Jülich, Germany

*Correspondence:

Gábor Závodszy
g.zavodszy@uva.nl

Specialty section:

This article was submitted to
Computational Physiology and
Medicine,

a section of the journal
Frontiers in Physiology

Received: 20 April 2017

Accepted: 19 July 2017

Published: 02 August 2017

Citation:

Závodszy G, van Rooij B, Azizi V and
Hoekstra A (2017) Cellular Level
In-silico Modeling of Blood Rheology
with An Improved Material Model for
Red Blood Cells.
Front. Physiol. 8:563.
doi: 10.3389/fphys.2017.00563

1. INTRODUCTION

On the cellular level, blood is a dense suspension of various types of cells. Red blood cells (RBC) form the primary component with an approximate volume fraction of 42% (Davies and Morris, 1993) determining the bulk blood rheology. They have a biconcave shape and a typical diameter of $8 \mu\text{m}$. Platelets (PLTs), the second most numerous component with typically 1 PLT for every 10 RBCs (Björkman, 1959) form the link between transport dynamics and vital biochemical processes related to thrombus formation. In their unactivated state PLTs have a rigid ellipsoidal form. The collective behavior of RBCs and PLTs can provide explanation to the most fundamental transport phenomena in blood, for instance the non-Newtonian viscosity (Merrill and Pelletier, 1967), the margination of platelets (Beck and Eckstein, 1980; Tilles and Eckstein, 1987), the Fåhræus effect (Barbee and Cokelet, 1971), the appearance of a cell-free layer (Maeda et al., 1996; Kim et al., 2009), or the scaling of shear-induced diffusion of RBCs (Mountrakis et al., 2016). The necessity to accurately reproduce these effects grows as the typical length-scale of the examined system reaches below $\approx 200 \mu\text{m}$, at which point the macroscopic description no longer yields accurate local dynamics (Popel and Johnson, 2005). With the development of modern medical devices more and more elements reside in the micrometer domain, such as the strut structure of flow-diverters

(Lubicz et al., 2010) or woven endobridge (WEB) flow disruptor devices (Ding et al., 2011). This together with additional complex phenomena that require detailed cellular modeling of the flow, for instance platelet aggregation (Nesbitt et al., 2009) or white blood cell (WBC) trafficking (Fay et al., 2016), triggers an increasing need to understand how the rheology and the transport of the RBCs and PLTs are influenced while acting over such small scales.

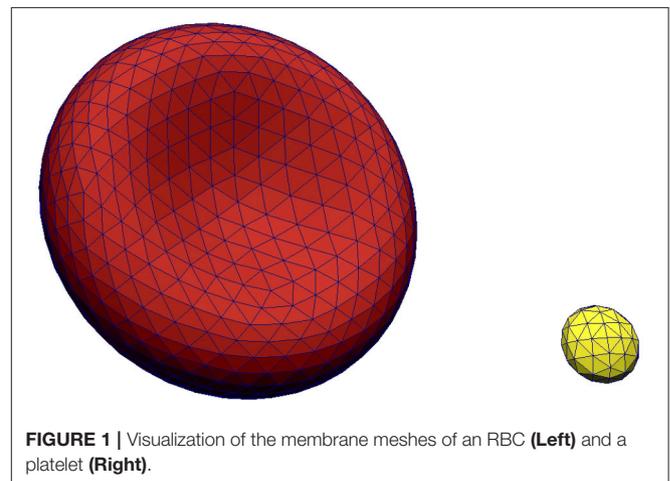
In the solutions targeting these questions the mechanical responses of the RBCs and PLTs are often expressed with constitutive models applied through their membranes accounting for the responses of the various structural elements (Ye et al., 2016). Some examples for these material models are the spectrin-link membrane model of Dao et al. (2006) or the energy model of Skalak et al. (1973). Fedosov et al. (2010a) employed the dissipative particle dynamics (DPD) method with a constitutive description gained by coarse-graining the model of Dao et al. (2006) to study various transport features of blood (Fedosov et al., 2010b, 2011b,c; Fedosov and Gompper, 2014; Yazdani and Karniadakis, 2016). A low dimensional RBC membrane model was developed by Pan et al. (2010) and was compared to the coarse-grained spectrin-link model of Fedosov et al. (2010a). More recently, a two-component RBC membrane model that consist of a separate lipid bilayer and spectrin network (Chang et al., 2016) was introduced to examine the difference in the deformation of healthy and infected RBCs. MacMeccan et al. (2009) developed a model that coupled the lattice Boltzmann method (LBM) to finite element method (FEM) based cell mechanics and investigated the viscosity behavior of blood in shear flows at various hematocrit levels. Later, Reasor et al. (2012) used the spectrin-link membrane representation rather than solving Cauchy's equation to model the trajectory and deformation of elastic deformable particles. This model was also used to study the margination of platelets (Mehrabadi et al., 2016). Moreover, Krüger et al. (2011) developed a combination of lattice Boltzmann method (LBM) and finite element RBC membrane model based on the energy model of Skalak et al. (1973), and used the immersed boundary method (IBM) to couple the fluid and the membrane. This model was used to study the tank-treading behavior of single RBCs next to the deformation behavior and the relative viscosity of RBC suspensions (Krüger et al., 2013; Gross et al., 2014; Krüger, 2016). In addition, Shi et al. used LBM in combination with the fictitious-domain method to couple the plasma to the spectrin-link membrane model. They studied the deformation of an RBC in capillary flows, during tank-treading motion and hydrodynamic interaction between two cells (Shi et al., 2014). Hashemi and Rahnama (2016) investigated the deformation of RBCs in capillary flows with an LBM-FEM based model with IBM coupling.

In this paper, our framework called Hemocell (High pErformance MicrOscopic CELLular Library)¹ is presented for modeling the flow of blood on a cellular level. Hemocell is designed to be easily extendible with additional cell-types and interactions and to provide the high computational performance

that enables applications up to macroscopic scales. Blood plasma is represented as a continuous fluid simulated with LBM, while the cells are represented as discrete element method (DEM) membranes coupled to the fluid flow by the immersed boundary method. Furthermore, two different material models for the RBC membrane mechanics have been investigated. One is the aforementioned coarse-grained spectrin-link model of Fedosov et al. (2010a) and a new one that addresses several shortcomings of the former. The validation of Hemocell, in combination with our new RBC material model, is presented through single-cell mechanical experiments (i.e., stretching and shearing cases). We demonstrate that the proposed new material model reproduces both the single-cell mechanical responses and the collective transport dynamics in very good agreement with experiments, as well as it provides an accurate mechanical response and an increased structural stability under higher shear forces and strong deformations. The later is necessary, since it is known from recent high-field-strength MRI measurements of Bouvy et al. (2016) that pulsation effects are significant even on the mesoscopic level of smaller arterioles. Moreover, it can also enable simulations of transport mechanisms in micro-fluidic settings or in the vicinity of micro-medical devices, where strong deformations and high shear values and gradients can be expected. Hemocell also forms a fundamental component in building versatile multi-scale models of arterial health and diseases (Hoekstra et al., 2016).

2. METHODS

The solvent (blood plasma) in Hemocell is modeled as an incompressible Newtonian fluid using the lattice Boltzmann method implemented in the Palabos library (Lagrava et al., 2012) which is known to be capable of producing accurate flow results in vascular settings (Závodszy and Paál, 2013; Anzai et al., 2014). The surfaces of RBCs and PLTs are described as boundary layers immersed into the plasma. These layers are discretized using N_v vertices which are connected by N_e edges yielding N_t surface triangles (see **Figure 1** for an example in case of an RBC and a PLT). The connectivity and symmetries are similar to the structure of the cytoskeleton as imaged by



¹<https://www.hemocell.eu/> (Accessed July 25, 2017).

atomic force microscopy (Swihart et al., 2001; Liu et al., 2003). In our simulations the membrane of each RBC consisted of $N_v = 642$ vertices, $N_e = 1,920$ edges, and $N_t = 1,280$ faces. The mechanical behavior of a cell is expressed using this discrete membrane structure. The response to deformations is formulated as a set of forces acting on the cell membrane, which is coupled to the plasma flow through a validated in-house immersed-boundary implementation (Mountrakis et al., 2014; Mountrakis, 2015) that has an efficient parallel design. Mountrakis et al. (2015) demonstrated that the framework can be scaled up to 10^6 cells executing on 8,192 cores without significant loss of parallel efficiency.

2.1. Description Of The Coarse-Grained spectrin-Link Membrane Model

In Hemocell, two distinct constitutive model have been implemented for RBCs to act on the membrane mesh. The first one is based on the systematic coarse-graining of the model of Dao et al. (2006). For the detailed derivation we refer to the work of Fedosov et al. (2010a). The model is briefly outlined below:

The free-energy of a cell is described as

$$U_{total} = U_{in-plane} + U_{bend} + U_{volume} + U_{area} + U_{visc}.$$

The location x_i of each vertex on the membrane mesh is updated according to the force:

$$F_i = \frac{\partial U_{total,i}}{\partial x_i}.$$

The total free-energy is composed of the following elements:

1. The in-plane potential models the compression response of the underlying cytoskeletal network along the membrane surface. The edges of the surface triangles represent the cumulative behavior of the local spectrin links using the wormlike chain (WLC) nonlinear spring description:

$$U_{in-plane} = \sum_{i=1..N_e} U_{WLC} + \sum_{j=1..N_t} \frac{C_q}{A_k^q},$$

$$U_{WLC} = \frac{k_B T l_m}{4p} \frac{3r_i^2 - 2r_i^3}{1 - r_i},$$

$$C_q = \frac{\sqrt{3} A_0^2 k_B T (4r_0^2 - 9r_0 + 6)}{4p l_m (1 - r_0)^2},$$

where p, l_m are the persistence length and the maximum length of the spectrin links, $r_i = l_i/l_m \in [0, 1]$, l_0 is the average length of links, $r_0 = l_0/l_m$ and $A_0 = \sqrt{3} l_0^2/4$.

2. The potential to account for bending rigidity:

$$U_{bend} = \sum_{i=1..N_e} \tilde{\kappa} [1 - \cos(\theta_i - \theta_0)],$$

where $\tilde{\kappa} = 2\kappa/\sqrt{3}$, θ_i is the instantaneous, θ_0 is the equilibrium angle between neighboring faces sharing an edge, and κ is the bending constant.

3. The volume conservation energy is a fictitious potential which accounts for the forces arising from the change of volume:

$$U_{volume} = \frac{k_V (V - V_0)^2}{2V_0},$$

where V is the current, and V_0 is the equilibrium volume of the cell.

4. The area conservation potential is similarly a non-physical term representing the inextensible nature of the bilipid layer:

$$U_{area} = \frac{k_A (A - A_0)^2}{2A_0} + \sum_{k=1..N_t} \frac{k_{A_l} (A_k - A_{0,k})^2}{2A_{0,k}},$$

where A, A_0 are the global and $A_k, A_{0,k}$ are the local actual and equilibrium surface areas, respectively.

5. The additional term to correct membrane viscosity:

$$U_{visc} = \sum_{i=1..N_e} -\frac{1}{2} \eta_m v_{m,n}^2,$$

where $v_{m,n}$ denotes the relative velocity of the vertices m and n connected by edge i and membrane viscosity $\eta_m = 22 \times 10^{-3}$ Pas is chosen such that RBCs yield realistic tank-treading and tumbling frequencies (Fedosov et al., 2014).

The free parameters of this model ($\kappa = 100 k_B T$, $k_V = 6,000$, $k_A = 5,900$, $k_{A_l} = 100$) were adopted from (Fedosov et al., 2010a) with the exception of the maximum link extension ratio ($r_0 = \frac{l_0}{l_m^{max}} = 2.6$), which was fine-tuned for our current discretization. The usefulness of this model was demonstrated in a series of publications (Fedosov et al., 2010a, 2011a; Fedosov and Gompper, 2014; Mehrabadi et al., 2016). However, it also has a few shortcomings. The bending response is based on Helfrich's model (Helfrich, 1973) which only accounts for the properties of the lipid bilayer and not the underlying structures. Furthermore, the coarse-graining of the bending rigidity for the triangulated mesh is based on the work of Gompper and Kroll (1996) which assumes small angles and equilateral triangles, both of which are often not fulfilled for sheared RBCs. As a consequence, the bending energy in this model yields a sinusoidal force-response that has a sub-linear response for angles over $\frac{\pi}{4}$, which even decays further for larger angles. This can lead to insufficient force responses and consequently to acute angles or collapse of neighboring faces. The resulting problems can often be mitigated by using a linear bending response that fits the slope of the sinusoidal at low angles. Moreover, the global surface conservation potential can lead to unphysical responses, since a local stretch of the membrane instantly causes the contraction of the rest of the membrane forcing the surface points to move toward the center of the cell.

2.2. Description Of The New Constitutive Model

We propose a new material model in the form of a set of forces acting on the same triangulated cellular membrane. The initial assumption for this model is that during small deformations

all these forces present a linear regime with different slopes as the response types correspond to different components of the cell and are independent of each other. However, for large enough deformations the cytoskeleton adds contribution to all of them, resulting in qualitatively similar behavior. For instance, a response for small bending is assumed to be dominated by the curvature rigidity of the bilipid membrane resulting in a term linear in angle for the DEM membrane, while for larger deformation the underlying cytoskeleton deforms as well yielding an additional quickly diverging term. In the following we describe this model in two steps by separating the phenomenological description and the implementation.

1. The link force acts along links between surface points and represents the response to stretching and compression of the underlying spectrin-network beneath the representative link. The formulation of the force is similar in spirit to the worm-like-chain potential model often used to mimic the mechanical properties of polypeptide chains. It presents a linear part which corresponds to smaller deformations and a fast-diverging non-linear part which represents the limits of the material toward this type of deformation by quickly increasing the force response as the stretch approaches the persistence-length.

$$F_{link} = -\frac{\kappa_l dL}{p} \left[1 + \frac{1}{\tau_l^2 - dL^2} \right],$$

where $dL = \frac{L_i - L_0}{L_0}$ is the normal strain defined as the relative deviation from the equilibrium length L_0 with $\tau_l = 3.0$ is chosen based on the assumption that the represented spectrin-network reaches its persistence length at the relative expansion ratio of 3. The persistence-length of a spectrin filament was taken as $p = 7.5 \text{ nm}$ (Li et al., 2005).

2. The bending force acts between two neighboring surface elements representing the bending response of the membrane arising primarily from the non-zero thickness of the spectrin-network. On each surface it points along the normal direction of that surface. As opposed to the previous model in which bending is expressed by modeling the bending rigidity of the bilipid membrane (Helfrich, 1973), the form of the employed force term here is similar to the form of the previous link force to account for increased resistances coming for additional sources, such as the connection of the membrane to the underlying cytoskeleton.

$$F_{bend} = -\frac{\kappa_b d\theta}{L_0} \left[1 + \frac{1}{\tau_b^2 - d\theta^2} \right],$$

where $d\theta = \theta_i - \theta_0$. From simple geometric considerations it follows that the limiting angle τ_b scales with the discretization length of the surface elements (L_0). We fix the smallest representable curvature $r_{min} = \frac{L_0}{2} \sin \frac{\tau_b}{2}$. From the micro pipette aspiration images of Mohandas and Evans (1994) a rough approximation for the necessary curvature radius of $0.2 \mu\text{m}$ can be inferred by examining the membrane curvature at the pipette neck. For the currently employed resolution

($L_0 = 0.5 \mu\text{m}$) the limiting angle is chosen to be $\tau_b = \frac{\pi}{6}$. This choice prevents unrealistic sharp surface edges while allowing curvature radii as small as $0.18 \mu\text{m}$ to be represented.

3. The local surface conservation force acts locally on surface elements (i.e., triangles) and has the same form. It represents the combined surface response of the supporting spectrin-network and the lipid bilayer of the membrane to stretching and compression. This force is applied to all three vertices of each face and it points toward the centroid of the corresponding surface triangle.

$$F_{area} = -\frac{\kappa_a dA}{L_0} \left[1 + \frac{1}{\tau_a^2 - dA^2} \right],$$

where $dA = \frac{A_i - A_0}{A_0}$. Strong-deformation experiments of erythrocytes show that at around 40% of surface area change the membrane of most cells is damaged permanently (Li et al., 2013). We set $\tau_a = 0.3$, thus prohibiting surface area changes larger than 30%.

4. The volume conservation force is the only global term. It is used to maintain quasi-incompressibility of the cell. It is applied at each node of each surface element and it points toward the normal of the surface.

$$F_{volume} = -\frac{\kappa_v dV}{L_0} \left[\frac{1}{\tau_v^2 - dV^2} \right],$$

where $dV = \frac{V - V_0}{V_0}$, $\tau_v = 0.01$ and $\kappa_v = 20 k_B T$ is chosen to be a large but numerically still stable constant.

This constitutive model has three free parameters for RBC modeling: κ_l , κ_b , and κ_a . These are chosen to satisfy mechanical single-cell experimental results.

2.3. Implementation Of the Constitutive Forces For The New Model

The proposed forces can be realized in multiple ways on the given DEM structure, thus the implementation method is an inseparable part of the model. **Figure 2I** aids this description by showing a notation for two neighboring surface elements.

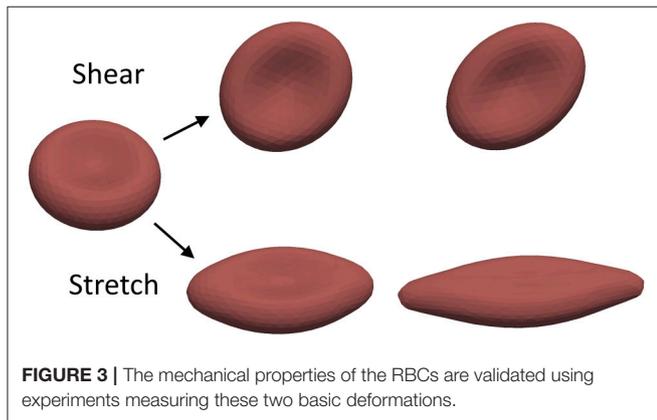
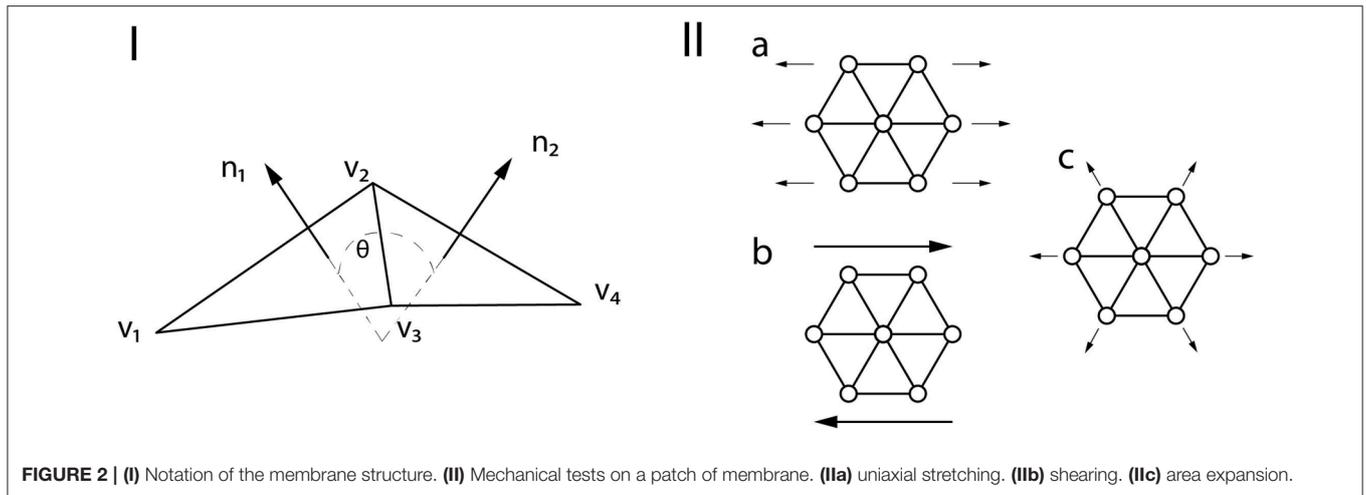
1. For each edge $\vec{e}_i, i \in [1..N_e]$ the link force \vec{F}_{link} is added to the total force acting on the end nodes of that edge (i.e., the IBM particles). Following the notation of **Figure 2I** for the edge between the nodes \vec{v}_1 and \vec{v}_2 , the resulting link forces are:

$$\vec{F}_{link_{v_1}} = F_{link} * \frac{\vec{v}_2 - \vec{v}_1}{\|\vec{v}_2 - \vec{v}_1\|} = -\vec{F}_{link_{v_2}}.$$

2. The bending force is applied for each edge $\vec{e}_i, i \in [1..N_e]$ on the four nodes of the two connecting surface elements. For the edge between the nodes \vec{v}_1 and \vec{v}_3 :

$$\vec{F}_{bend_{v_k}} = -F_{bend} * \vec{n}_k, k \in [1, 2]$$

$$\vec{F}_{bend_{v_l}} = F_{bend} * \frac{\vec{n}_1 + \vec{n}_2}{2}, l \in [3, 4].$$



- The local surface conservation force acts on each $\vec{f}_j, j \in [1..N_t]$ surface elements. For the face with the normal vector \vec{n}_1 and centroid \vec{C} :

$$\vec{F}_{area,m} = F_{area} * (\vec{C} - \vec{v}_m), m \in [1, 2, 3].$$

- The volume conservation force is applied on the three nodes of each surface element:

$$\vec{F}_{volume_j} = F_{volume} * \frac{S_j}{S_{avg}} * \vec{n}_j,$$

where S_j is the surface area of the j -th element and S_{avg} is the average surface area.

2.4. Validation Of the Mechanical RBC Responses

The free parameters of our mechanical RBC membrane model are fit to match the results of the optical-tweezer stretching experiments (Mills et al., 2004) and the Wheeler shear experiment (Yao et al., 2001). The single-cell deformation types during the measurements are shown in Figure 3.

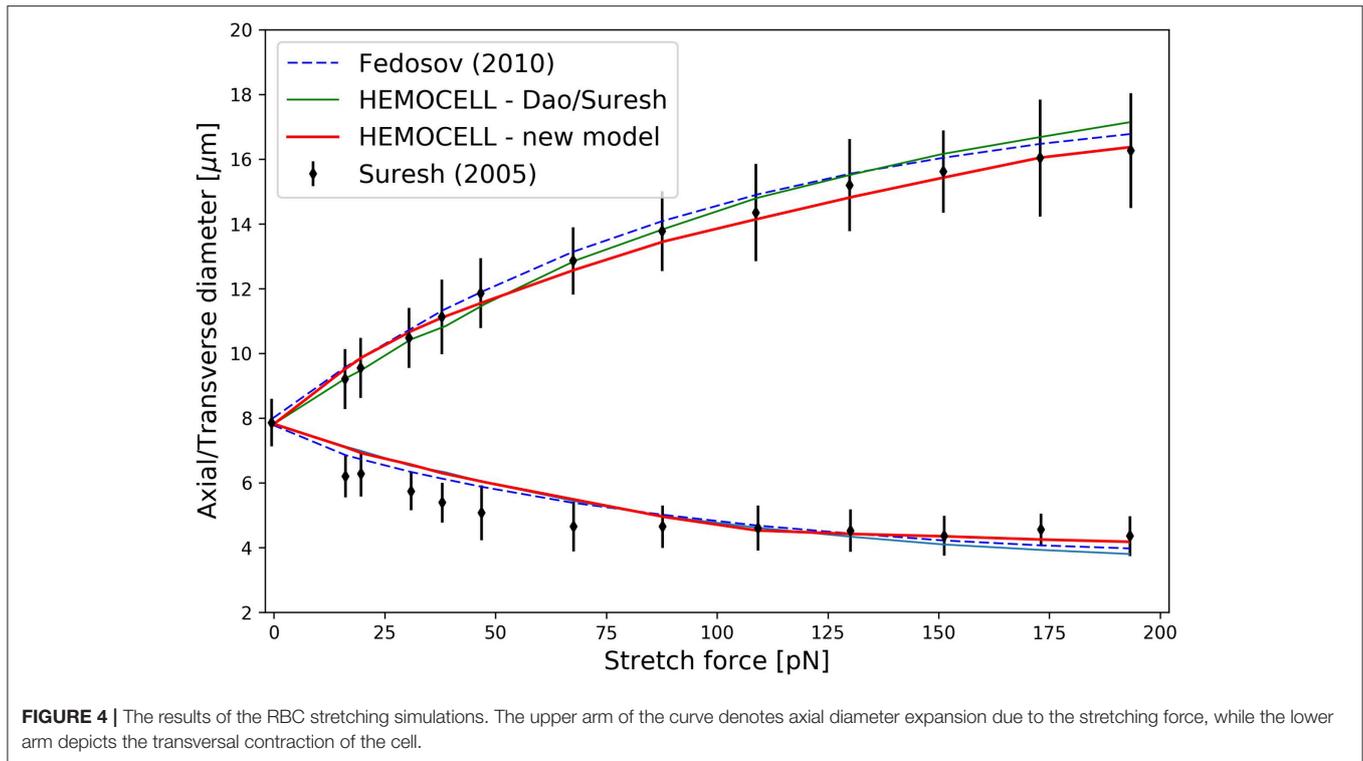
In the optical-tweezers experiment small silica beads are attached on the opposing sides of the RBC. One is then fixed to the wall of the experimental container while the other is moved away by a focused laser beam. The arising forces result in a stretching of the RBC along the longitudinal axis and contraction along the transversal axes. In our simulation the same force magnitudes are used as in the experiment. They are applied on five percent of the membrane area on the opposing ends of the RBC. These areas represent the attachment surfaces of the silica beads.

The stretching curves of the two material models implemented in Hemocell are compared to the experiment of Mills et al. (2004) and to the results of Fedosov et al. (2010a) and are shown in Figure 4. Both constitutive membrane models can reproduce the stretching behavior of a single RBC in the given forcing regime with good accuracy. However, since the responses of the different force types are more balanced in our new model (i.e., it is less likely, that one force will dominate over the others during deformation), while the spectrin-link model is over-dominated by the in-plane link-force, our model captures the transversal contraction at higher stretches with more accuracy.

In the wheeler experiment performed by Yao et al. (2001) an RBC is positioned in shear flow such that the axis of symmetry of the cell lies in the plane of the shear and is perpendicular to the flow velocity. The deformation of the RBCs is then inferred from measuring its laser diffraction pattern in the flow. We numerically compute the behavior of a single RBC placed in pure shear flow with shear rates between 17 and 200 s^{-1} , in accordance with the experiment. The deformation index of the RBC is defined as given in Yao et al. (2001):

$$DI = \frac{(D_{max}/D_0)^2 - 1}{(D_{max}/D_0)^2 + 1},$$

where D_0 is the original diameter of the RBC (7.82 μm) and D_{max} is the maximal diameter during the deformation at a constant shear rate value. The results are compared to the experimental results and to simulated results of MacMeccan et al. (2009) in Figure 5. Both material models give a deformation



index that are in agreement with Yao's experiment and with the simulations of MacMeccan et al. (2009). It is important to point out that the numerical accuracy of this simulation is more sensitive to the fluid-structure coupling compared to the previous stretching scenario, therefore, a close match with the measurements implies an accurate coupling between the plasma flow and the cell membranes. Additionally, the numerical limit of shear rate is tested for both material models with this setting. In our implementation the new material model could resist higher sustained shear rates ($\dot{\gamma}_{max} = 2,500 \text{ s}^{-1}$) than the spectrin-link model ($\dot{\gamma}_{max} = 500 \text{ s}^{-1}$) before the RBC collapsed due to insufficient force response arising from numerical errors (for the onset of such an error see the inset image in **Figure 5**).

The fit to the experimental results yield $\kappa_l = 15 k_B T$, $\kappa_a = 5 k_B T$, and $\kappa_b = 80 k_B T$. These values are used throughout this work for the new constitutive model. Evans (1983) measured the bending modulus to be in the order of $50 k_B T$, not far from our κ_b . Additionally, with the selected κ_a value the local surface extensions under physiological flow conditions are smaller than the set limit of 30%, typically below 7%, which agrees with the literature (Fung, 1993). Please note that the selection of these parameters is not unique, other sets might exist that also fit the single-cell experimental results well with the proposed mechanical model. To infer further material characteristics of the model, we employed a simulation to deform a single hexagonal patch of the membrane (for an overview of the applied deformations see **Figure 2II**). The uniaxial stretching yields a surface Young modulus of $E_s = 27.82 \mu\text{N}/\text{m}$. Assuming that the major deformation response arise from the membrane (the bilipid layer, and the spectrin, actin filaments) and its width

in the range of $25 - 50 \text{ nm}$ (Gov and Safran, 2005; Yoon et al., 2009), the typical Young modulus for small deformations $E = 1 \text{ kPa}$ of healthy RBCs (Maciaszek and Lykotrafitis, 2011) gives the surface tensile modulus of $E_s = 25 - 50 \mu\text{N}/\text{m}$, in agreement with our results. The shear deformation of the patch yields $\mu = 10.87 \mu\text{N}/\text{m}$, close to the upper region of the measured ranges of $6 - 10 \mu\text{N}/\text{m}$ (Mohandas and Evans, 1994; Park et al., 2011). Finally, area expansion gives a compression modulus of $K = 21.88 \mu\text{N}/\text{m}$ near the reported range of $18 - 20 \mu\text{N}/\text{m}$ (Park et al., 2011). Assuming homogeneous isotropic linear behavior (that only holds for small deformations), the relation between the elastic constants yields a Poisson's ratio of 0.29, in the vicinity of the expected value of 1/3.

From the unique material properties an emergent ability of RBCs traveling in small, confined flows is their deformation to parachute-like shapes (Noguchi and Gompper, 2005). This behavior is necessary to pass small micro-capillaries of diameters below the diameter of the undeformed RBC (Tsukada et al., 2001). **Figure 6** shows an example of a simulated RBC that deforms toward this shape in a tight channel computed with the new material model.

2.5. Generating Cell Initial Conditions

An important component of simulating blood flows on a cellular level is the selection of initial conditions for the cells, such as position and orientation. These are far from trivial since, due to the biconcave shape of RBCs, their volume ($\approx 71 \mu\text{m}^3$) compared to the volume of their enclosing box ($\approx 224 \mu\text{m}^3$) is low. Using the densest possible packing along a regular grid thus yields a hematocrit of 32% which is often inadequate as it

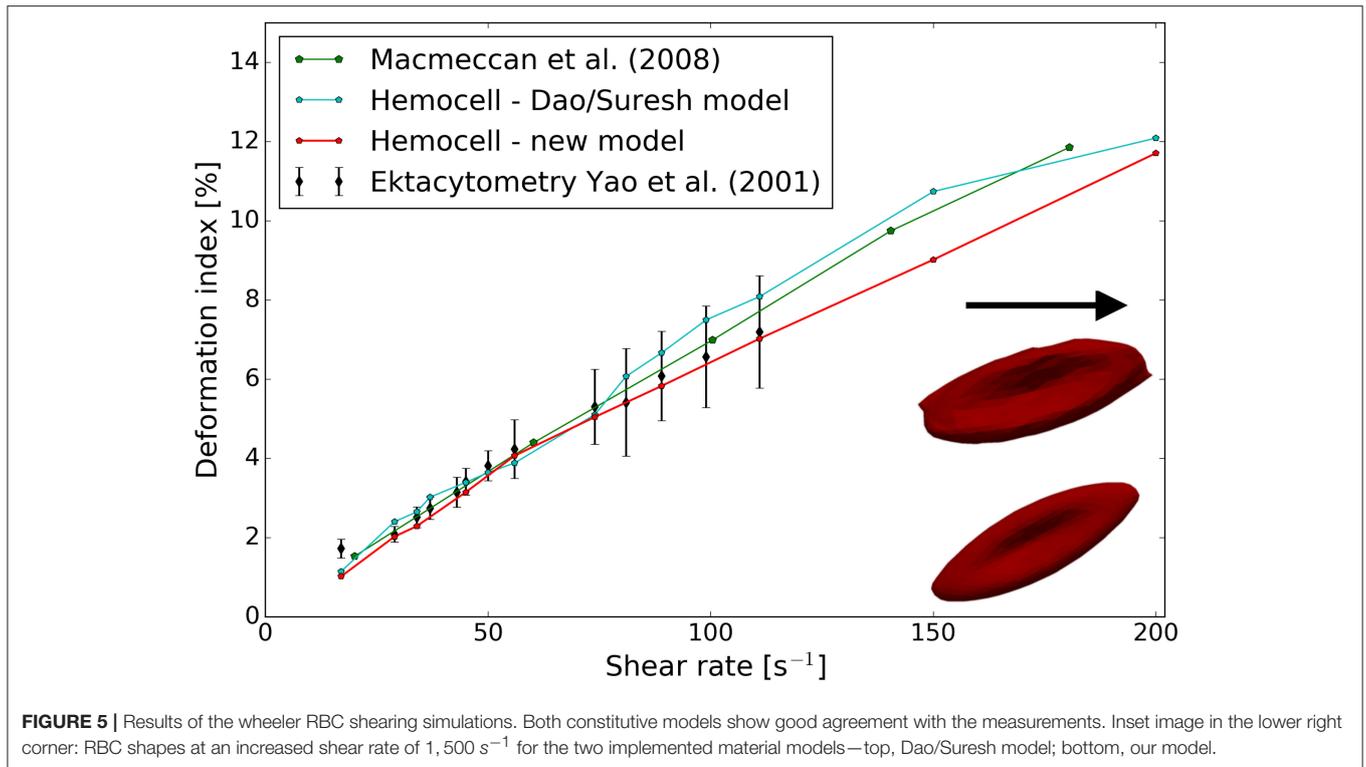


FIGURE 5 | Results of the wheeler RBC shearing simulations. Both constitutive models show good agreement with the measurements. Inset image in the lower right corner: RBC shapes at an increased shear rate of 1,500 s⁻¹ for the two implemented material models—top, Dao/Suresh model; bottom, our model.

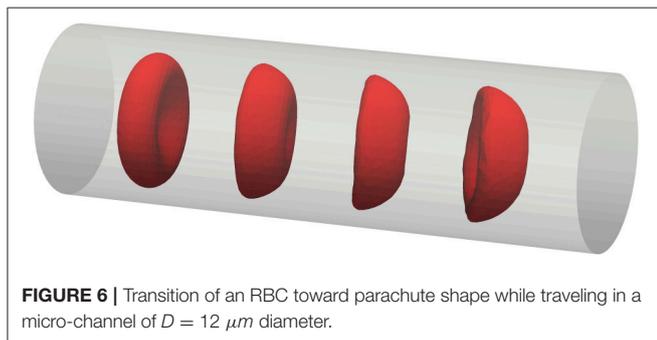


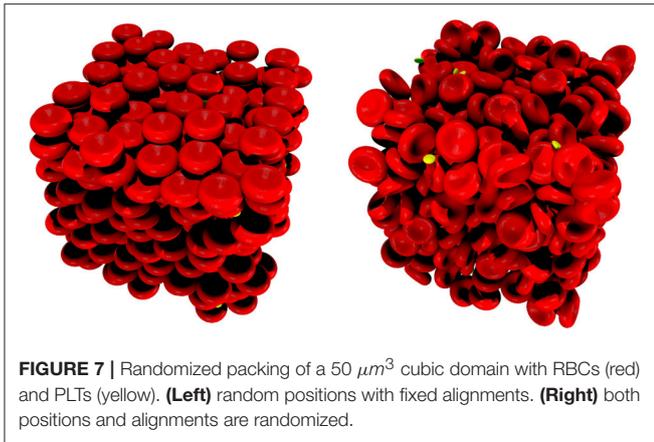
FIGURE 6 | Transition of an RBC toward parachute shape while traveling in a micro-channel of $D = 12 \mu\text{m}$ diameter.

does not reach the level of physiologic hematocrit of human blood. A further issue is the need for a randomized distribution to avoid initial artifacts originating from the regular positioning and orientations. To circumvent these difficulties an additional kinetic simulation was developed to compute realistic initial distributions even at high hematocrit values. Instead of the real biconcave shapes, the enclosing ellipsoid of the RBCs were used to execute a simple kinetic process for hard ellipsoid packing. The so-called the force-bias model (Mościński et al., 1989; Bargieł and Mościński, 1991; Bezrukov et al., 2002) was applied to these enclosing ellipsoids. The algorithm proceeds as follows. The positions of the center of the cells are randomly distributed in the simulation domain. Next, two scaling variables are defined for every cell type (e.g., RBC, platelet): d^{in} represents the possible largest scaling in the system without any overlap between the cells. While d^{out} is initially set so that the merged volume

(counting overlapping volumes only once) of all the ellipsoids scaled with it equals the total volume of the enclosing ellipsoids corresponding to the desired hematocrit level. Then, a repulsive force is applied between overlapping ellipsoids, proportional to the volume of the overlapping regions:

$$\vec{F}_{ij} = \delta_{ij} p_{ij} \frac{\vec{r}_j - \vec{r}_i}{|\vec{r}_j - \vec{r}_i|},$$

where δ_{ij} equals 1 if there is an overlap between particle i and j and 0 otherwise, while p_{ij} is a chosen potential function. In our case, the potential function was selected to be proportional to the overlapping volume of the d^{out} scaled particles. The positions are updated following Newtonian mechanics where mass is proportional to the particle scaling radius. This ensures that larger particles will move slower than smaller ones (i.e., an RBC will push away a platelet rather than the other way around). The rearrangement of the cells have a tendency of increasing d^{in} . As a final step the size of d^{out} is reduced every iteration according to a chosen contraction rate τ . The computation stops when $d^{out} \leq d^{in}$ at which point the system is force-free, since there are no overlaps. Using this method, we were able to push the initial hematocrit value up to 46% covering the physiological range. Additionally, we can fix the orientation of the cells by only allowing translation of their center of mass during this computation, thus predefining the alignment of the particles. This is beneficial for initializing higher velocity flows where the cells are expected to be lined up with the bulk flow direction. **Figure 7** presents two sample initial conditions generated with this method.



It is possible to initialize simulations of up to 10^6 cells efficiently this way. These simulations are free from regular-grid positioning artifacts from the beginning, which in turn reduces the computational time significantly. Though the actual computational cost it saves varies by geometry, hematocrit, flow velocity, etc., in our simulations the warm-up phase needed to allow the initially regularly placed cells to arrange more realistically amounted to 10–30% of the total computational time, while with the randomized initialization this whole phase could be omitted.

3. RESULTS

Our ultimate goal of accurate mechanical modeling of cellular membranes in blood flows is to allow for the resolution of the collective transport dynamics and coupling this to relevant biochemical processes. In the following, these transport properties are explored using the new constitutive model in the cases of a straight vessel sections of varying diameters. A snapshot from the simulation of the $D = 128 \mu\text{m}$ case is visualized in **Figure 8**. The RBCs close to the wall experience much larger deformations than those in the center of the channel. In every simulation PLTs are present as well in a physiologic concentration (around 1/10th of the RBC cell count). Since the elastic response of the unactivated platelets are at least an order of magnitude stronger for small deformations than the response of RBCs (Haga et al., 1998), the platelets are simulated with the same constitutive model as RBCs, however, the constants $\kappa_l, \kappa_a, \kappa_b$ are multiplied by 10. These simulations also benefit from the above mentioned randomized initialization of the cells.

The first fundamental transport property examined is the apparent viscosity. The results are compared to the experimental results collected by Pries et al. (1992). These experimental results are aggregated for hematocrit levels of 20, 45, and 60% after a correction for temperature and medium viscosity. Based on these data, an empirical formula is also derived in Pries et al. (1992) which was used in the current work to produce the expected results for the hematocrit level of 30%. It can be insightful to briefly overview the general measurement method of blood viscosity in experimental settings. The hematocrit level refers to the discharge hematocrit present in the blood tank, from

where the flow is directed through a tube of various diameters driven by hydrostatic pressure. The relation of the pressure and the appearing average flow velocity in the tube defines the viscosity. This is taken into account in the current simulations by translating the discharge hematocrit values to hematocrit values actually present in the tube during the measurements by applying Equation (8) from (Pries et al., 1992). The simulations are initialized with zero velocity in the whole domain after which the flow is started up and driven by external body force. The results together with the experimental results are shown in **Figure 9**.

The results show good agreement with the measurements. For the simulations of $H = 45\%$ after the initialization the undeformed cells create large clusters. In the current work, the notation of an RBC cluster refers to a group of RBCs having at least a single membrane point in touch with another RBC of the same cluster. In the initial phase of the simulations the elastic effects of these RBC clusters are perceivable as the viscosity during the first few milliseconds increases quickly, well above the expected values. This is caused by the deformation of the cells residing inside these large and dense clusters and this behavior is one of the major components that leads to yield-stress. After a critical threshold in shape deformation they loose these stable structures and the viscosity quickly settles back to the expected level. For more details see Sections 3.1 and 3.2.

Another distinctive feature of cellular suspension flows is the formation of a cell-free layer (CFL) close to the walls as a result of lift force acting on the cells. The width of the appearing cell-free zones are defined using the density distribution of cells. It is the distance from the wall at which point the density distribution averaged along the vessel section reaches 5%. The results are compared to the *in vitro* experiments at different hematocrit levels of Tateishi et al. (1994) in **Figure 10**.

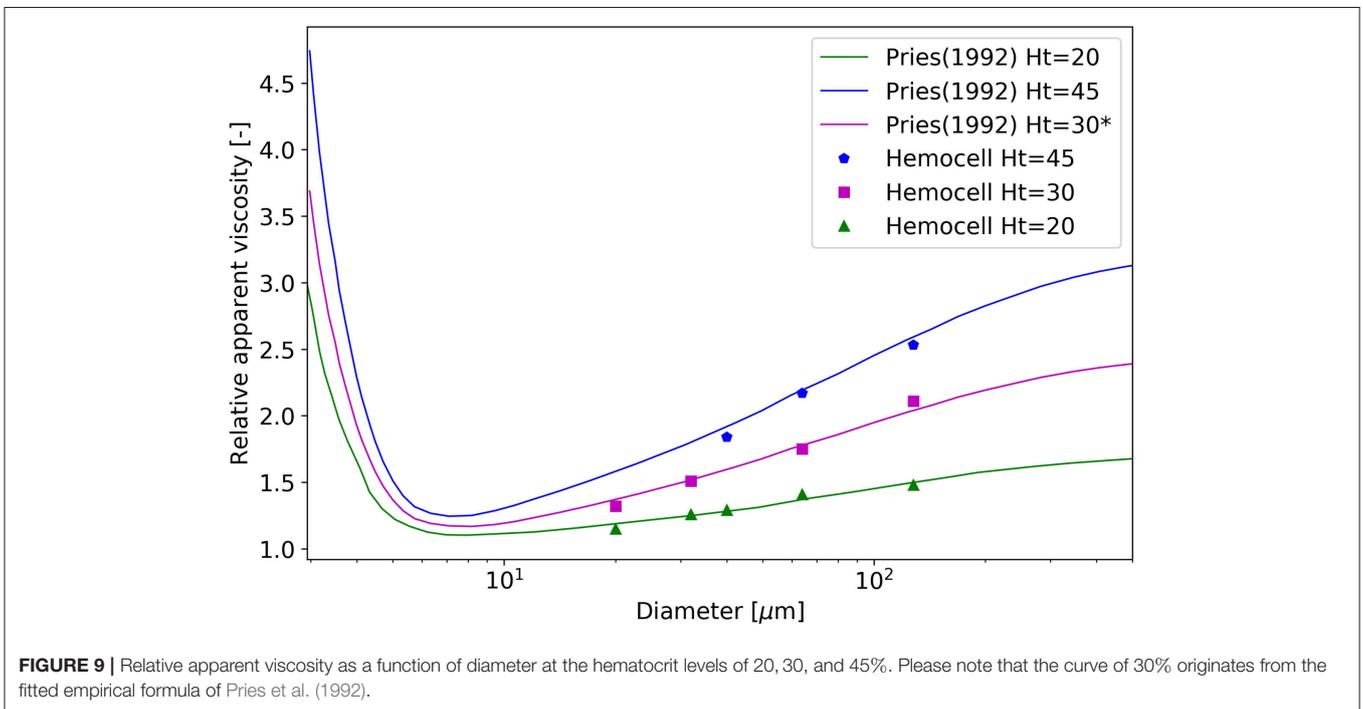
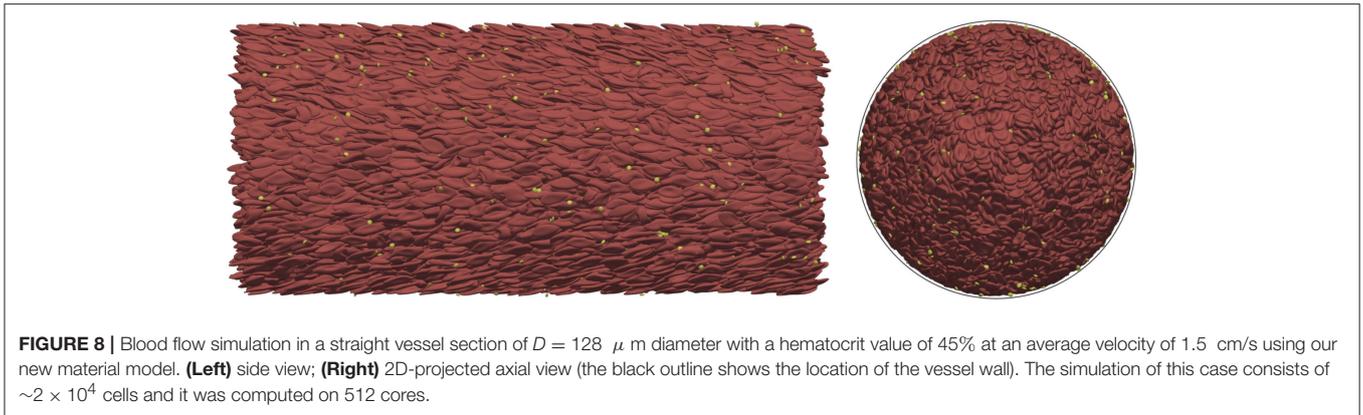
While our simulated diameter range surpasses the bounds of the experimental range, the overlapping region shows good agreement for the hematocrit level of 30 and 45%. The level of 20% does not have a directly corresponding measurement, however, it is situated between between the experimental results of 16% and 30%, as expected. For a given diameter the CFL decreases with the increase of hematocrit as more RBCs are packed into the same domain volume.

Finally, to validate the flow profile in stationary flow a straight, rectangular channel was set up to recreate the flow environment of the experimental work of Carboni et al. (2016). The hematocrit level was set to 35%, and the driving body-force was calibrated to have a volumetric flow rate matching the experiment. In **Figure 11**, the velocity profile was compared to the profile obtained from the PIV measurements.

The simulated profile fits the measurement well and has the same plug-shape along with similar widths of high-shear regions at the sides of the channel.

3.1. Break-up of the RBC Structures At Increasing Shear-Rates

It is a well-known phenomenon that toward low shear-rate values the viscosity of blood increases steeply (Chien, 1970). This is caused by the formation of dense clusters of RBCs including rouleaux structures. In our simulations, aggregation interactions



between cells were not included, thus, these structures arise from the various alignments and high density of the cells. This effect was investigated in the case of the $D = 128 \mu\text{m}$ vessel section at $H = 45\%$. The whole system is initialized to be still. Then, it is driven by a constant body force, and once the average velocity equilibrates (typically after a few hundred ms) the relative apparent viscosity is recorded. The shear is not constant along the radius of the vessel, however, for slow flows its local value scales approximately linearly with the average velocity. **Figure 12** shows the relative apparent viscosity of the whole vessel section at low average velocities.

The relation appears to be logarithmic (see the fitted exponential decay), which is in agreement with the literature (Baskurt and Meiselman, 1997). Around the average velocity of 1 cm/s the apparent viscosity of the vessel section already settles suggesting that the majority of the RBC structures are gone. The

further increase in velocity from this point on only results in a minor change of bulk viscosity.

3.2. Effects Of the Initial RBC Deformation

Due to the elastic deformable nature of RBCs, blood can exhibit yield-stress behavior if the hematocrit level is high enough (Picart et al., 1998). In such a dense suspension of cells under low shear-stress the clusters can behave similarly to deformable solids. The relative positions of the RBCs within these clusters remain the same while they deform. At a critical stress value the force required to further deform the cells becomes larger than the force needed to separate them, thus breaking the structure. From that point blood transitions to fluid-like behavior. The stability of these clusters is dependent on several variables, for instance the level of hematocrit and the concentration of fibrinogen in blood plasma (Baskurt and Meiselman, 1997). However, a weaker

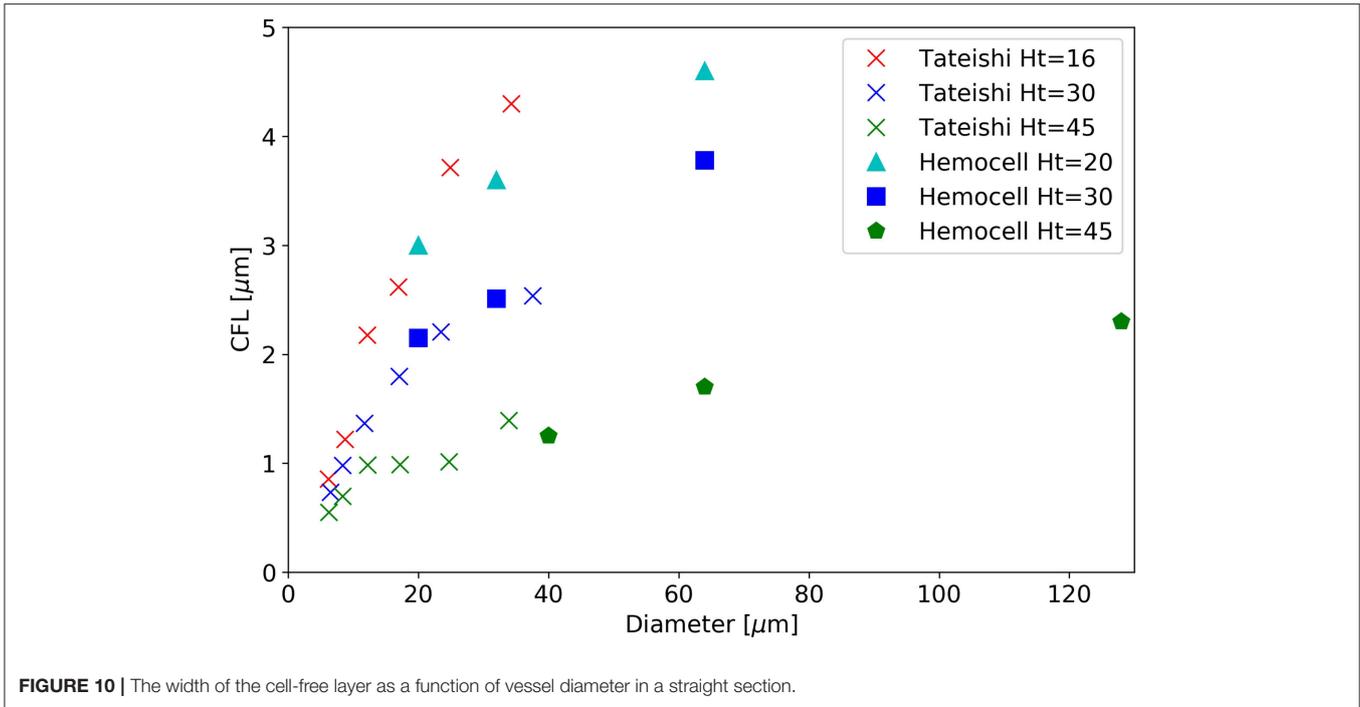


FIGURE 10 | The width of the cell-free layer as a function of vessel diameter in a straight section.

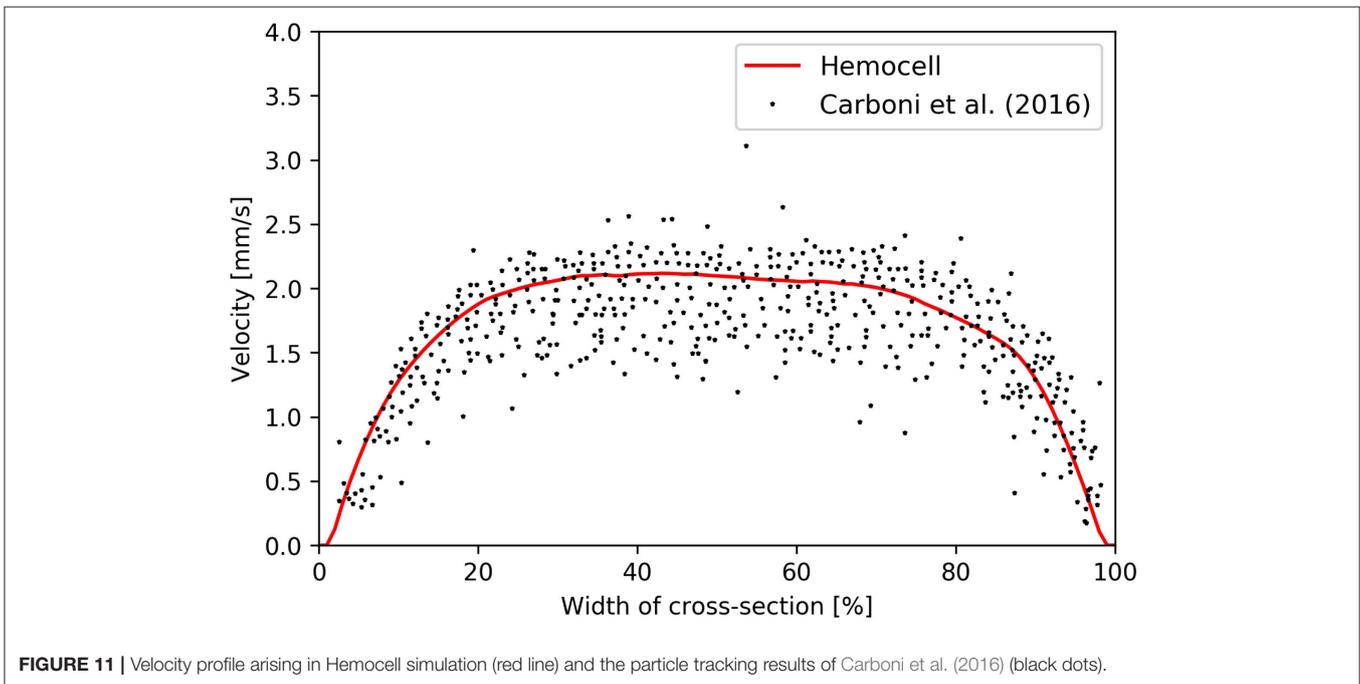


FIGURE 11 | Velocity profile arising in Hemocell simulation (red line) and the particle tracking results of Carboni et al. (2016) (black dots).

yield-stress behavior still arises in the absence of fibrinogen (and other endogen proteins) at high hematocrit values ($> 30\%$) (Blackshear et al., 1983; Morris et al., 1989). This effect is perceivable during some phases of the simulations, such as the initial start up of the flows in our straight vessel sections. To investigate it, the plasma was brought up to the stationary velocity driven by external body force without any cells. This is necessary

to separate the effects of initial cell deformations from the effects of initially driving the fluid up to the desired velocity. The velocity is set to a high enough value (e.g., $6 \frac{mm}{s}$ for the $D = 64 \mu m$, $H = 45\%$ case and $1.5 \frac{cm}{s}$ for the $D = 128 \mu m$, $H = 45\%$ case) for the large RBC structures to break. The undeformed cells with randomized positions and alignments are then placed into the flow while the driving force is kept constant. This moment

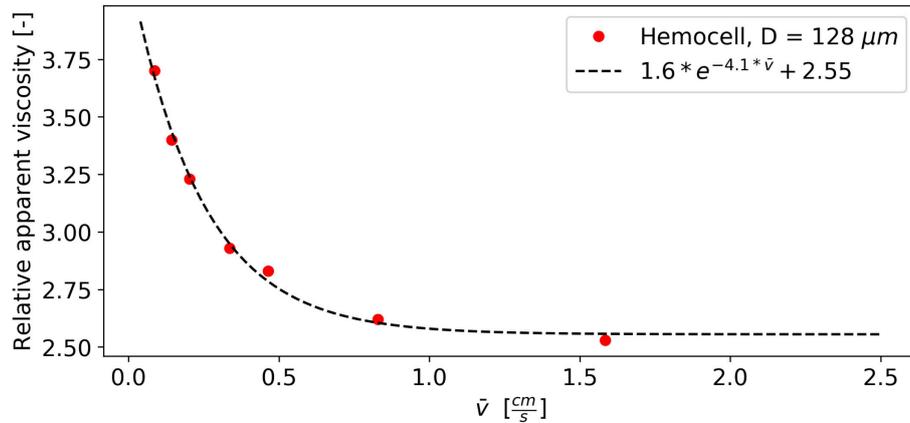


FIGURE 12 | Viscosity drops as clusters are breaking up with the increase of flow velocity and shear.

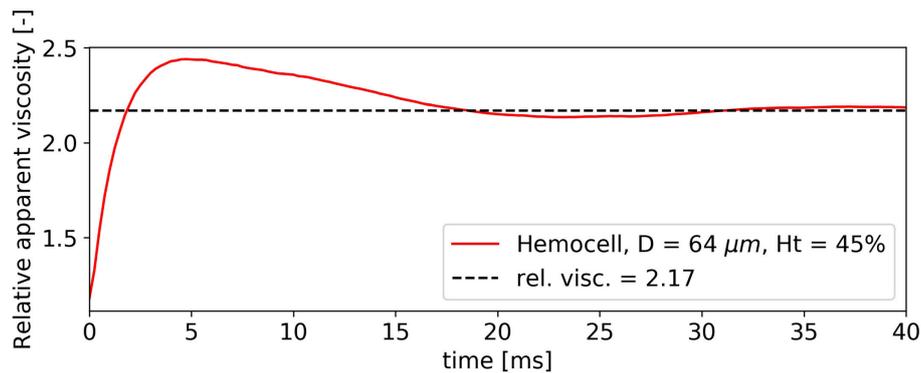


FIGURE 13 | Elastic behavior of dense RBC clusters cause an initial viscosity peak.

is denoted as $t = 0$ s. **Figure 13** shows the progression of the relative viscosity from this point in the case of $D = 64 \mu\text{m}$, $H = 45\%$.

During the first 3 ms the relative viscosity rises from the value of 1 steeply while the plasma flow slows down. At this stage, the RBCs do not flow but deform. The local velocity in the fluid corresponds to the deformation velocity of the cells. Around 4 ms, the relative viscosity reaches its peak value and the clusters start to break up, i.e., the relative positions of the RBCs start to change and the suspension no longer displays solid-like features. After this point blood quickly settles back to its stable final relative viscosity. The same viscosity pattern is observable for all simulations with $H = 45\%$ during the initial phase, however, for smaller diameters the phenomenon is less significant. It must be noted, however, that in our case both the surface and the cytoplasmic viscosity was the same as the plasma viscosity, while experimental results suggest higher values of 2–6 mPas for cytoplasm (Park et al., 2011) and $10^{-10} - 10^{-9}$ Ns/m for the bilipid membrane (Waugh, 1982; Evans and Yeung, 1994). This difference is likely to have a strong effect on the characteristic times of cell deformation that is not investigated here.

4. CONCLUSIONS

The novel material model produces results in good agreement with several experiments targeting both single-cell mechanics and collective transport behavior. It also performs well for higher shear rate values where the other investigated model might fall short. It is capable of capturing the emerging solid-like behavior of dense RBC suspensions under low shear-rates. Furthermore, since our RBC material model is able to handle strong deformations coupling it with the LBM method for the plasma flow which operates at very small time-steps (in the order of 10^{-8} s for the demonstrated flows) allows for small scale transient effects such as flow instabilities behind obstacles (e.g., stenosis or micro medical devices) to be simulated as well.

The framework itself is structured to be easy to extend with additional material models and cell types, e.g., white blood cells, and with other fields, such as concentrations of different chemical components as well as with new biophysical processes, for instance bond formations. The efficient highly parallel implementation is capable of handling large domain sizes, thus it is able to cover the range between cell-based micro-scale and macroscopic domains.

The demonstrated capabilities make this framework in combination with our constitutive model an ideal environment for exploring the transport effects of blood flows *in-silico*. It forms a solid ground for resolving accurate transport mechanics in vascular flows as a necessary component for modeling complex phenomena such as cell aggregation around micro-medical devices, thrombus formation and rheological response of diseases effecting RBC mechanical properties.

AUTHOR CONTRIBUTIONS

GZ conceived the research, designed the model and wrote 50% the paper; BvR collected and analyzed the experimental data, validated the Dao/Suresh model in our implementation, revised the new material model, and wrote 50% of the paper; VA

contributed to the technical realization of Hemocell and revised the final version of the paper; AH conceived and supervised the research, and revised the manuscript. All authors read and approved the final version of the manuscript.

FUNDING

This work was supported by the European Union Horizon 2020 research and innovation programme under grant agreement no. 675451, the CompBioMed project and grant agreement no. 671564, the ComPat project. This work was sponsored by NWO Exacte Wetenschappen (Physical Sciences) for the use of supercomputer facilities, with financial support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organization for Science Research, NWO).

REFERENCES

- Anzai, H., Falcone, J.-L., Chopard, B., Hayase, T., and Ohta, M. (2014). Optimization of strut placement in flow diverter stents for four different aneurysm configurations. *J. Biomech. Eng.* 136:061006. doi: 10.1115/1.4027411
- Barbee, J. H., and Cokelet, G. R. (1971). The fahraeus effect. *Microvasc. Res.* 3, 6–16. doi: 10.1016/0026-2862(71)90002-1
- Bargiel, M., and Mościński, J. (1991). C-language program for the irregular close packing of hard spheres. *Comput. Phys. Commun.* 64, 183–192. doi: 10.1016/0010-4655(91)90060-X
- Baskurt, O. K., and Meiselman, H. J. (1997). Cellular determinants of low-shear blood viscosity. *Biorheology* 34, 235–247. doi: 10.1016/S0006-355X(97)00027-9
- Beck, M. Jr., and Eckstein, E. (1980). Preliminary report on platelet concentration in capillary tube flows of whole blood. *Biorheology* 17:455.
- Bezrukov, A., Bargiel, M., and Stoyan, D. (2002). Statistical analysis of simulated random packings of spheres. *Particle Particle Syst. Charact.* 19:111. doi: 10.1002/1521-4117(200205)19:2<111::AID-PPSC111>3.0.CO;2-M
- Björkman, S. (1959). A new method for enumeration of platelets. *Acta Haematol.* 22, 377–379. doi: 10.1159/000205797
- Blackshear, P. L., Deetz, D. C., Morris, C., and Kayser, S. J. (1983). Yield stress in sedimented red blood cell layers. *Ann. New York Acad. Sci.* 416, 263–275. doi: 10.1111/j.1749-6632.1983.tb35193.x
- Bouvy, W. H., Geurts, L. J., Kuijff, H. J., Luijten, P. R., Kappelle, L. J., Biessels, G. J., et al. (2016). Assessment of blood flow velocity and pulsatility in cerebral perforating arteries with 7-T quantitative flow MRI. *NMR Biomed.* 29, 1295–1304. doi: 10.1002/nbm.3306
- Carboni, E. J., Bognet, B. H., Bouchillon, G. M., Kadilak, A. L., Shor, L. M., Ward, M. D., et al. (2016). Direct tracking of particles and quantification of margination in blood flow. *Biophys. J.* 111, 1487–1495. doi: 10.1016/j.bpj.2016.08.026
- Chang, H.-Y., Li, X., Li, H., and Karniadakis, G. E. (2016). MD/DPD multiscale framework for predicting morphology and stresses of red blood cells in health and disease. *PLoS Comput. Biol.* 12:e1005173. doi: 10.1371/journal.pcbi.1005173
- Chien, S. (1970). Shear dependence of effective cell volume as a determinant of blood viscosity. *Science* 168, 977–979. doi: 10.1126/science.168.3934.977
- Dao, M., Li, J., and Suresh, S. (2006). Molecularly based analysis of deformation of spectrin network and human erythrocyte. *Mater. Sci. Eng.* 26, 1232–1244. doi: 10.1016/j.msec.2005.08.020
- Davies, B., and Morris, T. (1993). Physiological parameters in laboratory animals and humans. *Pharmaceut. Res.* 10, 1093–1095. doi: 10.1023/A:1018943613122
- Ding, Y., Lewis, D., Kadirvel, R., Dai, D., and Kallmes, D. F. (2011). The woven endobridge: a new aneurysm occlusion device. *Amer. J. Neuroradiol.* 32, 607–611. doi: 10.3174/ajnr.A2399
- Evans, E., and Yeung, A. (1994). Hidden dynamics in rapid changes of bilayer shape. *Chem. Phys. Lipids* 73, 39–56. doi: 10.1016/0009-3084(94)90173-2
- Evans, E. A. (1983). Bending elastic modulus of red blood cell membrane derived from buckling instability in micropipet aspiration tests. *Biophys. J.* 43, 27–30. doi: 10.1016/S0006-3495(83)84319-7
- Fay, M. E., Myers, D. R., Kumar, A., Turbyfield, C. T., Byler, R., Crawford, K., et al. (2016). Cellular softening mediates leukocyte demargination and trafficking, thereby increasing clinical blood counts. *Proc. Natl. Acad. Sci. U.S.A.* 113, 1987–1992. doi: 10.1073/pnas.1508920113
- Fedosov, D., Caswell, B., Suresh, S., and Karniadakis, G. (2011a). Quantifying the biophysical characteristics of plasmodium-falciparum-parasitized red blood cells in microcirculation. *Proc. Natl. Acad. Sci. U.S.A.* 108, 35–39. doi: 10.1073/pnas.1009492108
- Fedosov, D. A., Caswell, B., and Karniadakis, G. E. (2010a). Systematic coarse-graining of spectrin-level red blood cell models. *Comput. Methods Appl. Mech. Eng.* 199, 1937–1948. doi: 10.1016/j.cma.2010.02.001
- Fedosov, D. A., Caswell, B., Popel, A. S., and Karniadakis, G. E. (2010b). Blood flow and cell-free layer in microvessels. *Microcirculation* 17, 615–628. doi: 10.1111/j.1549-8719.2010.00056.x
- Fedosov, D. A., and Gompper, G. (2014). White blood cell margination in microcirculation. *Soft Matter* 10, 2961–2970. doi: 10.1039/C3SM52860J
- Fedosov, D. A., Lei, H., Caswell, B., Suresh, S., and Karniadakis, G. E. (2011b). Multiscale modeling of red blood cell mechanics and blood flow in malaria. *PLoS Comput. Biol.* 7:e1002270. doi: 10.1371/journal.pcbi.1002270
- Fedosov, D. A., Noguchi, H., and Gompper, G. (2014). Multiscale modeling of blood flow: from single cells to blood rheology. *Biomech. Model. Mechanobiol.* 13, 239–258. doi: 10.1007/s10237-013-0497-9
- Fedosov, D. A., Pan, W., Caswell, B., Gompper, G., and Karniadakis, G. E. (2011c). Predicting human blood viscosity *in silico*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11772–11777. doi: 10.1073/pnas.1101210108
- Fung, Y.-C. (1993). *Biomechanics: Mechanical Properties of Living Tissues*. New York, NY: Springer-Verlag.
- Gompper, G., and Kroll, D. M. (1996). Random surface discretizations and the renormalization of the bending rigidity. *J. Phys.* 6, 1305–1320. doi: 10.1051/jp1:1996246
- Gov, N., and Safran, S. (2005). Red blood cell membrane fluctuations and shape controlled by atp-induced cytoskeletal defects. *Biophys. J.* 88, 1859–1874. doi: 10.1529/biophysj.104.045328
- Gross, M., Krüger, T., and Varnik, F. (2014). Rheology of dense suspensions of elastic capsules: normal stresses, yield stress, jamming

- and confinement effects. *Soft Matter* 10, 4360–4372. doi: 10.1039/c4sm00081a
- Haga, J. H., Beaudoin, A. J., White, J. G., and Strony, J. (1998). Quantification of the passive mechanical properties of the resting platelet. *Ann. Biomed. Eng.* 26, 268–277. doi: 10.1114/1.118
- Hashemi, Z., and Rahnama, M. (2016). Numerical simulation of transient dynamic behavior of healthy and hardened red blood cells in microcapillary flow. *Int. J. Numer. Methods Biomed. Eng.* 32:e02763. doi: 10.1002/cnm.2763
- Helfrich, W. (1973). Elastic properties of lipid bilayers: theory and possible experiments. *Zeitschrift für Naturforschung C* 28, 693–703. doi: 10.1515/znc-1973-11-1209
- Hoekstra, A. G., Alowayyed, S., Lorenz, E., Melnikova, N., Mountrakis, L., van Rooij, B., et al. (2016). Towards the virtual artery: a multiscale model for vascular physiology at the physics–chemistry–biology interface. *Philos. Trans. R. Soc. A* 374:20160146. doi: 10.1098/rsta.2016.0146
- Kim, S., Ong, P. K., Yalcin, O., Intaglietta, M., and Johnson, P. C. (2009). The cell-free layer in microvascular blood flow. *Biorheology* 46, 181–189.
- Krüger, T. (2016). Effect of tube diameter and capillary number on platelet margination and near-wall dynamics. *Rheologica Acta* 55, 511–526. doi: 10.1007/s00397-015-0891-6
- Krüger, T., Gross, M., Raabe, D., and Varnik, F. (2013). Crossover from tumbling to tank-treading-like motion in dense simulated suspensions of red blood cells. *Soft Matter* 9, 9008–9015. doi: 10.1039/c3sm51645h
- Krüger, T., Varnik, F., and Raabe, D. (2011). Efficient and accurate simulations of deformable particles immersed in a fluid using a combined immersed boundary lattice boltzmann finite element method. *Comput. Math. Appl.* 61, 3485–3505. doi: 10.1016/j.camwa.2010.03.057
- Lagrava, D., Malaspinas, O., Latt, J., and Chopard, B. (2012). Advances in multi-domain lattice boltzmann grid refinement. *J. Comput. Phys.* 231, 4808–4822. doi: 10.1016/j.jcp.2012.03.015
- Li, F., Chan, C. U., and Ohl, C. D. (2013). Yield strength of human erythrocyte membranes to impulsive stretching. *Biophys. J.* 105, 872–879. doi: 10.1016/j.bpj.2013.06.045
- Li, J., Dao, M., Lim, C., and Suresh, S. (2005). Spectrin-level modeling of the cytoskeleton and optical tweezers stretching of the erythrocyte. *Biophys. J.* 88, 3707–3719. doi: 10.1529/biophysj.104.047332
- Liu, F., Burgess, J., Mizukami, H., and Ostafin, A. (2003). Sample preparation and imaging of erythrocyte cytoskeleton with the atomic force microscopy. *Cell Biochem. Biophys.* 38, 251–270. doi: 10.1385/CBB:38:3:251
- Lubicz, B., Collignon, L., Raphaeli, G., Pruvot, J.-P., Bruneau, M., De Witte, O., et al. (2010). Flow-diverter stent for the endovascular treatment of intracranial aneurysms. *Stroke* 41, 2247–2253. doi: 10.1161/STROKEAHA.110.589911
- Maciaszek, J. L., and Lykotraftis, G. (2011). Sickle cell trait human erythrocytes are significantly stiffer than normal. *J. Biomech.* 44, 657–661. doi: 10.1016/j.jbiomech.2010.11.008
- MacMeccan, R. M., Clausen, J., Neitzel, G., and Aidun, C. (2009). Simulating deformable particle suspensions using a coupled lattice-boltzmann and finite-element method. *J. Fluid Mech.* 618:13. doi: 10.1017/S0022112008004011
- Maeda, N., Suzuki, Y., Tanaka, J., and Tateishi, N. (1996). Erythrocyte flow and elasticity of microvessels evaluated by marginal cell-free layer and flow resistance. *Am. J. Physiol. Heart Circul. Physiol.* 271, H2454–H2461.
- Mehrabadi, M., Ku, D. N., and Aidun, C. K. (2016). Effects of shear rate, confinement, and particle parameters on margination in blood flow. *Phys. Rev. E* 93:023109. doi: 10.1103/PhysRevE.93.023109
- Merrill, E. W., and Pelletier, G. A. (1967). Viscosity of human blood: transition from newtonian to non-newtonian. *J. Appl. Physiol.* 23, 178–182.
- Mills, J. P., Qie, L., Dao, M., Lim, C. T., and Suresh, S. (2004). Nonlinear elastic and viscoelastic deformation of the human red blood cell with optical tweezers. *Mech. Chem. Biosyst.* 1, 169–180.
- Mohandas, N., and Evans, E. (1994). Mechanical properties of the red cell membrane in relation to molecular structure and genetic defects. *Ann. Rev. Biophys. Biomol. Struct.* 23, 787–818. doi: 10.1146/annurev.bb.23.060194.004035
- Morris, C. L., Rucknagel, D. L., Shukla, R., Gruppo, R. A., Smith, C. M., and Blackshear, P. (1989). Evaluation of the yield stress of normal blood as a function of fibrinogen concentration and hematocrit. *Microvasc. Res.* 37, 323–338. doi: 10.1016/0026-2862(89)90050-2
- Mościński, J., Bargieł, M., Rycerz, Z., and Jacobs, P. (1989). The force-biased algorithm for the irregular close packing of equal hard spheres. *Mol. Simulat.* 3, 201–212. doi: 10.1080/08927028908031373
- Mountrakis, L. (2015). *Transport of Blood Cells Studied with Fully Resolved Models*. Ph.D. thesis, University of Amsterdam.
- Mountrakis, L., Lorenz, E., and Hoekstra, A. (2016). Scaling of shear-induced diffusion and clustering in a blood-like suspension. *EPL (Europhys. Lett.)* 114:14002. doi: 10.1209/0295-5075/114/14002
- Mountrakis, L., Lorenz, E., and Hoekstra, A. G. (2014). Validation of an efficient two-dimensional model for dense suspensions of red blood cells. *Int. J. Modern Phys. C* 25:1441005. doi: 10.1142/S0129183114410058
- Mountrakis, L., Lorenz, E., Malaspinas, O., Alowayyed, S., Chopard, B., and Hoekstra, A. G. (2015). Parallel performance of an ib-lbm suspension simulation framework. *J. Comput. Sci.* 9, 45–50. doi: 10.1016/j.jocs.2015.04.006
- Nesbitt, W. S., Westein, E., Tovar-Lopez, F. J., Tolouei, E., Mitchell, A., Fu, J., et al. (2009). A shear gradient-dependent platelet aggregation mechanism drives thrombus formation. *Nat. Med.* 15, 665–673. doi: 10.1038/nm.1955
- Noguchi, H., and Gompper, G. (2005). Shape transitions of fluid vesicles and red blood cells in capillary flows. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14159–14164. doi: 10.1073/pnas.0504243102
- Pan, W., Caswell, B., and Karniadakis, G. E. (2010). A low-dimensional model for the red blood cell. *Soft Matter* 6, 4366–4376. doi: 10.1039/c0sm00183j
- Park, Y., Best, C. A., Kuriabova, T., Henle, M. L., Feld, M. S., Levine, A. J., et al. (2011). Measurement of the nonlinear elasticity of red blood cell membranes. *Phys. Rev. E* 83:051925. doi: 10.1103/PhysRevE.83.051925
- Picart, C., Piau, J.-M., Galliard, H., and Carpentier, P. (1998). Human blood shear yield stress and its hematocrit dependence. *J. Rheol.* 42, 1–12. doi: 10.1122/1.550883
- Popel, A. S., and Johnson, P. C. (2005). Microcirculation and hemorrheology. *Annu. Rev. Fluid Mech.* 37, 43–69. doi: 10.1146/annurev.fluid.37.042604.133933
- Pries, A., Neuhaus, D., and Gaetgens, P. (1992). Blood viscosity in tube flow: dependence on diameter and hematocrit. *Amer. J. Physiol. Heart Circul. Physiol.* 263, H1770–H1778.
- Reasor, D. A., Clausen, J. R., and Aidun, C. K. (2012). Coupling the lattice-boltzmann and spectrin-link methods for the direct numerical simulation of cellular blood flow. *Int. J. Numer. Methods Fluids* 68, 767–781. doi: 10.1002/flid.2534
- Shi, X., Zhang, S., and Wang, S.-L. (2014). Numerical simulation of hemodynamic interactions of red blood cells in micro-capillary flow. *J. Hydrodyn.* 26, 178–186. doi: 10.1016/S1001-6058(14)60020-2
- Skalak, R., Tozeren, A., Zarda, R., and Chien, S. (1973). Strain energy function of red blood cell membranes. *Biophys. J.* 13, 245–264. doi: 10.1016/S0006-3495(73)85983-1
- Swihart, A., Mikrut, J., Ketterson, J., and Macdonald, R. (2001). Atomic force microscopy of the erythrocyte membrane skeleton. *J. Microscopy* 204, 212–225. doi: 10.1046/j.1365-2818.2001.00960.x
- Tateishi, N., Suzuki, Y., Soutani, M., and Maeda, N. (1994). Flow dynamics of erythrocytes in microvessels of isolated rabbit mesentery: cell-free layer and flow resistance. *J. Biomech.* 27, 1119–1125. doi: 10.1016/0021-9290(94)90052-3
- Tilles, A. W., and Eckstein, E. C. (1987). The near-wall excess of platelet-sized particles in blood flow: its dependence on hematocrit and wall shear rate. *Microvasc. Res.* 33, 211–223. doi: 10.1016/0026-2862(87)90018-5
- Tsukada, K., Sekizuka, E., Oshio, C., and Minamitani, H. (2001). Direct measurement of erythrocyte deformability in diabetes mellitus with a transparent microchannel capillary model and high-speed video camera system. *Microvasc. Res.* 61, 231–239. doi: 10.1006/mvre.2001.2307
- Waugh, R. E. (1982). Surface viscosity measurements from large bilayer vesicle tether formation. II. Experiments. *Biophys. J.* 38, 29–37. doi: 10.1016/S0006-3495(82)84527-X
- Yao, W., Wen, Z., Yan, Z., Sun, D., Ka, W., Xie, L., et al. (2001). Low viscosity Ektacytometry and its validation tested by flow chamber. *J. Biomech.* 34, 1501–1509. doi: 10.1016/S0021-9290(01)00109-9
- Yazdani, A., and Karniadakis, G. E. (2016). Sub-cellular modeling of platelet transport in blood flow through microchannels with constriction. *Soft Matter* 12, 4339–4351. doi: 10.1039/C6SM0154H

- Ye, T., Phan-Thien, N., and Lim, C. T. (2016). Particle-based simulations of red blood cells" a review. *J. Biomech.* 49, 2255–2266. doi: 10.1016/j.jbiomech.2015.11.050
- Yoon, Y.-Z., Hong, H., Brown, A., Kim, D. C., Kang, D. J., Lew, V. L., et al. (2009). Flickering analysis of erythrocyte mechanical properties: dependence on oxygenation level, cell shape, and hydration level. *Biophys. J.* 97, 1606–1615. doi: 10.1016/j.bpj.2009.06.028
- Závodszy, G., and Paál, G. (2013). Validation of a lattice Boltzmann method implementation for a 3D transient fluid flow in an intracranial aneurysm geometry. *Int. J. Heat Fluid Flow* 44, 276–283. doi: 10.1016/j.ijheatfluidflow.2013.06.008

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Závodszy, van Rooij, Azizi and Hoekstra. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

HPC for the rest of us: porting real-world research applications to an HPC environment

Alessandro Melis^{a,*}, Enrico Dall'Ara^b, Alberto Marzo^a, Mariano Vazquez^c,
Gavin J. Pringle^d, Marco Viceconti^a, CompBioMed Consortium

^a*Department of Mechanical Engineering and Insigneo Institute for in silico Medicine,
University of Sheffield, Sheffield S1 3JD, United Kingdom*

^b*Department of Oncology and Metabolism and Insigneo Institute for in silico Medicine,
University of Sheffield, Sheffield S1 3JD, United Kingdom*

^c*Barcelona Supercomputing Centre*

^d*EPCC, University of Edinburgh*

Abstract

Computational workflows aimed at clinical decision support consist of complex orchestration of multiple codes, often under commercial license, whose results should be obtained in the shortest time possible. These requirements differ from those of traditional HPC applications, therefore, the use of HPC systems in computational medicine is not as widespread as in other scientific research domains. In this paper, the porting of one musculoskeletal exemplar application from a Tier-3 to a Tier-1 system is described. The workflow is tested for a representative range of use cases while monitoring execution time and memory consumption. The results indicate that there is not an immediate benefit in porting the application as it is on the Tier-1 system. Finally, recommendations on optimisation and improvement of computational medicine workflows are proposed.

Keywords: computational medicine, digital volume correlation, high performance computing

*Corresponding author.

Address: Department of Mechanical Engineering and INSIGNEO Institute for *in silico* Medicine, Sir Frederick Mappin Building and Pam Liversidge Building, University of Sheffield, Mappin Street, Sheffield S1 3JD, United Kingdom.

Telephone: +44 (0) 0114-222-6173

Email address: a.melis@sheffield.ac.uk (Alessandro Melis)

1. Introduction

In spite of good intentions, the use of High Performance Computing (HPC) remains limited to relatively few scientific research domains. Historically, HPC systems were used for specific applications where the computational cost could not be reduced on normal systems, and that carried enough scientific, social, and/or economic impact to justify the high cost of porting the codes, optimise them for HPC, and run on what were very expensive computer systems. Over the years these “killer applications” were for example weather forecasting, fluid dynamics, nuclear test simulations, and molecular dynamics. Still today, if we look at the projects awarded with computational time over the systems of the PRACE European Tier-0 initiative¹, we see that Engineering is dominated by fluid dynamics problems, while the “Biochemistry, Bioinformatics and Life sciences” area mostly involves molecular dynamics problems.

However, there is a growing expectation that HPC should play a more pervasive role in science, as “All scientific disciplines are becoming “computational” today”, “Industry and SMEs are increasingly relying on the power of supercomputers to invent innovative solutions”, and “HPC is part of a global race”, as the H2020 HPC home page states². The European HPC strategy in the Horizon 2020 research framework aims to three macroscopic objectives, one of which is indeed “achieving excellence in HPC application delivery and use”, with the establishment of Centres of Excellence (CoEs) to “focus and coordinate support to the application of HPC in scientific or industrial domains that are most important for Europe”. One of these CoE is the CompBioMed, dedicated to *Computational Medicine*³.

There are considerable differences between traditional HPC applications and those developed in the framework of computational medicine, for example for

¹<http://www.prace-ri.eu/projectaccessawarded/>

²<https://ec.europa.eu/programmes/horizon2020/en/h2020-section/>

[high-performance-computing-hpc](http://www.compbiomed.eu)

³<http://www.compbiomed.eu>

27 Subject-Specific Modelling (SSM) for clinical decision support:

- 28 a) Traditional HPC applications tend to be very complex numerical mathemat-
29 ics, but quite simple in the structure, whereas most SSM applications involve
30 a complex workflow where multiple codes, each performing a specialised task,
31 are orchestrated;
- 32 b) Traditional HPC applications are entirely developed by the researchers, and
33 are normally available in source code format, with limited dependencies and
34 with no usage restriction; SSM workflow frequently use commercial software
35 components to shorten the development cycle, and to focus on the added
36 value;
- 37 c) Traditional HPC applications are motivated almost exclusively by the need
38 to solve a model which order cannot be reduced, and which execution time is
39 unacceptably long on conventional, less performant computer architectures,
40 or by the need to simulate larger models which require the vast memory HPC
41 offers; SSM models seek HPC support for a broad range of reasons including
42 for example the need to inform surrogate models, or the problem of coping
43 with the combinatory explosion in multiscale models and strongly-coupled
44 models;
- 45 d) Traditional HPC is dominated by fundamental research, whereas SSM mod-
46 els have a life cycle that will bring them ideally in the hands of clinical and
47 industrial users;
- 48 e) Traditional HPC applications run in batch, and require HPC in the stricter
49 sense (i.e. they are CPU-bounded); SSM applications are orchestrations of
50 multiple codes, some CPU-, some memory-, and some I/O bounded. In some
51 cases, there is a requirement to run as soon as possible (Urgent Computing),
52 and/or the need for some interactivity.

53 The aim of the present study is to explore the process of porting of one
54 exemplary SSM application (biomechanical characterisation of bone biopsies at
55 the tissue scale) from a conventional Tier-3 cluster to a national Tier-1 system.
56 The goal is to document the specific barriers that prevented a straightforward

57 porting, and the optimisation work required to better exploit the additional
58 computational power Tier-1 systems offer, with the long-term of objective of
59 learning lessons that can change the business practices of HPC centres in favour
60 of a broader portfolio of applications.

61 **2. Materials and Methods**

62 *2.1. Porting process*

63 We considered our starting point a complete workflow running on our lo-
64 cal Tier-3 system (hereinafter referred as version 1). A first revision cycle was
65 linked to the portability. All the hardware/software dependencies of the work-
66 flow were identified, and their compatibility with the Tier-1 system verified.
67 The first software revision included all changes that were mandatory to have a
68 version of the workflow running on the Tier-1 system (version 2). At this stage,
69 both version 1 (running on the Tier-3 system) and version 2 (running on the
70 Tier-1 system), were profiled using three input sets of growing size (and compu-
71 tational cost) representing the normal, an intermediate size, and the largest size
72 of the problems to be analysed. The main bottlenecks in the version-2 workflow
73 were identified, and revised to perform better and/or scale better on the Tier-1
74 system. From these results, further recommendations for a future release were
75 drawn.

76 *2.2. Description of the BoneDVC workflow*

77 The BoneDVC workflow analyses two stacks of CT images of the same bone
78 tissue specimen [1, 2]. The former is fixed (reference configuration) and the
79 latter is imaged under compression (displaced configuration). The linear dis-
80 placement field mapping the second volume to the first is calculated by means
81 of the Digital Volume Correlation (DVC) method. The strain field in the com-
82 pressed specimen is obtained by differentiation of the computed displacement
83 field.

84 The BoneDVCv1 workflow was staged in the following steps (Figure 1):

- 85 1. The μ CT images were uploaded to the HPC facility from an external data
86 repository;
- 87 2. The displacement field was calculated given the images and a set of con-
88 figuration parameters for the DVC software;
- 89 3. The displacement field from the registration step was used as input for
90 the calculation of the strain field;
- 91 4. The workflow output, in form of the displacement and strain fields, was
92 downloaded back to the external repository.

93 *2.3. Portability analysis*

94 The BoneDVCv1 was developed as a composition of commercial, licensed
95 and open-source software. The images registration was performed by means of
96 the Sheffield Image Registration Toolkit (ShIRT), an in-house software devel-
97 oped in C and MATLAB (MATLAB R2016a, TheMathWorks, Inc.) at The
98 University of Sheffield, UK [3, 4]. The strain field calculation was performed via
99 ANSYS post-processing tools (ANSYS Mechanical APDL v14.0, Ansys, Inc.,
100 USA). The registration and differentiation steps were orchestrated by means of
101 the workflow management system Taverna Standalone 2.5 [5], whose only de-
102 pendency is Java 7. The workflow and the functions for data upload/download
103 were glued together in a Python 2.7 wrapper which equipped the final user with
104 a command line interface. It should be noted that all software was serial at this
105 point.

106 *2.4. Profiling*

107 A small pair of μ CT images (30 MB each) was used as test case and the
108 profiling was performed by collecting during runtime the timestamp at the be-
109 ginning and ending of each main block in the BoneDVC workflow (Figure 2).
110 As expected, the majority of the computational time was employed by the im-
111 age registration process (ShIRT). The time required to compute the strain field
112 (ANSYS-post) was less than 1% of the total running time. Interestingly, the
113 compilation of the MATLAB scripts calling ShIRT, improved the system I/O

114 overhead and the total running time decreased by 9.3%. In both versions, the
115 Taverna overhead was negligible.

116 **3. Results**

117 *3.1. Changes required to port to Tier-1 system*

118 To avoid commercial licenses on the tier-1 platform, the strain field calcula-
119 tion was replaced by a MATLAB script. The script was built in a standalone
120 executable via the MATLAB compiler (R2016a (9.0.1)). This executable can be
121 run through the MATLAB Compiler Runtime (mcr 9.0.1) which is free-to-use
122 on the Tier-1 platform and does not require a license to run.

123 *3.2. Tier-1 vs Tier-3 performance comparison*

124 The BoneDVCv2 was tested on both Tier-1 and Tier-3 infrastructures in
125 serial mode, i.e., only one node was employed for the workflow execution, as
126 ShIRT is not currently parallelised and its execution would not benefit from
127 the allocation of multiple nodes. The Tier-1 serial node consisted of a 2.0 GHz
128 10-core Intel Xeon E7-4850 (Westmere) processor with 1 TB physical memory
129 available; the Tier-3 node had a 2.4 GHz 8-core Intel Xeon E5-2630 (Haswell)
130 series processor and 64 GB RAM. The performance was assessed in terms of wall-
131 clock time and maximum virtual memory allocation (Figure 3) on three datasets
132 of increasing size (10, 100, and 1000 MB in size for each μ CT 3D image). The
133 workflow was run three times for each dataset and the results were reported in
134 terms of mean and standard deviation values.

135 The computational time (Figure 3a) increases with the input image size for
136 both machines, and the workflow execution is in all the cases faster on the Tier-
137 3. The Tier-1 execution time is 3.2, 2.3, and 2.2 times the Tier-3 one for the
138 three datasets, respectively. The memory consumption is reported with respect
139 to the single image size for each dataset (Figure 3b). In the case of the smallest
140 dataset, the Tier-1 allocates three times the maximum virtual memory used
141 by the Tier-3. The performance is similar for the 1GB image dataset as both

142 machines allocate about 80 GB of virtual memory. In the case of the largest
143 image dataset, the Tier-3 peak requirement is of 90 GB and it is almost twice
144 the Tier-1 memory allocation.

145 *3.3. Changes to improve Tier-1 performance*

146 The profiling of BoneDVCv2 shows a clear bottleneck with the current ShIRT
147 library. This is a custom-code written in C in early 2000 with methods to
148 circumvent the severe memory limitations that personal computers had back
149 then. A preliminary analysis of the source code highlighted several problems:

- 150 – The code was developed with little concern for software sustainability, as
151 it was initially intended to be used only by its own developer. There is
152 virtually no documentation, and the source code has very few comments;
- 153 – To save memory, the code implements complex bitmask operations that
154 make the code very complex (and unnecessarily so nowadays);
- 155 – The code is not designed to work on large 3D images (which at that time
156 did not exist) or to use parallel computing to speed-up calculations;
- 157 – Some intellectual property issues exist on the source that makes impossible
158 to license it as open source in the present form.

159 Following this analysis, it was concluded that the only viable option to im-
160 prove BoneDVC on the Tier-1 system is to re-write the ShIRT library into a
161 modern, well documented, and parallel code. It is estimated that the operation
162 will require at least one person-year, including clean room re-write, debugging,
163 testing and optimisation.

164 **4. Discussion**

165 The aim of the present study was to explore the process of porting of an
166 exemplary SSM application (biomechanical characterisation of bone biopsies at
167 the tissue scale) from a conventional Tier-3 cluster to a national Tier-1 system.

168 In order to have the BoneDVC workflow running on the Tier-1 system, it
169 was necessary to eliminate all dependencies from commercial software. This
170 was achieved by replacing the strain field calculation done in ANSYS-post with
171 a MATLAB numerical code. Following, MATLAB code was compiled and run
172 through the free-to-use runtime environment.

173 On the Tier-3, BoneDVCv2 runs faster than on the Tier-1 machine. This
174 is because the Tier-3 is equipped with more powerful and recent CPUs that, in
175 the case of a serial job, outperform the Tier-1 hardware. However, the Tier-
176 1 serial node is optimised for post-processing purposes in which the ability of
177 handling big amount of data is favoured to the detriment of execution time.
178 Therefore, the Tier-1 memory consumption is half the Tier-3 one for the largest
179 image dataset. But these comparisons are scarcely significant; Tier-1 systems
180 are intended for the execution of codes that are optimised for parallel execution;
181 profiling them for mono-thread codes makes little sense, if not to highlight the
182 need to intervene on the code, in order to achieve scalability.

183 The profiling suggested that the main bottleneck resides in ShIRT, where
184 the image registration equation is iteratively solved for each voxel. This is
185 a time-consuming process because the voxels are analysed one-at-a-time in a
186 serial loop, and a serial execution does not properly exploit neither Tier-1 or
187 Tier-3 hardware configurations. The registration algorithm treats each voxel as
188 independent by its neighbours, therefore a first improvement can be achieved
189 by simply computing the potentially 10,000s of different voxels in parallel.

190 An analysis of ShIRT source code confirmed that parallelisation of the cur-
191 rent version would be impossible. Thus, we decided, also to remove the license
192 limitations, to perform a clean-room rewrite of ShIRT in different programming
193 language, and with the structure that enable parallelisation and code optimisa-
194 tion with very large input sets. The new code, tentatively named OpenShIRT,
195 will be made available as open source code, and then ported and optimised for
196 execution on ARCHER.

197 In conclusion, the lesson learned from this porting exercise can be sum-
198 marised as follow:

- 199 I) In complex computational medicine workflows, it is usually a single compo-
200 nent of the workflow that would benefit from executing on an HPC system;
201 this calls for a deployment model where conventional server nodes can be
202 configured in a safe execution environment, to run all non-computationally-
203 intensive parts of the workflow, while the computationally-intensive code
204 execution is spawned to the HPC system. This approach would also make
205 easier to retain in such complex workflows the use of commercial software,
206 that can save considerable time in the their development;
- 207 II) The performance on a Tier-1 platform is dependent on the size of the sim-
208 ulation, which is best for very large systems requiring 100s if not 1000s of
209 parallel tasks. Smaller models will run poorly on Tier-1 platforms and, as
210 such, are more suited to less performant platforms, such as Tier-3 systems.
211 This was found to be the case for the system investigated here. However,
212 we plan to implement a parallel version of ShIRT which, in conjunction
213 with a much larger dataset, can exploit Tier-1 platforms efficiently;
- 214 III) Most research software for computational medicine is written without con-
215 sidering future needs for scalability. We recommend that sustainability
216 guidelines such as those proposed by the Software Sustainability Institute⁴;
217 but more specifically, it is important to acknowledge that after a gener-
218 ation of pioneering application, good part of the computational medicine
219 software we write today will continue to be used for years to come, and the
220 growing data footprint in our like in many other domains will inevitably
221 pose scalability issues sooner or later in the life of the software artefact.
222 Plan for HPC should become the motto of computational medicine soft-
223 ware developers.

⁴<https://www.software.ac.uk>

224 **Acknowledgements**

225 The authors would like to thank William Griffith, who deployed the BoneDVC
226 workflow on the Tier-3 system, and assisted in the porting to the Tier-1 system.

227 **Conflict of Interest**

228 The authors acknowledge the partial support from the CompBioMed Cen-
229 tre of Excellence (H2020-EINFRA-2015-1-675451), the EPSRC funded project
230 MultiSim (Grant N. EP/K03877X/1), the UK NC3R funded project (Grant N.
231 113629); and the FP7 European projects CHIC (grant N. 600841). The authors
232 declare that they do not have any financial or personal relationships with other
233 people or organisations that could have inappropriately influenced this study.

234 **References**

- 235 [1] E. Dall’Ara, D. C. Barber, M. Viceconti, About the inevitable compromise
236 between spatial resolution and accuracy of strain measurement for bone
237 tissue: A 3d zero-strain study, *Journal of biomechanics* 47 (12) (2014) 2956–
238 2963.
- 239 [2] M. Palanca, G. Tozzi, L. Cristofolini, M. Viceconti, E. Dall’Ara, Three-
240 dimensional local measurements of bone strain and displacement: compari-
241 son of three digital volume correlation approaches, *Journal of biomechanical*
242 *engineering* 137 (7) (2015) 071006.
- 243 [3] D. C. Barber, D. R. Hose, Automatic segmentation of medical images us-
244 ing image registration: diagnostic and simulation applications, *Journal of*
245 *medical engineering & technology* 29 (2) (2005) 53–63.
- 246 [4] D. C. Barber, E. Oubel, A. F. Frangi, D. R. Hose, Efficient computational
247 fluid dynamics mesh generation by image registration, *Medical image anal-*
248 *ysis* 11 (6) (2007) 648–662.

249 [5] K. Wolstencroft, R. Haines, D. Fellows, A. Williams, D. Withers, S. Owen,
250 S. Soiland-Reyes, I. Dunlop, A. Nenadic, P. Fisher, et al., The taverna work-
251 flow suite: designing and executing workflows of web services on the desktop,
252 web or in the cloud, *Nucleic acids research* 41 (W1) (2013) W557–W561.

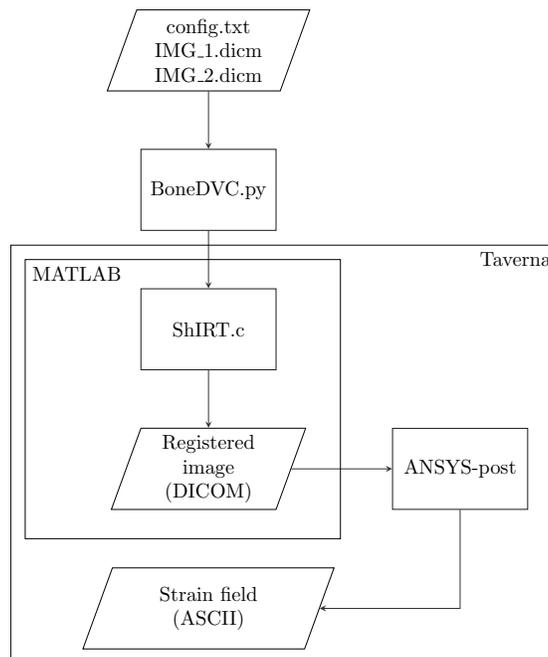


Figure 1: BoneDVC workflow. A Python script collects the pipeline input (configuration file and DICOM images) and initialises the Taverna workflow which orchestrates the registration and differentiation steps. The final output consists in the computed strain field and the registered image.

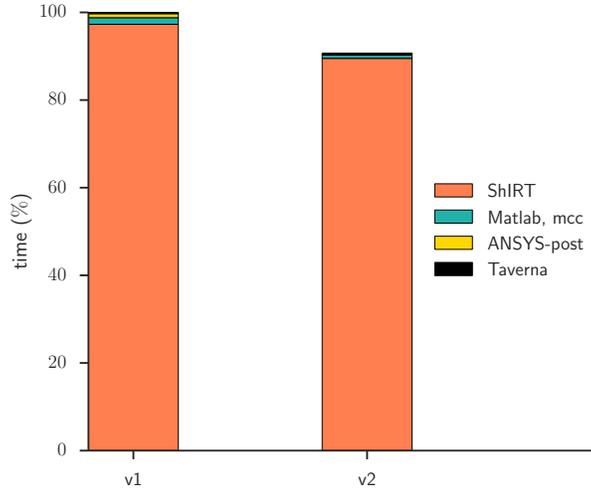


Figure 2: Time profiling for the two versions of BoneDVC developed. Elapsed time is normalised with respect to the v1 total running time. Each bar considers both the run time and the overhead due to system I/O between different codes within the workflow.

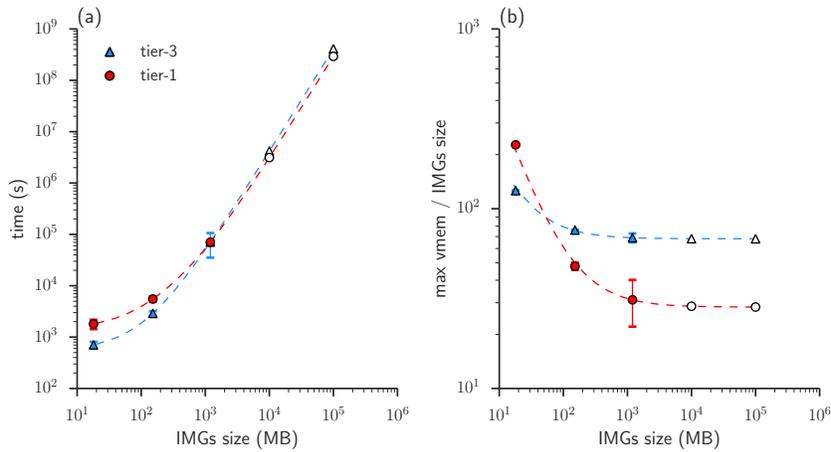


Figure 3: (a) Wall-clock time required for the execution of BoneDVCv2 workflow for the three datasets of increasing size. (b) Normalised virtual memory requirement for each registration task. The memory requirement is scaled and reported with respect to the dataset size. Execution time and memory consumption for images larger than 1000 MB are linearly extrapolated from benchmark results (blank markers).

An Ensemble-Based Protocol for the Computational Prediction of Helix–Helix Interactions in G Protein-Coupled Receptors using Coarse-Grained Molecular Dynamics

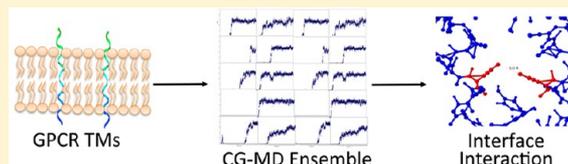
Nojood A. Altwaijry,^{†,§} Michael Baron,[†] David W. Wright,^{‡,¶} Peter V. Coveney,^{‡,¶} and Andrea Townsend-Nicholson^{*,†,¶}

[†]Institute of Structural and Molecular Biology, Research Department of Structural and Molecular Biology, Division of Biosciences, University College London, London, WC1E 6BT, United Kingdom

[‡]Centre for Computational Science, Department of Chemistry, University College London, London WC1H 0AJ, United Kingdom

[§]King Saud University, Riyadh, Kingdom of Saudi Arabia

ABSTRACT: The accurate identification of the specific points of interaction between G protein-coupled receptor (GPCR) oligomers is essential for the design of receptor ligands targeting oligomeric receptor targets. A coarse-grained molecular dynamics computer simulation approach would provide a compelling means of identifying these specific protein–protein interactions and could be applied both for known oligomers of interest and as a high-throughput screen to identify novel oligomeric targets. However, to be effective, this *in silico* modeling must provide accurate, precise, and reproducible information. This has been achieved recently in numerous biological systems using an ensemble-based all-atom molecular dynamics approach. In this study, we describe an equivalent methodology for ensemble-based coarse-grained simulations. We report the performance of this method when applied to four different GPCRs known to oligomerize using error analysis to determine the ensemble size and individual replica simulation time required. Our measurements of distance between residues shown to be involved in oligomerization of the fifth transmembrane domain from the adenosine A_{2A} receptor are in very good agreement with the existing biophysical data and provide information about the nature of the contact interface that cannot be determined experimentally. Calculations of distance between rhodopsin, CXCR4, and β_1 AR transmembrane domains reported to form contact points in homodimers correlate well with the corresponding measurements obtained from experimental structural data, providing an ability to predict contact interfaces computationally. Interestingly, error analysis enables identification of noninteracting regions. Our results confirm that GPCR interactions can be reliably predicted using this novel methodology.



1. INTRODUCTION

We need to understand how proteins behave in order to manipulate them successfully. The means by which to achieve accurate, precise, and reproducible predictions of the key properties of therapeutically relevant proteins is a fundamental question in computational biology. Molecular dynamics (MD) simulations have been used to study complex biomolecular systems, but it is not possible to define how a system behaves from a single trajectory; single trajectory systems behave as Gaussian random processes, making the attainment of accurate predictions from a single run not a realistic proposition. Accurate predictions that correlate well with experimental data have been achieved with the use of multiple short MD simulations to enhance the sampling of conformational space and hence the convergence of observable properties.^{1–7} These ensemble-based fully atomistic MD studies have primarily focused on ligand-protein binding free energies, where there exists a wealth of experimental data with which to compare computational findings. In this paper, we take our first steps to assess the reliability and reproducibility of analogous CG-MD simulations. For this work, we have elected to examine the molecular nature of protein–protein interactions between G

protein-coupled receptors (GPCRs). This is a biological system with which we are familiar experimentally.^{8–14}

GPCRs are a particularly well-studied family of membrane proteins. Not only are they a large and important group of signaling proteins, they are also the targets for approximately 40% of all therapeutic compounds in clinical use. Although over 800 human proteins are classified as GPCRs, drugs have been developed against fewer than 10% of these targets.¹⁵ Thus, there is huge potential to expand the number of targets for which new therapies can be designed. Novel therapeutic design is also important if one of the goals of personalized medicine, to develop new drugs for patient-specific variations of GPCRs, is to be achieved. Inclusion of functional GPCR homomers and heteromers in drug discovery programs also provides a means of expanding the range of novel targets for the development of therapeutic agents.¹⁶ Originally believed to function as monomeric proteins, many functional GPCR oligomers have now been identified. Early examples include the obligate heteromeric assembly of GABA_BR1 and GABA_BR2 required to form a

Received: December 23, 2016

Published: April 6, 2017

Table 1. Computational Methods Used for Modeling Mammalian GPCR Dimers

type	GPCR dimers	method	force field	interface	number of replicas	time scale	ref	
homodimers	Rho/Rho	docking	CVFF	TM4,5/TM4,5		100 ps	Filipek et al., 2004 ⁶³	
				TM1,2/TM1,2		ND ^a	Han et al., 2009 ⁶⁴	
		MD	Gromos-87 OPLSAA	TM4,5/TM4,5	1	45 ns	Kaczor et al., 2013 ⁶⁵	
					1	0.1 μ s	Cordomí and Perez, 2009 ⁶⁷	
				Amber/parm99	2	300 ns	Neri et al., 2010 ⁶⁸	
				CG-MD	Martini	1 ^b	8 μ s	Periole et al., 2007 ³⁶
		CG-MD	Martini	TM1,2, H8	10	100 μ s	Periole et al., 2012 ⁶⁹	
				TM4,5				
				TM6,7				
				H8/H8	2	5 μ s	Ghosh et al., 2014 ⁴¹	
	β_2 -/ β_2 -adrenergic	CG-MD	Martini	TM1/TM1	4	\sim 200 μ s	Prasanna et al., 2014 ⁴⁰	
				TM4,5/TM4,5	1 ^c	\sim 18 μ s	Mondal et al., 2013 ³⁹	
				TM5/TM5				
		CG-MD	Martini	TM6/TM6				
				TM3/TM3				
	β_1 -/ β_1 -adrenergic	CG-MD	Martini	TM1/TM1	3 ^d (2 runs; different starting point)	2 μ s	Mondal et al., 2013 ³⁹	
	α_{1B} -/ α_{1B} -adrenergic	docking		TM5/TM5			ND	Fanelli et al., 1999 ⁷⁰
				TM6/TM6				
	5-HT ₄ /5-HT ₄	docking		TM7/TM7				
				TM2,4/TM2,4			ND	Soulier et al., 2005 ⁷¹
5-HT _{1A} /5-HT _{1A}	docking	CHARMM	TM4,5/TM4,5			15 ns	Russo et al., 2007 ⁷²	
			TM4,5,IL2/TM4,5			ND	Berthouze et al., 2007 ⁷³	
CXCR4/CXCR4	docking	MD	OPLSAA	TM3/TM4,5	1	50 ns	Gorinski et al., 2012 ⁷⁴	
				TM5/TM5			ND	Kaczor et al., 2013 ⁶⁵
NTSR ₁ /NTSR ₁	docking	CHARMM	TM1/TM4			ND	Rodriguez et al., 2012 ⁷⁵	
			TM4/TM4					
δ -OR/ δ -OR	CG-MD	Martini	TM2,3,4/TM2,3,4	1 ^e	250 ns	Casciari et al., 2008 ⁷⁶		
			TM4/TM4 favored over TM4,5/TM4,5	2	1.5 μ s	Provasi et al., 2010 ³⁷		
κ -OR/ κ -OR	docking	CHARMM	TM1/TM2			ND	Johnston et al., 2011 ⁷⁷	
			A _{2A} R/A _{2A} R	TM1,2,3/TM1,2,3	TM1/TM1		ND	Kaczor et al., 2013 ⁶⁵
	docking			TM1,4/TM1,4			Fanelli and Feline, 2011 ⁷⁸	
				TM2,3/TM2,3				
A ₃ R/A ₃ R	MD	Amber7 FF99	TM6,7/TM6,7					
			H8,I3/TM6					
TXA ₂ /TXA ₂	docking		TM4,5/TM4,5	1	500 ps	Kim and Jacobson, 2006 ⁷⁹		
			TM1/TM1			ND	Fanelli et al., 2011 ⁸⁰	
D ₂ R/D ₂ R	Monte Carlo		TM1/TM2,EL2					
			H8/H8					
SSTR ₁ /SSTR ₁	Monte Carlo		ND			ND	Woolf and Linderman, 2004 ⁸¹	
			ND			ND	Woolf and Linderman, 2004 ⁸¹	
LHR-LHR	docking	MD	CHARMM	TM4/TM6,7			Fanelli 2007 ⁸⁰	
				TM4/TM4	1	1 ns	Fanelli 2007 ⁸⁰	
				TM4/TM6				
				TM5/TM6				
A _{2A} R/D ₂ R	docking			TM4/TM1,3				
				TTM3,4/TM5,6			ND	Canals et al., 2003 ²⁵
mGluR ₂ /5-HT _{2A}	docking	MD	CHARMM 22/27	TM3,4,5/TM4,5				
				TM4,5/TM3,4,5			ND	Bruno et al., 2009 ³³
μ -OR/ δ -OR	docking			TM4,5/TM4,5	1	40 ns	Bruno et al., 2009 ³³	
				TM4,5/TM4,5				
	MD	GROMOS87		TM6,7/TM4,5			ND	Liu et al., 2009 ⁸²
				TM1,7/TM4,7				
				TM1,7/TM4,5	1	5 ns	Liu et al., 2009 ⁸²	
				TM4,7/TM4,5				

Table 1. continued

type	GPCR dimers	method	force field	interface	number of replicas	time scale	ref
homotetramer	(V ₂ R) ₄	MD	CHARMM 22/27	TM3,4/TM4,7 TM4,5/TM4,5	1	5 ns	Witt et al., 2007 ³²

^aND: not defined, IL: intracellular loop, EL: extracellular loop. ^bFour different structures. ^cNine different structures. ^dTwo runs; different starting point. ^eUmbrella sampling of 43 different starting points.

functional GABA_B receptor¹⁷ and heterodimerization of the delta and kappa opioid receptor subtypes to form an opioid receptor with the κ_2 receptor subtype pharmacology.¹⁸ The archetypal class A GPCR rhodopsin forms structural dimers organized in paracrystalline arrays in membranes¹⁹ and in the model crystal structure of this GPCR (1N3M).²⁰ For the design of cost-effective “designer” drugs for individuals that target receptor oligomers, it will be necessary to develop a powerful and sophisticated computational method for understanding the interactions involved in the formation of GPCR oligomers.

Biological methods for studying GPCR oligomers in native cells and tissues or in recombinant mammalian expression systems include coimmunoprecipitation, Western blot analyses, cross-linking studies, yeast two hybrid experiments, bimolecular fluorescence complementation via GFP reconstitution (BiFC), energy transfer-based methods (FRET and BRET), functional cross-talk, and activation by dimeric/bivalent ligands.²¹ Unfortunately, these methods frequently allow for alternative interpretations of the results and therefore do not provide unequivocal answers regarding multimerization occurrence between candidate pairs of GPCRs nor do they yield specific details of the interface(s) between interacting GPCRs. Structural methods such as X-ray crystallography and atomic force microscopy could provide some of this information, but only three Class A GPCR dimer structures have been solved^{22–24} and tend to describe “contact areas” rather than specific molecular interfaces. For the development of an accurate computational model for analyzing GPCR interfaces, it is essential to have good experimental data with which to validate the model. Such data are made available for the A_{2A} adenosine receptor subtype, which has been shown to participate in the formation of both heteromeric²⁵ and homomeric GPCRs.²⁶ The identification of homomeric A_{2A} receptors provided an opportunity to identify the transmembrane domain (TMS) involved in self-association by far-UV CD spectroscopy and SDS-PAGE using synthetic peptides corresponding to the different transmembrane domains.²⁷ A subsequent study²⁸ mapped TM helix interactions in the A_{2A} receptor for 31 different peptide pairs. We have previously worked with the A_{2A} receptor gene²⁹ and are interested in identifying patient-specific variations within this and related nucleoside and nucleotide receptor subtypes.

There have been many computational studies of GPCR interactions (see Table 1). The methodologies for modeling these have, in general, adopted one of two approaches: (i) molecular dynamics simulations using models based on homology with the nearest related GPCR for which structural data exist or (ii) docking.^{30,31} Initial GPCR MD studies were performed using CHARMM and AMBER, which were subsequently integrated into NAMD^{32,33} and GROMACS.^{34,35} Although there is no established standard protocol for MD simulations of GPCRs, a number have used GROMACS with the Martini force field,^{36–41} which is designed specifically for lipids and membranes and allows the lipid composition most suited to the receptor in question to be incorporated into the simulation. The more recent of the GPCR dimer modeling studies have been conducted using coarse-grained simulations, which take

Table 2. Sequences of the A_{2A}R Helices Used in Ensemble Simulation Sets

A _{2A} R helices	sequence ^a
TM2-wild-type	FVVSLLAAAD52 ⁵⁰ IAVGVLAI ⁵¹ PAIT
TM5-wild-type	MNYMVFYFNFFACVLVP189 ^{55,50} LLMLGVYLRI
TMS-M177 ^{5,38} A	MNY <u>AV</u> YFNFFACVLVP189 ^{55,50} LLMLGVYLRI
TMS-M193 ^{5,54} A	MNYMVFYFNFFACVLVP189 ^{55,50} LL <u>AL</u> GVYLRI
TMS-M193 ^{5,54} I	MNYMVFYFNFFACVLVP189 ^{55,50} LL <u>IL</u> GVYLRI

^aResidues suggested to play a key role in the dimerization of the A_{2A}R TMS helix²⁷ are indicated in bold; mutated residues are underlined and italicized. The conserved amino acid for each TM helix is italicized and is numbered using both the Ballesteros and Weinstein nomenclature (superscript) and by residue number.

less compute time and therefore provide an opportunity to perform a substantial number of replicas for each set of simulation conditions.

When we began our studies, approximately 30 computational GPCR dimer models had been published (Table 1). Of these, two are Monte Carlo-derived, 15 are based on docking, and nine have been generated using atomistic MD simulations. The rest are CG-MD models. Historically, docking was the earliest method to be employed and has been used regularly; its current use is widespread. Alternative methods of modeling began with Monte Carlo methods, moving to fully atomistic MD and a subsequent shift to CG-MD, which is the predominant MD method currently in use for GPCRs. CG-MD is popular as it is cheaper and faster and has been shown, when CG models are subsequently converted to atomistic representations, to produce similar results to models generated by atomistic MD.^{38,42,43} CG-MD simulations have also been used to study TM helix–helix dimers for non-GPCR types of cell surface receptors such as Glyphorin A and ErbB dimers.^{44,45}

GPCRs exhibit thermodynamic equilibrium states and therefore are “mixing” in the language of ergodic dynamical systems theory.⁵ Neighboring trajectories diverge exponentially, and only probabilistic descriptions are meaningful. For these intrinsic reasons, collections of trajectories differing only in their initial conditions, known as ensembles, are the best means of studying the properties of such systems. Each individual system in the ensemble is referred to as a replica. As an additional benefit, performing such ensemble-based molecular dynamics simulations provides close control of errors and uncertainties in predictions. In this paper, we present the development of a robust and rapid method of this kind for identifying helix–helix interactions in GPCRs.

2. METHODS

Here, we aim to develop a consistent, rapid, reproducible CG-MD methodology for the study of interacting helices. This method involves placing two GPCR transmembrane helices (a simulation set) in a membrane and running simulations with the hope of identifying interactions between the helices. In these simulations, we will be using distance as a means of

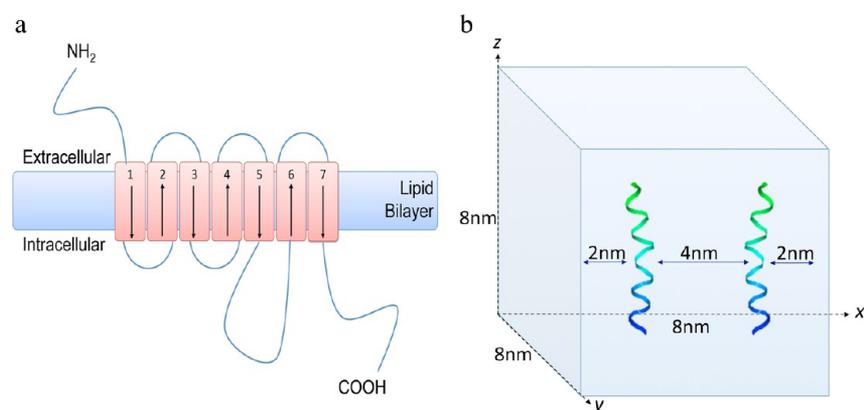


Figure 1. Experimental system showing (a) a schematic representation of the A_{2A} receptor structure indicating the directionality of the TM helices within the lipid bilayer and (b) placement of the TMs within the simulation box prior to the addition of lipid and water.

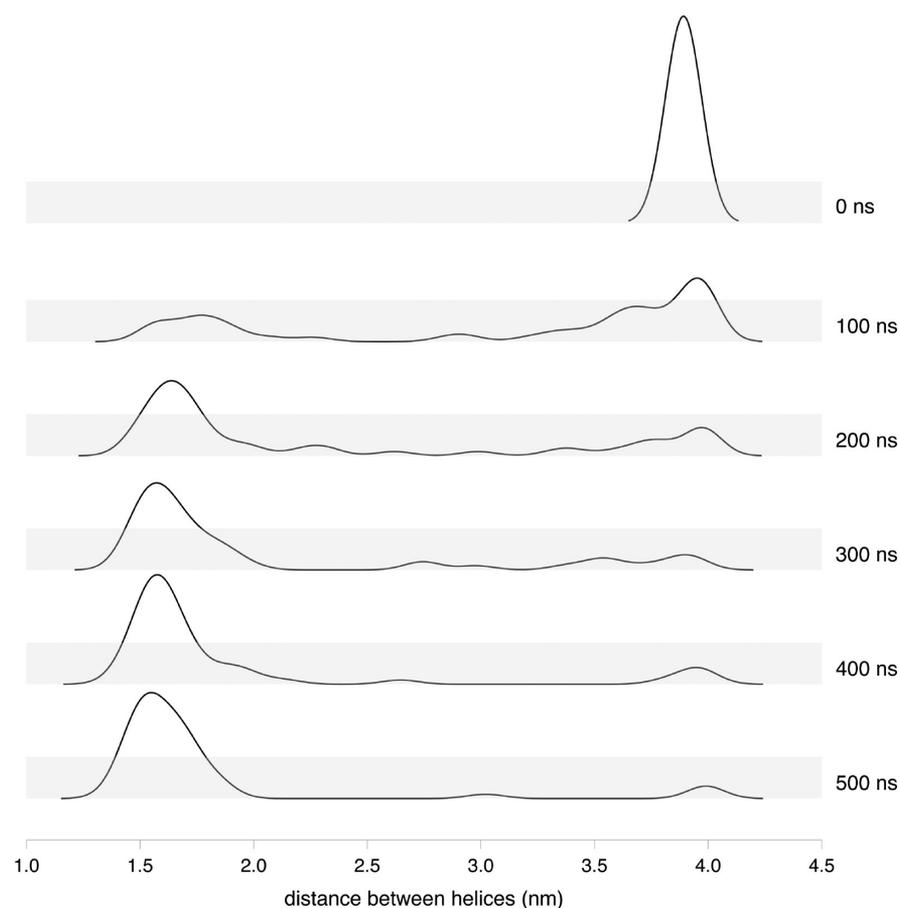


Figure 2. Distribution of the mean distance between the two TMS–TMS wild type helices at 0, 100, 200, 300, 400, and 500 ns in all 50 replicas.

identifying two different types of interactions: interactions between helices and interactions between amino acid residues on each helix. For the successful identification of both types of interactions, it is necessary to specify the number of replicas (identical independent simulations other than for the initial velocity seeds assigned to the particles) and the run time needed to achieve converged results and see how well they reproduce experimental results. The number of replicas must be sufficient to achieve a reproducible result as evidenced by a sufficiently small error estimate.

We will use the terms “stable dimer” and “dimerization” to refer to interactions between helices. A 10 Å truncation cutoff

(backbone to backbone) has been set for dimerization, as it has been shown experimentally that a unique FRET signal is generated when two labeled peptides are located within 10 Å of each other and form an excited state dimer.⁴⁶ The term “specific interactions” will be used to refer to interactions between amino acid side chains on the dimerized helices. Specific interactions will be identified from contact matrices (heat maps). Although a 12 Å truncation cutoff had previously been used to analyze these interactions,⁴⁷ we will set our interaction cutoff to 10 Å because the existence of hydrogen bond (C_α-H[⋯]O) contacts as a function of the interhelical axial distance is between 6 and 12 Å. Side chain to side chain

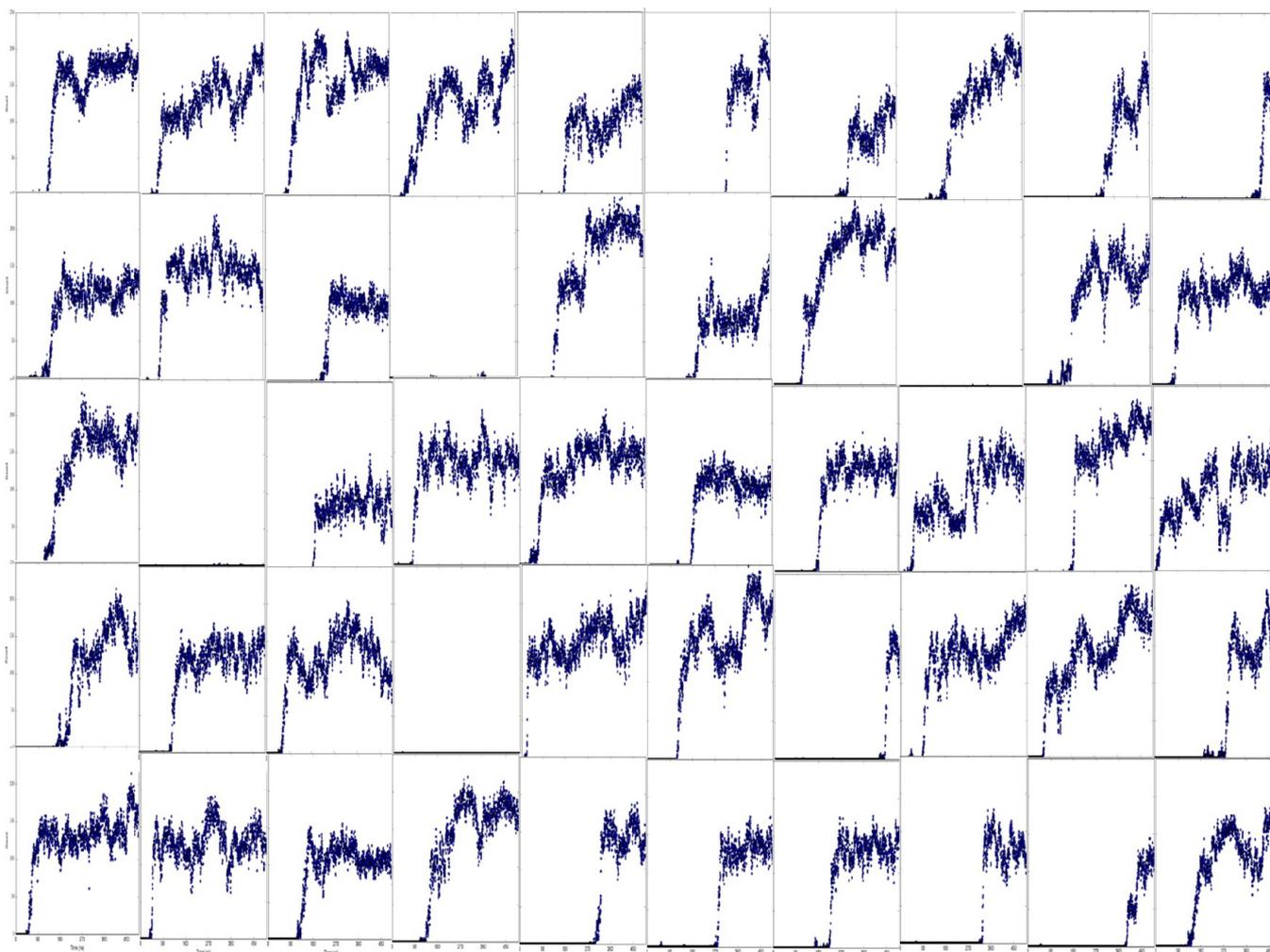


Figure 3. Number and timing of pairwise interactions for each of the 50 replicas within the wild-type TMS–TMS dimer ensemble are shown. The x and y axes are linear and represent run length from 0 to 500 ns and the number of interaction events from 0 to 250 counts, respectively.

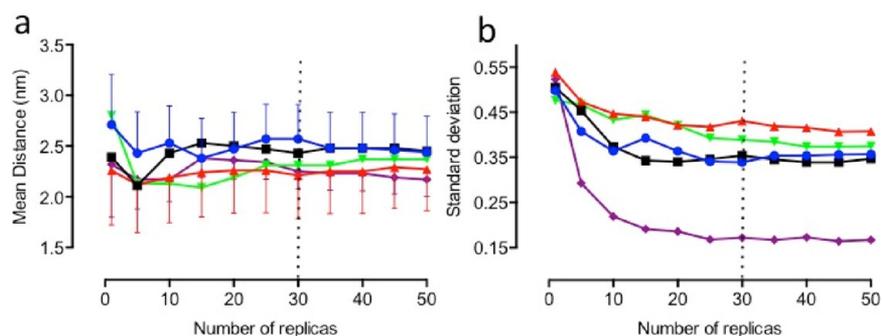


Figure 4. Variation in (a) the mean distance between TM helices and (b) the error (standard deviation) is shown as a function of the number of replicas performed for the following simulation sets: (blue circle) wild-type TMS–TMS helices, (black square) M177A-mutated TMS–TMS helices, (red triangle) M193A-mutated TMS–TMS helices, (green inverted triangle) M193I-mutated TMS–TMS helices, and (purple diamond) wild-type TM2–TM2 helices.

distances consistent with this are used to identify specific interactions with distances of 5–7 Å reflecting stronger interactions.

From Table 1, it can be seen that the longest total simulation time for atomistic MD is 0.1 μ s and for CG simulations is 200 μ s. The formation of a long-lasting helix dimer was identified within a few hundred nanoseconds in CG-MD studies of Glucophorin A, a non-GPCR model for studying TM

membrane protein structure.²³ The number of replicas performed in these different studies varies tremendously but is never greater than 10. Excellent agreement has been obtained between computed binding free energies and experimental data when ensembles of up to 50 replicas are used.¹ We therefore selected 500 ns for the run time and 50 replicas as starting parameters for these studies. These calculations were run on Legion and Grace, two high-performance Research Computing

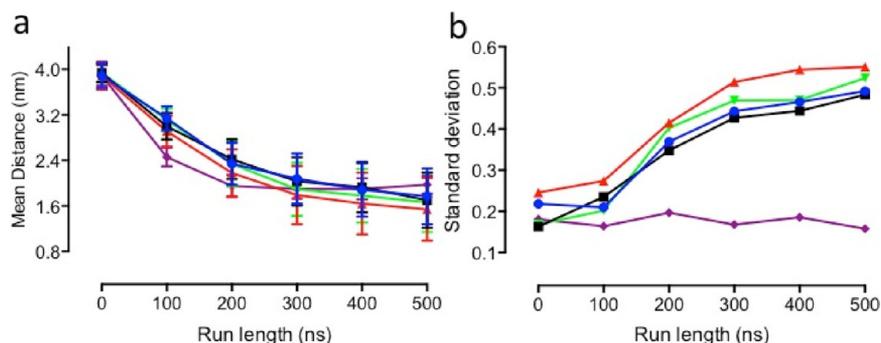


Figure 5. Variation in (a) the mean distance between TM helices and (b) the error (standard deviation) is shown as a function of the run length for the following simulation sets: (blue circle) wild-type TMS–TMS helices, (black square) M177A-mutated TMS–TMS helices, (red triangle) M193A-mutated TMS–TMS helices, (green inverted triangle) M193I-mutated TMS–TMS helices, and (purple diamond) wild-type TM2–TM2 helices.

clusters at University College London (UCL) (details of the machines used can be found at https://wiki.rc.ucl.ac.uk/wiki/RC_Systems#Legion_technical_specs and https://wiki.rc.ucl.ac.uk/wiki/RC_Systems#Grace_technical_specs). Our preliminary tests showed that CG simulations (one ensemble) of 500 ns run on Legion completed within approximately 150 h. CG simulations (one ensemble) of 500 ns run on Grace completed within approximately 72 h.

2.1. CG Simulations. All CG-MD simulations were performed in GROMACS (version 4.6.4) (www.gromacs.org). The temperature was equilibrated for all three groups: protein, lipid bilayer, and solvent (water) with ions to remove the center of mass motion relative to the bilayer and protein. The thermalization run was carried out for 100 ps. The simulations were then run at 310 K (the human physiological temperature), which is below the phase transition temperature of pure DPPC (315 K). The system output of the temperature was evaluated to ensure that it stabilized at the required temperature (310 K) before continuing until pressure equilibration was attained. An ensemble of 50 replicas for each simulation box (see [Tables 2 and 4](#)) was performed. Each simulation was run for 500 ns. CG atom velocities were drawn from a Maxwell–Boltzmann distribution at $T = 310$ K, but all other variables were kept constant; standard deviation was used to compare differences in mean distance outputs. Each simulation was run independently with the initial configurations differing by only the starting velocity; they were performed under the NPT ensemble (i.e., constant temperature, pressure, and particle number) using the Martini 2.2 force field.⁴⁸ The temperatures of the protein and lipid were coupled using the velocity-rescaling (modified Berendsen) thermostat at 310 K (human physiological body temperature) with a coupling constant of $T_t = 1$ ps. The system pressure was semi-isotropic using the Berendsen algorithm at 1 bar with a coupling constant of $T_p = 1$ ps and a compressibility of 1×10^{-4} bar⁻¹. An integration time step of 30 fs was chosen, and the coordinates were saved every 10000 subsequent steps for further analysis. The electrostatic interactions were shifted to zero between 1.0 nm. The Lennard-Jones (LJ) potential was shifted to zero between 0.9 and 1.2 nm to reduce the cutoff noise. The neighbor list for pairwise nonbonded interactions was determined by the Verlet cutoff scheme at 1.4 nm and updated every 10 steps.

2.2. Dimer Analysis. Interhelix distance matrices were calculated for the helix–helix dimer formation and contact maps used to identify specific interactions between residues were generated using the GROMACS tool `g_mdmat`. The individual helix–helix contacts from each replica were examined by

calculating the resulting interhelix distance matrices from the initial simulation starting distance of 4 nm (40 Å) to assess the reproducibility, number of replicas, and run time needed to achieve convergence through a locally written code. In those runs where dimerization was observed, the trajectories were combined and examined in greater depth by calculating the averaged interhelix distance matrices with three different truncation distances of 10, 12, and 15 Å to determine the dimerization properties between the helices. A cutoff distance of 10 Å was then applied to identify residues involved in helix–helix interactions. To investigate the influence of the number of replica simulations on the reliability of our results, we calculated mean interaction distances for ensembles of varying size. For evaluations of run length, mean distance output for the entire ensemble was calculated at 100 ns increments.

Representative atomistic structures of the different CG dimers were generated through use of the “backward” Python script⁴⁹ and the `g_cluster` tool in GROMACS using the `gromos` algorithm at a cutoff of 2.5 nm.⁵⁰ Visualization was performed using VMD.⁵¹ Approximate distances between the atomistic residues in interacting helices were measured using Jmol (www.jmol.org). Amino acid positions have been described using amino acid number in conjunction with the Ballasteros–Weinstein nomenclature⁵² (in superscript). Pairwise combinations used in the analyses were obtained from a matrix of the number of amino acid residues in helix 1 multiplied by the number of amino acid residues in helix 2. In the A_{2A} receptor, there are 729 possible pair combinations between the two 27 residue long TMS helices; for example, combination 552 specifies the combination of residue 23 (helix 1) with residue 24 (helix 2), representing the V196^{5,57}–Y197^{5,58} interaction.

2.3. Construction of A_{2A} R TM Helices and Preparation of the Simulation Box. Initial simulations were performed using TMS of the human A_{2A} adenosine receptor, which has been shown experimentally to form a homodimer.²⁸ TM2 of the A_{2A} receptor was used as a negative control as it was unable to form a homodimer under the same experimental conditions. M193^{5,54}, identified experimentally as being involved in the helical interface and located within a PXXXM motif, and M177^{5,38}, which we identified as residing in a previously unidentified upstream PXXXM motif, were mutated in silico (M177^{5,38}A, M193^{5,54}A, and M193^{5,54}I), to permit simulation of the experimental condition in which M193^{5,54} had been mutated and the biological properties of the mutated protein compared with wild type.

The five A_{2A} TM helices shown in [Table 2](#) were generated using MODELLER 9.12 following the procedure detailed^{53,54}

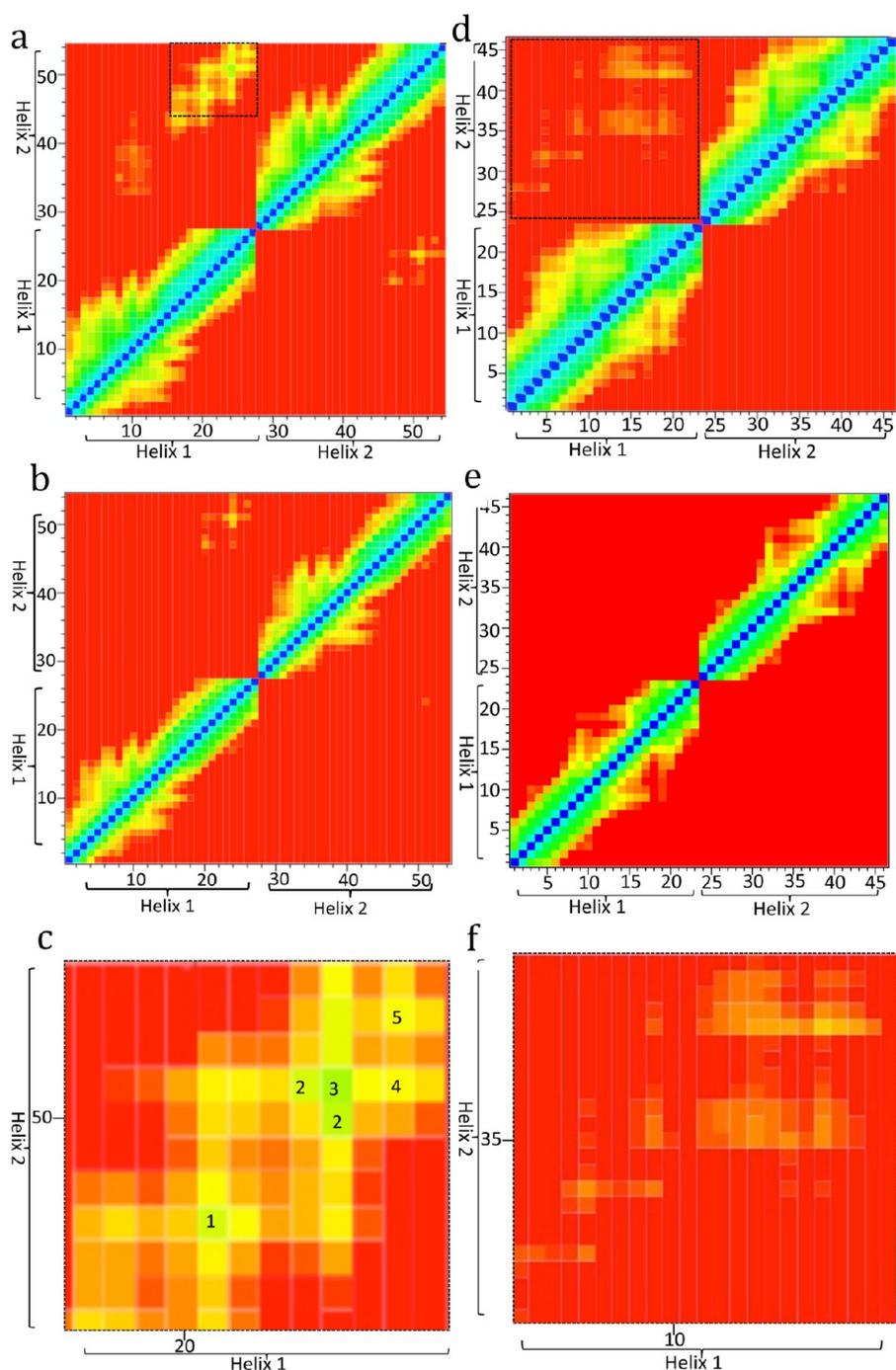


Figure 6. Contact matrices (heat maps) showing specific interactions between residues, as measured by distance, between two A_{2A} helices (“helix 1” and “helix 2”) for the wild-type TMS–TMS simulation (a–c) and the TM2–TM2 negative control (d–f). Results shown are the average for each ensemble. Interhelical distances at the 15 Å cutoff are shown in the top left quarter of panels (a) and (d). Interhelical distances at the 12 Å cutoff are shown in the top left corner of (b) and (e) and in the lower right quarter of panels (a) and (d). Interhelical distances at the 10 Å cutoff are shown in the lower right quarter of panels (a) and (e). The region shown in the black rectangle in (a) and (d) is magnified in (c) and (f), respectively. The five numbered interactions shown in (c) are identified in Table 3. The color scale indicates distance between helices: blue corresponds to 0 Å (superposition of the two helical backbones at all cutoffs); green corresponds to 5 Å (10 Å cutoff), 6 Å (12 Å cutoff), and 7.5 Å (15 Å cutoff); yellow corresponds to 7 Å (10 Å cutoff), 8 Å (12 Å cutoff), and 12 Å (15 Å cutoff); red corresponds to the cutoff distances applied (10, 12, or 15 Å).

using the crystal structure of the A_{2A} receptor (PDB accession number 3EML; GI: 209447557).⁵⁵ The atomistic helices were subsequently converted to CG models using the “martinize” Python script.⁴³ A simulation box of dimensions 8 nm × 8 nm × 8 nm was constructed containing two wild-type TMS helices (Figure 1). The helices were placed 4 nm apart and aligned in a parallel orientation mimicking the natural positioning of the

helix in the membrane (see Figure 1a) with their long axes parallel to the z -axis of the box (see Figure 1b). The TM helices were separated by 4 nm at the beginning of the simulation to rule out any initial interhelix interactions. Water and lipids were then added. Approximately 190 molecules of the 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC) lipid bilayer and additional water molecules (~2660–2690) were added in

Table 3. Number of Interactions (Hits) for Specific Interacting Residues Identified in the Contact Matrices for the Wild-Type TMS–TMS Simulation at the 10 Å Cutoff

figure label	interacting residues	replica number										mean distance \pm standard deviation (in Å)	
		1	5	10	15	20	25	30	35	40	45		50
1	M193 ^{5.54} –M193 ^{5.54}	0	0	1	1	1	2	2	3	5	6	6	7.59 \pm 2.89
2	V196 ^{5.57} –Y197 ^{5.58}	0	1	2	2	2	4	4	6	9	12	13	9.16 \pm 2.5
3	Y197 ^{5.58} –Y197 ^{5.58}	0	1	1	1	1	1	2	4	6	7	8	9.11 \pm 2.85
4	Y197 ^{5.58} –R199 ^{5.60}	0	1	1	1	1	1	1	3	5	6	6	9.83 \pm 3.57
5	R199 ^{5.60} –R199 ^{5.60}	0	1	1	1	1	1	1	3	4	5	5	8.06 \pm 2.99

coarse-grained form in a 3-dimensional cuboid box with periodic boundary conditions using the “insane” Python script.⁵⁶ For neutralization of the net charge on the protein, water molecules were replaced by counterions (either Na⁺ or Cl⁻, as appropriate, depending on the amino acid composition of the helices).

3. RESULTS

Our aim is to investigate the computational parameters required to obtain converged results, to identify whether these results match the experimentally obtained data for the self-association of the TMS helices of the A_{2A} adenosine receptor, and if they do, to further validate these parameters using structural biology data from class A GPCRs that have been experimentally shown to form a dimeric biological unit.

3.1. Internal Sampling, Convergence, and Reproducibility. In our simulations, the TM helices were observed to diffuse freely in the lipid bilayer. The kernel density estimation of the mean distance between the two wild type TMS helices at $t = 0$ and at increments of 100 ns up to completion of the simulation at 500 ns across the 50 replica ensemble is shown in Figure 2. At $t = 0$, the two helices are at their starting positions 40 Å (4.0 nm) apart. At $t = 500$ ns, the mean distance between the helices has adopted a normal distribution with a mean distance of ~ 16 Å between them. The intermediate time points show the redistribution of the distance from the starting point at $t = 0$ to the final mean distance between the helices at 500 ns. A graphical representation of the number and timing of interactions observed in each replica within the ensemble of 50 replicas for the wild-type TMS–TMS simulation is shown in Figure 3. Four of the 50 replicas showed no contact between the helices, which gives rise to the small peak at 40 Å in Figure 2. Three of the replicas began to show contact toward the end of the run, which corresponds to the smaller peak seen at 30 Å in Figure 2.

3.1.1. Optimal Replica Number Required. Five different ensembles, one for each A_{2A} receptor simulation set, were run independently in CG-MD simulations for the total run time of 500 ns. These data, which included both wild type and mutated helix sequences, were used to investigate whether variations in the optimal replica number required would occur between different simulation sets. This information was used to identify the minimum replica number required to achieve convergence for any given simulation set.

Figure 4a reveals that there is no statistically significant difference in the mean distance as a function of replica number. However, a decrease in the error of the mean is observed with increasing ensemble size. From Figure 4b it can be seen that the rate of decrease in the error slows after approximately 15 replicas are included in the ensemble. For each of the five sets, larger ensembles provide less variation in the error of the mean, and an ensemble of 30 replicas represents a good compromise

between computational effort and minimization of the error in the mean distance calculated. We conclude that an ensemble of 30 replicas is sufficient to achieve convergence.

3.1.2. Minimum Run Length Required. The effect of run time on the average distance between the helices was examined by calculating the mean distance and the standard deviation within the 50 replicas for simulations of varying duration (0, 100, 200, 300, 400, and 500 ns). Figure 5a shows a significant effect of run length on both mean distance and standard deviation, confirming the results of Figure 3 and reflecting the time required for interactions to take place. For four of the five simulation sets, the standard deviation increases as a function of time with the rate of increase slowing as the run length becomes longer. In contrast, no change in the standard deviation over time is seen in the TM2–TM2 set (Figure 5b). Interestingly, TM2 homodimers could not be detected experimentally.⁵⁷ The absence of an increase in error in the mean distance as a function of time may serve as an indicator of an absence of interaction between two helices within an ensemble. We conclude that an ensemble run for a simulation time of 300 ns is sufficient to achieve convergence.

3.2. Interacting Interfaces. The final mean distance between the two helices in the ensemble of 50 replicas was used to identify the specific interactions between the A_{2A} homodimers for each simulation set tested. Following application of the 10 Å cutoff, 26% of the ensemble formed stable dimers in the wild-type TMS–TMS simulations. In the mutated TMS-M177^{5.38}A and TMS-M193^{5.54}I simulation sets, 16% of the ensemble formed stable dimers, whereas in TMS-M193^{5.54}A, dimers were detected in 28% of the ensemble. For all four of the TMS simulation sets, the detected interactions took place at the same position within the helices, indicating that a defined orientation is needed to establish a specific interaction. In the negative control (the TM2–TM2 simulation set), 24% of the ensemble resulted in the formation of stable dimers, but there were no specific interactions identified between residues. For all simulations, we combined the trajectories of those pairwise combinations in which dimerization was identified after the cutoff of 10 Å had been applied and compared the results with heat maps of interactions observed at 12 and 15 Å (see Figure 6). The location of the contact interface was then mapped by comparison with the crystal structure of A_{2A}R (3EML).

3.2.1. Identification of Contact Interface for the Wild-Type TMS–TMS Homodimer. Figure 6 shows the average interhelical contact distance between the two wild-type TMS–TMS helices (Figure 6a–c) or between the negative control TM2–TM2 helices (Figure 6d–f). The proximity of the wild-type helices is best visualized at 15 Å (Figure 6a and c). The interacting residues in the wild-type TMS–TMS simulation are found in the bottom third of the C terminal end of TMS. From the averaged interhelix contact matrices, the specific interactions were found to be within the experimentally identified

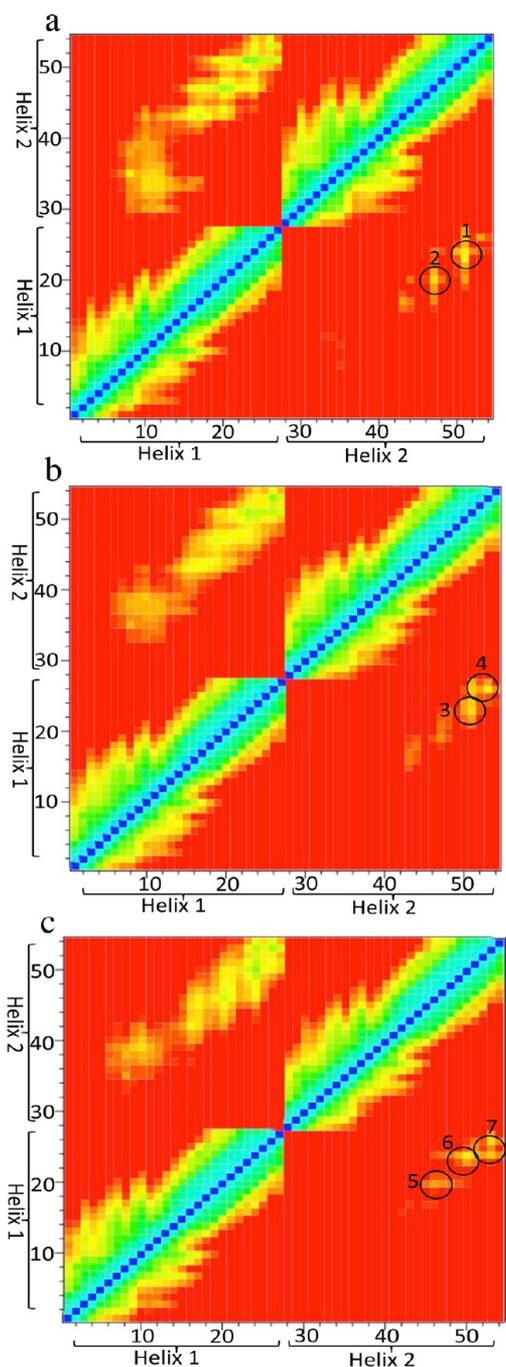


Figure 7. Contact matrices (heat maps) showing specific interactions between two mutated A_{2A} TMS helices (“helix 1” and “helix 2”) with the following residues mutated: M177A (a), M193^{5.54}A (b), and M193^{5.54}I (c). Results shown are the average for each ensemble. Interhelical distances at the 15 and 12 Å cutoff distances are shown in the top left and lower right quarter of panels (a–c), respectively. The color scale is as indicated in Figure 6. Circles indicate areas with key interhelical contacts. The identified amino acid interactions are numbered as follows: (1) M193^{5.54} with M193^{5.54}; (2, 3) V196^{5.57} with Y197^{5.58} and Y197^{5.58} with Y197^{5.58}; (4) Y197^{5.58} with I200^{5.61} and R199^{5.60} with R199^{5.60}; (5) L192^{5.53} with I193^{5.54}, V196^{5.57} with Y197^{5.58}, and Y197^{5.58} with R199^{5.60}; (6) Y197^{5.58} with I200^{5.61} and Y197^{5.58} with R199^{5.60}; and (7) R199^{5.60} with R199^{5.60}.

M193^{5.54}xxVY197^{5.58} motif at an interhelical distance of ~ 8 – 9 Å. The methionine at position 193^{5.54} of helix 1 interacts with the methionine at the same position on helix 2, reinforcing the

Table 4. Sequences of the Rhodopsin, CXCR4, and β_1 AR Receptor Helices Used in Ensemble Simulation Sets

receptor	helices	sequences ^a
rhodopsin	TM1	QFSMLAAYMFLMLGFPIN ^{1.50} (55)FLTLVYTVQ
	TM2	NYILLNLAVAD ^{2.50} (83)LFMVFGGFTTTLTYTSLH
	TM4	ENHAIMGVAFTW ^{4.50} (161)VMALACAAPPL
	TM5	NESFVIYMFVVHFIP ^{5.50} (215)LIVIFFCYGQ
	CXCR4	VVVFQFQHIMVGLLP ^{5.50} (211)GIVIL
β_1 AR	TM1	QWEAGMSLLMALVLLVIVAGN ^{1.50} (59)VLVIAAIG
	TM2	NLFITSLACAD ^{2.50} (87)LVMGLLVVPPGATLVV
	TM4	ARAKVICTVW ^{4.50} (166)AISALVSFLPIMM
	TM5	AYAIASSIISFYIP ^{5.50} (219)LLIMIFVYLRVY

^aThe conserved amino acid for each TM helix is shown in italics and is numbered using both the Ballesteros and Weinstein nomenclature (superscript) and by residue number.

suggestion²⁷ of its importance in the formation of the TMS homodimer. From Figure 6d it can be seen that the distance between TM2–TM2 is close enough to form potential specific interactions; however, none were detected in the combined trajectories for this negative control. Results obtained at the 15 Å cutoff (Figure 6f) were random and nonspecific, supporting the selection of a minimum cutoff distance of 12 Å. It should also be noted that there was no increase in the standard deviation over time for the TM2–TM2 simulation (Figure 5b), whereas there was an increase in this quantity for all simulation sets in which specific interactions occurred, indicating that the change in error over time may be a useful indicator of helix–helix interactions. The frequency of specific interactions identified in the wild-type TMS–TMS ensemble was determined by calculating the mean distance for each frame of every replica individually. Table 3 shows that the five most prominently occurring interactions were between M193^{5.54}–M193^{5.54}, V196^{5.57}–Y197^{5.58}, Y197^{5.58}–R199^{5.60}, R199^{5.60}–R199^{5.60}, and R199^{5.60}–I200^{5.61}.

These findings are consistent with the experimental results²⁷ identifying that the interaction between two wild-type A_{2A} TM5 peptide sequences involved amino acid residue M193^{5.54}. Our findings are also consistent with experimental data showing the formation of A_{2A} receptor homodimers using bioluminescence resonance energy transfer (BRET)²⁶ and bimolecular fluorescence complementation (BiFC).⁵⁸ The presence of specific interactions between TM2 helices was experimentally investigated, and none were detected.^{27,57} Our CG-MD simulations produced the same results as the experimentally obtained findings with the formation of wild-type TMS–TMS dimers involving the M193^{5.54} residue, and no specific interaction was detected between TM2–TM2 helices in silico.

3.2.2. Mutated TMS Interacting Interfaces. Identification of the presence of M193^{5.54} in the contact interface suggested that this residue may play a significant role in how the two TMS helices interact. To investigate this possibility, we introduced sets of ensemble simulations that included mutated helices (see Table 2). Two types of point mutations were used: substitution of (i) methionine to alanine and (ii) methionine to isoleucine. Investigation of the TMS peptide sequence revealed that two separate PxxxM motifs existed within the same helix with a methionine residue present at M177^{5.38} as well the methionine residue identified at M193^{5.54}. Each of these methionine residues was mutated to alanine. We also mutated M193^{5.54} to isoleucine because a conserved PxxxI motif is found in the related family of P2Y receptors at the same location as the originally identified PxxxM motif in A_{2A} R.

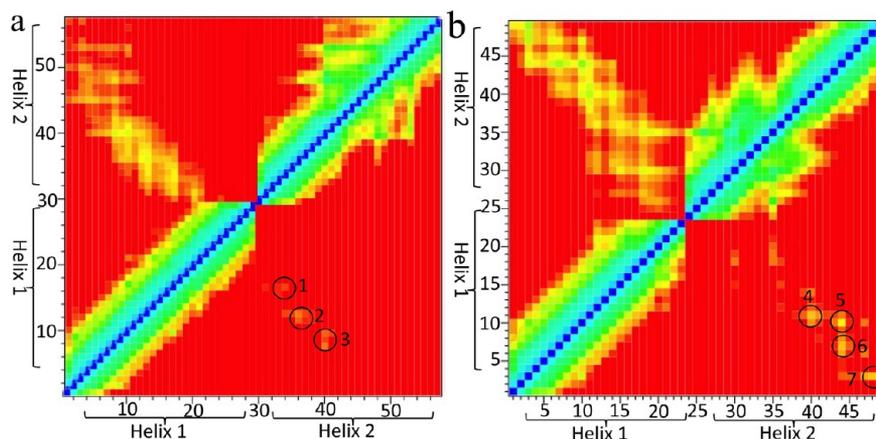


Figure 8. Contact matrices (heat maps) between two rhodopsin helices, showing specific interactions between TM1 (helix 1) and TM2 (helix 2) (a) and between TM4 (helix 1) and TM5 (helix 2) (b). Results shown are the average for each ensemble. The color scale is as indicated in Figure 6. Circles indicate areas with key interhelical contacts. The identified amino acid interactions are numbered as follows: (1) F127^{1.47} with L218^{2.44}; (2) L122^{1.42} with L220^{2.46}, I123^{1.43} with L218^{2.44}, and N219^{2.45} with L220^{2.46}; (3) Y118^{1.38} with D224^{2.50}, M119^{1.39} with D224^{2.50}, F120^{1.40} with D224^{2.50}, and F120^{1.40} with L225^{2.51}; (4) F418^{4.48} with L521^{5.51} and T419^{4.49} with L521^{5.51}; (5) F418^{4.48} with F525^{5.55}; (6) M414^{4.44} with F525^{5.55}, G415^{4.45} with F525^{5.55}, and V416^{4.46} with F525^{5.55}; (7) H411^{4.41} with Y528^{5.58}, H411^{4.41} with G529^{5.59}, and H411^{4.41} with Q530^{5.60}.

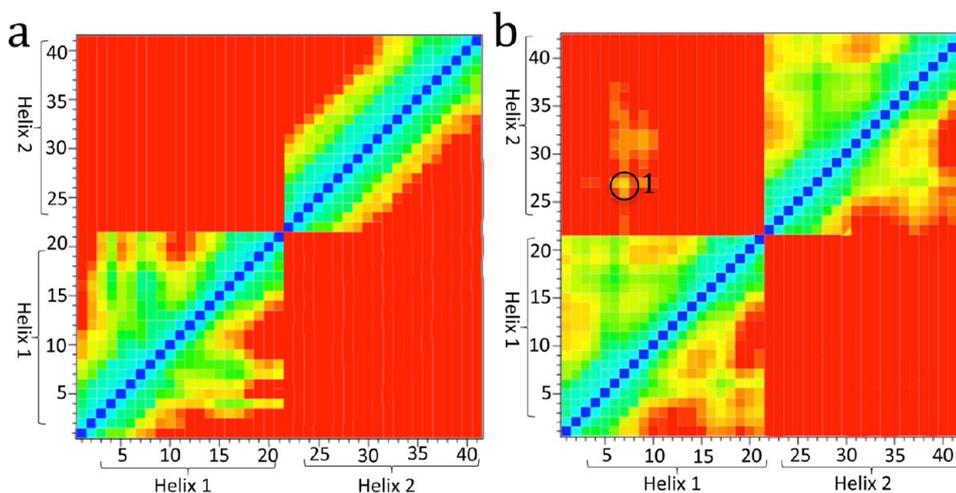


Figure 9. Contact matrices (heat maps) between two CXCR4 helices, showing specific interactions (a) between TMS5 (Helix 1) and TMS6 (Helix 2) and (b) between TMS5 (Helix 1) and TMS5 (Helix 2). The identified amino acid interactions are numbered as follows: 1) F201^{5.40} with V198^{5.37}, Q200^{5.39}, F201^{5.40}, Q202^{5.41}, I204^{5.43} and M205^{5.44}. Table 5 shows a comparison of the distances between specific atoms in interacting residues of the representative structures and the distances between the same atoms in the model structure.

Specifically interacting residues in the TMS5-M177A simulation set were identical to those identified in wild type TMS5–TMS5 dimers (Figure 7a) and included the M193^{5.54}xxxVY197^{5.58} motif. M177^{5.38} was not directly involved in the dimerization between the two helices in any simulation. The specific interactions observed in the TMS5-M193I^{5.54} simulation set were almost identical to those found in the wild-type but included I193^{5.54} in the interaction despite the loss of the methionine at position 193 (Figure 7b). In contrast, the TMS5-M193A^{5.54} mutation completely changed the contact interface of the helices (Figure 7c), and the key interacting residues were identified at a similar distance but contained within a novel V196^{5.57}YxR199^{5.60} motif. This provides a molecular explanation for the finding that mutation of the full-length A_{2A}R at position M193^{5.54} noticeably alters the monomer:dimer ratio as observed with SDS-PAGE.²⁷ Mutation of M193A^{5.54} causes a change in the way in which the two helices come together that prevents formation of TMS5 homodimers, emphasizing the importance of the M193^{5.54} residue in the specificity of TMS5–TMS5 dimer formation in vivo.

3.3. Comparison with Experimental Structural Data.

For assessing the validity of our method, it is necessary to compare our results with experimental values. Our computational results closely match the experimental biophysical data of A_{2A} receptor homodimers and provide information regarding the nature of the contact interface between the two helices that cannot be determined experimentally. We then wished to determine if we could obtain findings in agreement with experimentally obtained structural data. Dimerization in Class A GPCRs involves the transmembrane domains, as opposed to Class C GPCRs, where dimerization is mediated by the large N terminal domain of the protein.⁵⁹ We identified three additional dimeric Class A GPCRs in the PDB database that fulfilled the following criteria: (1) the crystallographic asymmetric unit is a dimer; (2) the software-determined (PISA) quaternary structure is a dimer; and (3) the dimeric quaternary structure has been confirmed functionally. Rhodopsin, the CXCR4 chemokine receptor, and the β₁ adrenergic receptor were chosen for further study; their corresponding TM helices (listed in Table 4)

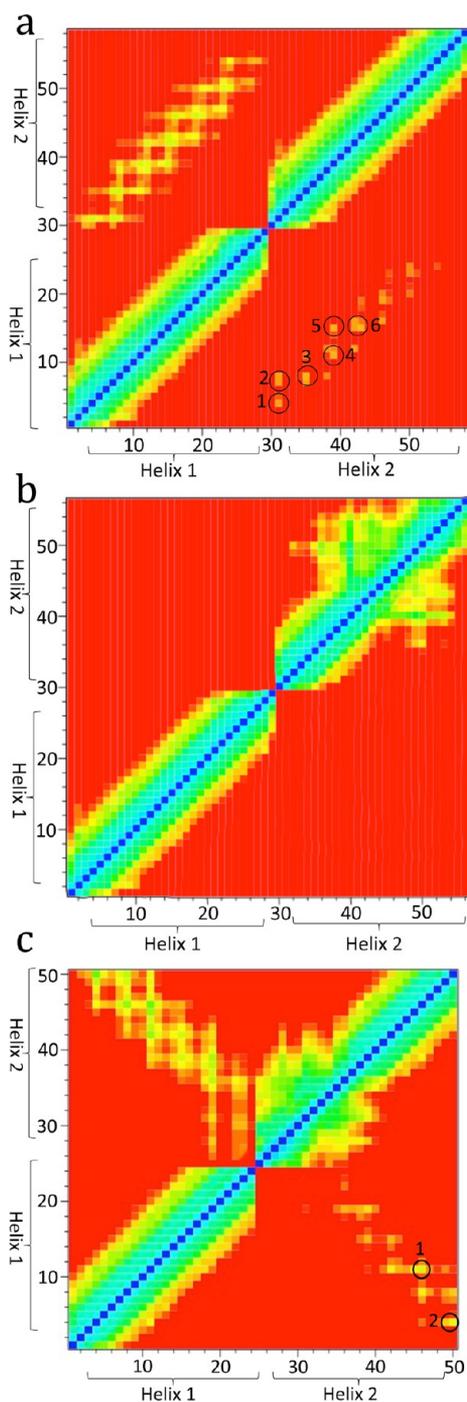


Figure 10. Contact matrices (heat maps) between two β_1 AR helices showing specific interactions between (a) TM1 (helix 1) and TM2 (helix 2), (b) TM1 (helix 1) and TM1 (helix 2), and (c) between TM4 (helix 1) and TM5 (helix 2) (c). Results shown are the average for each ensemble. The color scale is as indicated in Figure 6. Circles indicate areas with key interhelical contacts. The identified amino acid interactions are numbered as follows: in (a) (1) W40^{1.31} with A42^{1.33}, S45^{1.36}, and L46^{1.37}; (2) M44^{1.35} with L46^{1.37}; (3) A49^{1.39} with M48^{1.38}; (4) L53^{1.44} with M48^{1.38}; (5) L53^{1.44} with M48^{1.38}, V51^{1.40}, and V52^{1.41}; (6) L54^{1.45} with V51^{1.40}; (b) (1) K159^{4.43} with Y231^{5.58} and (2) W166^{4.50} with Y227^{5.62}. Table 5 shows a comparison of the distances between specific atoms in interacting residues of the representative structures and the distances between the same atoms in the model structure.

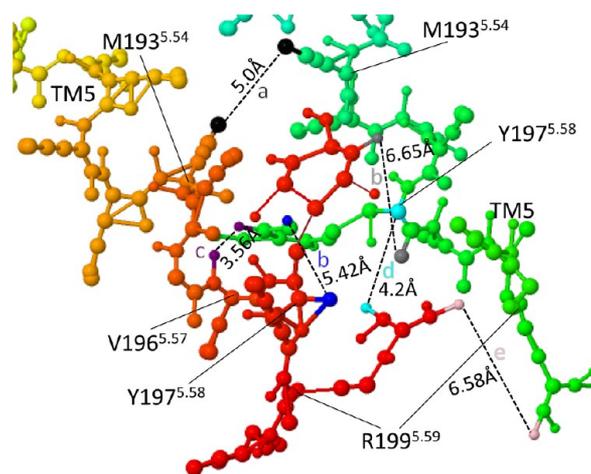


Figure 11. Atomistic representation of the pairwise interactions identified from the wild-type TMS–TMS ensemble. The representative mean distance is shown in the figure, and the mean distance \pm SD for all hits detected per pair is shown in Table 3. All distances between interacting amino acids are calculated from side chain to side chain.

were constructed as described in section 2.3 and used in ensemble-based simulations.

3.3.1. Rhodopsin (1N3M). Rhodopsin has been shown to exist in a native oligomeric form,²⁰ and an atomic model of the rhodopsin dimer has been proposed as a working model for G protein-coupled receptors.¹⁹ Three contact points between the rhodopsin monomers have been reported. The first is considered to be the strongest with the largest contact area (578 Å²) and is located between TM4 and TM5. The second exhibits a contact area of 333 Å² and is located between TM1 and TM2. The third contact point is considered the weakest interaction and is found between rows of dimers at the extracellular ends of TM1 with a contact area of 146 Å².^{19,22} We ran two heterologous simulations between rhodopsin helices TM1 and TM2 and between rhodopsin helices TM4 and TM5 (see Table 4) to identify whether contact interfaces could be identified for either. Figure 8 shows that, for both simulation sets, stable dimers were established, confirming that our computational method is able to produce results in agreement with structural data. In each case, the mean distance between helices was \sim 7.6–8 Å. The mean distance between specific interacting residues in the TM1–TM2 simulation (Figure 8a) is further apart than the mean distance between specific interacting residues in the TM4–TM5 simulation (Figure 8b).

3.3.2. CXCR4 (3ODU). The crystal structure of the CXCR4 chemokine receptor bound to an antagonist small molecule IT1t has been reported and reveals a homodimer with an interface involving TM helices 5 and 6.²³ We investigated interactions between helices in the CXCR4 receptor and identified the formation of stable dimers with specific interactions (Figure 9). We first ran a heterologous simulation between TMS5 and TM6 and were unable to identify any interactions. CXCR4 is able to form homodimers in the absence of ligand⁶⁰ that are unable to be dissociated by a peptide derived from TM6,⁶¹ suggesting that in unliganded CXCR4, the dimer interface may reside between TMS5 and TMS5 in a manner analogous to the A_{2A} receptor. We then ran a CXCR4 TMS–TMS simulation and identified, from the averaged interhelix contact matrices, the formation of dimers with specific interactions between F201^{5.40} and the following six residues: V198^{5.37}, Q200^{5.39}, F201^{5.40}, Q202^{5.41}, I204^{5.43}, and M205^{5.44}.

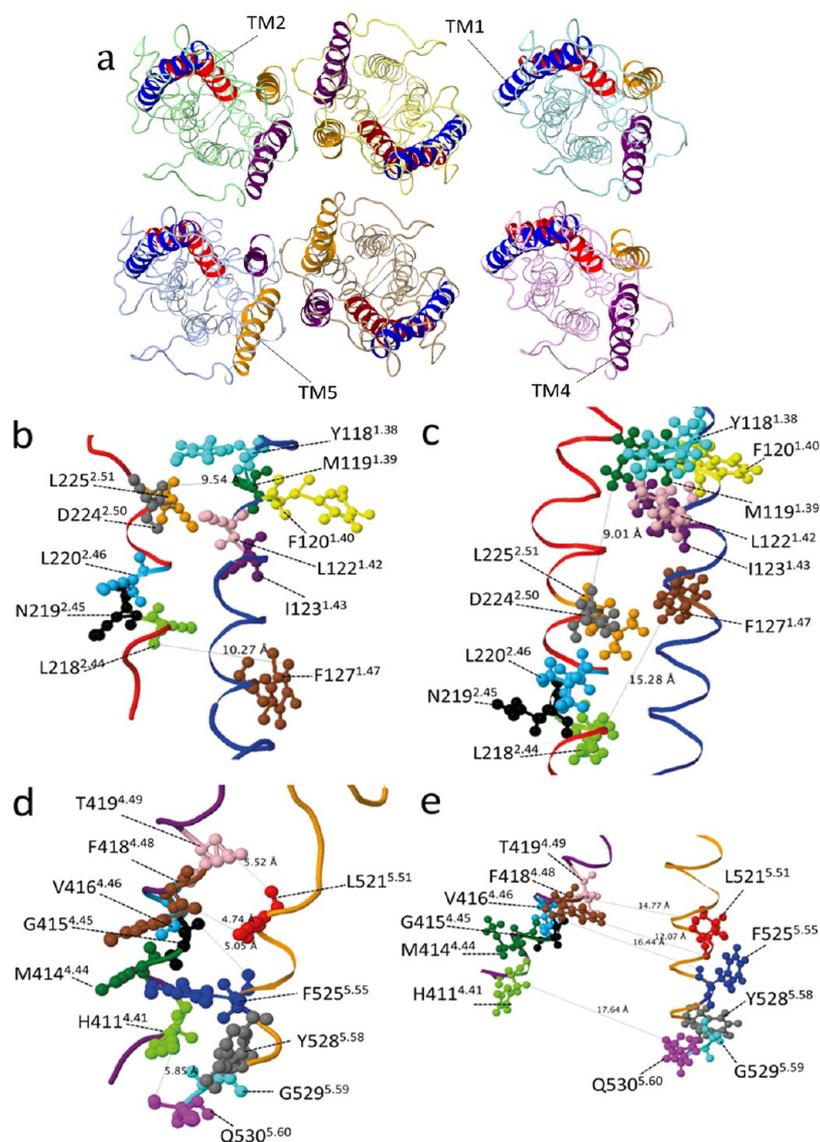


Figure 12. (a) Atomistic structure of the rhodopsin dimer model (1N3M) viewed from above with the TMs used for simulations identified by color as follows: TM1 (blue), TM2 (red), TM4 (purple), and TM5 (orange). Representative TM structures were obtained from the means of all replicas in which interactions were detected. The representative and model structures of TM1–TM2 are shown in (b) and (c), respectively. The representative and model structures of TM4–TM5 are shown in (d) and (e), respectively. Specific interactions were identified in TM1–TM2 simulations (M119^{1.39} with D224^{2.50} and F127^{1.47} with L218^{2.44}) and in TM4–TM5 simulations (H411^{4.41} with Q530^{5.60}, G415^{4.45} with F525^{5.55}, F418^{4.48} with L521^{5.51}, and T419^{4.49} with L521^{5.51}). Table 5 shows a comparison of the distances between specific atoms in interacting residues of the representative structures and the distances between the same atoms in the model structure.

3.3.3. β_1 -Adrenergic Receptor (4GPO). Two alternating dimer interfaces have been proposed from the crystal structure of the ligand-free basal state of the β_1 adrenergic receptor (β_1 AR). The first involves TM1, TM2, extracellular loop 1, and the C-terminal H8; the second involves TM4 and TM5.²⁴ We ran two heterologous simulations between β_1 AR helices TM1 and TM2 and between β_1 AR helices TM4 and TM5 (see Table 4) to identify whether contact interfaces could be identified for either. No stable dimers were formed in the TM1–TM2 simulation (Figure 10a). We investigated the possibility that the contact interface was formed between the two TM1 helices. We ran a TM1–TM1 simulation and identified a stable dimer in only one replica in the ensemble (Figure 10b). Stable dimers were formed between TM4 and TM5 with specific interactions identified between L159^{4.43} and Y231^{5.58} and between W166^{4.50} and Y277^{5.62} (Figure 10c).

3.4. Atomistic Representation and Proposed Nature of Interactions. In CG simulations, a small group of atoms is treated as a single particle (in a 4:1 ratio), a representation that lacks the specific details needed to describe the nature and type of interactions that might take place when the two TM helices are within 10 Å of each other. Representative atomistic structures were generated from CG models to enable a measurement of distance between atoms,⁴⁹ allowing hypotheses to be drawn regarding the molecular nature and possible role of the interactions between dimeric helices.

Figure 11 shows a representation of the converted atomistic wild-type A_{2A} TMS dimer. Using this atomistic representation, the presence of possible electrostatic interactions or hydrogen bonding was investigated by measuring the distance between the specific interacting residues. The interaction between the two methionine residues (M193^{5.54}–M193^{5.54}) and between

Table 5. Comparison of Distance of the Identified Interacting Residues^a from Contact Matrix Graphs of the Rhodopsin, CXCR4, and β_1 AR Helices and the Crystal Rhodopsin Dimer (1N3M), CXCR4 Dimer (4GPO), and the β_1 AR Dimer (3ODU)

receptor	helices	interacting residues	crystal structure distance (Å)	mean distance (Å) \pm standard deviation
rhodopsin	TM1–TM2	F127 ^{1.47} –L218 ^{2.44}	15.28	14.2 \pm 4.07
		M119 ^{1.39} –D224 ^{2.50}	9.01	10.32 \pm 3.1
	TM4–TMS	F418 ^{4.48} –L521 ^{5.51}	12.07	9.18 \pm 3.24
		T419 ^{4.49} –L521 ^{5.51}	14.77	8.13 \pm 1.96
		G415 ^{4.45} –F525 ^{5.55}	16.44	7.5 \pm 2.3
CXCR4	TM5–TMS	H411 ^{4.41} –Q530 ^{5.60}	17.64	9.06 \pm 3.34
		F201 ^{5.40} –V198 ^{5.37}	7.37	15.55 \pm 3.34
		F201 ^{5.40} –Q200 ^{5.39}	11.03	13.7 \pm 1.46
		F201 ^{5.40} –F201 ^{5.40}	7.91	13.6 \pm 2.89
		F201 ^{5.40} –Q202 ^{5.41}	8.72	14.93 \pm 3.13
		F201 ^{5.40} –I204 ^{5.43}	12.14	13.11 \pm 3.37
		F201 ^{5.40} –M205 ^{5.44}	10.6	12.6 \pm 4.96
β_1 AR	TM1–TM1	W40 ^{1.31} –A42 ^{1.33}	12.21	15.28 ^b
		W40 ^{1.31} –S45 ^{1.36}	12.41	17.5 ^b
		W40 ^{1.31} –L46 ^{1.37}	13.84	18.3 ^b
		M44 ^{1.35} –L46 ^{1.37}	9.8	13.46 ^b
		A49 ^{1.39} –M48 ^{1.38}	8.86	5.39 ^b
		L53 ^{1.44} –M48 ^{1.38}	12.19	5.01 ^b
		L53 ^{1.44} –V51 ^{1.40}	10.75	4.9 ^b
		L53 ^{1.44} –V52 ^{1.41}	11.13	5.2 ^b
		L54 ^{1.45} –V51 ^{1.40}	13.9	5.09 ^b
		TM4–TMS	K159 ^{4.43} –Y231 ^{5.58}	ND ^c
	W166 ^{4.50} –Y227 ^{5.62}		ND ^c	7.9 \pm 1.99

^aDistances are measured from backbone to backbone. ^bInteractions were detected in only one replica in the ensemble. ^cNot determined (ND): The distances between TM4 and TMS could not be measured due to the orientation of the dimer in the 4GPO crystal structure, which is submitted showing the TM1–TM2 dimer interface.

valine and tyrosine (V196^{5.57}–Y197^{5.58}) is likely to correspond to van der Waals interactions. Y197^{5.58} in helix 1 and Y197^{5.58} in helix 2 each interact as hydrogen donor and acceptor in the dimer, forming bonds between the peptide backbone and the tyrosine side chain (see Figure 11). As the measurement of these distances is longer than the optimal hydrogen bond distance, 2.7 Å, such hydrogen bonds are more likely to be formed backbone-to-side-chain because their interhelical distance of 8 Å is above the 7.6 Å limit of backbone-to-backbone interactions.⁴⁷

The rhodopsin dimer model (1N3M), shown in Figure 12a, reveals that there is a greater interface area between TM4 and TMS than between TM1 and TM2. The specific interacting residues identified from the atomistic representation obtained using our computational method are distributed throughout the length of TM1 and TM2 but restricted to the bottom third of TM4 and TMS with respect to the intracellular face of the receptor (Figure 12b–e). A comparison of our results with the TM1 and TM2 contact interface of 1N3M is shown in Figure 12b and c, respectively. The measured distance between the hydrogen on the COOH group of M119^{1.39} and the double-bonded oxygen of the COOH group on the side chain of D224^{2.50} is 10.32 \pm 3.21 Å in our model (Figure 12b), similar to 9.01 Å in 1N3M (Figure 12c). Measurement of the distance between F127^{1.47} and L218^{2.44} is 14.20 \pm 4.07 Å in our model and 15.28 Å in 1N3M. F127^{1.47} and L218^{2.44} are located toward the bottom of their respective helices, a position that is constrained by the first intracellular loop of rhodopsin in 1N3M but not in our model. Similar conservation of distance was identified between interacting residues in TM4 and TMS (see Table 5).

Our studies of CXCR4 identified novel interactions in the homodimer between TMS and TMS (Figure 13). This is similar to what was seen for A_{2A}, but the interacting residues in CXCR4 are closer to the extracellular side of the membrane than in A_{2A}. A comparison of the mean distance between interacting residues obtained from the simulations with the distance measured between the same residues in the crystal structure shows a similar conservation of distance, particularly between interacting residues further down the helix. This suggests a contribution of the loops for influencing interactions toward the ends of the helices, as was seen for rhodopsin.

Like rhodopsin, contact interfaces between TM1 and TM2 (Figure 14a,b) and between TM4 and TMS had been proposed for the β_1 adrenergic receptor. However, using our method it is possible to identify a contact interface between TM1 and TM1 rather than between TM1 and TM2. Our measurements of distance are in agreement with those of the crystal structure. Our data suggest that the TM4–TMS contact interface, and the four specific amino acids identified within it, may constitute the principal dimer interface in β_1 AR homodimers (Figure 14c). It was not possible to compare the distances obtained in the TM4–TMS simulation with those measured in the crystal structure 4GPO, which had been submitted with the orientation of the dimer showing the proposed TM1–TM2 interface.

4. CONCLUSIONS

In the present study, we have developed and assessed a method of ensemble-based coarse-grained classical molecular dynamics that we have used to predict protein–protein interactions between TM helices of dimeric GPCRs. We applied our method to four different homomeric GPCRs for which

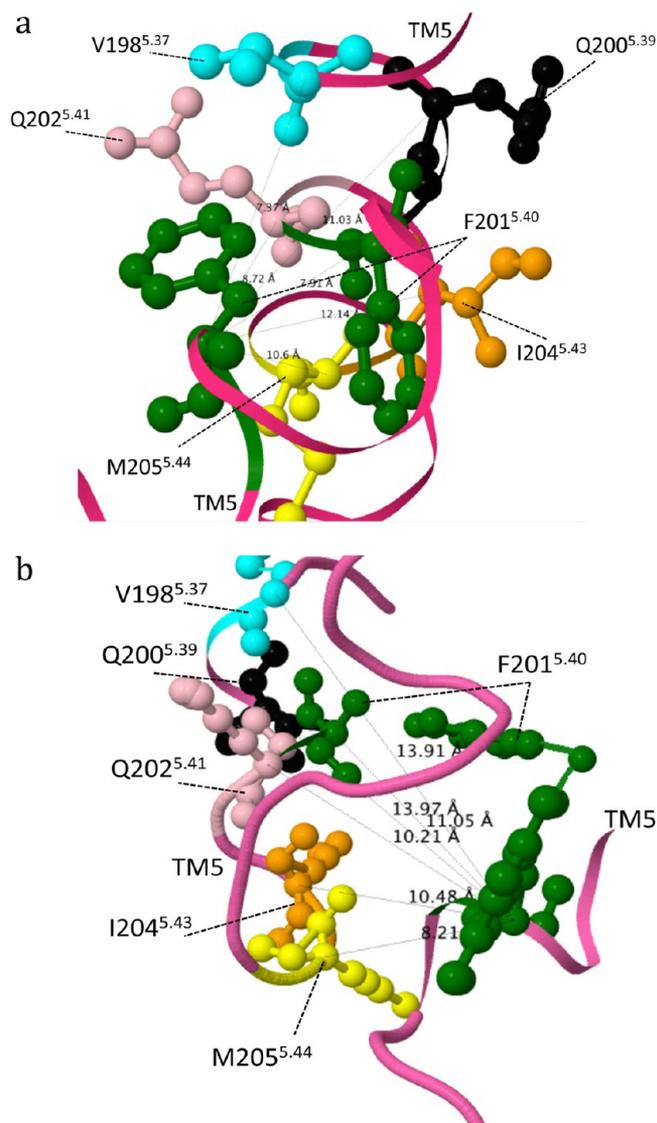


Figure 13. (a) Atomistic structure of the CXCR4 dimer model (3ODU) with the TMs used for simulations identified by color where TMs5 is pink. Representative TM structures were obtained from the means of all replicas in which interactions were detected. The representative and model structures of TMs5–TMs5 are shown in (a) and (b), respectively. Specific interactions were identified in TMs5–TMs5 simulations (F201^{5.40} with V198^{5.37}, Q200^{5.39}, F201^{5.40}, Q202^{5.41}, I204^{5.43}, and M205^{5.44}). Table 5 shows a comparison of the distances between specific atoms in interacting residues of the representative structures and the distances between the same atoms in the model structure.

experimental data exist and compared our predicted results with published experimental data. We have found that, in each case, the ensemble-based CG-MD methodology provides a reproducible measurement of the distance between interacting helices that corresponds well with the experimental data and is within the range of distances at which protein–protein interactions occur.

The first case was that of the A_{2A} adenosine receptor, which had been shown experimentally to form homodimeric receptors through interactions between the TMs5 helices of the two monomers. Our results identified specific interactions involving the PxxxM motif of TMs5 and, specifically, at the M193^{5.54} residue within that motif. Our method accurately identified

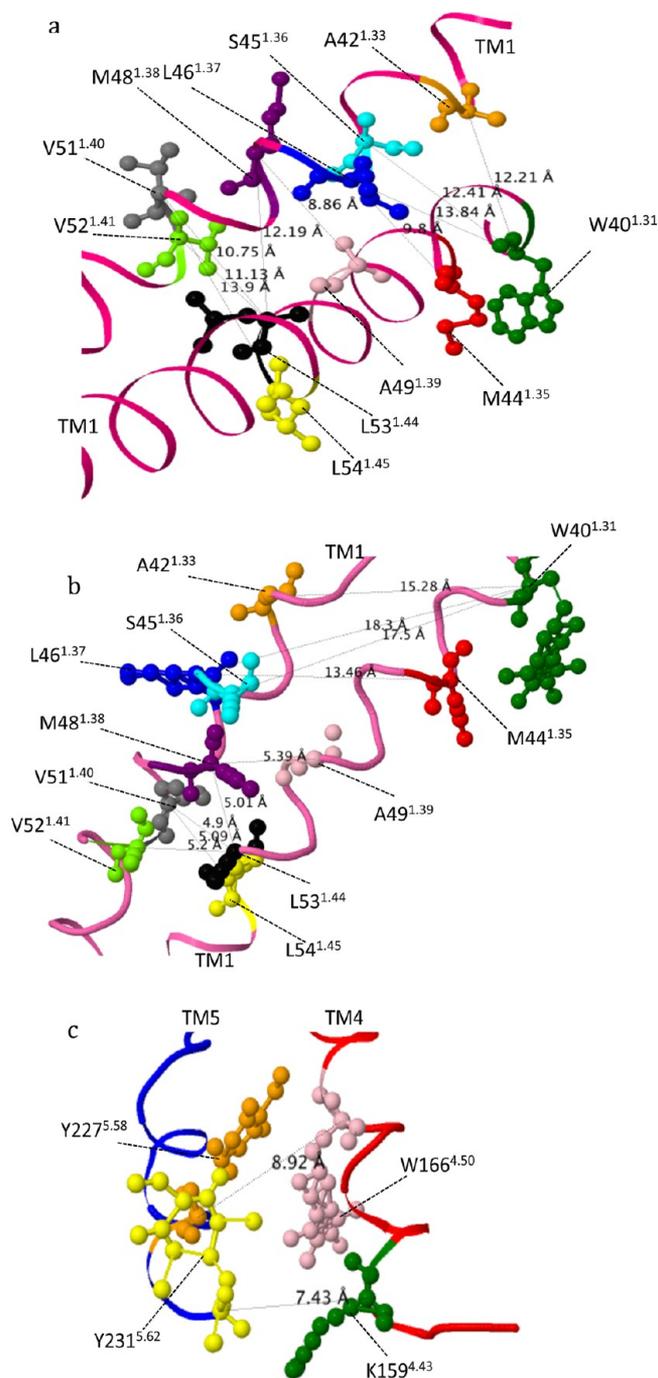


Figure 14. (a) Atomistic structure of the β_2 AR dimer model (4GPO) with the TMs used for simulations identified by color as follows: TM1 (pink), TM4 (red), and TM5 (blue). Representative TM structures were obtained from the means of all replicas in which interactions were detected. The representative and model structures of TMs1–TMs1 are shown in (a) and (b), respectively. The representative and model structures of TMs4–TMs5 are shown in (c). Specific interactions were identified in TMs1–TMs1 simulations (W40^{1.31} with A42^{1.33}, S45^{1.36}, and L46^{1.37}; M44^{1.35} with L46^{1.37}; A49^{1.39} with M48^{1.38}; L53^{1.44} with M48^{1.38}, L53^{1.44} with M48^{1.38}, V51^{1.40}, and V52^{1.41}; L54^{1.45} with V51^{1.40}) and in TMs4–TMs5 simulations (K159^{4.43} with Y231^{5.58}, W166^{4.50} with Y227^{5.62}). Table 5 shows a comparison of the distances between specific atoms in interacting residues of the representative structures and the distances between the same atoms in the model structure.

residues shown experimentally to be involved in TMs5 homodimerization. In parallel with work done experimentally,

we investigated the role of M193^{S54} by characterizing the M193^{S54}A mutation. From this, we identified that the contact interface of the helices was completely changed and that the key interacting residues identified in the wild-type conformation had moved to a new position, preventing the formation of TMS homodimers. Our results provide a molecular explanation for the experimental finding that the M193^{S54}A mutation alters the monomer:dimer ratio at a level of detail that could not be determined biophysically and would require structural biology studies to confirm experimentally. The second case we examined was that of the rhodopsin dimer for which crystallographic data had identified contact interfaces between TM1 and TM2 and between TM4 and TM5. Ensemble CG-MD confirmed dimerization and the identification of specific interactions within each of these heterologous TM pairs. There is a striking convergence between the distances predicted computationally and those calculated from 1N3M, particularly for specific interactions between TMs 1 and 2, showing that our method is able to provide accurate and precise predictions in agreement with experimental findings. Our method is also able to identify novel interfaces as seen in the third (CXCR4) and fourth (β_1 AR) cases we studied, where we identified a novel interface in CXCR4 between TMS and TMS and a novel interface in β_1 AR between TM1 and TM1, in addition to confirming the previously identified contact interface between TM4 and TMS in β_1 AR. The β_1 AR has been shown to form transient interactions, whereas the β_2 adrenergic receptor can form stable oligomers.⁶² Our ability to detect a stable dimer of TM1–TM1 in the β_1 AR shows the value of ensemble-based simulations for the identification of transient interactions.

We note that, in the four cases we studied, there appears to be a pattern emerging of the nature and location of the contact interfaces. We observe either a single interface, at TMS in both A_{2A} and CXCR4, or two contact interfaces, as seen in rhodopsin and β_1 AR, one of which involves TM1 and the other which is between TM4 and TM5. Interestingly, interactions in TMS are observed in both cases. As a greater number of dimeric GPCR crystal structures with corresponding biophysical and functional data become available, the conservation of the pattern we have detected should become clearer.

Our results unequivocally demonstrate that sufficient conformational sampling is required in coarse-grained MD to obtain reproducible and reliable results. In our simulations, we identified that several of the replicas within the ensemble failed to show any interactions and that a number of others began to interact late in the simulation at a point when accurate estimates of distance could no longer be achieved. A single trajectory simulation, particularly if either of these circumstances were to occur, would give inaccurate and potentially misleading results. Indeed, as we have repeatedly emphasized, ensembles are required to obtain accurate and precise results. We used error analysis to determine appropriate choices for ensemble size and run length. For ensemble size, we observed that the rate of change in the standard deviation of the mean distance between helices decreased with increasing replica size and found that approximately 30 replicas were sufficient per ensemble to obtain reproducible results. For run length, we observed that the rate of increase in the standard deviation of the mean distance between helices increased with increasing run length, but that the rate of increase slowed substantially after approximately 300 ns. Interestingly, the negative control we included in our simulations showed no variation in the standard deviation of the mean distance between helices as a

function of run length and a low standard deviation with a very rapid decrease to a constant value at an ensemble size of ~ 15 replicas. This behavior was notably different from simulations in which interactions were identified and provides a means of confirming the absence of interaction.

In conclusion, we have provided a systematic, reproducible, and reliable protocol for determining the specific points of interaction between GPCR dimers. Our method discriminates between residues in TM helices that form specific interactions and residues that are in close proximity but do not interact. Our work extends the recent findings of ensemble-based fully atomistic MD studies, which have shown that an ensemble-based approach is required to generate predictions of protein properties that correlate well with experimental data.⁸³ Our method, which is similar in spirit to a recent publication by Wassenaar et al.,⁸⁴ is of great utility in further understanding GPCR function and also has broad applicability to many different types of membrane proteins, including receptor tyrosine kinases, ion channels, transporters, and oligomeric complexes of their various combinations.

AUTHOR INFORMATION

Corresponding Author

*E-mail: a.townsend-nicholson@ucl.ac.uk

ORCID

David W. Wright: 0000-0002-5124-8044

Peter V. Coveney: 0000-0002-8787-7256

Andrea Townsend-Nicholson: 0000-0002-7250-2208

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge the support of EU H2020 CompBioMed project (<http://www.compbioimed.eu/>, 675451), the MRC Medical Bioinformatics project (MR/L016311/1), and funding from the UCL Provost. N.A.A. was supported by a King Saud University Studentship. The authors acknowledge the use of the UCL Legion and Grace High Performance Computing Facilities (Legion@UCL, Grace@UCL) and associated support services for the completion of this work.

REFERENCES

- (1) Sadiq, S. K.; Wright, D. W.; Kenway, O. A.; Coveney, P. V. Accurate ensemble molecular dynamics binding free energy ranking of multidrug-resistant HIV-1 proteases. *J. Chem. Inf. Model.* **2010**, *50* (5), 890.
- (2) Wan, S.; Coveney, P. V. Rapid and accurate ranking of binding affinities of epidermal growth factor receptor sequences with selected lung cancer drugs. *J. R. Soc., Interface* **2011**, *8* (61), 1114.
- (3) Wright, D. W.; Hall, B. A.; Kenway, O. A.; Jha, S.; Coveney, P. V. Computing Clinically Relevant Binding Free Energies of HIV-1 Protease Inhibitors. *J. Chem. Theory Comput.* **2014**, *10*, 1228.
- (4) Wan, S.; Knapp, B.; Wright, D. W.; Deane, C. M.; Coveney, P. V. Rapid, Precise, and Reproducible Prediction of Peptide-MHC Binding Affinities from Molecular Dynamics That Correlate Well with Experiment. *J. Chem. Theory Comput.* **2015**, *11* (7), 3346.
- (5) Coveney, P. V.; Wan, S. On the calculation of equilibrium thermodynamic properties from molecular dynamics. *Phys. Chem. Chem. Phys.* **2016**, *18*, 30236.
- (6) Bhati, A. P.; Wan, S. Z.; Wright, D. W.; Coveney, P. V. Rapid, Accurate, Precise, and Reliable Relative Free Energy Prediction Using Ensemble Based Thermodynamic Integration. *J. Chem. Theory Comput.* **2017**, *13* (1), 210.

- (7) Wan, S.; Bhati, A. P.; Zasada, S. J.; Wall, I.; Green, D.; Bamborough, P.; Coveney, P. V. Rapid and Reliable Binding Affinity Prediction of Bromodomain Inhibitors: a Computational Study. *J. Chem. Theory Comput.* **2017**, *13*, 784.
- (8) King, B. F.; Townsend-Nicholson, A. Involvement of P2Y1 and P2Y11 purinoceptors in parasympathetic inhibition of colonic smooth muscle. *J. Pharmacol. Exp. Ther.* **2008**, *324* (3), 1055.
- (9) Devader, C.; Drew, C. M.; Geach, T. J.; Tabler, J.; Townsend-Nicholson, A.; Dale, L. A novel nucleotide receptor in *Xenopus* activates the cAMP second messenger pathway. *FEBS Lett.* **2007**, *581* (27), 5332.
- (10) Dalpiaz, A.; Townsend-Nicholson, A.; Beukers, M. W.; Schofield, P. R.; IJzerman, A. P. Thermodynamics of full agonist, partial agonist, and antagonist binding to wild-type and mutant adenosine A1 receptors. *Biochem. Pharmacol.* **1998**, *56* (11), 1437.
- (11) Atkinson, M. R.; Townsend-Nicholson, A.; Nicholl, J. K.; Sutherland, G. R.; Schofield, P. R. Cloning, characterisation and chromosomal assignment of the human adenosine A3 receptor (ADORA3) gene. *Neurosci. Res.* **1997**, *29* (1), 73.
- (12) Townsend-Nicholson, A.; Schofield, P. R. A threonine residue in the seventh transmembrane domain of the human A1 adenosine receptor mediates specific agonist binding. *J. Biol. Chem.* **1994**, *269* (4), 2373.
- (13) Townsend-Nicholson, A.; Shine, J. Molecular cloning and characterisation of a human brain A1 adenosine receptor cDNA. *Mol. Brain Res.* **1992**, *16* (3–4), 365.
- (14) Selbie, L. A.; Townsend-Nicholson, A.; Iismaa, T. P.; Shine, J. Novel G protein-coupled receptors: a gene family of putative human olfactory receptor sequences. *Mol. Brain Res.* **1992**, *13* (1–2), 159.
- (15) Shonberg, J.; Kling, R. C.; Gmeiner, P.; Löber, S. GPCR crystal structures: Medicinal chemistry in the pocket. *Bioorg. Med. Chem.* **2015**, *23*, 3880.
- (16) George, S. R.; O'Dowd, B. F.; Lee, S. P. G-protein-coupled receptor oligomerization and its potential for drug discovery. *Nat. Rev. Drug Discovery* **2002**, *1* (10), 808.
- (17) Jones, K. A.; Borowsky, B.; Tamm, J. A.; Craig, D. A.; Durkin, M. M.; Dai, M.; Yao, W. J.; Johnson, M.; Gunwaldsen, C.; Huang, L. Y.; Tang, C.; Shen, Q. R.; Salon, J. A.; Morse, K.; Laz, T.; Smith, K. E.; Nagarathnam, D.; Noble, S. A.; Branchek, T. A.; Gerald, C. GABA(B) receptors function as a heteromeric assembly of the subunits GABA(B)R1 and GABA(B)R2. *Nature* **1998**, *396* (6712), 674.
- (18) Jordan, B. A.; Devi, L. A. G-protein-coupled receptor heterodimerization modulates receptor function. *Nature* **1999**, *399* (6737), 697.
- (19) Liang, Y.; Fotiadis, D.; Filipek, S.; Saperstein, D. A.; Palczewski, K.; Engel, A. Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes. *J. Biol. Chem.* **2003**, *278* (24), 21655.
- (20) Fotiadis, D.; Liang, Y.; Filipek, S.; Saperstein, D. A.; Engel, A.; Palczewski, K. Atomic-force microscopy: Rhodopsin dimers in native disc membranes. *Nature* **2003**, *421* (6919), 127.
- (21) Bai, M. Dimerization of G-protein-coupled receptors: roles in signal transduction. *Cell. Signalling* **2004**, *16*, 175.
- (22) Fotiadis, D.; Jastrzebska, B.; Philippsen, A.; Muller, D. J.; Palczewski, K.; Engel, A. Structure of the rhodopsin dimer: a working model for G-protein-coupled receptors. *Curr. Opin. Struct. Biol.* **2006**, *16* (2), 252.
- (23) Wu, B.; Chien, E. Y.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **2010**, *330* (6007), 1066.
- (24) Huang, J.; Chen, S.; Zhang, J. J.; Huang, X. Y. Crystal structure of oligomeric beta1-adrenergic G protein-coupled receptors in ligand-free basal state. *Nat. Struct. Mol. Biol.* **2013**, *20* (4), 419.
- (25) Canals, M.; Marcellino, D.; Fanelli, F.; Ciruela, F.; de Benedetti, P.; Goldberg, S. R.; Neve, K.; Fuxe, K.; Agnati, L. F.; Woods, A. S.; Ferre, S.; Lluis, C.; Bouvier, M.; Franco, R. Adenosine A2A-dopamine D2 receptor-receptor heteromerization: qualitative and quantitative assessment by fluorescence and bioluminescence energy transfer. *J. Biol. Chem.* **2003**, *278* (47), 46741.
- (26) Canals, M.; Burgueno, J.; Marcellino, D.; Cabello, N.; Canela, E. I.; Mallol, J.; Agnati, L.; Ferre, S.; Bouvier, M.; Fuxe, K.; Ciruela, F.; Lluis, C.; Franco, R. Homodimerization of adenosine A2A receptors: qualitative and quantitative assessment by fluorescence and bioluminescence energy transfer. *J. Neurochem.* **2004**, *88* (3), 726.
- (27) Thevenin, D.; Lazarova, T.; Roberts, M. F.; Robinson, C. R. Oligomerization of the fifth transmembrane domain from the adenosine A2A receptor. *Protein Sci.* **2005**, *14* (8), 2177.
- (28) Thevenin, D.; Lazarova, T. Stable interactions between the transmembrane domains of the adenosine A2A receptor. *Protein Sci.* **2008**, *17*, 1188.
- (29) Le, F.; Townsend-Nicholson, A.; Baker, E.; Sutherland, G. R.; Schofield, P. R. Characterization and chromosomal localization of the human A2a adenosine receptor gene: ADORA2A. *Biochem. Biophys. Res. Commun.* **1996**, *223* (2), 461.
- (30) Filizola, M.; Weinstein, H. The study of G-protein coupled receptor oligomerization with computational modeling and bioinformatics. *FEBS J.* **2005**, *272*, 2926.
- (31) Simpson, L. M.; Taddese, B.; Wall, I. D.; Reynolds, C. A. Bioinformatics and molecular modelling approaches to GPCR oligomerization. *Curr. Opin. Pharmacol.* **2010**, *10* (1), 30.
- (32) Witt, M.; Ślusarz, M. J.; Ciarkowski, J. Molecular Modeling of Vasopressin V2 Receptor Tetramer in Hydrated Lipid Membrane. *QSAR Comb. Sci.* **2008**, *27* (6), 684.
- (33) Bruno, A.; Entrena Guadix, A. E.; Costantino, G. Molecular Dynamics Simulation of the Heterodimeric mGluR2/5HT2A Complex. An Atomistic Resolution Study of a Potential New Target in Psychiatric Conditions. *J. Chem. Inf. Model.* **2009**, *49*, 1602.
- (34) Vanni, S.; Neri, M.; Tavernelli, I.; Rothlisberger, U. Observation of "Ionic Lock" Formation in Molecular Dynamics Simulations of Wild-Type $\beta 1$ and $\beta 2$ Adrenergic Receptors. *Biochemistry* **2009**, *48*, 4789.
- (35) Rodríguez, D.; Ranganathan, A.; Carlsson, J. Strategies for Improved Modeling of GPCR-Drug Complexes: Blind Predictions of Serotonin Receptors Bound to Ergotamine. *J. Chem. Inf. Model.* **2014**, *54*, 2004.
- (36) Periole, X.; Huber, T.; Marrink, S. J.; Sakmar, T. P. G Protein-Coupled Receptors Self-Assemble in Dynamics Simulations of Model Bilayers. *J. Am. Chem. Soc.* **2007**, *129*, 10126.
- (37) Provasi, D.; Johnston, J. M.; Filizola, M. Lessons from Free Energy Simulations of δ -Opioid Receptor Homodimers Involving the Fourth Transmembrane Helix. *Biochemistry* **2010**, *49*, 6771.
- (38) Periole, X.; Marrink, S. J. The Martini coarse-grained force field. *Methods Mol. Biol.* **2013**, *924*, 533.
- (39) Mondal, S.; Johnston, J. M.; Wang, H.; Khelashvili, G.; Filizola, M.; Weinstein, H. Membrane Driven Spatial Organization of GPCRs. *Sci. Rep.* **2013**, *3*, 2909.
- (40) Prasanna, X.; Chattopadhyay, A.; Sengupta, D. Cholesterol Modulates the Dimer Interface of the $\beta 2$ -Adrenergic Receptor via Cholesterol Occupancy Sites. *Biophys. J.* **2014**, *106*, 1290.
- (41) Ghosh, A.; Sonavane, U.; Joshi, R. Mutiscale modelling to understand the self-assembly mechanism of human beta2-adrenergic receptor in lipid bilayer. *Comput. Biol. Chem.* **2014**, *48*, 29.
- (42) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111* (27), 7812.
- (43) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. The MARTINI coarse-grained force field: Extension to proteins. *J. Chem. Theory Comput.* **2008**, *4* (5), 819.
- (44) Psachoulia, E.; Fowler, P. W.; Bond, P. J.; Sansom, M. S. Helix-helix interactions in membrane proteins: coarse-grained simulations of glycophorin a helix dimerization. *Biochemistry* **2008**, *47* (40), 10503.
- (45) Prakash, A.; Janosi, L.; Doxastakis, M. Self-association of models of transmembrane domains of ErbB receptors in a lipid bilayer. *Biophys. J.* **2010**, *99* (11), 3657.

- (46) Sahoo, D.; Weers, P. M.; Ryan, R. O.; Narayanaswami, V. Lipid-triggered conformational switch of apolipoprotein III helix bundle to an extended helix organization. *J. Mol. Biol.* **2002**, *321* (2), 201.
- (47) Senes, A.; Ubarretxena-Belandia, I.; Engelman, D. M. The C-H...O hydrogen bond: A determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 9056.
- (48) De Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J. Chem. Theory Comput.* **2013**, *9* (1), 687.
- (49) Wassenaar, T. A.; Pluhackova, K.; Böckmann, R. A.; Marrink, S. J.; Tieleman, D. P. Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. *J. Chem. Theory Comput.* **2014**, *10*, 676.
- (50) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide folding: When simulation meets experiment. *Angew. Chem., Int. Ed.* **1999**, *38* (1–2), 236.
- (51) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14* (1), 33–38.
- (52) Ballesteros, J. A.; Weinstein, H. Integrated Methods for the Construction of Three-Dimensional Models and Computational Probing of Structure-Function Relations in G Protein-Coupled Receptors. *Methods Neurosci.* **1995**, *25*, 366.
- (53) Sali, A.; Potterton, L.; Yuan, F.; van Vlijmen, H.; Karplus, M. Evaluation of comparative protein modeling by MODELLER. *Proteins: Struct., Funct., Genet.* **1995**, *23* (3), 318.
- (54) Fiser, A.; Sali, A. Modeller: generation and refinement of homology-based protein structure models. *Methods Enzymol.* **2003**, *374*, 461.
- (55) Jaakola, V. P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y.; Lane, J. R.; Ijzerman, A. P.; Stevens, R. C. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **2008**, *322* (5905), 1211.
- (56) Wassenaar, T. A.; Ingólfsson, H.; Böckmann, R. A.; Tieleman, D. P.; Marrink, S. J. Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.* **2015**, *11*, 2144.
- (57) Thevenin, D.; Roberts, M. F.; Lazarova, T.; Robinson, C. R. Identifying interactions between transmembrane helices from the adenosine A2A receptor. *Biochemistry* **2005**, *44*, 16239.
- (58) Vidi, P. A.; Chen, J.; Irudayaraj, J. M. K.; Watts, V. J. Adenosine A2A receptors assemble into higher-order oligomers at the plasma membrane. *FEBS Lett.* **2008**, *582*, 3985.
- (59) Levitz, J.; Habrian, C.; Bharill, S.; Fu, Z.; Vafabakhsh, R.; Isacoff, E. Y. Mechanism of Assembly and Cooperativity of Homomeric and Heteromeric Metabotropic Glutamate Receptors. *Neuron* **2016**, *92* (1), 143.
- (60) Babcock, G. J.; Farzan, M.; Sodroski, J. Ligand-independent dimerization of CXCR4, a principal HIV-1 coreceptor. *J. Biol. Chem.* **2003**, *278* (5), 3378.
- (61) Percherancier, Y.; Berchiche, Y. A.; Slight, I.; Volkmer-Engert, R.; Tamamura, H.; Fujii, N.; Bouvier, M.; Heveker, N. Bioluminescence resonance energy transfer reveals ligand-induced conformational changes in CXCR4 homo- and heterodimers. *J. Biol. Chem.* **2005**, *280* (11), 9895.
- (62) Dorsch, S.; Klotz, K. N.; Engelhardt, S.; Lohse, M. J.; Bunemann, M. Analysis of receptor oligomerization by FRAP microscopy. *Nat. Methods* **2009**, *6* (3), 225.
- (63) Filipek, S.; Krzysko, K. A.; Fotiadis, D.; Liang, Y.; Saperstein, D. A.; Engel, A.; Palczewski, K. A concept for G protein activation by G protein-coupled receptor dimers: the transducin/rhodopsin interface. *Photochem. Photobiol. Sci.* **2004**, *3* (6), 628.
- (64) Han, Y.; Moreira, I. S.; Urizar, E.; Weinstein, H.; Javitch, J. A. Allosteric communication between protomers of dopamine class A GPCR dimers modulates activation. *Nat. Chem. Biol.* **2009**, *5* (9), 688.
- (65) Kaczor, A. A.; Selent, J.; Sanz, F.; Pastor, M. Modeling complexes of transmembrane proteins: systematic analysis of protein-protein docking tools. *Mol. Inf.* **2013**, *32*, 717.
- (66) Filizola, M.; Wang, S. X.; Weinstein, H. Dynamic models of G-protein coupled receptor dimers: indications of asymmetry in the rhodopsin dimer from molecular dynamics simulations in a POPC bilayer. *J. Comput.-Aided Mol. Des.* **2006**, *20* (7–8), 405.
- (67) Cordomi, A.; Perez, J. J. Structural rearrangements of rhodopsin subunits in a dimer complex: a molecular dynamics simulation study. *J. Biomol. Struct. Dyn.* **2009**, *27* (2), 127.
- (68) Neri, M.; Vanni, S.; Tavernelli, I.; Rothlisberger, U. Role of aggregation in rhodopsin signal transduction. *Biochemistry* **2010**, *49* (23), 4827.
- (69) Periole, X.; Knepp, A. M.; Sakmar, T. P.; Marrink, S. J.; Huber, T. Structural determinants of the supramolecular organization of G protein-coupled receptors in bilayers. *J. Am. Chem. Soc.* **2012**, *134* (26), 10959.
- (70) Fanelli, F.; Menziani, C.; Scheer, A.; Cotecchia, S.; De Benedetti, P. G. Theoretical study on receptor-G protein recognition: New insights into the mechanism of the alpha 1b-adrenergic receptor activation. *Int. J. Quantum Chem.* **1999**, *73* (2), 71.
- (71) Soulier, J. L.; Russo, O.; Giner, M.; Rivail, L.; Berthouze, M.; Onger, S.; Maigret, B.; Fischmeister, R.; Lezoualc'h, F.; Sicsic, S.; Berque-Bestel, I. Design and synthesis of specific probes for human 5-HT4 receptor dimerization studies. *J. Med. Chem.* **2005**, *48* (20), 6220.
- (72) Russo, O.; Berthouze, M.; Giner, M.; Soulier, J. L.; Rivail, L.; Sicsic, S.; Lezoualc'h, F.; Jockers, R.; Berque-Bestel, I. Synthesis of specific bivalent probes that functionally interact with 5-HT4 receptor dimers. *J. Med. Chem.* **2007**, *50* (18), 4482.
- (73) Berthouze, M.; Rivail, L.; Lucas, A.; Ayoub, M. A.; Russo, O.; Sicsic, S.; Fischmeister, R.; Berque-Bestel, I.; Jockers, R.; Lezoualc'h, F. Two transmembrane Cys residues are involved in 5-HT4 receptor dimerization. *Biochem. Biophys. Res. Commun.* **2007**, *356* (3), 642.
- (74) Gorinski, N.; Kowalsman, N.; Renner, U.; Wirth, A.; Reinartz, M. T.; Seifert, R.; Zeug, A.; Ponimaskin, E.; Niv, M. Y. Computational and Experimental Analysis of the Transmembrane Domain 4/5 Dimerization Interface of the Serotonin 5-HT1A Receptor. *Mol. Pharmacol.* **2012**, *82* (3), 448.
- (75) Rodriguez, D.; Gutierrez-de-Teran, H. Characterization of the homodimerization interface and functional hotspots of the CXCR4 chemokine receptor. *Proteins: Struct., Funct., Genet.* **2012**, *80* (8), 1919.
- (76) Casciari, D.; Dell'Orco, D.; Fanelli, F. Homodimerization of neurotensin 1 receptor involves helices 1, 2, and 4: insights from quaternary structure predictions and dimerization free energy estimations. *J. Chem. Inf. Model.* **2008**, *48* (8), 1669.
- (77) Johnston, J. M.; Aburi, M.; Provasi, D.; Bortolato, A.; Urizar, E.; Lambert, N. A.; Javitch, J. A.; Filizola, M. Making structural sense of dimerization interfaces of delta opioid receptor homodimers. *Biochemistry* **2011**, *50* (10), 1682.
- (78) Fanelli, F.; Felling, A. Dimerization and ligand binding affect the structure network of A2A adenosine receptor. *Biochim. Biophys. Acta, Biomembr.* **2011**, *1808*, 1256.
- (79) Kim, S. K.; Jacobson, K. A. Computational prediction of homodimerization of the A3 adenosine receptor. *J. Mol. Graphics Modell.* **2006**, *25* (4), 549.
- (80) Fanelli, F. Dimerization of the lutropin receptor: insights from computational modeling. *Mol. Cell. Endocrinol.* **2007**, *59*, 260.
- (81) Woolf, P. J.; Linderman, J. J. An algebra of dimerization and its implications for G-protein coupled receptor signaling. *J. Theor. Biol.* **2004**, *229*, 157.
- (82) Liu, X.; Ming Kai, M.; Jin, J.; Wang, R. Computational study of the heterodimerization between l and d receptors. *J. Comput.-Aided Mol. Des.* **2009**, *23*, 321.
- (83) Wan, S.; Bhati, A.; Skerratt, S.; Omoto, K.; Shanmugasundaram, V.; Bagal, S.; Coveney, P. V. Evaluation and Characterization of Trk Kinase Inhibitors for the Treatment of Pain: Reliable Binding Affinity Predictions. *J. Chem. Inf. Model.* **2017**, Available Online DOI: [10.1021/acs.jcim.6b00780](https://doi.org/10.1021/acs.jcim.6b00780).
- (84) Wassenaar, T. A.; Pluhackova, K.; Moussatova, A.; Sengupta, D.; Marrink, S. J.; Tieleman, D. P.; Böckmann, R. A. High-Throughput Simulations of Dimer and Trimer Assembly of Membrane Proteins. The DAFT Approach. *J. Chem. Theory Comput.* **2015**, *11*, 2278.

Evaluation and Characterization of Trk Kinase Inhibitors for the Treatment of Pain: Reliable Binding Affinity Predictions from Theory and Computation

Shunzhou Wan,[†] Agastya P. Bhati,[†] Sarah Skerratt,[‡] Kiyoyuki Omoto,[§] Veerabahu Shanmugasundaram,^{||} Sharan K. Bagal,[‡] and Peter V. Coveney^{*,†}

[†]Centre for Computational Science, Department of Chemistry, University College London, London WC1H 0AJ, United Kingdom

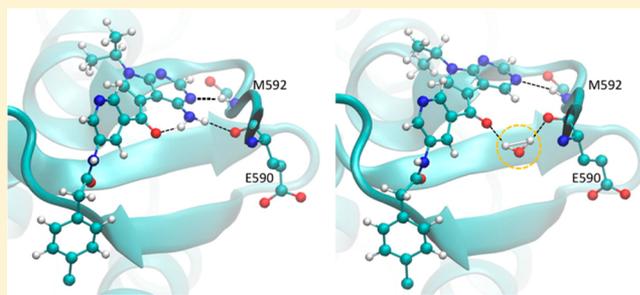
[‡]Pfizer Global R&D, The Portway Building, Granta Park, Cambridge CB21 6GS, United Kingdom

[§]Pfizer Worldwide Research and Development, 610 Main Street, Cambridge, Massachusetts 02139, United States

^{||}Pfizer Worldwide Research and Development, Groton Laboratories, Groton, Connecticut 06340, United States

Supporting Information

ABSTRACT: Optimization of ligand binding affinity to the target protein of interest is a primary objective in small-molecule drug discovery. Until now, the prediction of binding affinities by computational methods has not been widely applied in the drug discovery process, mainly because of its lack of accuracy and reproducibility as well as the long turnaround times required to obtain results. Herein we report on a collaborative study that compares tropomyosin receptor kinase A (TrkA) binding affinity predictions using two recently formulated fast computational approaches, namely, Enhanced Sampling of Molecular dynamics with Approximation of Continuum Solvent (ESMACS) and Thermodynamic Integration with Enhanced Sampling (TIES), to experimentally derived TrkA binding affinities for a set of Pfizer pan-Trk compounds. ESMACS gives precise and reproducible results and is applicable to highly diverse sets of compounds. It also provides detailed chemical insight into the nature of ligand–protein binding. TIES can predict and thus optimize more subtle changes in binding affinities between compounds of similar structure. Individual binding affinities were calculated in a few hours, exhibiting good correlations with the experimental data of 0.79 and 0.88 from the ESMACS and TIES approaches, respectively. The speed, level of accuracy, and precision of the calculations are such that the affinity predictions can be used to rapidly explain the effects of compound modifications on TrkA binding affinity. The methods could therefore be used as tools to guide lead optimization efforts across multiple prospective structurally enabled programs in the drug discovery setting for a wide range of compounds and targets.



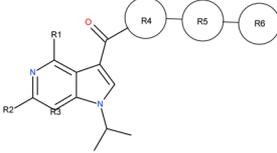
INTRODUCTION

The availability of computational methods that can reliably, rapidly, and accurately predict the binding affinities of ligands to a target protein of interest would greatly facilitate drug discovery programs by enabling project teams to more effectively triage design ideas and therefore synthesize only those compounds with a high probability of being pharmacologically active. The overall effect of this approach would be to reduce the number of “design–synthesis–test” cycles needed to generate compounds of sufficient bioactivity to progress to the clinic. Until recently, *in silico* methods of binding affinity prediction have not been regarded as reliable enough to produce such actionable results. The American Chemical Society’s Cross-Pharmaceutical Industry group has been discussing this particular challenge of conducting elaborative studies and has come out with an opinion piece recently, in which it is specifically noted that “coordinated, blinded prediction challenges offer the best opportunity to develop

broad understanding and broadly applicable methods”.¹ The Coveney group has now developed a suite of computational methods that deliver rapid, accurate, precise, and reliable binding affinity predictions. We report here a prospective computational study on a then-ongoing project at Pfizer² to assess the effectiveness of these methods based on the use of a Binding Affinity Calculator (BAC) software tool and associated services.³ The approach makes use of an automated workflow⁴ running in a high-performance computing environment that builds models, runs large numbers of calculations, and analyzes the output data in order to place reliable error bounds on predicted ligand binding affinities. In order to assess the reliability of these methods, the predictions were performed blind at UCL and subsequently compared with the experimentally determined TrkA binding affinity data.² For a

Received: December 21, 2016

Published: March 20, 2017

Table 1. Ligands Considered in This Study, Numbered as Per the Order in the Files Provided by Pfizer^a


Compound	R1	R2	R3	R4	R5	R6	IC ₅₀ (nM)	ΔG (kcal/mol)
1	NH ₂	H	N				6.97 4.25	-11.20 -11.49
3	NH ₂	H	N				24.53	-10.45
4	H	H	N				10.86	-10.93
6	NH ₂	H	N				8.67	-11.07
7	H	H	CH				44.25	-10.09
8	H	H	CH				735.79	-8.42
11	H	H	N				11.04 5.45	-10.92 -11.34
12	H	H	N				14.18	-10.77
13	H	H	N				224.74	-9.13
16	H	H	N				9800.00	-6.87
17	H	NH ₂	N				6.84 3.73	-11.21 -11.57
22	OH	H	N				1756.36	-7.90
23	H	NHCH ₃	N				9.55	-11.01
24	H	OH	N				249.42	-9.06
26	H	H	N				20.53	-10.55
30	H	H	N				201.70 47.87	-9.19 -10.05

^aAll of the ligands have the same net neutral charge. The experimental TrkA inhibitory values (IC₅₀) and the binding free energies derived from them are shown. Experimental IC₅₀ measurements were conducted independently in two separate laboratories using an identical protocol;² Pfizer, Sandwich (U.K.) IC₅₀ values are shown in black; TCG Lifescience (India) IC₅₀ values are shown in blue.

new method being pursued by an industrial computational chemist, these types of live blinded individual project team collaborations are the best kinds of pilot studies that can be performed.

Tropomyosin receptor kinase A (TrkA) is used as the target protein in this study. TrkA is the cognate receptor of the nerve growth factor (NGF) neuropeptide, a neurotrophic factor involved in the regulation of growth, maintenance, proliferation, and survival of certain target neurons.⁵ Preclinical and clinical studies have identified a crucial role for NGF in the pathogenesis of pain; the clinical efficacy of anti-NGF monoclonal antibody (mAb) therapies against several pain end points is well-documented, and preclinically the inhibition of the TrkA kinase domain by small-molecule kinase inhibitors has been shown to reverse the effects of NGF-mediated pain transduction.^{6,7} In view of the key role of NGF in modulating pain, there is significant interest in the clinical development of

small-molecule TrkA inhibitors to complement anti-NGF mAb treatment options.

The UCL group has recently introduced new approaches termed Enhanced Sampling of Molecular dynamics with Approximation of Continuum Solvent (ESMACS) and Thermodynamic Integration with Enhanced Sampling (TIES)⁸ for the reliable prediction of ligand–protein free energies. The ESMACS approach centers on the molecular mechanics Poisson–Boltzmann surface area (MMPBSA) method,⁹ which employs a continuum approximation for the aqueous solvent, while TIES is based on thermodynamic integration. The approaches emphasize the necessity of invoking ensemble-based sampling to reliably compute macroscopic quantities using microscopic modeling methods.¹⁰ The BAC tool³ is employed to automate much of the complexity of running and marshalling the required molecular dynamics simulations as well as collecting and analyzing data. The BAC tool utilized to perform the ESMACS and TIES studies

constitutes a computational pipeline built from an array of software tools and services and requires access to suitable computing resources. Integration and automation are central to the reliability of the method, ensuring that the results are reproducible and can be delivered rapidly.

■ COMPUTATIONAL METHODS

The TrkA binding affinities of the compounds in Table 1 were calculated using ESMACS and TIES. These compounds were picked as representatives of the full series previously studied experimentally,² covering the dynamic range of the TrkA pharmacology assay and containing the structural features deemed as key determinants of TrkA activity. (A second series of Pfizer pan-Trk inhibitors, structurally dissimilar to those highlighted in Table 1, also showed good agreement between their ESMACS-predicted and experimentally determined TrkA binding affinities. Studies of those inhibitors will be published in due course.) These congeneric compounds all have the same net neutral charge. In ESMACS, the MMPBSA.py.MPI¹¹ module of the AMBER12 package¹² was employed for the free energy calculations of the complex (G_{complex}), the receptor (G_{receptor}), and the compound (G_{ligand}) (eq 1). The polar solvation free energy was calculated using the Poisson–Boltzmann equation with a grid spacing of 0.5 Å and dielectric constants of 1 and 80 for the protein and solvent, respectively. The energetic analyses, including the configurational entropy calculations, were conducted on 50 snapshots of complexes, protein, and compounds extracted evenly from each 4 ns production run in the ensemble MD simulations of complexes only (1-traj and 2-traj; see below) or from separate ensemble simulations of complexes and compounds (3-traj). In TIES, the energy derivatives $\partial V/\partial\lambda$ (eq 2) were recorded every 2 fs during the simulations, and a stochastic integration^{10,13} was performed using a trapezoidal method. Hydrogen-bond analyses were performed for the ensemble simulation trajectories. A hydrogen bond is considered to be formed when the distance between a hydrogen-bond acceptor and a hydrogen-bond donor is less than a defined distance cutoff and the acceptor–hydrogen–donor angle is greater than an angle cutoff. The cutoffs used in the current study are 3.0 Å for the distance and 135° for the angle. The protein and the compounds can have intramolecular and intermolecular hydrogen bonds, of which the intermolecular ones contribute significantly to the binding affinities. Some “bridging” water molecules, which are hydrogen-bonded to the protein and the compounds at the same time, also play an important role in binding of ligands to proteins, in addition to their direct interactions. In ESMACS predictions, the inter- and intramolecular interactions of the compounds and the protein are calculated explicitly, while the interactions with water, including such bridging water molecules, are taken into account implicitly.

In ESMACS, the free energy is evaluated approximately on the basis of the extended MMPBSA method,^{8,14} including configurational entropy and the free energy of association.¹⁵ Free energy changes ($\Delta G_{\text{binding}}$) are determined for the molecules in their solvated states. The binding free energy change is then calculated as

$$\Delta G_{\text{binding}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}} \quad (1)$$

where G_{complex} , G_{receptor} , and G_{ligand} are the free energies of the complex, the receptor, and the ligand, respectively.

The three terms on the right-hand side of eq 1 can be generated from single simulations of the complexes or from separate simulations of complexes, receptor(s), and ligand(s). The former is the so-called one-trajectory (1-traj) method, in which the trajectories of the receptor(s) and ligand(s) are extracted from those of the complexes. The latter is the so-called three-trajectory (3-traj) method when the energies are derived from independent simulations of the three components or the two-trajectory (2-traj) method when the energy is derived from independent simulations for only the receptor or the ligand. In the drug development field, binding is usually investigated for a set of ligands bound to the same protein target. The free energy of the receptor, G_{receptor} , is then the same for all of ligands when it is derived from an independent simulation of the receptor and hence can be treated as a constant. In the current study, we employ three approaches to calculate $\Delta G_{\text{binding}}$: (a) employing only simulations of the complexes (1-traj); (b) the same as (a) but using the average of the receptor free energies $\langle G_{\text{receptor}} \rangle$ from the 1-traj method (denoted as 2-traj); and (c) the same as (b) but also invoking separate ligand simulations for the derivation of G_{ligand} (denoted as 3-traj).

In our more recent TIES method,¹³ an “alchemical transformation”¹⁶ for the mutated entity, either the ligand or the protein (in this study it is always the ligand), is used in both aqueous solution and within the ligand–protein complex. The relative free energy changes for the alchemical mutation processes, $\Delta G_{\text{aq}}^{\text{alch}}$ and $\Delta G_{\text{complex}}^{\text{alch}}$, are calculated as

$$\Delta G^{\text{alch}} = \int_0^1 \left\langle \frac{\partial V(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (2)$$

where λ ($0 \leq \lambda \leq 1$) is a coupling parameter such that $\lambda = 0$ and $\lambda = 1$ correspond to the initial and final thermodynamic states, $V(\lambda)$ is the potential energy of an intermediate state λ , and $\langle \dots \rangle_{\lambda}$ denotes an ensemble average over configurations representative of the state λ . The relative binding free energy difference is then calculated as

$$\Delta \Delta G^{\text{binding}} = \Delta G_{\text{aq}}^{\text{alch}} - \Delta G_{\text{complex}}^{\text{alch}} \quad (3)$$

Compared with ESMACS, the TIES approach, like all other alchemical-based free energy methods, is usually more accurate in comparing one congeneric ligand to another. However, the nature of these alchemical methods implies that their applications will be limited when diverse sets of compounds are of interest. A recent publication¹⁷ in which the authors claimed that their implementation of free energy perturbation (FEP) calculations successfully predicted a number of more active compounds (a correlation coefficient of 0.71 between the predicted and experimental binding affinities was achieved) has the same limitations. ESMACS, however, has no such limitations. Our previous studies⁸ have shown that ESMACS can be applied to sets of compounds that are highly diverse in terms of both the number of atoms and the net charge. In addition, ESMACS is able to make assessments of the effect that an individual ligand has on the protein to which it binds. In particular, the application of two- and three-trajectory versions of ESMACS allows one to probe the extent to which both the ligand and protein adjust their conformations on binding, broadly according to a “lock and key” or “induced fit” recognition mechanism.

In the current study, we apply the ESMACS approach for all of TrkA compounds highlighted in Table 1 and consider TIES

for a selected subset. In the evaluation and characterization of TrkA binding affinity predictions from the Pfizer compound set, the ESMACS and TIES protocols were conducted as recently described.^{8,13} TrkA compounds were optimized at the Hartree–Fock level with the 6-31G* basis (HF/6-31G*) in Gaussian 03¹⁸ and parametrized using Antechamber and RESP in AmberTools 12 with the general AMBER force field (GAFF).¹⁹ The Amber ff99SBildn force field²⁰ was used for the protein. In ESMACS, we used 25 replicas in an ensemble calculation for each ligand. In TIES, five replicas were used for each selected pair of ligands. The MD package NAMD2.9²¹ was used throughout the equilibration and production simulations with periodic boundary conditions. We used a protocol established in our previous publications^{13,14} in which 2 ns of equilibration and 4 ns of production were conducted for each replica. Energy analyses showed that a 4 ns production run was sufficient to have convergence results for the current molecular systems (see Figure S1 in the Supporting Information). In ESMACS, the MMPBSA.py.MPI¹¹ module of the Amber package¹² was used to extract the free energies of the complexes, the protein, and the compounds (eq 1), including configurational entropy from normal mode calculations.

ESMACS was performed on SuperMUC, two separate high-end supercomputers with a total of ca. 245 000 cores in a combination of various processor technologies (<https://www.lrz.de/services/compute/supermuc/systemdescription/>), at the Leibniz Rechenzentrum (Leibniz Supercomputing Centre) in Garching, Germany. We were able to run our calculations on them as if it were a single supercomputer.²² TIES calculations were performed on ARCHER, a Cray XC30 supercomputer (equipped with ca. 110 000 cores), the U.K.'s National High Performance Computing Service located in Edinburgh. A single ESMACS study of one ligand–protein interaction and a TIES study of the binding free energy difference for two congeneric ligands can both be completed in about 8 wall-clock hours on 1200 and 3120 CPU cores, respectively, on ARCHER (see details in our previous publication^{8,13}). (It is possible to further reduce the time required by use of GPU accelerators.) The ESMACS and TIES approaches are both scalable, allowing a large number of calculations to be performed concurrently, depending only on the computing resources available.

RESULTS

We report a collaborative study performed between academic computational chemists and a pharmaceutical company on a data set of 16 pan-Trk ligands (Table 1).² Of the TrkA cocrystal structures provided by Pfizer,² we elected to utilize the cocrystal structure of TrkA and **1** (Figure 1, PDB code 5JFV) as it contained the highest number of crystallographically defined TrkA residues. The missing residues were built by ModLoop,²³ while the initial structures of the complexes were constructed by a docking method employing UCSF DOCK.²⁴ Nine of the 16 compounds were successfully docked into the binding site of the TrkA protein with orientations agreeing with those from available X-ray structures. Compounds that failed to generate suitable cocomplexes via docking were manually positioned on the basis of X-ray and/or modeled structures of similar compounds.

Comparison of Experimental TrkA Binding Free Energies with ESMACS Predictions. The comparison of the TrkA binding free energies obtained from ESMACS predictions and those determined experimentally is shown in Figure 2. The predicted binding free energies from the 1-traj

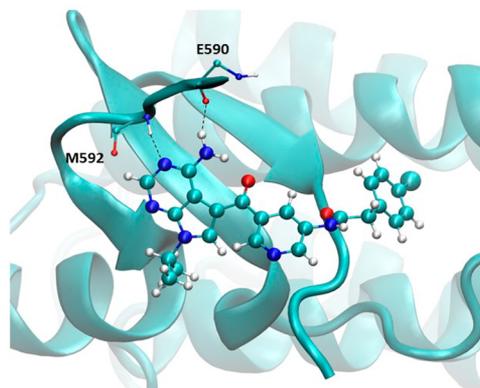


Figure 1. Crystal structure of **1** bound to TrkA, viewed from the N-lobe to the C-lobe of the kinase. Hydrogen bonds are displayed by dashed lines. The protein is shown in cyan cartoon, and ligand atoms are colored by element: hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue. For clarity, the N-lobe is not shown.

approach exhibit a moderate Pearson correlation with the experimental data (with a coefficient of 0.42; Figure 2a). The 2-traj approach, in which the same free energy of the receptor averaged from the 1-traj method is used, significantly improves the correlation between the predictions and experimental data (with a coefficient of 0.76; Figure 2b). Incorporating the free energies of the ligands from their individual simulations in water (the 3-traj approach) generates a correlation similar to that from the 2-traj approach (with a coefficient of 0.79; Figure 2c). The improvements are achieved by including the adaptation energies⁸ of the receptor in the current study, and possibly of the ligands in general cases, which quantify the free energy changes of a molecule between its bound and free states. The 3-traj ESMACS results are in good agreement with the experimental measurements. It should be noted that the ESMACS method is capable of comparing the binding affinities of the entire set of ligands, notwithstanding the relatively significant structural differences. It should also be noted that the experimental TrkA binding free energies were approximated from half-maximal inhibitory concentration (IC_{50}) values, which provide only semiquantitative estimates. In addition, experimental IC_{50} values for individual compounds vary between intralab measurements, and their average IC_{50} values vary between 1.6- and 4.2-fold when measured independently across the two laboratories (Table 1 and Table S1 in the Supporting Information). A root-mean-square (RMS) error of ~24% has recently been inferred from interlaboratory variations of reported binding affinities by Chodera and Mobley.²⁵

Improvement of Docking Predictions by the ESMACS Approach. For the nine compounds that were successfully docked into the TrkA protein, the binding free energies were also calculated using the structures after docking and energy minimization. The minimization process optimizes the geometry of a collection of atoms so that the net interatomic force on each atom is acceptably close to zero. The remaining seven compounds were not included in these calculations because the manually constructed structures exhibit some close contacts between the protein and the compounds, to which the binding free energy estimations are very sensitive. When the compounds could be docked into a single protein structure, the ranking of their binding affinities was reasonably predicted from the docked structures (Figure 3a). However, the results from

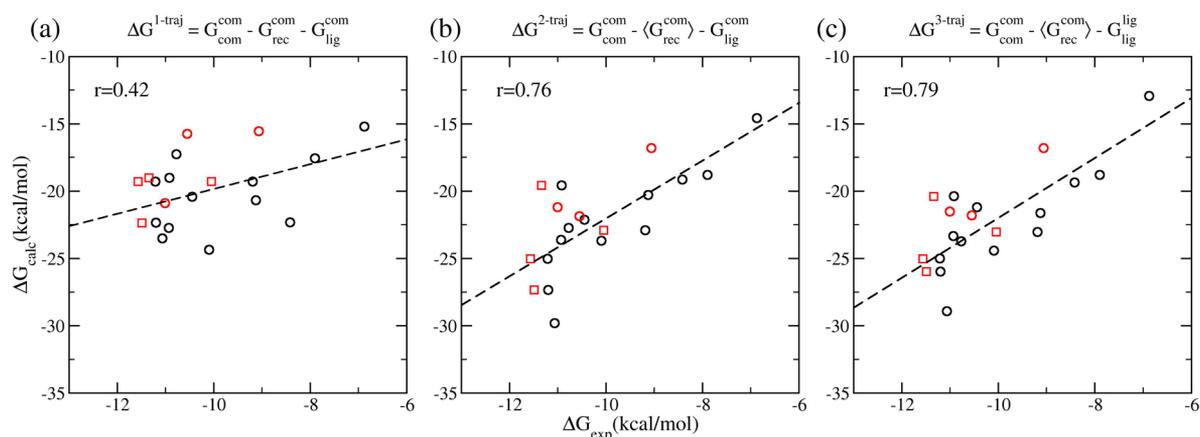


Figure 2. Comparisons of the TrkA experimental data and the calculated binding free energies from (a) 1-traj, (b) 2-traj, and (c) 3-traj ESMACS approaches. The equation in each panel indicates the calculations used in each case. The term $G_{\text{com/rec/lig}}^{\text{com/lig}}$ represents the free energy of the component (subscript: complex, receptor, or ligand) obtained from simulation of the same or different component (superscript: complex or free ligand in aqueous solution). The calculated binding free energies are associated with standard errors of ca. 0.6 kcal/mol in the 1-traj approach and ca. 3 kcal/mol in the 2- and 3-traj approaches (see Table S2), which are not shown in the figures for reasons of clarity. The experimental data from two sites (Pfizer, Sandwich and TCG Lifescience) are displayed in black and red, respectively. The correlation coefficients shown in the figure were calculated using the averages of the calculated binding free energies and the experimental data from Pfizer, Sandwich (black circles) and TCG Lifescience (red circles) where the former are not available. Large uncertainties are associated with the correlation coefficients because of the large error bars of the calculated and experimental binding affinities. Further analyses with bootstrapping resampling (see details in the Supporting Information) generate correlation coefficients of 0.39 ± 0.26 , 0.60 ± 0.26 , and 0.62 ± 0.27 for the 1-, 2-, and 3-traj approaches, respectively. It is evident that the 2- and 3-traj methods improve the ranking provided by the 1-traj method.

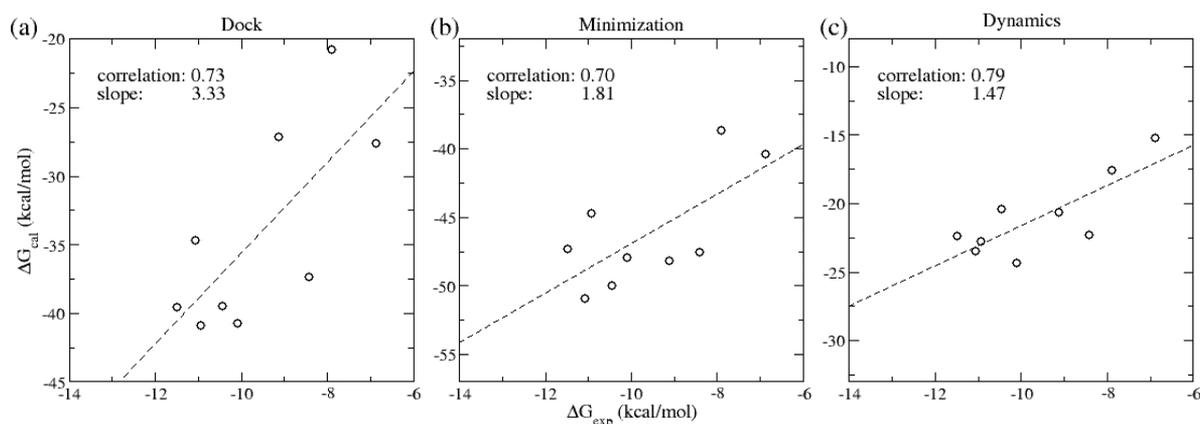


Figure 3. Calculated binding free energies for 1, 3, 4, 6, 7, 8, 13, 16, and 22 (a) from single structures after docking, (b) from 25 structures for each compound after 11 000-step minimization with a sophisticated conjugate-gradient method, and (c) from ensemble averages from three-trajectory ESMACS studies. Compounds 1, 3, 4, 6, 7, 8, 13, 16, and 22 are highly structurally similar (see Table 1).

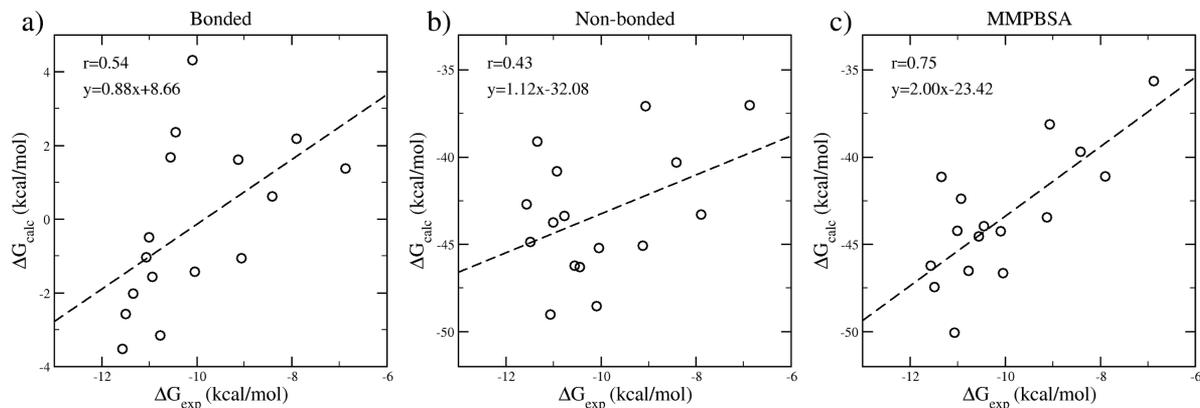


Figure 4. Correlations of free energy components and the experimental data from the 3-traj approach. Both (a) bonded and (b) nonbonded energy terms contribute to the ranking of binding affinities, with similar correlation coefficients between the calculations and experimental data. Their combination, the MMPBSA energy (c), exhibits better correlations with the experimental data than the components themselves.

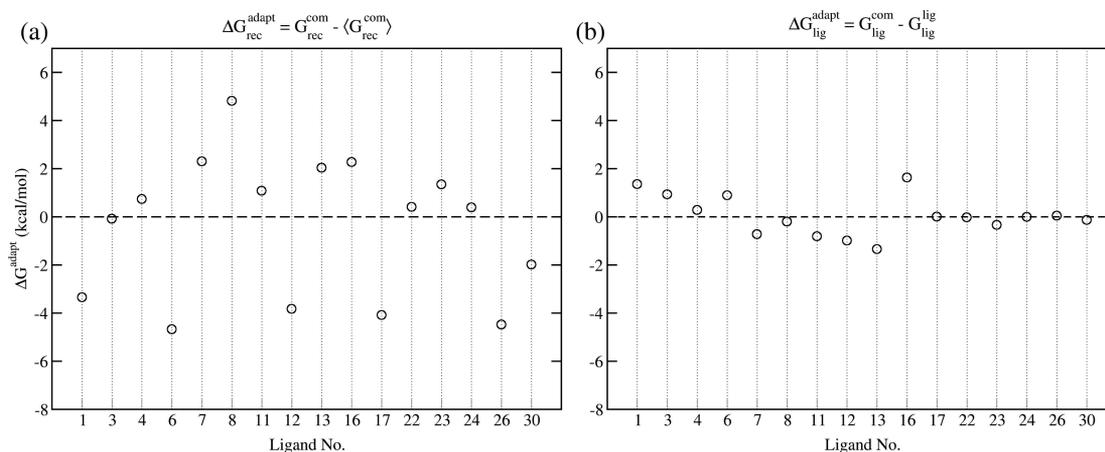


Figure 5. Adaptation free energies of (a) the receptor and (b) the compounds showing the binding free energy changes between the 1- and 3-traj approaches. The terms $G_{\text{rec/lig}}^{\text{com/lig}}$ are the free energies of receptor or ligands (subscript) calculated from simulations performed for the complex or the ligands (superscript).

the docking structures overestimate the free energy differences among the compounds, as shown in Figure 3a, in which the large slope of the regression line (3.33) indicates that the predicted energy differences are on average more than 3 times larger than the experimentally measured ones. The energy minimization significantly decreases the overestimations (Figure 3b, with a regression slope of 1.81). The ESMACS study further decreases the overestimation and makes the points on the scatter plot more concentrated around the regression line (Figure 3c). The docking method ranks the binding affinities reasonably well only for the compounds with closely related structures. Its results are also strongly dependent on the quality and choice of the structure(s) of the target protein.²⁶ The ESMACS approach, however, is not sensitive to the initial structure and achieves a similar correlation coefficient between the calculated and experimental binding affinities for all of the compounds irrespective of whether they can be successfully docked (Figure 2c).

ESMACS Binding Affinities: Contributions from Energy Components. The components of the free energy calculations (see the Supporting Information for more details) may provide insight into the mechanism of compound binding.²⁷ In this study, no apparent correlations could be found for the individual energy components of the internal van der Waals, electrostatic, or electrostatic solvation energies. The total bonded energies, including bond, angle, and dihedral interactions, manifest a moderate correlation with the experimental binding free energies in the 3-traj approach, with a Pearson coefficient of 0.54 (Figure 4a). The nonbonded interactions, including the internal electrostatic, solvation electrostatic, and van der Waals interactions, have a slightly weaker correlation (Pearson coefficient of 0.43) than the bonded energies. The combination of bonded and nonbonded interactions significantly improves the correlation, with a Pearson coefficient of 0.75 (Figure 4c). The bonded and nonbonded energies are the two components of the total adaptation energy (see below), which is associated with the conformational changes upon compound binding. It is therefore not surprising that both of them show correlations with the binding affinities. The configurational entropy components do not exhibit a significant correlation with the experimental binding affinity (Figure S2). They vary in a range of 4.6 kcal/mol; the MMPBSA energies vary in a range of 14.41

kcal/mol by comparison. Inclusion of the contribution from configurational entropy into the MMPBSA energy does not improve the Pearson correlation coefficients significantly (see Figure 2c, which includes configurational entropy, and Figure 4c, which does not).

Adaptation Energy: A Measure of Conformational Change upon Binding. The improvement obtained with the 3-traj approach compared with the 1-traj approach is achieved by relaxing the assumption that the receptor and the compounds sample similar conformations in both the free and bound states. Unfavorable adaptation energies⁸ are usually induced when the conformations within the free state are significantly perturbed upon compound binding. The adaptation energies from the 3-traj approach indeed provide more insights into the mechanism of compound binding. They are related to the structural modifications made to the compounds and the protein and shed important light on the optimization. As the ranking of the ligand binding is the main concern of the study and the relative adaptation energies of the protein are energetically as informative as the absolute ones in the binding affinity comparison, no attempt has been made here to compute accurate absolute adaptation energies for the protein. All of the adaptation energies for the protein are relative ones in the current paper. Non-negligible adaptation energies are associated with conformational changes upon compound binding, meaning that the binding involves “induced fit” recognition. Compound 8, for example, introduces the largest adaptation energy within the protein, while 16 has the largest adaptation energy within the compound (Figure 5). The binding of 8 induces a conformational change within the protein because there is a reluctance to accommodate the methoxy group in the linker region of 8 into the binding site; the arrangement of the amide group in 16 forces the compound to adopt a high-energy conformation in the protein’s binding site (see details below).

Comparison of Experimental TrkA Binding Free Energies with TIES Predictions. The TIES method provides a more accurate approach to estimate relative binding affinities, but the scope for its use is rather tightly circumscribed, as it is applicable only for pairs of compounds that do not present significant structural differences. Our previous studies have demonstrated that TIES offers more quantitative accuracy in its predictions than ESMACS.^{13,28} Here we apply TIES to the

TrkA data set to highlight the accuracy of the approach. In TIES, studies are performed for 14 pairs of TrkA compounds (Figure 6). The calculated binding free energy differences

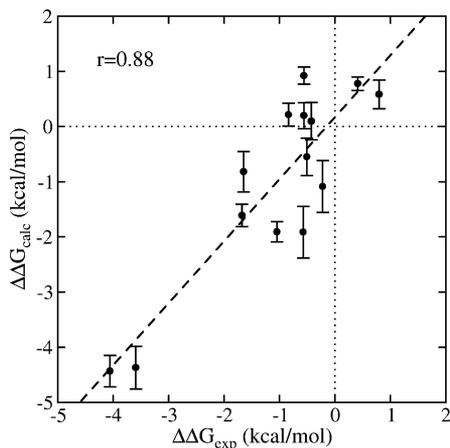


Figure 6. Correlation between TIES-predicted relative binding affinities and experimental data. The black line is the correlation line, while the dotted lines ($x = 0$ and $y = 0$) create four quadrants. Ten out of the 14 data points are in quadrants I ($x > 0$ and $y > 0$) and III ($x < 0$ and $y < 0$), meaning that the calculated binding free energy differences have the same sign as those from the experimental data.

correlate well with the experimental data, with a Pearson correlation of 0.88 (Figure 6), compared with 0.79 from the 3-traj ESMACS study (Figure 2c). A directional agreement is achieved if a prediction has the same sign as the experimental observation; otherwise, the result is deemed to be directional disagreement. For 10 out of the 14 pairs studied, TIES successfully achieves directional agreements, meaning that the results predict the direction of the change in the binding affinity correctly (Figure 6). As shown in the figure, all but one of the TIES predictions that directionally disagree lie on the border of quadrants II and III within their calculated error bars. When the error bars of the experimental $\Delta\Delta G$ are also taken into account, a better agreement might be achieved than that reported here.

DISCUSSION

The calculations, especially the ESMACS ones, can provide further insights for binding affinity variation between structurally similar ligands. In order to analyze the results, it is necessary to consider the differences between the one-, two-, and three-trajectory methods. The 1-traj simulation calculates $\Delta G_{\text{binding}}$ assuming there is no energetic penalty for the adaptation of the protein or ligand to the binding conformation; the 2-traj method takes the change in protein energy into consideration, and the 3-traj method accounts for the changes in both protein and ligand energies. The simulations enable the exploration of key differences in binding modes for structurally related ligands, as we now discuss.

Hinge Binding Group Modifications: Compounds 1, 4, and 22. Initial compounds in the pyrrolopyrimidine series, such as **1**, contain an amino group at the 4-position of the hinge binding group (Table 2). Compounds such these, while TrkA-active, required optimization of their kinase selectivity profile.²⁹ It is known within the scientific literature that minimizing the number of hydrogen-bonding interactions made between a kinase inhibitor and the kinase hinge binding region can lead to an enhanced kinome selectivity profile.^{30,31} In order to design compounds with an optimal kinase selectivity profile, an analysis of the crystal structure of **1** bound to TrkA and the nonliganded TrkA apo structure was undertaken.²⁹ The aminopyrrolopyrimidine motif of **1** makes a two-point hydrogen-bonding interaction with hinge residues E590 and M592. When **1** is superpositioned onto the TrkA apo crystal structure, the $-\text{NH}_2$ motif of the aminopyrrolopyrimidine group overlays directly with a conserved water molecule observed in the apo structure. The $-\text{NH}_2$ group was therefore removed to generate compounds such as **4**, in which a bridging water molecule forms hydrogen bonds with the backbone carbonyl oxygen at E590 and the ketone oxygen of the TrkA ligand. The hydrogen bond between the ligand and the backbone N–H of hinge residue M592 is maintained. The deletion of the $-\text{NH}_2$ group had only a minimal effect on the TrkA activity, as **1** and **4** had broadly similar IC_{50} values in the TrkA pharmacology assay.

ESMACS calculations show that compounds without the $-\text{NH}_2$ group at the 4-position (all compounds except **1**, **3**, **6**, and **22**; Table 2) have a bridging water molecule between the

Table 2. Structures and Properties of 1, 4, and 22^a

Compound	1	4	22
IC_{50} (nM)	4.25/6.97	10.86	1756.36
ΔG_{exp} (kcal/mol)	-11.49/-11.20	-10.93	-7.91
$\Delta\Delta G_{\text{exp}}$ (kcal/mol)	0.00	0.56/0.27	3.58/3.29
$\Delta\Delta G_{\text{ESMACS}}$ (kcal/mol)	0.00	2.63	7.18
$\Delta\Delta G_{\text{TIES}}$ (kcal/mol)	0.00	-0.20	4.37

^aTrkA free energy activities derived from experimental IC_{50} values are denoted as ΔG_{exp} . The relative binding free energies from experiment and ESMACS and TIES calculations are denoted as $\Delta\Delta G_{\text{exp}}$, $\Delta\Delta G_{\text{ESMACS}}$, and $\Delta\Delta G_{\text{TIES}}$, respectively. Experimental data from Pfizer, Sandwich (U.K.) are shown in black, and those from TCG Lifescience (India) are shown in blue.

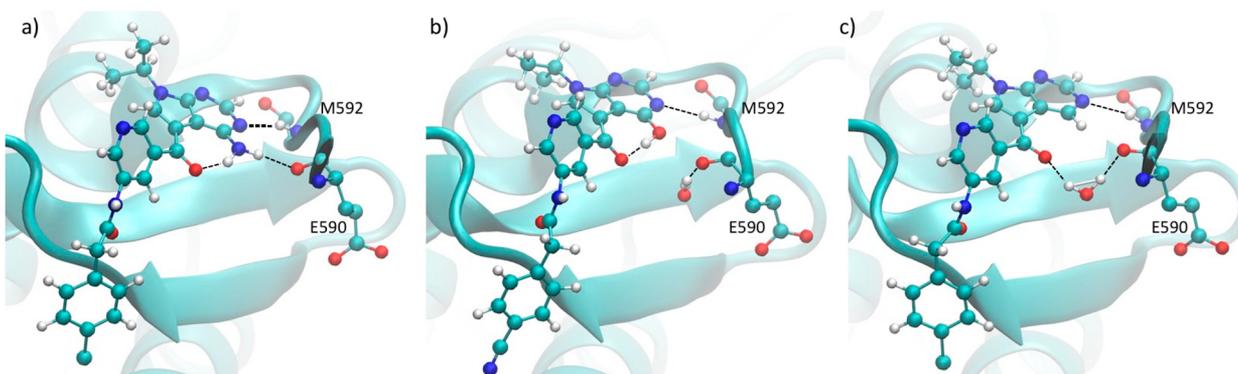


Figure 7. ESMACS simulations of (a) **1**, (b) **22**, and (c) **4** bound to TrkA. Final conformations of 4 ns production runs from all 25 replicas are overlapped and smoothed by averaging over 10 frames. Key residues E590 and M592 (for clarity, side-chain atoms are not shown) at the TrkA hinge region are highlighted. Hydrogen bonds are shown as dashed lines. The protein is shown in cyan cartoon, and ligand atoms are colored by element: hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue. For reasons of clarity, the N-lobe of the protein has been removed.

Table 3. Structures and Properties of 12, 17, 23, and 24^a

Compound	12	17	23	24
IC ₅₀ (nM)	14.18	3.73/6.84	9.55	249.42
ΔG_{exp} (kcal/mol)	-10.77	-11.57/-11.21	-11.01	-9.06
$\Delta\Delta G_{\text{exp}}$ (kcal/mol)	0.80/0.44	0.00	0.56/0.20	2.51/2.15
$\Delta\Delta G_{\text{ESMACS}}$ (kcal/mol)	1.29	0.00	3.49	8.21
$\Delta\Delta G_{\text{TIES}}$ (kcal/mol)	0.58	0.00	-0.92	-

^aTrkA free energy activities derived from experimental IC₅₀ values are denoted as ΔG_{exp} . The relative binding free energies from experiment and ESMACS and TIES calculations are denoted as $\Delta\Delta G_{\text{exp}}$, $\Delta\Delta G_{\text{ESMACS}}$, and $\Delta\Delta G_{\text{TIES}}$, respectively. Experimental data from Pfizer, Sandwich (U.K.) are shown in black, and those from TCG Lifescience (India) are shown in blue.

ketone oxygen of the ligand and the backbone carbonyl oxygen of hinge residue E590, with an average occupancy of $41 \pm 9\%$. The occupancy indeed underscores the presence of a water molecule at the location as the molecule dynamically moves into and out of the cutoff distance and angle that designate the occurrence of a bridging water molecule. Simulations of compounds containing the $-\text{NH}_2$ group at the 4-position (**1**, **3**, and **6**; Table 2) show the same bridging water to be present but at a far lower frequency (occupancy values of just $2 \pm 0\%$ for **1**, **3**, and **6**) (Figure 7).

ESMACS confirms that the bridging water molecule highlighted in Figure 7 helps mediate and thus satisfy key protein–compound interactions, rendering the overall binding free energies of compounds such as **4** less affected by the absence of the H-bond when $-\text{NH}_2$ is not present. This capability is reflected in the close similarity between the calculated TrkA free energy values for **1** and **4** and those derived from TrkA pharmacology experiments (Table 2). Compound **22** was synthesized as part of an effort to characterize putative compound metabolites. As highlighted in Table 2, **22** is only weakly active in the TrkA pharmacology (IC₅₀ = 1756 nM). ESMACS simulations show a low average occupancy of the aforementioned bridging water molecule (average occupancy $10 \pm 6\%$) due to the presence of the $-\text{OH}$

group at the 4-position (Figure 7). The calculations also show that the $-\text{OH}$ group of **22** is more likely to form an intramolecular H-bond with the carbonyl group of the ligand than to form an intermolecular H-bond with hinge residue E590 of the protein. Within the 25-member ensemble, only one ESMACS replica has an intermolecular H-bond forming with an occupancy of 84%, compared with $<5\%$ for all of the others. The $-\text{OH}$ group occupies the space of the bridging water molecule and does not have a direct hydrogen-bonding interaction with E590. The intramolecular H-bond of the $-\text{OH}$ group places the $-\text{OH}$ oxygen in a position where an electrostatic repulsion arises with the backbone carbonyl oxygen of E590 (Figure 7b). The result of having a nonsatisfied hydrogen-bonding group ($-\text{OH}$) in **22** is a reduction in overall ligand binding affinity. This is reflected not only in the measured activity data ($\Delta G_{\text{exp}} = -7.1$ kcal/mol) but also in the relative binding affinities calculated by ESMACS and TIES (Table 2).

Hinge Binding Group Modifications: Compounds 12, 17, 23, and 24. The TrkA cocrystal structure of compounds such as **12** indicated that a polar atom capable of making an additional hydrogen-bonding interaction with M592 might be accommodated at the 2-position of the hinge binding group (Table 3). Compound **17** was synthesized initially to test this

theory and proved to be more active in the TrkA pharmacology model than the parent molecule **12**. Compound **23** was then synthesized to see whether adding a methyl group to the amine would further boost the potency through the introduction of hydrophobic interactions with the lipophilic side chain of Y591. Compound **23** was potent in the TrkA assay, albeit slightly less so than **12**. Compound **24** was synthesized as a putative metabolite of **12** and was found to be >15-fold less active at TrkA.

ESMACS shows that addition of polar groups at the 2-position increases the total occupancy of hydrogen bonds with M592 from $23 \pm 4\%$ (for **12**) to $85 \pm 4\%$, $80 \pm 4\%$, and $91 \pm 19\%$ (for **17**, **23**, and **24**, respectively). This includes the occupancy of the hydrogen bond made between the pyrimidyl nitrogen and the $-\text{NH}$ group of M592, which is common to all of the compounds. The increased occupancies of **17**, **23**, and **24** result in the additional hydrogen bonds between the polar groups at the 2-position of the compounds and the $\text{C}=\text{O}$ group of M592. The simulations also show that addition of an $-\text{NH}_2$ group (**17**) does not introduce any unfavorable steric interaction within the protein or the ligand (see Figure 8). The

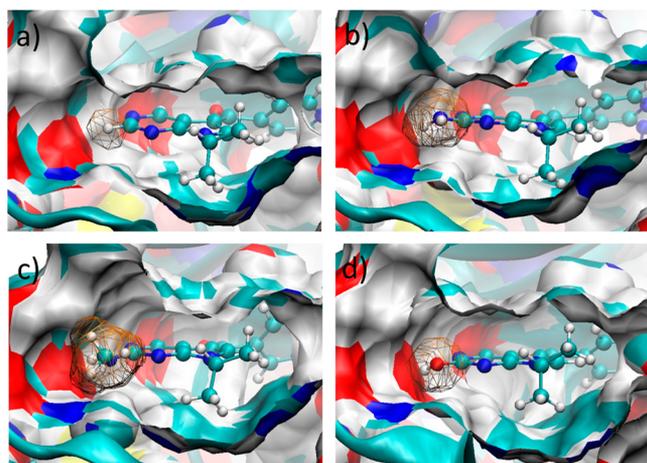


Figure 8. ESMACS simulations of (a) **12**, (b) **17**, (c) **23**, and (d) **24** bound to TrkA. Representative conformations are displayed using the final conformations of 4 ns production runs of one replica from each 25-member ensemble. The protein is shown by the surface representation and the ligand by the ball-and-stick representation with hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue. The surfaces of the R2-position groups are shown in the wireframe representation.

slightly improved binding affinity of **17** versus **12** stems from the favorable electrostatic interaction (including the solvation electrostatic component from the Poisson–Boltzmann calculation) between the compound and the protein. The addition of the methyl group on the amine (**23**), however, induces an unfavorable adaptation energy within the protein (see the comparison of relative adaptation energies for **12/17** and **23** in Figure 5a) due to steric hindrance between the protein and **23** (Figure 8c). The presence of an $-\text{NH}_2$ group at the 2-position (**17**) does not induce an unfavorable adaptation energy. The introduction of an $-\text{OH}$ group (**24**), however, introduces a significant adaptation energy in the protein because the $-\text{OH}$ group forms the strongest hydrogen bond with M592 among the subgroup of compounds **12**, **17**, **23**, and **24**, which induces steric hindrance within the complex and reduces the overall binding affinity (Figure 8).

Linker Group Modifications: Compounds **1**, **4**, and **22**.

The TrkA cocrystal structure of compounds such as **1** highlighted a conserved water molecule that formed hydrogen-bonding interactions with the $\text{N}-\text{H}$ group of the amide linker and protein residues K544 and E560 (Table 4 and Figure 9). A methyl group was added to the amide nitrogen to assess the effect of displacing the conserved water molecule (and the potentially altered amide conformation relative to the pyridyl ring) on the TrkA activity. The effect of transposing the amide $\text{N}-\text{H}$ to the other side of the carbonyl group was also investigated, as this, if the activity were retained, could open up new parallel chemistry opportunities within the chemotype. Compound **3**, an example of the *N*-methylamide, remained active at TrkA, although with reduced affinity versus its desmethyl congener **1**. Compound **6**, an example of the transposed $\text{N}-\text{H}$ amide, also remained active at TrkA with almost identical affinity as **1**.

ESMACS shows that the modifications highlighted in Table 4 do not alter the key hydrogen-bonding interaction between the carbonyl oxygen of the amide group of **1**, **3**, and **6** and the backbone $\text{N}-\text{H}$ of D668, with occupancies of $56 \pm 7\%$, $54 \pm 3\%$, and $51 \pm 6\%$, respectively (Figure 9). The altered position of the amide group in **1** and **6** slightly affects the appearance of the bridging water molecules between the ligands and residue E560, with occupancies of $15 \pm 6\%$ and $20 \pm 5\%$ for **1** and **6**, respectively. Minimal direct hydrogen-bonding interactions are predicted between the amide $-\text{NH}$ group of **1** and **6** and residue K544 or E560. The replacement of the amide $-\text{NH}$ group of **1** by the more hydrophobic *N*-methyl group is predicted to displace the adjacent bridging water molecule (Figure 9a,b) and overall make binding to TrkA energetically less favorable. There is a good correlation between the ESMACS-predicted and experimentally derived free energies of TrkA binding for **1**, **3**, and **6** (Table 4).

Linker Group Modifications: Compounds **7** and **8**.

Although the linker group of the pyrrolopyrimidine/pyrrolopyridine series binds to a relatively narrow region of the ligand binding site, between the two adjacent ATP and DFG pockets, the tolerance of the TrkA protein toward bulkier linker groups was briefly investigated as part of the TrkA program. An example from this work is **8**, in which a methoxy group has been added to the linker-group carbonyl. As highlighted in Table 5, a significant reduction in TrkA activity is observed for **8** versus **7**.

ESMACS calculations indicate that the oxygen atom of the methoxy group of **8** is located proximal ($3.47 \pm 0.31 \text{ \AA}$) to the carbonyl group of residue D668. The occupancy of a bridging water molecule between the methoxy group and the carbonyl is only $5 \pm 2\%$, indicating that the orientation and the space between these groups preclude a bridging water molecule being available to quench any electrostatic repulsion (Table 5 and Figure 10). The chirality of the methoxy group also prevents the polar oxygen from having any direct or favorable water-bridged interactions with E560 or D668. The binding of **8** induces the largest adaptation energy (Figure 5a) in the TrkA protein of all compounds assessed. Energetic analyses show that although the bonded energy and the van der Waals interactions are favorable for **8** compared with **7** (by 3.70 and 2.29 kcal/mol, respectively), a significantly unfavorable electrostatic energy (10.59 kcal/mol), mainly from the interactions between the polar oxygen of the methoxy group and the polar/charged groups of the protein (carbonyl group of D668 and carboxylate

Table 4. Structures and Properties of 1, 3, and 6^a

Compound	1	3	6
IC ₅₀ (nM)	4.25/6.97	24.53	8.67
ΔG _{exp} (kcal/mol)	-11.49/-11.20	-10.45	-11.07
ΔΔG _{exp} (kcal/mol)	0.00	1.04/0.75	0.42/0.13
ΔΔG _{ESMACS} (kcal/mol)	0.00	4.78	-2.94
ΔΔG _{TIES} (kcal/mol)	0.00	1.91	-0.10

^aTrkA free energy activities derived from experimental IC₅₀ values are denoted as ΔG_{exp}. The relative binding free energies from experiment and ESMACS and TIES calculations are denoted as ΔΔG_{exp}, ΔΔG_{ESMACS}, and ΔΔG_{TIES}, respectively. Experimental data from Pfizer, Sandwich (U.K.) are shown in black, and those from TCG Lifescience (India) are shown in blue.

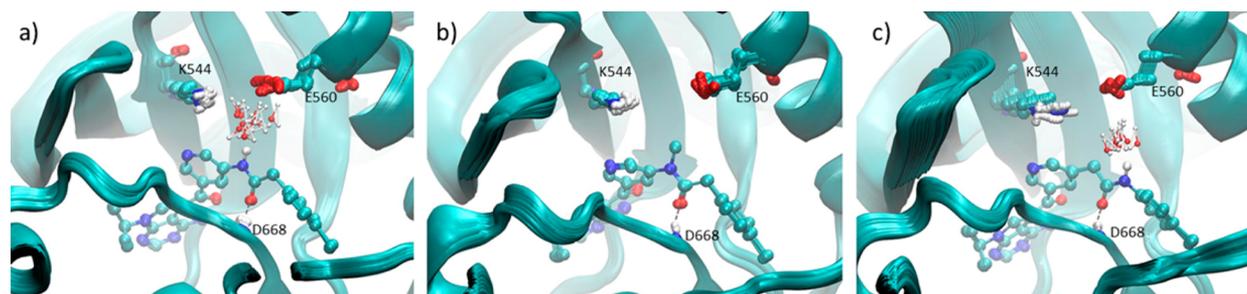


Figure 9. ESMACS simulations of (a) 1, (b) 3, and (c) 6 bound to TrkA. Final conformations of 4 ns production runs from all 25 replicas are overlapped and smoothed by averaging over 10 frames. The key TrkA protein residues D668, K544, and E560 are highlighted. The protein is shown in cyan cartoon, and ligand atoms are colored by element: hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue. Hydrogen bonds between the linker group carbonyl oxygen of 1, 3, and 6 and the backbone –NH group of D668 are shown as dashed lines.

Table 5. Structures and Properties of 7 and 8^a

Compound	7	8
IC ₅₀ (nM)	44.25	735.79
ΔG _{exp} (kcal/mol)	-10.09	-8.42
ΔΔG _{exp} (kcal/mol)	0.00	1.67
ΔΔG _{ESMACS} (kcal/mol)	0.00	5.06
ΔΔG _{TIES} (kcal/mol)	0.00	1.61

^aTrkA free energy activities derived from experimental IC₅₀ values are denoted as ΔG_{exp}. The relative binding free energies from experiment and ESMACS and TIES calculations are denoted as ΔΔG_{exp}, ΔΔG_{ESMACS}, and ΔΔG_{TIES}, respectively. Experimental data were obtained from Pfizer, Sandwich (U.K.) for these two compounds.

group of E560; Figure 10), is introduced by the inclusion of the (S)-OMe group.

The calculated free energies of TrkA binding for 7 and 8 correlate well with those obtained experimentally. Hence,

utilization of ESMACS to predict the TrkA binding affinity of 8 during the Pfizer TrkA project might well have influenced the medicinal chemistry team to deprioritize this compound and redirect synthetic efforts in other directions.

Linker Group Modifications: Compounds 4 and 16.

The hydrogen bond between the linker-group carbonyl and the backbone N–H of D668 is a key interaction across the pyrrolopyrimidine/pyrrolopyridine series. During the Pfizer TrkA program, reversing the amide group was modeled to assess whether the usual ligand binding mode could be adapted to maintain this interaction in the new amide arrangement. Docking of 16 into TrkA using a Pfizer docking protocol (data not shown) suggested that the reverse amide would likely not be able to adapt a conformation in which this interaction would be retained. However, 16 was synthesized as an example of a reverse amide system to challenge and/or verify the docking result. As can be seen in Table 6, reversing the amide group led to a significant reduction in TrkA activity.

ESMACS-based component binding analysis shows that 16 has the least favorable bonding and nonbonding interactions of the compounds in Table 1. The reversed positions of the –NH and C=O groups prevent formation of the hydrogen bond between the ligand and residue D668 (Figure 11), with an H-bond occupancy of <1% compared with occupancies of 30–56% for the rest of the compound set (except for 26, whose

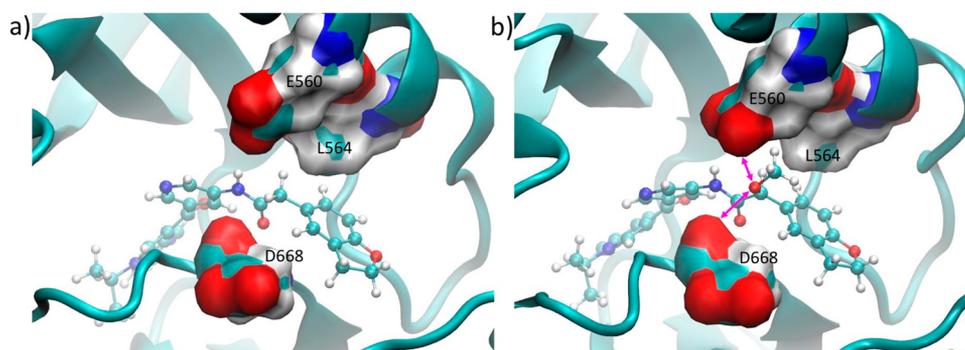


Figure 10. ESMACS simulations of (a) 7 and (b) 8 bound to TrkA. Final conformations of 4 ns production runs from all 25 replicas are overlapped and smoothed by averaging over 10 frames. The key TrkA protein residues D668, E560, and L564 are highlighted. The protein is shown in cyan cartoon, and ligand atoms are colored by element: hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue. The unfavorable electrostatic interactions between the methoxy group of 8 and the carbonyl group of D668 and between the methoxy group and the carboxylate group of E560 are highlighted by pink arrows.

Table 6. Structures and Properties of 4 and 16^a

Compound	4	16
IC ₅₀ (nM)	10.86	9800.00
ΔG_{exp} (kcal/mol)	-10.93	-6.87
$\Delta\Delta G_{\text{exp}}$ (kcal/mol)	0.00	4.06
$\Delta\Delta G_{\text{ESMACS}}$ (kcal/mol)	0.00	10.42
$\Delta\Delta G_{\text{TIES}}$ (kcal/mol)	0.00	4.43

^aTrkA free energy activities derived from experimental IC₅₀ values are denoted as ΔG_{exp} . The relative binding free energies from experiment and ESMACS and TIES calculations are denoted as $\Delta\Delta G_{\text{exp}}$, $\Delta\Delta G_{\text{ESMACS}}$, and $\Delta\Delta G_{\text{TIES}}$, respectively. Experimental data were obtained from Pfizer, Sandwich (U.K.) for these two compounds.

occupancy is 14% as a result of the competition from an extra hydrogen bond between the polar oxygen at R4 and residue K544 (Table 1). Indeed, ESMACS calculations reveal that the -NH and C=O groups of 16 align in a noncomplementary fashion with the -NH and C=O groups of residue D668, introducing a large unfavorable electrostatic interaction (Figure

11). Compound 16 has the largest ligand adaptation energy, indicating that the switching of N-H and C=O does not enable an alternate binding mode to be adopted, and therefore, 16 must adopt a high-energy conformation to complex with the protein. As is clear in Table 6, ESMACS correctly predicts the reduced TrkA potency of 16 compared with 4.

CONCLUSIONS

Methodologies that can predict the binding affinities of molecules ahead of synthesis represent a key area of interest in the pharmaceutical industry. The ability to effectively prioritize prospective compounds on the basis of their likely activity in a key pharmacology assay offers the potential to rapidly improve lead optimization timelines in drug discovery programs by reducing the number of “design–synthesis–screen” cycles needed to identify candidate-quality molecules suitable for clinical testing.³²

Despite the relative diversity of the compounds in this study, the 3-trajectory version of the ESMACS approach provides good agreement between the theoretically predicted binding free energies and those derived from experimentally measured activities for the entire data set. TIES generates even better agreement between the calculated and experimental binding free energy differences for a selected subgroup of ligands. This is manifest in the better correlation coefficient obtained in TIES (0.88) than in ESMACS (0.79) and accurate binding free energy differences from the former. This is the case for all of

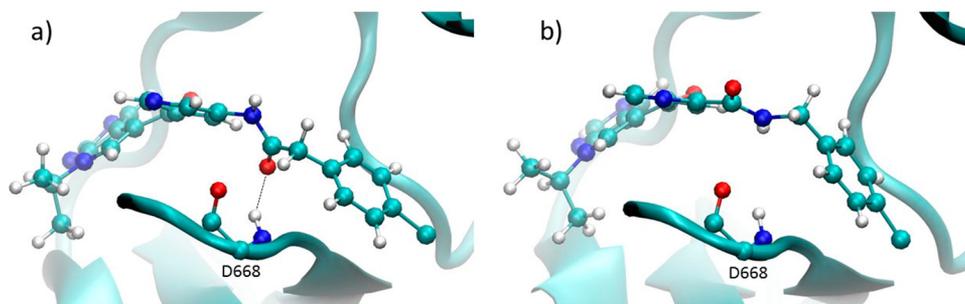


Figure 11. ESMACS simulations of (a) 4 and (b) 16 bound to TrkA. Final conformations of 4 ns production runs from all 25 replicas are overlapped and smoothed by averaging over 10 conformations. The key TrkA protein residue D668 is highlighted. The protein is shown in cyan cartoon, and ligand atoms are colored by element: hydrogen in white, carbon in cyan, oxygen in red, and nitrogen in blue. The hydrogen bond between the linker-group carbonyl oxygen of 4 and the backbone -NH group of D668 is shown as a dashed line. As can be seen in (b), switching of the N-H and C=O groups prevents the formation of a hydrogen bond between the carbonyl oxygen of 16 and the backbone -NH group of D668.

molecular systems we have studied.^{13,28,33} ESMACS also provides structural and energetic insight into the binding of individual ligand–protein systems: some compounds bind by a “lock and key” recognition mechanism, for which no significant conformational adjustment is required by the protein and the ligands; others bind according to an “induced fit” recognition mechanism involving significant conformational changes associated with non-negligible adaptation energies.

The Binding Affinity Calculator (BAC) used in this study automates the workflows for ESMACS and TIES calculations, making the use of these approaches easy and user-friendly. With powerful computing resources now widely available, these robust approaches should become more routine for industrial groups in the field of structure-based drug design. The results described herein, based on an analysis of TrkA ligands synthesized as part of the TrkA program at Pfizer, suggest that the binding affinity calculations have the potential to be successfully applied in real-time prospective ligand design.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.6b00780.

Detailed description of the energy decomposition, the energy convergence, and the error analyses, experimental measurements of TrkA inhibitory activity, and the predicted binding free energies (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: p.v.coveney@ucl.ac.uk.

ORCID

Peter V. Coveney: 0000-0002-8787-7256

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge the support from the EPSRC via the 2020 Science Programme (<http://www.2020science.net/>), EP/I017909/1), the EU H2020 Projects ComPat (<http://www.compat-project.eu/>, 671564) and CompBioMed (<http://www.compbioimed.eu/>, 675451), the Qatar National Research Fund (7-1083-1-191), the MRC Medical Bioinformatics Project (MR/L016311/1), and the UCL Provost. We are grateful to the Hartree Centre for access to the BlueWonder2 computer and for the help of its scientific support staff. We acknowledge the Leibniz Supercomputing Centre for providing access to SuperMUC and the assistance of its scientific support staff. We also made use of ARCHER, the U.K.'s National High Performance Computing Service, funded by the Office of Science and Technology through the EPSRC's High-End Computing Programme. Access to ARCHER was provided through the 2020 Science Programme. This research was also partially supported by the PLGrid Infrastructure, through which access to Prometheus, the Polish supercomputer run by ACK Cyfronet AGH in Krakow, was provided. A.P.B. is supported by an Overseas Research Scholarship from UCL and an Inlaks Scholarship from the Inlaks Shivdasani Foundation.

■ REFERENCES

(1) Sherborne, B.; Shanmugasundaram, V.; Cheng, A. C.; Christ, C. D.; Desjarlais, R. L.; Duca, J. S.; Lewis, R. A.; Loughney, D. A.; Manas,

E. S.; McGaughey, G. B.; Peishoff, C. E.; van Vlijmen, H. Collaborating to Improve the Use of Free-Energy and Other Quantitative Methods in Drug Discovery. *J. Comput.-Aided Mol. Des.* **2016**, *30*, 1139–1141.

(2) Skerratt, S. E.; Andrews, M.; Bagal, S. K.; Bilsland, J.; Brown, D.; Bungay, P. J.; Cole, S.; Gibson, K. R.; Jones, R.; Morao, I.; Nedderman, A.; Omoto, K.; Robinson, C.; Ryckmans, T.; Skinner, K.; Stuppel, P.; Waldron, G. The Discovery of a Potent, Selective, and Peripherally Restricted Pan-Trk Inhibitor (PF-06273340) for the Treatment of Pain. *J. Med. Chem.* **2016**, *59*, 10084–10099.

(3) Sadiq, S. K.; Wright, D.; Watson, S. J.; Zasada, S. J.; Stoica, I.; Coveney, P. V. Automated Molecular Simulation Based Binding Affinity Calculator for Ligand-Bound HIV-1 Proteases. *J. Chem. Inf. Model.* **2008**, *48*, 1909–1919.

(4) Groen, D.; Bhati, A. P.; Suter, J.; Hetherington, J.; Zasada, S. J.; Coveney, P. V. FabSim: Facilitating Computational Research through Automation on Large-Scale and Distributed E-Infrastructures. *Comput. Phys. Commun.* **2016**, *207*, 375–385.

(5) Norman, B. H.; McDermott, J. S. Targeting the Nerve Growth Factor (NGF) Pathway in Drug Discovery. Potential Applications to New Therapies for Chronic Pain. *J. Med. Chem.* **2017**, *60*, 66–88.

(6) Hefti, F. F.; Rosenthal, A.; Walicic, P. A.; Wyatt, S.; Vergara, G.; Shelton, D. L.; Davies, A. M. Novel Class of Pain Drugs Based on Antagonism of NGF. *Trends Pharmacol. Sci.* **2006**, *27*, 85–91.

(7) Mantyh, P. W.; Koltzenburg, M.; Mendell, L. M.; Tive, L.; Shelton, D. L. Antagonism of Nerve Growth Factor-TrkA Signaling and the Relief of Pain. *Anesthesiology* **2011**, *115*, 189–204.

(8) Wan, S.; Knapp, B.; Wright, D. W.; Deane, C. M.; Coveney, P. V. Rapid, Precise, and Reproducible Prediction of Peptide-MHC Binding Affinities from Molecular Dynamics That Correlate Well with Experiment. *J. Chem. Theory Comput.* **2015**, *11*, 3346–3356.

(9) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33*, 889–897.

(10) Coveney, P. V.; Wan, S. On the Calculation of Equilibrium Thermodynamic Properties from Molecular Dynamics. *Phys. Chem. Chem. Phys.* **2016**, *18*, 30236–30240.

(11) Miller, B. R., 3rd; McGee, T. D., Jr.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. Mmpbsa.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321.

(12) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

(13) Bhati, A. P.; Wan, S.; Wright, D. W.; Coveney, P. V. Rapid, Accurate, Precise, and Reliable Relative Free Energy Prediction Using Ensemble Based Thermodynamic Integration. *J. Chem. Theory Comput.* **2017**, *13*, 210–222.

(14) Wright, D. W.; Hall, B. A.; Kenway, O. A.; Jha, S.; Coveney, P. V. Computing Clinically Relevant Binding Free Energies of HIV-1 Protease Inhibitors. *J. Chem. Theory Comput.* **2014**, *10*, 1228–1241.

(15) Swanson, J. M.; Henchman, R. H.; McCammon, J. A. Revisiting Free Energy Calculations: A Theoretical Connection to MM/PBSA and Direct Calculation of the Association Free Energy. *Biophys. J.* **2004**, *86*, 67–74.

(16) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. Alchemical Free Energy Methods for Drug Discovery: Progress and Challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150–160.

(17) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a

Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* **2015**, *137*, 2695–2703.

(18) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. J.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2004.

(19) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(20) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber FF99SB Protein Force Field. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1950–1958.

(21) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(22) Highfield, R. Supercomputer bid to create the first truly personalised medicine. <https://blog.sciencemuseum.org.uk/supercomputer-bid-to-create-the-first-truly-personalised-medicine/> (accessed Feb 22, 2017).

(23) Fiser, A.; Sali, A. Modloop: Automated Modeling of Loops in Protein Structures. *Bioinformatics* **2003**, *19*, 2500–2501.

(24) Allen, W. J.; Balias, T. E.; Mukherjee, S.; Brozell, S. R.; Moustakas, D. T.; Lang, P. T.; Case, D. A.; Kuntz, I. D.; Rizzo, R. C. Dock 6: Impact of New Features and Current Docking Performance. *J. Comput. Chem.* **2015**, *36*, 1132–1156.

(25) Chodera, J. D.; Mobley, D. L. Entropy-Enthalpy Compensation: Role and Ramifications in Biomolecular Ligand Recognition and Design. *Annu. Rev. Biophys.* **2013**, *42*, 121–142.

(26) Bissantz, C.; Folkers, G.; Rognan, D. Protein-Based Virtual Screening of Chemical Databases. 1. Evaluation of Different Docking/Scoring Combinations. *J. Med. Chem.* **2000**, *43*, 4759–4767.

(27) Gohlke, H.; Kiel, C.; Case, D. A. Insights into Protein-Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-RaiGDS Complexes. *J. Mol. Biol.* **2003**, *330*, 891–913.

(28) Bunney, T. D.; Wan, S.; Thiyagarajan, N.; Sutto, L.; Williams, S. V.; Ashford, P.; Koss, H.; Knowles, M. A.; Gervasio, F. L.; Coveney, P. V.; Katan, M. The Effect of Mutations on Drug Sensitivity and Kinase Activity of Fibroblast Growth Factor Receptors: A Combined Experimental and Theoretical Study. *EBioMedicine* **2015**, *2*, 194–204.

(29) Andrews, M. D.; Bagal, S. K.; Gibson, K. R.; Omoto, K.; Ryckmans, T.; Skerratt, S. E.; Stuppel, P. A. Pyrrolo[2,3-*d*]pyrimidine Derivatives as Inhibitors of Tropomyosin-Related Kinases and Their Preparation and Use in the Treatment of Pain. WO2012137089A1, 2012.

(30) Xing, L.; Klug-Mcleod, J.; Rai, B.; Lunney, E. A. Kinase Hinge Binding Scaffolds and Their Hydrogen Bond Patterns. *Bioorg. Med. Chem.* **2015**, *23*, 6520–6527.

(31) Morphy, R. Selectively Nonselective Kinase Inhibition: Striking the Right Balance. *J. Med. Chem.* **2010**, *53*, 1413–1437.

(32) Ou-Yang, S. S.; Lu, J. Y.; Kong, X. Q.; Liang, Z. J.; Luo, C.; Jiang, H. Computational Drug Discovery. *Acta Pharmacol. Sin.* **2012**, *33*, 1131–1140.

(33) Wan, S.; Bhati, A. P.; Zasada, S. J.; Wall, I.; Green, D.; Bamborough, P.; Coveney, P. V. Rapid and Reliable Binding Affinity Prediction of Bromodomain Inhibitors: A Computational Study. *J. Chem. Theory Comput.* **2017**, *13*, 784–795.

Rapid and Reliable Binding Affinity Prediction of Bromodomain Inhibitors: A Computational Study

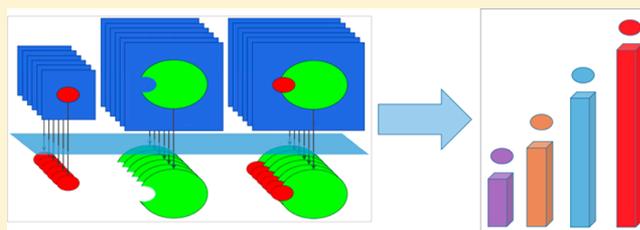
Shunzhou Wan,[†] Agastya P. Bhati,[†] Stefan J. Zasada,[†] Ian Wall,[‡] Darren Green,[‡] Paul Bamborough,[‡] and Peter V. Coveney^{*,†,‡}

[†]Centre for Computational Science, Department of Chemistry, University College London, London WC1H 0AJ, United Kingdom

[‡]GlaxoSmithKline, Gunnels Wood Road, Stevenage, Hertfordshire, SG1 2NY, United Kingdom

S Supporting Information

ABSTRACT: Binding free energies of bromodomain inhibitors are calculated with recently formulated approaches, namely ESMACS (enhanced sampling of molecular dynamics with approximation of continuum solvent) and TIES (thermodynamic integration with enhanced sampling). A set of compounds is provided by GlaxoSmithKline, which represents a range of chemical functionality and binding affinities. The predicted binding free energies exhibit a good Spearman correlation of 0.78 with the experimental data from the 3-trajectory ESMACS, and an excellent correlation of 0.92 from the TIES approach where applicable. Given access to suitable high end computing resources and a high degree of automation, we can compute individual binding affinities in a few hours with precisions no greater than 0.2 kcal/mol for TIES, and no larger than 0.34 and 1.71 kcal/mol for the 1- and 3-trajectory ESMACS approaches.



1. INTRODUCTION

Computational chemistry has been an established tool in drug discovery for a number of years. The number of crystal structures available in the public domain and within pharmaceutical companies on which to base computational studies continues to rise rapidly. Despite the increase in resources applied and in experimental data on which to base the studies, industrial structure-based design approaches have evolved very little in recent times.¹ In particular, the approaches are largely qualitative and largely dependent on the experience and knowledge of the practitioner.^{2,3} Attempts to quantify protein–ligand binding affinities are rare. Expert practitioners have little confidence in existing tools to make robust predictions and certainly not to do so on a time scale that can substantially impact medicinal chemistry programmes.

Recently, there has been a renewed interest in the use of free energy calculations in drug discovery programmes. In particular, the FEP+ implementation of Free Energy Perturbation (FEP) has shown potential to improve the ability to predict protein–ligand binding affinities on an industrially relevant time scale.⁴ Research is ongoing to understand how broadly applicable the method is, and how accurate its predictions are when applied to active drug discovery programmes.

Although FEP+ applies replica exchange solute tempering (REST) in which exchange moves are made between different λ windows, its predictions, like those from many other approaches, are generated from a single output for each pair of mutations. Advances in high-end computing capabilities offer the opportunity to run vast numbers of calculations in parallel. The application of these computational capabilities to free

energy calculations allows results to be returned quickly and multiple replicas of simulations⁵ to be run, leading to tighter control of standard errors. If such an approach could be validated and implemented in an industrial setting it would represent a major step forward in structure-based design capabilities. The first step in this process is to validate the performance on an industrially relevant data set.

Depending on the reliability, rapidity, and throughput of these calculations, they might find application at various stages of the drug discovery and development process across the wider pharmaceutical industry. As these methods require significant compute resources beyond existing in-house industrial capacity, assessment (and any subsequent adoption) of the methodology requires access to high performance computing resources.

Research into epigenetic proteins is currently a major and rapidly evolving focus for the pharmaceutical industry.^{6–8} Bromodomain-containing proteins, and in particular the four members of the BET (bromodomain and extra terminal domain) family, each containing two bromodomains, have been widely studied. Small molecule inhibitors able to competitively antagonize the binding of acetylated histone tails to these modules have been shown to exert profound effects on gene expression and have shown promising preclinical efficacy in pathologies ranging from cancer to inflammation. Indeed, several compounds are progressing through early stage clinical trials and are showing exciting early results.⁹ Most inhibitors reported show similar binding

Received: August 11, 2016

Published: December 22, 2016

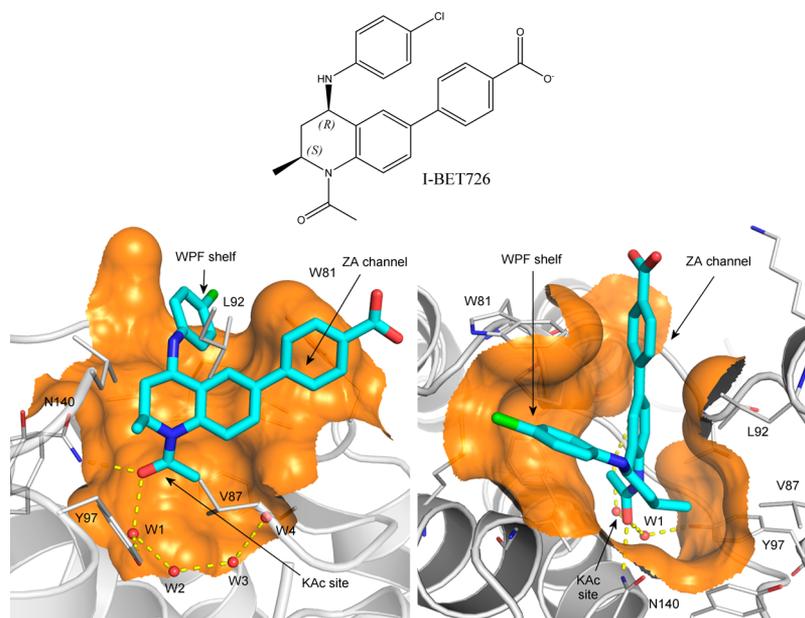


Figure 1. Bromodomain inhibitor I-BET726 and its binding mode in BRD4-BD1. Two views are displayed for the binding mode (PDB ID: 4BJX¹⁵), in which I-BET726¹⁶ is represented as stick in cyan/blue/red/green, the protein is shown as cartoon in silver, the crystallographic water molecules are shown as red balls, and clipped protein surfaces are shown in orange.

potencies to all BET family bromodomains. A representative inhibitor-protein structure is given in Figure 1, showing the key elements for the inhibitor binding. This study will concentrate on the first bromodomain of bromodomain-containing protein 4 (BRD4-BD1) for which extensive crystallographic and ligand binding data are available.^{10–12}

The purpose of the present study is to assess the potential for rapid, accurate, precise, and reproducible binding affinity calculations based on the use of a Binding Affinity Calculator (BAC)¹³ software tool and associated services including a Python-based toolkit, FabSim,¹⁴ to automate data transfer and job submission. The approach is based on the use of high performance computing in an automated workflow which builds models, runs large numbers of replica calculations, and analyzes the output data in order to place reliable standard error bounds on predicted binding affinities.

2. COMPUTATIONAL SECTION

Models. In this study, chemical structures of 16 BRD4-BD1 ligands based on a single tetrahydroquinoline (THQ) template¹⁶ were provided by GSK (Table 1). The compound set was designed to represent a range of chemical functionality and binding affinities, but also to contain sets of closely related compounds with key SAR trends. Specifically, there are two growth vectors which cause a drop off in potency, a growth vector where substantial structural variation is tolerated and an enantiomeric pair where one isomer is significantly more potent than the other. These will be described below. The calculations were then performed by the UCL group, initially blind, to investigate the ability of BAC to reproduce the SAR trends. The experimental binding affinities were not released to the UCL group until all of the computations had been completed and the estimation of binding free energies were reported (see Figure 3 below). These predictions were subsequently compared with the independently measured experimental free energies of binding (Table 1) as part of the assessment of the reliability of the method.

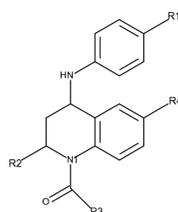
The X-ray crystal structure used to represent BRD4-BD1 (published coordinates PDB ID: 4BJX¹⁵) is a complex with a THQ (I-BET726¹⁶) chosen in order to reproduce the likely conditions for which free-energy calculations might be used in a real drug discovery program. The binding mode of I-BET726 is shown in Figure 1. Both of the significant protein–ligand hydrogen-bonding interactions take place through the THQ N1-acetamide carbonyl group, which interacts directly with the Asn140 side chain of the acetyl-lysine binding site and also through the W1 water to the side chain of Tyr97. The W1 water forms part of a chain of water molecules buried within the site (W2–W4, Figure 1). Substituents larger than acetamide at the N1-position extend into the water-filled part of the site. Our ligand data set included a small series of increasing size at N1 (1, 8, and 9, see Table 1) to probe whether the computer methodology can accurately predict the outcome of growing into this region.

The THQ R2-position (*S*)-methyl group occupies a small lipophilic site between the side chains of three residues of the BRD4-BD1 ZA-loop (Val87, Leu92, and Tyr97). Carbon–carbon contacts for all three lie within 4.25 Å, apparently offering little room for extension of this group without some structural reorganization. Our ligand data set includes a small number of analogues exploring larger substituents at the R2-position (10–14) to see whether the simulation can accurately predict the consequences of pushing on these residues.

The THQ 6-phenyl substituent fills the narrow “ZA channel” between Trp81 and Leu92. This ring can be substituted at the R3- and R4-positions, but R2-substitution is detrimental because this causes an increase in the inter-ring torsion. The ZA channel does not seem to accommodate the resulting wider 6-substituents. The remainder of our data set includes a variety of substituents probing electrostatic and steric changes at this position.

The THQ 4-anilino substituent projects onto the “WPF-shelf” subsite outside the acetyl-lysine pocket, close to Trp81. SAR around this ring has been published previously,¹⁶ showing

Table 1. Compounds Used in This Study, Ordered According to Their R4 Group, and Their Experimental IC₅₀s with Standard Deviations Converted into Binding Free Energies ΔG (kcal/mol)



Compound ^a	R1	R2	R3	R4	pIC ₅₀ ^b	pIC _{50 n}	ΔG ^b
2	H	Me	Me	H	5.6 ± 0.06	6	-7.7 ± 0.08
3	H	Me	Me		6.8 ± 0.04	2	-9.3 ± 0.05
4	H	Me	Me		6.8 ± 0.09	3	-9.3 ± 0.12
5	H	Me	Me		7.9 ± 0.03	2	-10.8 ± 0.04
6	H	Me	Me		5.6 ± 0.01	2	-7.7 ± 0.01
7	H	Me	Me		5.8 ± 0.07	4	-8.0 ± 0.10
1	H	Me	Me		7.0 ± 0.05	3	-9.6 ± 0.07
8	H	Me	Et		6.5 ± 0.01	2	-8.9 ± 0.01
9	H	Me	i-Pr		< 4.3 ^c	4	> -5.9
10	H	Me	Me		7.6 ± 0.1	16	-10.4 ± 0.14
16 ^d	H	Me	Me		5.4 ± 0.28	5	-7.4 ± 0.38
11	H	Et	Me		6.8 ± 0.3	4	-9.3 ± 0.41
12	H	Pr	Me		5.5 ± 0.01	4	-7.5 ± 0.01
13	H	Pr	Me		5.4 ± 0.01	2	-7.4 ± 0.01
14	H	Et	Me		6.7 ± 0.2	4	-9.2 ± 0.27
15	Cl	Me	Me		7.8 ± 0.1	44	-10.7 ± 0.14

^aCompounds 1–9 are electrostatically neutral, compounds 10–12 and 16 are positively charged, and compounds 13–15 are negatively charged.

^bStatistical errors were calculated from repeated IC₅₀ measurements. ^cThere was no activity at the highest concentration (50 μM) tested. ^dAll compounds are the 2-(S) 4-(R) isomers (Figure 1) except compound 16 which is 2-(R) 4-(S).

that substitution had small effects on potency, probably because one edge of the aniline ring is solvent-exposed. Therefore, we did not attempt to vary substituents at this position in our data set, preferring to keep it as constant as possible while exploring changes elsewhere.

In this series the 2-(S) 4-(R) isomers (see Figure 1) are significantly more potent than their alternative *trans* enantiomers. We included an enantiomeric pair of 2-(S) 4-(R)

(compound 10) and 2-(R) 4-(S) (compound 16) isomers to explore whether the simulations were capable of distinguishing between them.

The ligand-protein complexes were constructed by replacing the ligand in the PDB file with the ligands of interest (see structures in Supporting Information). For the congeneric compounds studied here, it is plausible to assume that they bind in the same mode in which all key compound-protein

interactions are preserved (Figure 1). For compounds with significant structural differences, computational docking may be required to generate reasonable complex structures.¹⁷ Preparation and setup of the simulations were implemented using BAC,¹³ including parametrization of the compounds, solvation of the complexes, electrostatic neutralization of the systems by adding counterions and generation of configurations files for the simulations. The AMBER ff99SB-ILDN¹⁸ force field was used for the protein, and TIP3P was used for water molecules. Standard protonation states were assigned to all titratable residues at pH 7, with histidines protonated on the ϵ position (HIE). Compound parameters were produced using the general AMBER force field (GAFF)¹⁹ with Gaussian 03²⁰ to optimize compound geometries and to determine electrostatic potentials at the Hartree–Fock level with 6-31G** basis functions. The restrained electrostatic potential (RESP) module in the AMBER package²¹ was used to calculate the partial atomic charges for the compounds. All systems were solvated in orthorhombic water boxes with a minimum extension from the protein of 14 Å.

Some of the ligands could adopt several rotamers for the R4 group (Table 1) in solution, which due to the constrained environment of the active site are unlikely to be sampled in a single bound simulation run with the protocol used (see Simulations subsection). To decide which rotamer(s) with which to start, metadynamics simulations²² were employed which used a history-dependent biasing potential to explore the conformational space of the chosen degrees of freedom, here the rotatable bond(s) involving groups R2 and R4 (see details in the Supporting Information). Five replicas were used for each metadynamics study to have a reasonable convergence of the potential profile of the chosen dihedral angle(s) of a ligand in complexed form. The calculated potential of mean force (PMF) was used to determine the most favorable rotamer(s) from which ESMACS simulations initiate. When unambiguous results ensued, the energetically most favorable rotamer was chosen. For some ligands, more than one rotamer showed similar free energies. In these cases, multiple initial structures were prepared using the rotamers suggested by the metadynamics study. The ligands with multiple rotamers were: 4, 7, 10, 11, 12, and 16. There are two rotamers for each of these, except for 7 for which there are three generated by the flip and twist between the two ring planes. This resulted in a total number of 23 complexes being simulated in the ESMACS study.

For each ligand with multiple rotamers, the energetic properties were analyzed and the most favorable rotamer was chosen as the final result of the ESMACS study (see Results section) and was used as the initial structure in the TIES study. For the latter, three subsets of the ligands were selected, within which relative binding free energies were calculated with the TIES approach. The subsets of the ligands were as follows: set 1 including ligands 1–9; set 2 including ligands 10, 11, and 12 which are positively charged; and set 3 including ligands 13, 14, and 15 which are negatively charged. The division of the full set of the cognate ligands into subsets is necessary because TIES, just as any other TI and FEP based method, encounters specific difficulties owing to major adjustments in long-range electrostatic interactions for the “congeners” where the net charges change.²³ For each subset, ligands were paired based on their similarities, and TIES calculations were performed to achemically mutate one ligand to another (see Theoretical Basis below).

Theoretical Basis. The UCL group have recently introduced new protocols for binding free energy calculations, termed “enhanced sampling of molecular dynamics with approximation of continuum solvent” (ESMACS)²⁴ and “thermodynamic integration with enhanced sampling” (TIES).²⁵ ESMACS is based on the molecular mechanics Poisson–Boltzmann surface area method (MMPBSA)²⁶ which makes a continuum approximation for the aqueous solvent, while TIES centers, as the abbreviation indicates, on thermodynamic integration (TI). Although the approaches are built around the standard MMPBSA and TI methodologies, our abbreviations are used to emphasize the central importance of the ensemble-based nature of the protocols employed as well as, in the case of ESMACS the wider generality and flexibility of the protocol.^{5,24} The size of the statistical mechanical ensembles is determined systematically so that predictions are accurate, precise, and reliable.^{24,27} Moreover, the term “MMPBSA” is used to mean very different things²⁸ in numerous journal articles and textbooks, including calculations based on single docked structures⁴ or on simulation trajectories, calculations with or without the inclusion of configurational entropies, and almost wholly using the so-called 1-trajectory approach. In the ESMACS protocol, we always mean a fully converged, ensemble-based determination of the free energy of binding from either a one-, two-, or three-trajectory method: this includes both the configurational entropy and the association free energy,^{27,29} and (where appropriate) the relative or absolute adaptation energy.²⁴ Statistical analyses³⁰ were performed throughout the study for all of the quantities obtained. A Binding Affinity Calculator (BAC)¹³ was used to perform ESMACS and TIES studies. BAC constitutes a computational pipeline built from selected software tools and services (see Supporting Information), and relies on access to a range of computing resources. It automates much of the complexity of building, running, and marshalling the molecular dynamics simulations, as well as collecting and analyzing data. Integration and automation are central to the reliability of the method, ensuring that the results are reproducible and can be delivered rapidly.

In ESMACS, the free energy is evaluated approximately based on the extended MMPBSA method,^{24,27} including configurational entropy, and the free energy of association.²⁹ Free energy changes ($\Delta G_{\text{binding}}$) are determined for the molecules in their solvated states. The binding free energy change is then calculated as

$$\Delta G_{\text{binding}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}} \quad (1)$$

where G_{complex} , G_{receptor} , and G_{ligand} are free energies of the complex, the receptor and the ligands, respectively. The Poisson–Boltzmann calculation is performed here using the Amber built-in PBSA solver³¹ with radii taken from the parameter/topology files, and the configurational entropy is calculated by a normal mode (NMODE) method.

The three terms on the right-hand side of eq 1 can be generated from single simulations of the complexes or from separated simulations of complexes, receptor(s), and ligand(s). The former is the so-called 1-trajectory (1-traj) method of which the trajectories of the receptor(s) and the ligand(s) are extracted from those of the complexes. The latter is the so-called 3-trajectory (3-traj) method of which the trajectories are generated from separate simulations of the three components. In the drug development field, binding is usually investigated for a set of ligands bound to the same protein target. The free

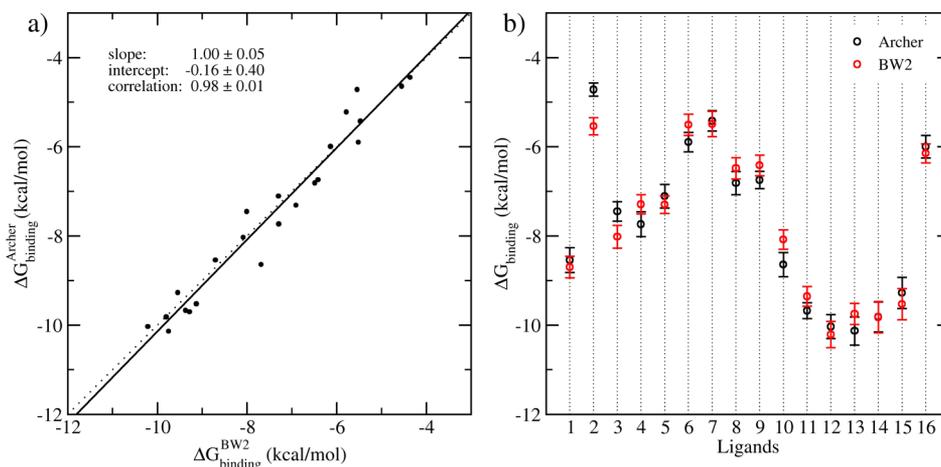


Figure 2. Correlation and standard errors of the calculated binding free energies from two independent studies of the BRD4-ligand models performed on BlueWonder2 and ARCHER. (a) Correlation of the predictions, including all rotamers, from 1-traj calculations performed on BlueWonder2 (BW2, horizontal axis) and ARCHER (vertical axis). Solid line, regression of the data using means of the calculated free energies; dotted line, 1:1 ideal regression. (b) the averages and their standard errors from the two separate calculations. One rotamer is used for each ligand.

energy of the receptor G_{receptor} will then be the same for all ligands in the 3-traj method. In the current study, we employ three approaches to calculate the binding free energies $\Delta G_{\text{binding}}$: (A) using single simulations of the complexes (1-traj); (B) same as (A) but using the average of the receptor free energies $\langle G_{\text{receptor}} \rangle$ from the 1-traj method (denoted as 2-traj); and (C) the same as (B) but invoking separated ligand simulations for the derivation of G_{ligand} (denoted as 3-traj).

In TIES, an alchemical transformation for the mutated entity, either the ligand or the protein, is used in both aqueous solution and within the ligand–protein complex. The free energy changes of the alchemical mutation processes, $\Delta G_{\text{aq}}^{\text{alch}}$ and $\Delta G_{\text{complex}}^{\text{alch}}$ are calculated by

$$\Delta G^{\text{alch}} = \int_0^1 \left\langle \frac{\partial V(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (2)$$

and the binding free energy difference is calculated from

$$\Delta \Delta G^{\text{binding}} = \Delta G_{\text{aq}}^{\text{alch}} - \Delta G_{\text{complex}}^{\text{alch}} \quad (3)$$

Here λ ($0 \leq \lambda \leq 1$) is a coupling parameter such that $\lambda = 0$ and $\lambda = 1$ correspond to the initial and final thermodynamic states, and $V(\lambda)$ is the potential energy of an intermediate state λ . $\langle \dots \rangle_{\lambda}$ denotes an ensemble average over configurations representative of the state λ . To avoid the well-known “endpoint catastrophe”,³² a soft-core potential is used, along with an approach to decouple at different rates the electrostatic and van der Waals interactions of the perturbed atoms with their environment (for more details see the [Supporting Information](#)). The foregoing is a textbook account of thermodynamic integration; TIES differs by virtue of performing ensemble based calculations,^{25,33} in the same manner as for ESMACS, thereby providing control of standard errors based on length of simulation and number of ensembles used. Indeed, it is important to note that the energy derivatives (eq 2) calculated for all λ windows are well described by Gaussian random processes, which makes it possible to draw on the theory of stochastic calculus to compute relevant properties reliably.⁵

Simulations. The ESMACS and TIES studies follow the protocol developed recently^{24,25} (see [Supporting Information](#)). The Amber package²¹ was used for the setup of the systems and analyses of the results, and the MD package NAMD2.9³⁴

was used throughout the equilibration and production runs of all simulations, including the metadynamics. An ensemble simulation was performed for each molecular system with identical atomic coordinates for all replicas. Energy minimizations were first performed with heavy protein atoms restrained at their initial positions. The initial velocities were then generated independently from a Maxwell–Boltzmann distribution at 50 K, and the systems were heated up to and kept at 300 K within 60 ps. A series of equilibration runs, totalling 2 ns, were conducted, while the restraints on heavy atoms were gradually reduced. Finally, 4 ns production simulations were run for each replica for all ESMACS and TIES simulations. In ESMACS, we used 25 replicas in an ensemble calculation for each ligand. In TIES, 5 replicas were used for each pair of the ligands. Simulation of each replica consists of 2 ns for equilibration and 4 ns for production run. Previous studies^{24,27,30,35} have shown that the combination of the simulation length and the size of the ensemble provides a trade off between computational cost and precision.

Two independent ESMACS runs were performed to assess the reproducibility of the calculations. One was run on ARCHER, a Cray XC30 supercomputer (equipped with ca 118 000 cores), the UK’s National High Performance Computing Service located in Edinburgh; the other run on the BlueWonder2 supercomputer, an IBM NextScale Cluster (8640 cores) located in Science and Technology Facilities Council’s (STFC’s) Hartree Centre. A comparison between the two runs is very instructive, both in terms of performance and time to solution as well as providing an opportunity to conduct a valuable reproducibility study. Thanks to the large number of cores on ARCHER, we ran the entire set of ESMACS binding affinity calculations concurrently and had the findings available within 7 h. On BlueWonder2, we were hampered by a number of issues, resulting in the calculations taking significantly longer than on ARCHER. On BlueWonder2 and ARCHER, about 440 000 and 335 000 core hours were consumed in the course of this study, including 300 000 and 225 000 core hours for production molecular dynamics simulations and 140 000 and 110 000 core hours for free energy calculations, respectively.

The TIES study was performed entirely on ARCHER. TIES simulations of 7 out of 12 pairs were packed together (the currently enforced limit of maximum tasks for one job in

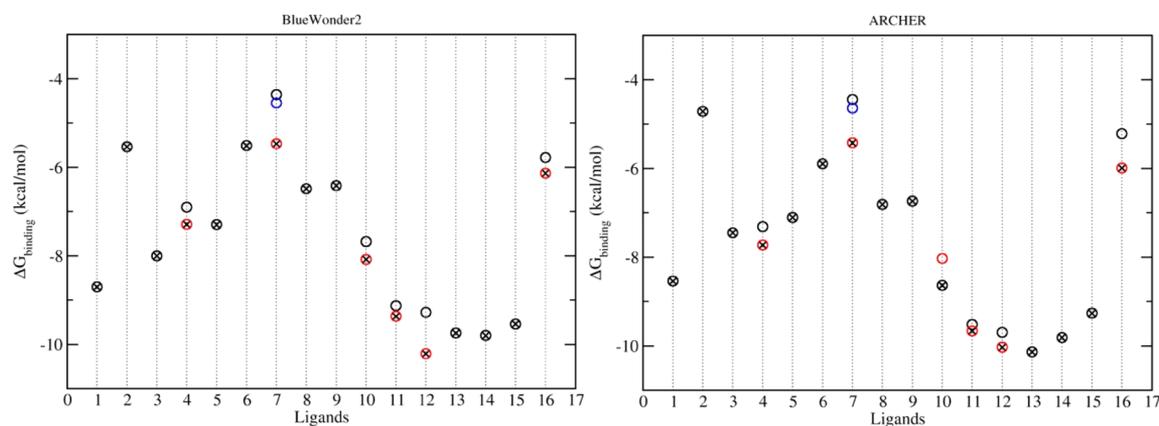


Figure 3. Calculated binding free energies from simulations on BlueWonder2 and ARCHER. The ligands are numbered as per Table 1. Circles with red/blue colors are the results based on studies with different rotamers. The circles with crosses are the final results with selected rotamers which are chosen on the basis of the sum of energies G_{ligand} and $\Delta G_{\text{binding}}$ (see eq 1). All of the calculated binding free energies are associated with standard errors of less than 1.7 kcal/mol, and are not shown in the figures for reasons of clarity.

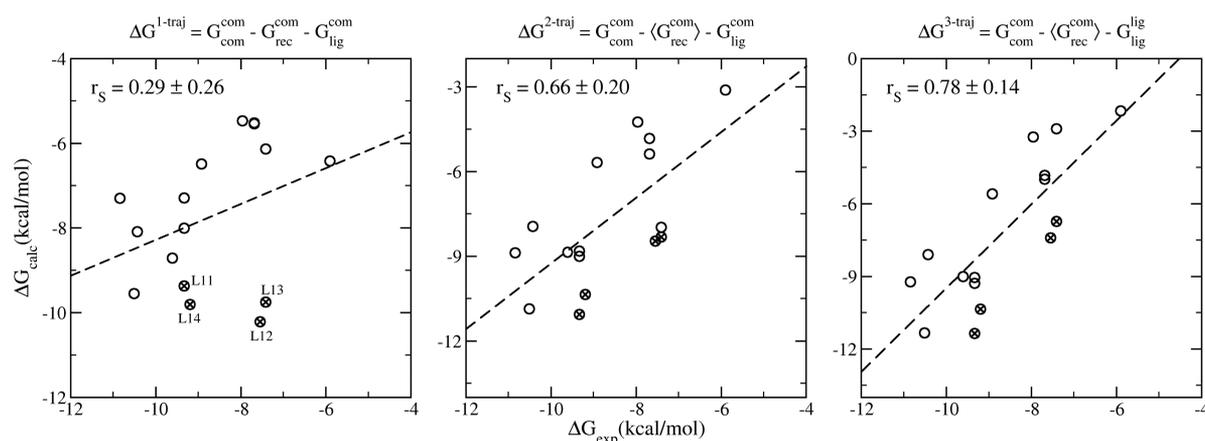


Figure 4. Spearman ranking correlations of the calculated binding free energies and the experimental data from 1-traj (left panel), 2-traj (center), and 3-traj (right panel) ESMACS approaches. The equations on the subfigures indicate the calculations used in each case. The subscripts (com/rec/lig) and the superscripts (com/lig) in the equations indicate the components (complexes, receptor, and ligands) and the simulations (complexes and free ligands), respectively. The ligands with modifications at the R2-position of the tetrahydroquinoline are marked with crosses; they are all significantly improved in the 2- and 3-trajectory version. The standard errors, which are 0.19–0.34 kcal/mol for the 1-traj and 1.02–1.71 kcal/mol for the 2- and 3-traj approaches, are not shown for reasons of clarity. They are calculated using a bootstrapping method (see Supporting Information). The 2- and 3-traj approaches have similar errors because the energy of the receptor is treated as a constant and hence the uncertainties are dominated by the energies of the complexes.

ARCHER) and submitted as one single job, requiring 43 680 cores, which has high priority and executes rapidly on the machine (waiting time in the ARCHER queue was around 8 h). This made it possible to complete the entire TIES study of seven pairs within 1 day including queuing time. Further speed up is feasible on supercomputers with hybrid architectures which would permit for example GPU acceleration for some parts of these calculations.

3. RESULTS

To assess the accuracy and precision of the method, we evaluated the binding affinities of the ligands (Table 1) to BRD4 and compared the computed results with experimental data. ESMACS was used for the full set of the ligands, including the stereoisomer which cannot be considered as a perturbation of any others. TIES was applied to three subsets of the ligands, of which each includes congeners with the same net charge.

Reproducibility. It is well-known that many complex systems exhibit sensitive dependence on initial conditions.^{5,36}

The differences of the initial velocities among individual simulations lead to rapid divergence of trajectories. The calculated thermodynamic properties from individual simulations will therefore inevitably differ. Two sets of ESMACS simulations were performed for the complexes independently on ARCHER and BlueWonder2 (see the Computational Section above). The results are compared by linear regression. Figure 2 shows the variances and correlation of the calculated binding free energies from the two sets of simulations. The correlation coefficient r is 0.98 ± 0.01 , with no statistical differences between slopes and intercepts of the calculated regression line and the ideal line ($x = y$) (Figure 2a). The two studies produce consistent results, with an average difference of 0.07 ± 0.10 kcal/mol, and an average unsigned difference of 0.39 ± 0.06 kcal/mol. Only 3 out of 16 predictions do not overlap within their variances (Figure 2b). For the purpose of ranking compound selectivity, the Spearman correlation coefficient (r_s) is also calculated, which shows that the rankings from the two studies are very close to one another, with a

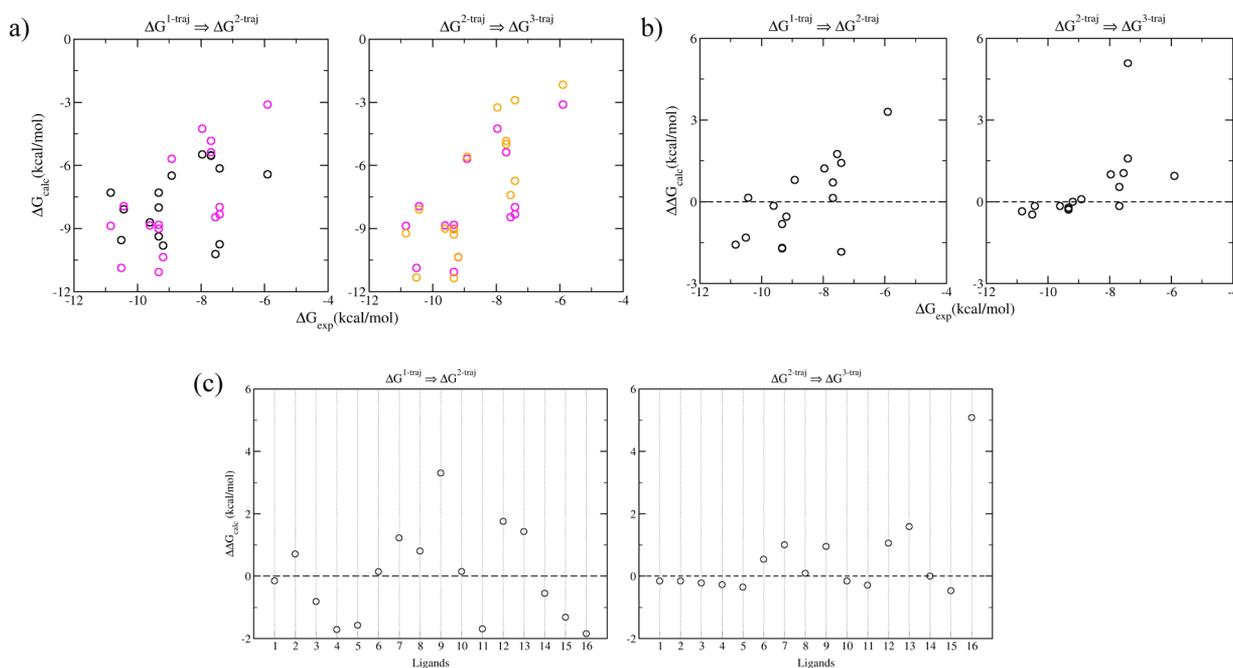


Figure 5. Improvement of the predictions by inclusion of the adaptation free energies of the receptor and the ligands: (a) the binding free energy changes between the 1-traj (black circles) and 2-traj (magenta circles) indicate the relative adaptation energies of the receptor; those between the 2-traj (magenta circles) and 3-traj (orange circles) show the adaptation energies of the ligands. The adaptation energies can be seen more clearly in panels b as a function of binding affinities, and in panels c for each ligand.

highly significant Spearman correlation of 0.98 ± 0.02 . In the following analyses, only the results from BlueWonder2 simulations are reported. The ARCHER-based calculations produced very similar results.

Choosing Rotamers. Some of the ligands (Table 1) would be able to adopt more than one rotamer in solution. When the ligands are complexed with the protein, they are usually trapped in a specific rotameric state. The energy barriers between rotamers are sufficiently high, especially for groups occupying the narrow ZA channel (Figure 1), that it is not feasible to achieve equilibrated populations of rotamers using the current protocol. We therefore initiated molecular dynamics simulations of complexes from all possible rotamers for six out of the 16 ligands. Simulations including different rotamers are only applied to the complexes, as the free ligands in water are able to fully sample all rotamers. The most favorable rotamer is selected (Figure 3) for each ligand in its complex form, based on the lowest ligand energy of the rotamer, which is approximated by $(G_{\text{ligand}} + \Delta G_{\text{binding}})$.

Comparison between ESMACS Calculations and Experiments. The predicted binding free energies from the 1-traj approach exhibit a weak Spearman correlation with the experimental data (Figure 4a). In the current study, as commonly employed in pharmaceutical drug development projects, a series of ligands are investigated for the same protein target. The energy of the protein therefore does not affect the ranking of the binding affinities (eq 1). Treating it as a constant, the 2-traj approach significantly improves the correlation between the predictions and experimental data (Figure 4b). Incorporating the free energies of the ligands from their individual simulations in free form, the 3-traj approach, further improves the correlation (Figure 4c). The most significant improvement appears in the compounds involving modifications at the R2 position of the tetrahydroquinoline (Table 1), for which the 1-traj approach predicts an increase

(compound 11 to 12) or no change (compound 14 to 13) in the binding affinities when an ethyl is replaced by an *n*-propyl. The 2- and 3-traj approaches correctly rank the binding potencies of the two pairs, and indeed of these four compounds. Although the improvement is most evident for the listed four compounds, it is indeed applied to the overall correlation for the whole set (Figure 4). It should be noted, however, that although the 2- and 3-traj methods significantly improve the correlations between predictions and experimental data, relatively large standard deviations, up to 1.7 kcal/mol, are associated with the calculated free energies. This indicates that the binding free energies must differ by no less than 3.4 kcal/mol to be statistically significant to a 95% confidence level. With the experimental free energy differences ranging from 0.0 to 4.9 kcal/mol, a larger ensemble would be required to reduce the relatively large uncertainties in order to render all the predictions statistically significant. It is important to mention here that experimental binding affinities usually contain errors to a greater extent³⁷ than are implied by Table 1. Generating reproducible binding free energies from experiments is as important as that from simulations to make the comparisons statistically significant.

The improvements are achieved by including the adaptation energies of the receptor and the ligands (Figure 5). While the adaptation energies for the ligands are calculated as the differences between the free energies in their bound and free states, those for the receptor are calculated relative to the average of the receptor free energies in the bound states. The absolute adaptation energies of protein for each ligand binding would require a converged ensemble simulation of protein in solvent, preferably initiated from structures of the protein in its free state. As the ranking of the ligand binding is the main concern of the study, and the relative adaptation energies of the protein are energetically as informative as the absolute ones in the binding affinity comparison, no attempt has been made

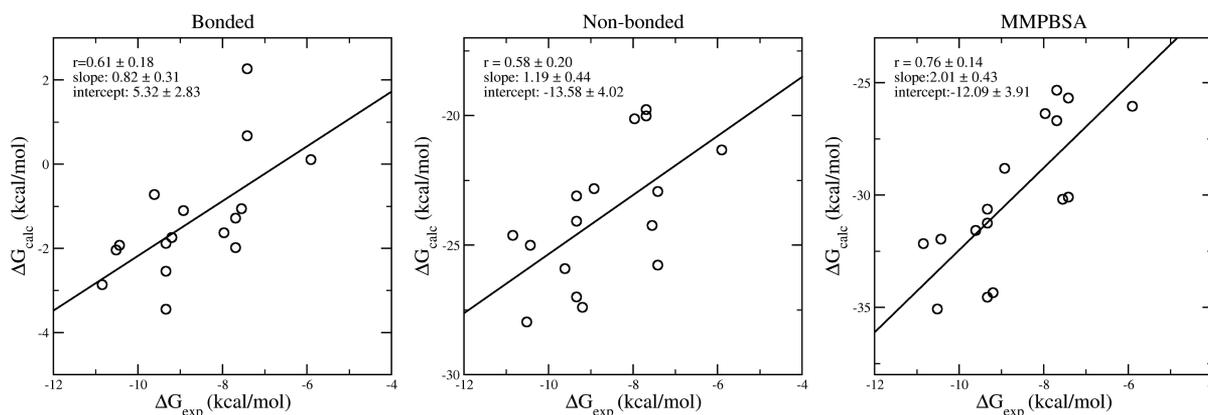


Figure 6. Correlations of free energy components and the experimental data from 3-traj approaches. Both bonded and nonbonded energy terms contribute to the ranking of binding affinities. Their combination (the MMPBSA energy) exhibits a better correlation with experimental data than the components themselves.

here to compute accurate absolute adaptation energies for the protein (all adaptation energies for the protein are relative ones in the current paper). Unfavorable adaptation energies are usually induced within the protein by the introduction of larger functional groups at R2 and R3 positions of the compounds (Figure 5c). It is interesting to note that, for the ligands with higher binding affinities ($\Delta G_{\text{binding}} < -8$ kcal/mol), only small adaptation energies arise for both the receptor and the ligands, which makes the binding apparently closer to a “lock and key” recognition mechanism; while for the ligands with lower binding affinities, non-negligible adaptation energies are found for both the receptor and the ligands (Figure 5b). The one exception is the stereoisomer ligand 16 (Table 1) whose binding induces the smallest relative adaptation energy of the receptor but the largest adaptation energy for the ligand (Figure 5c). This compound was included in the set as a negative control to investigate the ability of the computational procedure to highlight the weak activity of the enantiomer of compound 10. Compound 16 has a very high ligand strain, suggesting that this isomer has the wrong shape for the binding site and must adopt a high energy conformation to interact with the protein.

The ESMACS predictions are precise and accurate in terms of ranking the binding free energies, and are so in a reproducible way. However, ESMACS does not provide accurate absolute free energies. In our study, the 3-traj ESMACS protocol yields a much better correlation than the 1-traj ESMACS approach, while they have similar mean absolute deviations (1.97 ± 1.33 kcal/mol and 1.74 ± 0.29 kcal/mol, respectively). The mean absolute deviations can be significantly larger for more flexible ligands such as peptides binding to a major histocompatibility complex (MHC).²⁴ Some studies show that improved binding free energy prediction could be obtained by including important water molecules between ligands and protein.³⁸ Inclusion of the bridging water molecules (Figure 1) in the current ESMACS study, however, does not improve the correlation between the calculated binding free energies and the experimental measurements (see Figure S3 in the Supporting Information). Indeed, ours is a “generic” methodology which treats solvent implicitly and works well as such. Embarking on any approaches for treating water molecules explicitly would lead to a loss of simplicity in the method without conferring any benefit.

Component analyses of the binding free energies could provide insight into the mechanism of compound binding. In our study of peptide-MHC binding,²⁴ for example, the van der Waals interaction was shown to be the dominant component and to manifest a good correlation with the experimental binding free energies in the 3-trajectory ESMACS study. In the current molecular systems, however, there are no significant correlations between any energy component and the experimental data, except the bonded (including bond stretching, angle bending, torsions and improper torsions) and nonbonded (van der Waals, electrostatic and solvation) interactions as shown in Figure 6. The sum of the two interactions improves the correlation further, with a similar correlation coefficient as shown in Figure 4 for the 3-trajectory approach. It can therefore be deduced that the configurational entropy component does not contribute to the quality of the ranking prediction. As for the impact of the configurational entropy on the ranking of calculated binding affinities, different conclusions—improving,²⁷ worsening,³⁹ or having no effect²⁴—have been drawn for diverse protein–ligand complexes. The differences may indeed reflect the mechanism of protein–ligand binding which can be driven predominantly by enthalpy, by entropy, or by both.⁴⁰

Comparison between TIES Calculations and Experiments. Good agreement is found for the binding free energy differences between the TIES calculations and the experimental measurements. The TIES approach is both accurate and precise, yielding a Spearman ranking coefficient of 0.92 for the means of calculated and experimental binding free energy differences. The bootstrap analyses give a Spearman coefficient of 0.86 ± 0.15 , drawn from a non-normal distribution of resampling correlation coefficients (see Figure S4 in the Supporting Information). An average difference of 0.06 ± 0.26 kcal/mol, and an average unsigned difference of 0.75 ± 0.14 kcal/mol, are found between the predictions and experimental measurements (Figure 7). It should be noted that experimental binding affinities may have statistical errors of around 24% from isothermal titration calorimetry (ITC)³⁷ which is a more reliable method to measure binding affinities than the IC_{50} measurements used in this study. The variances observed in the study can be partially attributed to these experimental errors. The regression line is close to an ideal 1:1 regression, with a slope of 0.93 ± 0.24 and an intercept of -0.10 ± 0.33 kcal/mol. A similar level of accuracy was recently

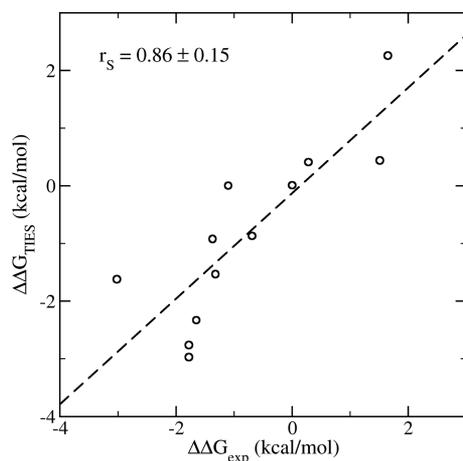


Figure 7. Correlations of the calculated binding free energy differences from the TIES study and from experimental measurement. The standard error bars from the TIES calculations are all no greater than 0.2 kcal/mol.

reported¹⁷ in an absolute binding free energy calculation for a set of drug-like molecules binding to BRD4, albeit little attention was paid to reproducibility in that work.

4. DISCUSSION

To analyze the results it is necessary to consider the differences between the 1-, 2- and 3-trajectory methods. The 1-trajectory simulation calculates the $\Delta G_{\text{binding}}$ assuming there is no energetic penalty for the adaptation of the protein or ligand to the binding conformation, the 2-trajectory method takes the change in protein energy into consideration and the 3-trajectory method accounts for the changes in both protein and ligand energy.

The test set for this evaluation was designed to investigate the ability of ESMACS to predict specific SAR trends. The first trend is the loss of potency observed as the hydrophobic part of the acetyl lysine mimetics is grown from methyl (1) to ethyl (8) to isopropyl (9). Figure 8 shows that, in the 1-traj simulation, compound 1 is predicted to be more potent than compounds 8 and 9, but no difference is predicted between 8 and 9. This is in contrast to the experimental results which show a modest loss of potency going from methyl to ethyl, but a substantial loss going from ethyl to isopropyl. On comparing the same three compounds in the 2-traj simulation, the predicted affinity of compound 9 is substantially reduced giving the correct ranking of the three compounds. The predictions for the three compounds are qualitatively the same for the 3-traj

calculation. These results indicate that the low activity of compound 9 is the result of the increased bulk of the isopropyl group causing strain in the system, which is manifested in an increased internal energy of the protein. In fact, compound 9 has the highest protein internal energy of the whole data set (Figure 6c). The TIES study gives the same ranking for the binding affinities as those from the 2- and 3-traj ESMACS, but with an enhanced correlation coefficient compared to that of ESMACS. Structurally, this result makes sense because the isopropyl group is rigid and so the strain caused by the increased bulk cannot be accommodated by the ligand and hence is transmitted to the protein.

The second SAR trend of interest is the effect of increasing bulk at the R2-position of the tetrahydroquinoline (THQ). Similar to the previous instance, increasing the size of the substituent at this position from methyl (10) to ethyl (11) to *n*-propyl (12) results in a sequential loss of activity. In this case the 1-traj method inverts the rank order of the three compounds (Figure 9), relative to experiment, predicting the *n*-propyl to be the tightest binder. The 2-traj calculation predicts a decrease of the potency of compound 12, consequently predicting it to be substantially weaker than compound 11, as observed experimentally. In the 3-traj method the potency of compound 12 drops further, relative to the other compounds. In all three ESMACS methods, the potency of compound 10 is under-predicted relative to the other compounds; the reason for this is unclear. The TIES study also predicts the binding of compound 12 to be the weakest in this subgroup, but could not distinguish the affinities of the compounds 10 and 11. A second pair of compounds also has the ethyl (14) to *n*-propyl (13) modification at the same position, and qualitatively the results parallel those described above. In this case, the *n*-propyl compound shows a correctly predicted loss of activity in both the TIES and 2- and 3-traj ESMACS methods, indicating the loss in potency for this bulky group is due to a combination of protein and ligand strain (Figure 6c). Again, this is consistent with the molecular structure of compound 12, because the *n*-propyl group is more flexible than the isopropyl group in the first example and consequently is more able to adopt a slightly strained ligand conformation.

The carbonyl group in all compounds forms a hydrogen bond with a conserved asparagine Asn140 in the BRD4. This is a key interaction observed in bromodomain-ligand complexes as it mimics the interaction made by the carbonyl of the acetyl lysine of the substrate. The occupancies of this hydrogen bond in all of the compounds are very similar, and do not appear to correlate with their binding affinities. Hence, when bulk is

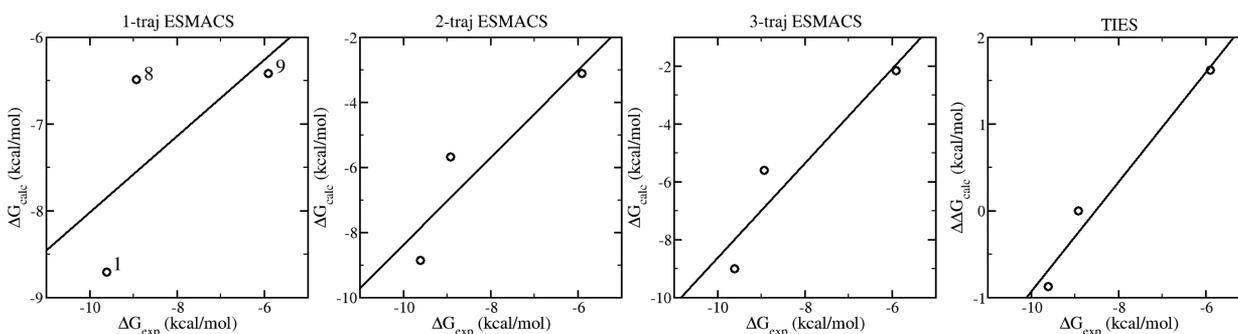


Figure 8. Calculated vs experimental binding free energies for ligands 1, 8, and 9 which are labeled in the 1-traj subfigure.

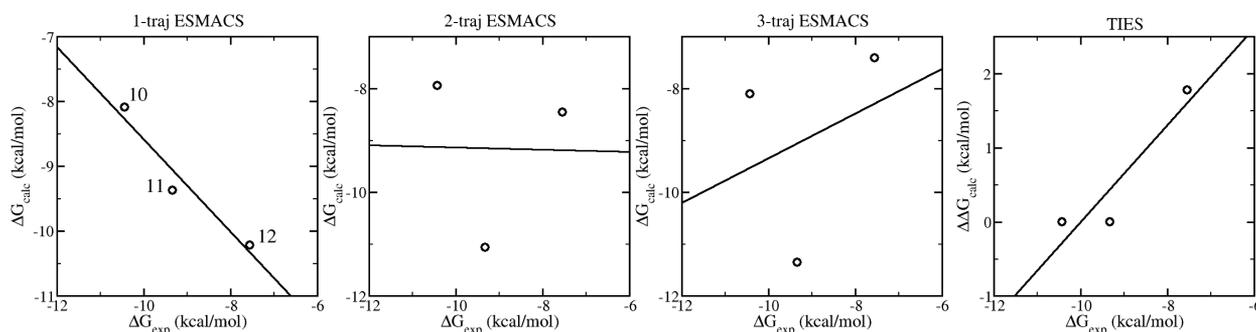


Figure 9. Calculated vs experimental binding free energies for ligands 10, 11, and 12 which are labeled in the 1-traj subfigure.

increased adjacent to this key binding motif, as described in the previous cases, the loss of potency does not result from the disruption of this hydrogen bond. Conversely, the interaction is maintained at the expense of creating internal strain in the system. This is consistent with the observation that this is a ubiquitous interaction in this protein and a key binding pharmacophore.

For compounds 13, 14, and 15, there is an extra hydrogen bond formed between the carboxylate group of the ligands and the Lys 91 of the protein. The replacement of *n*-propyl (13) with ethyl (14) and methyl (15) slightly changes the orientations of the compounds, making the carboxylate group better aligned for the formation of hydrogen bonds. The occupancies of this hydrogen bond are 16.6%, 31.7%, and 33.3% for the compounds 13, 14, and 15, respectively (the cutoff values are 3 Å and 120° for the donor–acceptor distance and donor–hydrogen–acceptor angle). The occupancies of the hydrogen bonds for these 3 compounds appear to correlate with their binding affinities (−7.41, −9.20, and −10.51 kcal/mol). However, the energetic contribution of this solvent exposed interaction is difficult to quantify, and the energetic analysis elsewhere in this section suggests that differences in intermolecular interactions in the complex contribute modestly to the differences in binding free energy compared to differences in internal energies. Further evidence for this is seen in the pairs of compounds 11,14 and 12,13, which are identical except for a carboxylic acid to piperidine (base) modification in this region. These pairs of compounds have very similar binding affinities both in terms of experimental and predicted values.

These data suggest that the internal energy of the protein and ligand are major contributors to the relative binding affinities of this series of compounds. The internal energy contributions are exemplified by plotting the difference between the binding free energies predicted using the different ESMACS methods (Figure 10).

Figure 10 plots the difference between the $\Delta G_{\text{binding}}$ calculated by the 3-traj method and the 1-traj method against the experimental $\Delta G_{\text{binding}}$. This is a plot of the internal energy of the protein and ligand against the experimental $\Delta G_{\text{binding}}$; it shows a strong correlation ($r = 0.87 \pm 0.13$). If this is contrasted with the 1 trajectory method—which essentially contains the intermolecular contributions to $\Delta G_{\text{binding}}$ (Figure 4a ($r = 0.29 \pm 0.26$))—one can conclude that, for this protein and set of ligands, the internal energy of the ligand and the protein differentiate between the ligands to a far greater extent than protein–ligand interaction energies.

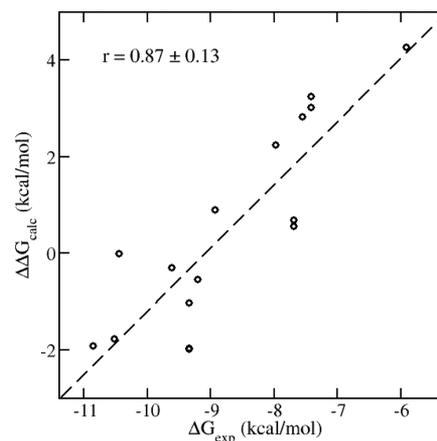


Figure 10. Correlations of the internal energy contributions to the calculated binding free energies and experimental measurement. The internal energy changes are calculated as the differences of the binding free energies between those from the 1-traj and 3-traj approaches: $\Delta\Delta G_{\text{calc}} = \Delta G_{\text{binding}}^{3\text{-traj}} - \Delta G_{\text{binding}}^{1\text{-traj}}$.

5. CONCLUSIONS

The Binding Affinity Calculator (BAC) software environment has been used to run ESMACS and TIES calculations against a series of inhibitors of BRD4 from the THQ chemical series. High performance computing was used in an automated workflow to build models, run multiple replicas of calculations and analyze the output, placing reliable standard error bounds on predicted binding affinities.

Despite the challenging data set used for this evaluation, good agreement is found for the binding free energy differences between the experimental measurements and the theoretical calculations, for both ESMACS and TIES. ESMACS is good as an “absolute” method for ranking an arbitrary set of ligands, which may be very diverse in terms of structures and electronic properties. To produce good rankings it is necessary in this case to use a 3-trajectory version, which has the further benefit of clearly distinguishing between ligands that adhere more closely to the “lock and key” or to the “induced fit” binding hypotheses. TIES performs well in determining relative binding free energies of congeners when there is no change of net charge, even when there are differences in structural features between pairs of compounds which one would not intuitively regard as minor.

Overall, ESMACS provides reliable binding affinity rankings and clear mechanistic insight into what factors drive binding processes in individual ligand–protein systems. TIES offers more quantitative accuracy in its predictions but, owing to its alchemical basis which necessarily keeps two congeneric ligands

in play at all times, is less able to provide similar structural and mechanistic insights into binding. The standard errors in ESMACS and TIES are fully controlled through simulation length and number of replicas used in the ensembles chosen.^{24,27}

The results also offer insight into possible design strategies for BRD4 ligands. In particular it has been noted that, within a chemical series, differences in binding affinity do not seem to result from differences in protein–ligand interaction energies. On the contrary, to avoid unfavorable (in this case, mainly steric) interactions, both ligands and protein adopt strained states resulting in large adaptation energies. Hence, during the optimization of a chemical series for BRD4 it would be prudent to carefully consider the shape complementarity of the ligand and the active site cavity. It should be noted, however, that this study did not include molecules in which unfavorable electrostatic protein–ligand interactions were introduced.

The results offer encouragement that BAC, and its underlying approach of running multiple replicas of each simulation to improve predictive power,⁵ can accurately rank ligands by their binding free energy. Further evaluations are planned to confirm this potential. As large and secure computing resources become more routinely available, for example through cloud computing, it will become increasingly easy for industrial groups to access approaches like the one outlined in this study. Consequently, the robust prediction of protein–ligand binding affinities in an industrial setting should become more routine and offer a long awaited development in the field of structure-based design.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.6b00794.

Detailed description of the methods used, additional energetic analyses of the metadynamics, free energy calculation with inclusion of explicit water molecules, alongside the atomic coordinates of the compound–protein complexes and experimental data on compound binding (PDF)

Structures of the studied compounds (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: p.v.coveney@ucl.ac.uk.

ORCID

Peter V. Coveney: 0000-0002-8787-7256

Funding

The authors would like to acknowledge the support of EPSRC via the 2020 Science programme (<http://www.2020science.net/>), EP/I017909/1), the Qatar National Research Fund (7-1083-1-191), the MRC Medical Bioinformatics project (MR/L016311/1), the EU H2020 CompBioMed (675451) and ComPat (671564) projects, and funding from the UCL Provost. We acknowledge use of Hartree Centre computer BlueWonder2 in this work and assistance from its scientific support staff. The STFC Hartree Centre is a research collaboratory in association with IBM providing High Performance Computing platforms funded by the UK's investment in e-Infrastructure. We made use of ARCHER, the UK's national High Performance Computing Service, funded by the Office of Science and Technology through EPSRC's High-End Comput-

ing Programme. Access to ARCHER was provided through the 2020 Science programme. This research was partially supported by the PLGrid Infrastructure through which access to Prometheus, the fastest Polish supercomputer run by ACK Cyfronet AGH in Krakow, was provided.

Notes

The authors declare no competing financial interest.

■ REFERENCES

- (1) Lounnas, V.; Ritschel, T.; Kelder, J.; McGuire, R.; Bywater, R. P.; Foloppe, N. Current progress in structure-based rational drug design marks a new mindset in drug discovery. *Comput. Struct. Biotechnol. J.* **2013**, *5*, e201302011.
- (2) Green, D. V.; Leach, A. R.; Head, M. S. Computer-aided molecular design under the SWOTlight. *J. Comput.-Aided Mol. Des.* **2012**, *26* (1), 51–56.
- (3) Kuhn, B.; Guba, W.; Hert, J.; Banner, D.; Bissantz, C.; Ceccarelli, S.; Haap, W.; Korner, M.; Kuglstatter, A.; Lerner, C.; Mattei, P.; Neidhart, W.; Pinard, E.; Rudolph, M. G.; Schulz-Gasch, T.; Woltering, T.; Stahl, M. A real-world perspective on molecular design. *J. Med. Chem.* **2016**, *59* (9), 4087–4102.
- (4) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J. Am. Chem. Soc.* **2015**, *137* (7), 2695–2703.
- (5) Coveney, P. V.; Wan, S. On the calculation of equilibrium thermodynamic properties from molecular dynamics. *Phys. Chem. Chem. Phys.* **2016**, *18* (44), 30236–30240.
- (6) Arrowsmith, C. H.; Bountra, C.; Fish, P. V.; Lee, K.; Schapira, M. Epigenetic protein families: a new frontier for drug discovery. *Nat. Rev. Drug Discovery* **2012**, *11* (5), 384–400.
- (7) Copeland, R. A. Epigenetic medicinal chemistry. *ACS Med. Chem. Lett.* **2016**, *7* (2), 124–127.
- (8) Copeland, R. A.; Olhava, E. J.; Scott, M. P. Targeting epigenetic enzymes for drug discovery. *Curr. Opin. Chem. Biol.* **2010**, *14* (4), 505–510.
- (9) Theodoulou, N. H.; Tomkinson, N. C.; Prinjha, R. K.; Humphreys, P. G. Clinical progress and pharmacology of small molecule bromodomain inhibitors. *Curr. Opin. Chem. Biol.* **2016**, *33*, 58–66.
- (10) Bamborough, P.; Diallo, H.; Goodacre, J. D.; Gordon, L.; Lewis, A.; Seal, J. T.; Wilson, D. M.; Woodrow, M. D.; Chung, C. W. Fragment-based discovery of bromodomain inhibitors part 2: optimization of phenylisoxazole sulfonamides. *J. Med. Chem.* **2012**, *55* (2), 587–596.
- (11) Chung, C. W.; Coste, H.; White, J. H.; Mirguet, O.; Wilde, J.; Gosmini, R. L.; Delves, C.; Magny, S. M.; Woodward, R.; Hughes, S. A.; Boursier, E. V.; Flynn, H.; Bouillot, A. M.; Bamborough, P.; Brusq, J. M.; Gellibert, F. J.; Jones, E. J.; Riou, A. M.; Homes, P.; Martin, S. L.; Uings, I. J.; Toum, J.; Clement, C. A.; Boullay, A. B.; Grimley, R. L.; Blandel, F. M.; Prinjha, R. K.; Lee, K.; Kirilovsky, J.; Nicodeme, E. Discovery and characterization of small molecule inhibitors of the BET family bromodomains. *J. Med. Chem.* **2011**, *54* (11), 3827–3838.
- (12) Chung, C. W.; Dean, A. W.; Woolven, J. M.; Bamborough, P. Fragment-based discovery of bromodomain inhibitors part 1: inhibitor binding modes and implications for lead discovery. *J. Med. Chem.* **2012**, *55* (2), 576–586.
- (13) Sadiq, S. K.; Wright, D.; Watson, S. J.; Zasada, S. J.; Stoica, I.; Coveney, P. V. Automated molecular simulation based binding affinity calculator for ligand-bound HIV-1 proteases. *J. Chem. Inf. Model.* **2008**, *48* (9), 1909–1919.
- (14) Groen, D.; Bhati, A. P.; Suter, J.; Hetherington, J.; Zasada, S. J.; Coveney, P. V. FabSim: Facilitating computational research through

automation on large-scale and distributed e-infrastructures. *Comput. Phys. Commun.* **2016**, *207*, 375–385.

(15) Wyce, A.; Ganji, G.; Smitheman, K. N.; Chung, C. W.; Korenchuk, S.; Bai, Y.; Barbash, O.; Le, B.; Craggs, P. D.; McCabe, M. T.; Kennedy-Wilson, K. M.; Sanchez, L. V.; Gosmini, R. L.; Parr, N.; McHugh, C. F.; Dhanak, D.; Prinjha, R. K.; Auger, K. R.; Tummino, P. J. BET inhibition silences expression of MYCN and BCL2 and induces cytotoxicity in neuroblastoma tumor models. *PLoS One* **2013**, *8* (8), e72967.

(16) Gosmini, R.; Nguyen, V. L.; Toum, J.; Simon, C.; Brusq, J. M.; Krysa, G.; Mirguet, O.; Riou-Eymard, A. M.; Boursier, E. V.; Trottet, L.; Bamborough, P.; Clark, H.; Chung, C. W.; Cutler, L.; Demont, E. H.; Kaur, R.; Lewis, A. J.; Schilling, M. B.; Soden, P. E.; Taylor, S.; Walker, A. L.; Walker, M. D.; Prinjha, R. K.; Nicodeme, E. The discovery of I-BET726 (GSK1324726A), a potent tetrahydroquinoline ApoA1 up-regulator and selective BET bromodomain inhibitor. *J. Med. Chem.* **2014**, *57* (19), 8111–8131.

(17) Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Accurate calculation of the absolute free energy of binding for drug molecules. *Chem. Sci.* **2016**, *7* (1), 207–218.

(18) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Genet.* **2010**, *78* (8), 1950–1958.

(19) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174.

(20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A., *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2004.

(21) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26* (16), 1668–1688.

(22) Laio, A.; Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.* **2008**, *71* (12), 126601.

(23) Lin, Y. L.; Aleksandrov, A.; Simonson, T.; Roux, B. An overview of electrostatic free energy computations for solutions and proteins. *J. Chem. Theory Comput.* **2014**, *10* (7), 2690–2709.

(24) Wan, S.; Knapp, B.; Wright, D. W.; Deane, C. M.; Coveney, P. V. Rapid, precise, and reproducible prediction of peptide-MHC binding affinities from molecular dynamics that correlate well with experiment. *J. Chem. Theory Comput.* **2015**, *11* (7), 3346–3356.

(25) Bhati, A. P.; Wan, S.; Wright, D. W.; Coveney, P. V. Rapid, accurate, precise and reliable relative free energy prediction using ensemble based thermodynamic integration. *J. Chem. Theory Comput.* **2016**, DOI: 10.1021/acs.jctc.6b00979.

(26) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33* (12), 889–897.

(27) Wright, D. W.; Hall, B. A.; Kenway, O. A.; Jha, S.; Coveney, P. V. Computing Clinically Relevant Binding Free Energies of HIV-1 Protease Inhibitors. *J. Chem. Theory Comput.* **2014**, *10* (3), 1228–1241.

(28) Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discovery* **2015**, *10* (5), 449–461.

(29) Swanson, J. M.; HENCHMAN, R. H.; McCAMMON, J. A. Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophys. J.* **2004**, *86*, 67–74.

(30) Genheden, S.; Ryde, U. How to obtain statistically converged MM/GBSA results. *J. Comput. Chem.* **2010**, *31* (4), 837–846.

(31) Luo, R.; David, L.; Gilson, M. K. Accelerated Poisson-Boltzmann calculations for static and dynamic systems. *J. Comput. Chem.* **2002**, *23* (13), 1244–1253.

(32) Beveridge, D. L.; Dicapua, F. M. Free-energy via molecular simulation - applications to chemical and biomolecular systems. *Annu. Rev. Biophys. Chem.* **1989**, *18*, 431–492.

(33) Bunney, T. D.; Wan, S.; Thiyagarajan, N.; Sutto, L.; Williams, S. V.; Ashford, P.; Koss, H.; Knowles, M. A.; Gervasio, F. L.; Coveney, P. V.; Katan, M. The effect of mutations on drug sensitivity and kinase activity of fibroblast growth factor receptors: A combined experimental and theoretical study. *EBioMedicine* **2015**, *2* (3), 194–204.

(34) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.

(35) Sadiq, S. K.; Wright, D. W.; Kenway, O. A.; Coveney, P. V. Accurate ensemble molecular dynamics binding free energy ranking of multidrug-resistant HIV-1 proteases. *J. Chem. Inf. Model.* **2010**, *50* (5), 890–905.

(36) Norman, G. E.; Stegailov, V. V. Stochastic theory of the classical molecular dynamics method. *Math. Models Comput. Simul.* **2013**, *5* (4), 305–333.

(37) Chodera, J. D.; Mobley, D. L. Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annu. Rev. Biophys.* **2013**, *42*, 121–142.

(38) Zhu, Y. L.; Beroza, P.; Artis, D. R. Including explicit water molecules as part of the protein structure in MM/PBSA calculations. *J. Chem. Inf. Model.* **2014**, *54* (2), 462–469.

(39) Rastelli, G.; Del Rio, A.; Degliesposti, G.; Sgobba, M. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *J. Comput. Chem.* **2010**, *31* (4), 797–810.

(40) Reynolds, C. H.; Holloway, M. K. Thermodynamics of ligand binding and efficiency. *ACS Med. Chem. Lett.* **2011**, *2* (6), 433–437.

HPC-BASED COMPACT INTERFACE QUASI-NEWTON ALGORITHM FOR FLUID-STRUCTURE INTERACTION*

ALFONSO SANTIAGO[†], MIGUEL ZAVALA-AKÉ[†], MARIANO VÁZQUEZ[‡], JUAN
CARLOS CAJAS[†], AND GUILLAUME HOUZEAUX[†]

Abstract. In this work we present a robust Compact Interface Quasi Newton (CIQN) algorithm to couple Fluid-Structure Interaction (FSI) problems. A MPI multi-code partitioned scheme, designed for computationally intensive fluid and solid problems is used. Thus, special care is taken to efficiently use the parallel computational resources, based in two main precepts. On one hand, memory allocation and computational operations for intermediate variables is avoided when possible. On the other hand, an efficient MPI communication scheme between parallel code must be set. With the proposed examples we show that the CIQN algorithm can efficiently deal with added mass instability and two-field coupled problems. The results are compared against the Aitken and Broyden relaxation schemes. A reduced simulation time, median and interquartile range for the distribution in coupling iterations, is obtained for the CIQN scheme. Finally, a parallel performance analysis is done using up to 4096 processors.

Key words. fluid structure interaction, partitioned scheme, coupling algorithm, high performance computing, biomechanics

AMS subject classifications. 65C20, 68U20, 68W10

1. Introduction. The fluid structure interaction problem (FSI) has gained great attention in the last decades due to the broad range of applications, like aerospace industry, manufacturing, wind energy production or biomechanics. These problems can be classified [?] in: class I, i.e. short-lived and small fluid displacement, as in explosions; class II, i.e. spread over time and small fluid displacements, as vibration in rockets or ultrasound; or class III, i.e. large relative motion and the interaction is dominated by the fluid, like parachute flutter or a beating heart.

From the algorithmic point of view, there are two main methods to solve this coupled problem: the monolithic scheme and partitioned scheme. In the monolithic case, one large matrix including the degrees of freedom for both fluid and solid must be inverted [13, 10, 21]. The partitioned scheme solves independently the fluid and the solid domains and exchanges the quantities of interest at certain synchronization points in the workflow [11, 27, 15, 29]. Both strategies have advantages and drawbacks. Monolithic scheme advocates claim larger application range (notably considering the densities ratio) and convergence properties [33, 13, 21]. Partitioned scheme defenders claim efficiency (notably when both fluid and structure problems are large) and flexibility [11, 27, 15, 29]. Robustness is in both cases an issue: in many cases the scheme is set as monolithic, but the resulting large algebraic system is solved iteratively by blocks. Indeed, it is usually observed that matrices coming from monolithic FSI solvers present severe convergence problems and proper preconditioning of the matrix is challenging [24, 33, 1].

If both fluid and solid problems are very large, the matrix resulting from a mono-

*Submitted to the editors August 14, 2017.

Funding: This work has been partially funded by the European Unions Horizon 2020 research and innovation programme through the project CompBioMed, grant agreement No 675451, under H2020-EU.1.4.1.3. - Development, deployment and operation of ICT-based e-infrastructures. Also, this work has been possible with the support of the Severo Ochoa program of the Spanish Government, SEV 2011 00067.

[†]Barcelona Supercomputing Center (BSC), Barcelona, Spain (alfonso.santiago@bsc.es).

[‡]Barcelona Supercomputing Center (BSC) - IIIA - CSIC.

lithic strategy could become very large. So, the kind of problems we target lead us, naturally, to a partitioned scheme. Therefore, a partitioned solver for Class III problems is presented here, which is derived from DeGroote’s [11] interface quasi Newton method (IQN), with a special care in the parallel implementation. The proposed scheme is implemented in the BSC’s in-house parallel multi-physics code Alya [20, 32, 16, 7]. We follow a multi-code strategy in which fluid and structure are solved in two different MPI-based parallel instances of Alya which interchange data in the contact surface. In [32], the authors show Alya’s almost linear scalability measured up to a hundred thousand cores and billions of elements in both fluid and solid mechanics problems. Therefore, Alya parallel multi-physics code is a natural option for the proposed partitioned scheme.

In this work, the target is to design a high-performance version of the interface quasi Newton algorithm. There are three main motivations to do so. First, as we stated before, we are looking for a partitioned solver where problems may be large on both the solid side, and the fluid side. Also, the IQN algorithm is capable to tackle with ease fluid structure interaction problems with more than one wet surface (n-field coupling) [3]. Finally it can deal with added mass instability, that may arise when the fluid and solid densities ratio is very close to one. While the monolithic schemes are supposedly well-suited to cope with this instability, partitioned schemes need to be carefully designed. This features are critical when modeling biomedical applications like the human heart, where two fluid chambers are interacting through a flexible thin wall. In this type of applications we have to deal with large problems on the solid side (3D finite-strain solid mechanics with non-linear anisotropic material models in blood vessels and contracting cardiac muscle, with deployed stents or flow diverters) and the fluid side (solving all the scales of highly transient flow in arterial aneurysms or cardiac chambers), multiple interfaces and added mass instability. Normally, each problem is solved using different numerical strategies and the partitioned scheme allows us to use the best solver for each one.

The rest of the paper is organized as follows. A brief description of the governing equations is outlined in [section 2](#). Then, an extensive description of the proposed algorithm is elaborated in [section 3](#). Three numerical experiments and a scalability test are presented and discussed in [section 4](#). Final conclusions are explained in [section 5](#).

Index notation convention. In order to better describe the implementation, the Einstein convention on repeated indices will be followed. This notation is particularly helpful on computational grounds, because it simultaneously describes the mathematics, the physics and the computational implementation aspects. Depending on the context, whether the problem is continuum or discretized, we use different labelling sets. For the continuum problem, the convention on repeated indices is the usual one, with indices labelling space dimensions. On the other hand, once the continuum problem is discretized and we obtain a system matrix, we identify different labelling sets. The lowercase greek alphabet $\alpha = 1, \dots, p$ labels the total number of degrees of freedom p , i.e. the matrix rows. The lowercase latin alphabet labels the matrix columns, $i = 1, \dots, q - 1$ where q is the last stored iteration. In the CIQN scheme matrices are not square, with $p \gg q$. Additionally, a capital latin subindex labels the IQN iteration $I = 1, \dots, q - 1$, where q is again the last stored iteration and, to simplify notation, the current iteration is not explicitly indexed. A final rule is how those indices operate: only those of the same kind contract. For instance, $Q_{\alpha i}^{I-1}$

is the Q matrix for iteration $I - 1$ with rows labelled α and columns i . When this matrix is multiplied by a certain vector B_i , it results in a given vector A_α^{I-1} :

$$(1) \quad A_\alpha^{I-1} = Q_{\alpha i}^{I-1} B_i = \sum_i^q Q_{\alpha i}^{I-1} B_i,$$

where latin index i contract.

2. Governing equations.

2.1. Physical and numerical modelling. The Newtonian fluid is modelled with incompressible Navier-Stokes equations using an Arbitrary Lagrangian-Eulerian (ALE) formulation:

$$(2) \quad \rho^f \frac{\partial u_i}{\partial t} + \rho^f (u_j - u_j^m) \frac{\partial u_i}{\partial x_j} + \frac{\partial}{\partial x_j} \left[+p\delta_{ij} - \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right] = +\rho^f f_i$$

$$(3) \quad \frac{\partial u_i}{\partial x_i} = 0,$$

where μ is the viscosity of the fluid, ρ^f the density, u_i the velocity, p is the mechanical pressure, f_i the force term and u_j^m is the fluid domain velocity. The numerical model is based on the Finite Element Method, using the Variational Multiscale Method [19] to stabilize convection and pressure. The resulting system is solved through a velocity-pressure splitting strategy. The Schur complement is obtained and solved with an Orthomin(1) algorithm [16]. To do so, the momentum equation is solved twice using GMRES method and the continuity equation is solved with the Deflated Conjugate Gradient algorithm. Fluid mechanics boundary conditions at the contact surface is set through the nodal speed by the ALE solver.

For the ALE formulation, the technique used is proposed in [6]. Mesh movement is solved through a Laplacian equation

$$(4) \quad \frac{\partial}{\partial x_j} \left([1 + \alpha^e] \frac{\partial b_i}{\partial x_j} \right) = 0,$$

where b_i are the components of the displacement in each point for the domain. The factor α^e is a diffusive term that, once discretized, is selected using the elements volume as $\alpha^e = V_{max}/V_e$ and controls the mesh distortion. In the last expression V_{max} and V_e are the maximum and current element size in the grid. This equation is solved with a Deflated Conjugate Gradient algorithm [26]. In this way, while small elements remain almost undeformed, large elements suffer the largest deformations. ALE boundary conditions at the contact surface is set through the nodal displacement from solid mechanics problem.

Finally, in Alya the solid mechanics problem is solved following a transient scheme and using a total Lagrangian formulation in finite strains [8]. The system is solved using GMRES method with Newmark time integration. The displacement d_i form of the linear momentum balance can be modeled as:

$$(5) \quad \rho^s \frac{\partial^2 d_i}{\partial t^2} = \frac{\partial P_{iJ}}{\partial X_J} + \rho^s B_i,$$

where ρ^s is the initial density of the body, B_i represents the body forces and P_{iJ} is the nominal stress tensor. Solid mechanics boundary conditions at the contact surface is set through the nodal forces from the fluid mechanics problem.

2.2. Coupling conditions. Let us label with “nsi” (after Navier-Stokes Incompressible) and “sld” the fluid and solid sides of a coupled fluid-structure interaction problem. Then, at the contact or “wet” surface, displacement and stresses must be continuous:

$$(6) \quad \text{nsi } d_i^{\Gamma_c} = \text{sld } d_i^{\Gamma_c}$$

$$(7) \quad \text{nsi } \sigma_{ij}^{\Gamma_c} = \text{sld } \sigma_{ij}^{\Gamma_c},$$

where $\text{nsi } d_i^{\Gamma_c}$ and $\text{sld } d_i^{\Gamma_c}$ are the deformation in the contact boundary for the fluid and for the solid respectively; and $\text{sld } \sigma_{ij}^{\Gamma_c}$ and $\text{nsi } \sigma_{ij}^{\Gamma_c}$ and are the stresses in the contact boundary. Both equations represent continuity of displacements and forces respectively. Finally, for Navier-Stokes $f_i = \text{nsi } \sigma_{ij}^{\Gamma_c} \cdot n_j$ where f_i is the traction in (2) and n_j is the normal to the surface.

It has been observed that when solid and fluid densities are similar and fluid and solid behave as incompressible or slightly compressible, the so called added mass instability may appear [12, 9, 14]. Also, this instability aggravates when any of these conditions happens:

- Time step decreases
- ρ^f / ρ^s ratio increases
- Viscosity of the fluid increase
- Rigidity of the solid decreases
- Temporal discretization order increases

In this paper, we focus on biomechanics applications. As tissue and blood have very similar densities, the added mass effect can affect convergence of the solution procedure in partitioned schemes. In order to deal with this instability, non-trivial relaxation schemes were developed, such as the Aitken algorithm [25, 5], the Broyden scheme [2] or the interface quasi Newton scheme [11].

If $d_\alpha^{I+1} = SLD(f_\alpha)$ is the displacement coming from the structure solver and $f_\alpha^{I+1} = NSI(d_\alpha)$ is the force coming from the fluid solver, the Dirichlet-Neumann fixed point algorithm for the coupled problem can be written as:

$$(8) \quad d_\alpha^{I+1} = SLD(NSI(d_\alpha)),$$

where I is the coupling iteration. The fixed point equation (8) can be solved through either Gauss-Seidel or Jacobi schemes [4]. In the former, the solvers run one after the other, improving convergence but penalizing performance. In the later both solvers run at the same time, improving performance but penalizing convergence [28]. In this work we use the Gauss-Seidel scheme (see [algorithm 1](#)) to ease convergence of the problem.

```

while Time loop do
   $d_\alpha = d_\alpha^{ini}$ 
  while Coupling loop do
     $f_\alpha = NSI(d_\alpha)$ 
     $d_\alpha = SLD(f_\alpha)$ 
     $d_\alpha^{I+1} = \varphi_{GS}(d_\alpha)$ 
  end
end

```

Algorithm 1: Gauss-Seidel scheme. φ_{GS} represents the relaxation algorithm.

3. Numerical techniques. In this section we firstly give a general description of the algorithm. Next, we propose several improvements to obtain the Compact Interface Quasi Newton (CIQN) iterative coupling scheme, especially designed for large-scale problems, where both solid and fluid involve similar (and potentially very large) computational effort.

3.1. General overview of the algorithm. In this context, the first implementation of the Interface Quasi Newton (IQN) algorithm is described in [11], which can be considered an improvement of the Aitken algorithm. Whereas the Aitken algorithm can be understood as a quasi Newton scheme where the Jacobian is enforced to be a scalar, in the IQN the Jacobian is approximated by a field defined in the contact surface and depending on the local residual variation over a given number of iterations [30].

Let us restate (8) as $\tilde{x}_\alpha = H(x_\alpha)$. If the residual at each iteration is:

$$(9) \quad r_\alpha = H(x_\alpha) - x_\alpha = \tilde{x}_\alpha - x_\alpha,$$

the problem is converged when $r_\alpha = 0$. To solve this problem using Newton-Raphson iterations, the derivative with respect to the variable must be computed:

$$(10) \quad \frac{\partial r_\alpha}{\partial x_\beta} \Delta x_\beta = -r_\alpha.$$

Solving this system, Δx_β is obtained and the next iteration unknown is computed as:

$$(11) \quad x_\alpha^{I+1} = x_\alpha + \Delta x_\alpha.$$

Generally $\partial r_\alpha / \partial x_\beta$ cannot be computed. But, the inverse Jacobian fulfills the secant equation:

$$(12) \quad \left(\frac{\partial r_\alpha}{\partial x_\beta} \right)^{-1} V_{\alpha i} \approx W_{\beta i},$$

where:

$$(13) \quad V_{\alpha i} = [\Delta r_\alpha^{I-1}, \Delta r_\alpha^{I-2}, \dots, \Delta r_\alpha^0]$$

$$(14) \quad W_{\alpha i} = [\Delta \tilde{x}_\alpha^{I-1}, \Delta \tilde{x}_\alpha^{I-2}, \dots, \Delta \tilde{x}_\alpha^0].$$

At each iteration I the columns of the increment matrices are built as follows:

$$(15) \quad \Delta r_\alpha^{I-1} = r_\alpha^{I-1} - r_\alpha \quad \text{for matrix } V_{\alpha i}$$

$$(16) \quad \Delta \tilde{x}_\alpha^{I-1} = \tilde{x}_\alpha^{I-1} - \tilde{x}_\alpha \quad \text{for matrix } W_{\alpha i}$$

where r_α and \tilde{x}_α are the values in the actual iteration and r_α^{I-1} and \tilde{x}_α^{I-1} are the values in the saved past iterations ($I = 1, \dots, q - 1$). The residual increment of the current iteration is approximated as a linear combination of the previous residuals increments:

$$(17) \quad \Delta r_\alpha = V_{\alpha i} \lambda_i,$$

where $\lambda_i \in \mathbb{R}^{q \times 1}$ is the solution of the optimization problem $\| \Delta r_\alpha - V_{\alpha i} \lambda_i \|$ described in [34]. It is worth to remark that matrices (13) and (14) are far from being square, because while the number of rows corresponds to the total number of degrees of

freedom, the number of columns is the stored number of Newton-Raphson iterations done.

Firstly the matrix $V_{\alpha i}$ is decomposed by a QR decomposition, where an orthogonal matrix $Q_{\alpha\beta} \in \mathbb{R}^{p \times p}$ and an upper triangular $U_{\alpha i} \in \mathbb{R}^{p \times q}$ are obtained:

$$(18) \quad V_{\alpha i} = Q_{\alpha\beta} U_{\beta i}.$$

As $U_{\alpha i}$ is upper triangular, only its first q rows are different from zero, which means that only the first q columns of the matrix Q are relevant to solve the system. Then a modified QR decomposition can be built with $U_{ij} \in \mathbb{R}^{q \times q}$ and $Q_{\alpha i} \in \mathbb{R}^{p \times q}$ such that:

$$(19) \quad V_{\alpha i} = Q_{\alpha k} U_{ki},$$

highly reducing the amount of memory and computing effort required, as described below in [subsections 3.2](#) and [3.3](#). After this decomposition, the vector λ_i can be built by backsubstitution of the upper triangular matrix U_{ij} :

$$(20) \quad U_{ij} \lambda_j = Q_{\alpha i} \Delta r_{\alpha}.$$

As $Q_{\alpha i}$ is orthogonal, the inverse is equal to the transpose, avoiding the inversion of this matrix. Also, as $\Delta r_{\alpha} = r_{\alpha}^{I-1} - r_{\alpha}$ and the objective is to get $\Delta r_{\alpha} = 0_{\alpha} - r_{\alpha}$, we can say:

$$(21) \quad U_{ij} \lambda_j = -Q_{\alpha i} r_{\alpha}.$$

Once λ_i is computed, the increment of the unknown Δx_{α} can be computed as $\Delta x_{\alpha} = W_{\alpha i} \lambda_i$, and the update of the unknown as:

$$(22) \quad x_{\alpha}^{I+1} = \tilde{x}_{\alpha} + W_{\alpha i} \lambda_i.$$

The interface quasi Newton coupling technique is summarized in [algorithm 2](#). In the first step, an initial guess for both fluid and solid solvers is computed and a first residual r_{α}^0 is computed. In the first iteration, a fixed relaxation ω_0 is used. After the first iteration, the loop continues until convergence is achieved. Once computed the second iteration, the increment matrices $V_{\alpha i}$ and $W_{\alpha i}$ can be built.

Each matrix has p rows (the total number of contact d.o.f.), and q columns (the number of past increments saved). Generally (see [subsection 3.2](#)) p is much larger than q . It is worth to remark that q is set as the number of *stored* iterations, being usually lower than the *performed* iterations, i.e. only the q newer past increments are used, discarding the older ones.

```

1 For each time step, solve:
2  $x_\alpha^0 = x_\alpha^{ini}$ 
3  $\tilde{x}_\alpha^0 = H(x_\alpha^0)$ 
4  $r_\alpha^0 = \tilde{x}_\alpha^0 - x_\alpha^0$ 
5 update  $x_\alpha^1 = \tilde{x}_\alpha^1 + \omega_0 r_\alpha^0$ 
   while problem not converged do
6    $\tilde{x}_\alpha = H(x_\alpha)$ 
7    $r_\alpha = \tilde{x}_\alpha - x_\alpha$ 
8   build  $V_{\alpha i} = [\Delta r_\alpha^{I-1}, \dots, \Delta r_\alpha^0]$ ; with  $\Delta r_\alpha^I = r_\alpha^I - r_\alpha$ 
9   build  $W_{\alpha i} = [\Delta \tilde{x}_\alpha^{I-1}, \dots, \Delta \tilde{x}_\alpha^0]$ ; with  $\Delta \tilde{x}_\alpha^I = \tilde{x}_\alpha^I - \tilde{x}_\alpha$ 
10  decompose  $V_{\alpha j} = Q_{\alpha i} U_{ij}$  (by QR decomposition)
11  solve  $U_{ij} \lambda_j = -Q_{\alpha i} r_\alpha$ 
12  update  $x_\alpha^{I+1} = \tilde{x}_\alpha + W_{\alpha i} \lambda_i$ 
   end

```

Algorithm 2: Interface quasi-newton algorithm overview.

To develop an efficient yet accurate implementation of the IQN algorithm, we should focus in the QR decomposition, which involves complex matrix operations.

3.2. Compact matrix-free QR decomposition algorithm. The QR decomposition by Householder reflections can be seen as a change of base, from a linearly independent set of vectors sorted in a matrix, to an orthogonal base of the original matrix. The goal is to obtain the orthogonal matrix $Q_{\alpha\beta}$ and the upper triangular matrix $U_{\alpha i}$, as follows:

$$(23) \quad Q_{\alpha\epsilon} = {}^1B_{\alpha\beta} {}^2B_{\beta\gamma} \dots {}^qB_{\gamma\epsilon}$$

$$(24) \quad U_{\alpha i} = {}^qB_{\alpha\beta} \dots {}^2B_{\beta\gamma} {}^1B_{\gamma\epsilon} V_{\epsilon i},$$

where $B_{\alpha\beta}$ are intermediate matrices obtained during the iterative decomposition. At each of these iterations, the matrix $V_{\alpha i}$ is modified column by column. We use a left supraindex to identify the corresponding iteration. We do not apply the repeated index convention to the left supraindices. This structure can be considered as a set of q ordered vectors:

$$(25) \quad V_{\alpha i} = \left[\begin{array}{c|c|c|c} \begin{bmatrix} v_{11} \\ v_{21} \\ \vdots \end{bmatrix} & \begin{bmatrix} v_{12} \\ v_{22} \\ \vdots \end{bmatrix} & \dots & \begin{bmatrix} v_{1q} \\ v_{2q} \\ \vdots \end{bmatrix} \end{array} \right] = [v_{\alpha 1}, v_{\alpha 2}, \dots, v_{\alpha q}] = {}^1V_{\alpha i}.$$

It is worth to remind that the total number of stored iterations q , i.e. the rows of $V_{\alpha i}$, will be always much lower than the total number of degrees of freedom p , i.e. the columns of $V_{\alpha i}$.

The QR decomposition iteratively makes each column orthogonal to the original base and to each other column in the matrix. The QR decomposition starts iteration j with a matrix ${}^jV_{\alpha i}$ obtained with data from iteration $j - 1$. To decompose the $j - th$ column of ${}^jV_{\alpha i}$, a unitary vector u_α has to be built:

$$(26) \quad u_\alpha = \frac{n_\alpha}{\|n_\alpha\|} \quad \text{with,} \quad n_\alpha = v_\alpha - \|v_\alpha\| {}^j e_\alpha,$$

where v_α is the column to decompose and ${}^j e_\alpha$ is a unitary vector with $j - th$ position

equal to 1 and to 0 otherwise. Then,

$$(27) \quad {}^j B_{\alpha\beta}^* = \delta_{\alpha\beta} - 2u_\alpha u_\beta$$

is the Householder matrix associated to the original plane, and $\delta_{\alpha\beta}$ is the identity matrix. If the matrix ${}^j V_{\alpha i}$ is premultiplied by ${}^j B_{\alpha\beta}^*$, a new matrix ${}^j B_{\alpha\beta}^* {}^j V_{\beta i}$ is obtained (recall that repeated left supraindices do not sum). The resulting matrix is upper triangular in all the j first columns; and dense everywhere else. For the decomposition of the first column of the matrix ${}^1 V_{\alpha\beta}$ this would look like:

$$(28) \quad {}^1 B_{\alpha\beta}^* {}^1 V_{\alpha i} = \begin{bmatrix} \|v_{\alpha 1}\| & \cdots & \cdots & \cdots \\ 0 & & & \\ \vdots & & {}^2 V_{\beta j} & \\ 0 & & & \end{bmatrix}$$

A new submatrix ${}^{j+1} V_{\beta j}$ is therefore defined as the remainder of ${}^1 V_{\alpha i}$ after erasing the first column and row. This process can be repeated until the initial matrix becomes upper triangular.

Once the algorithm is computed for all the columns of the original matrix ${}^1 V_{\alpha i}$, a set of q gradually smaller matrices ${}^1 B_{\alpha i}^* \in \mathbb{R}^{p \times p}$, ${}^2 B_{\alpha i}^* \in \mathbb{R}^{(p-1) \times (p-1)}$... ${}^j B_{\alpha i}^* \in \mathbb{R}^{(p-(j-1)) \times (p-(j-1))}$... ${}^q B_{\alpha i}^* \in \mathbb{R}^{1 \times 1}$ are obtained. In order to properly compute (23) and (24), matrices ${}^j B_{\alpha i}$ are filled with the identity:

$$(29) \quad {}^j B_{\alpha i} = \begin{bmatrix} I_{ij} & 0 \\ 0 & {}^j B_{\alpha i}^* \end{bmatrix}$$

where $I_{ij} \in \mathbb{R}^{j-1 \times j-1}$. Finally, the matrix $V_{\alpha i}$ is decomposed in an upper triangular U_{ij} and an orthogonal matrix $Q_{\alpha i}$, such that $V_{\alpha i} = Q_{\alpha i} U_{ij}$. The process is described in [algorithm 3](#).

```

1  ${}^1 V_{\alpha i} = V_{\alpha i}$ 
  for  $j=1 \dots q$  do
2   choose  $v_\alpha = {}^j V_{\alpha i}$  with  $\alpha = j \dots p$  and  $i = j$ 
3    $n_\alpha = v_\alpha - \|v_\alpha\| e_\alpha$ 
4    $u_\alpha = n_\alpha / \|n_\alpha\|$ 
5    $B_{\alpha\beta}^* = I - 2 u_\alpha u_\beta$ 
6    ${}^{j+1} V_{\alpha i} = B_{\alpha\beta} V_{\beta i}$ 
  end
7  $Q_{\alpha\epsilon} = {}^1 B_{\alpha\beta} {}^2 B_{\beta\gamma} \dots {}^q B_{\gamma\epsilon}$ 
8  $U_{\alpha i} = {}^q B_{\alpha\beta} {}^{q-1} B_{\beta\gamma} \dots {}^1 B_{\gamma\epsilon} V_{\epsilon i}$ 

```

Algorithm 3: overview of the QR decomposition algorithm.

In order to improve computing and memory cost, we propose some modifications, especially targeted to solving large-scale problems. As ${}^j B_{\alpha\beta}$ is obtained by the relation (27), the product ${}^j B_{\alpha\beta} {}^j V_{\beta i}$ can be expanded as:

$$(30) \quad {}^j B_{\alpha\beta} {}^j V_{\beta i} = (\delta_{\alpha\beta} - 2u_\alpha u_\beta) {}^j V_{\beta i} = {}^j V_{\alpha i} - 2u_\alpha u_\beta {}^j V_{\beta i}.$$

In this way, instead of computing and storing ${}^j B_{\alpha\beta}$ of size $p \times p$ for each iteration j , we compute and save q vectors u_α of size p . Moreover, $Q_{\alpha\beta}$, which is used in

the backsubstitution (see (21)), is never completely computed. Instead, the vector $-Q_{\alpha i}r_\alpha$ is directly computed, using a strategy similar to (30). The difference here, is that $Q_{\alpha\epsilon} = {}^1B_{\alpha\beta} \cdots {}^qB_{\gamma\epsilon}$, so after computing ${}^qB_{\alpha i}r_\alpha$, the rest of the matrices $B_{\alpha\beta} = \delta_{\alpha\beta} - 2u_\alpha u_\beta$ have to be premultiplied. For the first product ${}^qB_{\alpha\beta}r_\beta$ the expansion is:

$$(31) \quad {}^qB_{\alpha\beta}r_\alpha = (\delta_{\alpha\beta} - 2 {}^q u_\alpha {}^q u_\beta)r_\beta = r_\beta - 2 {}^q u_\alpha {}^q u_\beta r_\beta.$$

As the result of (31) is a vector, the rest of the premultiplications are computed similarly. Note that this algorithm not only avoids building the intermediate matrices $B_{\alpha\beta}$ but also the construction of the final matrix $Q_{\alpha\beta}$ of the QR decomposition. In this way, matrices B in (23) and (24) are never completely computed and stored. The proposed changes are described in algorithm 4.

```

1   ${}^1V_{\alpha i} = V_{\alpha i}$ 
   for  $j=1\dots q$  do
2     if  $j > 1$  then
        |  ${}^{j+1}V_{\alpha i} = {}^1V_{\beta i} - 2u_\alpha u_\beta {}^1V_{\beta i}$ 
       end
3     choose  $v_\alpha = {}^jV_{\alpha i}$  with  $\alpha = j\dots p$  and  $i = j$ 
4      $n_\alpha = v_\alpha - \|v_\alpha\|e_\alpha$ 
5      ${}^j u_\alpha = n_\alpha / \|n_\alpha\|$ 
   end
6   $U_{ij} = (\delta_{i\gamma} - 2 {}^q u_i {}^q u_\gamma) \cdots (\delta_{\gamma\alpha} - 2 {}^2 u_\gamma {}^2 u_\alpha) ({}^1V_{\alpha j} - 2 {}^1 u_\alpha {}^1 u_\beta {}^1V_{\beta j})$ 
7   $-Q_{\alpha i}r_\alpha = -(\delta_{i\gamma} - 2 {}^1 u_i {}^1 u_\gamma) (\delta_{\gamma\beta} - 2 {}^2 u_\gamma {}^2 u_\beta) \cdots (r_\beta - 2 {}^q u_\beta {}^q u_\alpha r_\alpha)$ 

```

Algorithm 4: Compact matrix-free QR decomposition.

3.3. Compact matrix-free QR decomposition parallel implementation.

When using domain decomposition methods as parallelization strategy [17], work and data are distributed among MPI tasks in partitions of the original discretization. Therefore, the contact surface will also be distributed among the MPI tasks and so the increment matrix $V_{\alpha i}$.

The proposed algorithm is a collection of matrix-vector and vector-vector products restricted to the contact or wet surface. An efficient parallelization of the algorithm requires a proper point-to-point MPI communication and a fusion of the compact QR decomposition and the backsubstitution step. Therefore, the input of the parallel algorithm will be the $V_{\alpha i}$ matrix and the residuals vectors to operate in the backsubstitution (see (21)), and the output will be the coefficient vector α_i . This is what we call the *Compact IQN algorithm* or CIQN. The whole sequence of steps is described in algorithm 7.

The mesh partitioner [23] divides the mesh minimizing the area between subdomains but without any requirements on the wet surface, as depicted in Figure 1. Then, in order to derive the upper triangular matrix from $V_{\alpha i}$ in an efficient way, we firstly look for a partition with a number of nodes on the wet surface that is at least equal to the amount of iterations stored q , which is usually a small number. We call this first partition the “leader” partition. With this information we renumber matrix $V_{\alpha i}$ so the first rows correspond to the leader partition and we start the algorithm.

Two critical operations to parallelize are $B_{\alpha\beta} {}^jV_{\beta i}$ and $Q_{\alpha i}r_\alpha$. As $B_{\alpha\beta} {}^jV_{\beta i}$ is

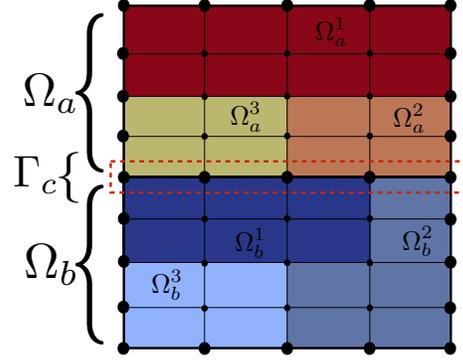


Fig. 1: Physical subdomains Ω_a and Ω_b in contact by the wet surface Γ_c . Each physical subdomain is subdivided in three computational subdomains (partitions). The wet surface, and therefore the vectors in matrix $V_{\alpha i}$ can be distributed along several partitions.

obtained by the relation (30), the multiplication sequence can be expanded as:

$$(32) \quad (\delta_{\alpha\beta} - 2u_\alpha u_\beta)^j V_{\beta i} = \delta_{\alpha\beta} {}^j V_{\beta i} - 2u_\alpha u_\beta {}^j V_{\beta i} = {}^j V_{\alpha i} - 2u_\alpha (u_\beta {}^j V_{\beta i}).$$

Now the product $u_\alpha {}^j V_{\alpha i}$ can be computed first, then compute $u_\alpha (u_\beta {}^j V_{\beta i})$ and finally subtract $I_{\alpha\beta} {}^j V_{\beta i} - 2u_\alpha u_\beta {}^j V_{\beta i}$. To compute U_{ij} (see (24)) we proceed similarly, but this time the first operation is $(\delta_{\alpha\beta} - 2{}^1 u_\alpha {}^1 u_\beta) V_{\beta i}$ and the rest of the matrices $(\delta_{\alpha\beta} - 2u_\alpha u_\beta)$ are premultiplied. Something similar happens with $Q_{\alpha i} r_\alpha$. The first multiplication, detailed in (31) can be expanded as:

$$(33) \quad (\delta_{\alpha\beta} - 2{}^q u_\alpha {}^q u_\beta) r_\beta = \delta_{\alpha\beta} r_\beta - 2{}^q u_\alpha {}^q u_\beta r_\beta = r_\alpha - 2{}^q u_\alpha ({}^q u_\beta r_\beta)$$

The product ${}^q u_\beta r_\beta$ is firstly computed and then $r_\alpha - 2{}^q u_\alpha ({}^q u_\beta r_\beta)$. The resulting vector is multiplied by $\delta_{\alpha\beta} - 2{}^{(q-1)} u_\alpha {}^{(q-1)} u_\beta$ until $\delta_{\alpha\beta} - 2{}^1 u_\alpha {}^1 u_\beta$. In this way, matrices ${}^q B_{\alpha i}$ are never completely computed, we only compute those entries that are needed. The steps to compute (32) and (33) are depicted in algorithms 5 and 6 respectively.

```

function SMat_x_mat()
1 | Data:  $v_\alpha$  and  $V_{\alpha i}$ 
2 | Result:  $AUX_{\alpha i}$ 

3 |  $aux_i = v_\alpha V_{\alpha i}$ 
4 | MPI_SUM(aux_i)
5 |  $AUX_{\alpha i} = V_{\alpha i} - 2v_\alpha aux_i$ 

6 | return  $AUX_{\alpha i}$ 

```

Algorithm 5: Algorithm for computing the product $B_{\alpha\beta} V_{\beta i} = (\delta_{\alpha\beta} - 2u_\alpha u_\beta)^j V_{\beta i}$. Line 3 is a vector-matrix product and line 5 has a vector-vector product and a subtraction.

```

function SMat_x_vec()
1 | Data:  $v_\alpha$  and  $r_\alpha$ 
2 | Result:  $aux_\alpha$ 

3 |  $a = v_\alpha r_\alpha$ 
4 | MPI_SUM(a)
5 |  $aux_\alpha = v_\alpha - 2v_\alpha a$ 

6 | return  $aux_\alpha$ 

```

Algorithm 6: Algorithm for computing the product $B_{\alpha\beta} r_\beta = (\delta_{\alpha\beta} - 2u_\alpha u_\beta)^j r_\beta$. Line 3 is a vector-vector product and line 5 has a scalar product and a subtraction.

The backsubstitution is computed once U_{ij} and $Q_{\alpha i} r_{\alpha}$ are built in parallel. As all the non-zero elements of the upper triangular matrix U_{ij} are located in the leader partition, this operation is computed only in this MPI task, with no extra communication. Finally the λ_i vector is scattered to the rest of subdomains, obtaining the result of the parallel algorithm. The complete CIQN algorithm is depicted in [algorithm 7](#). We remark again that thanks to the fusion between the QR decomposition and the preoperations for the backsubstitution, intermediate matrices (such as ${}^q B_{\alpha i}$) are never completely computed and stored.

4. Numerical experiments. In this section, we present a validation case, three numerical examples, and a parallel performance analysis. [Subsection 4.2](#) shows the behavior of the coupling schemes in a simple 3D cases. In [subsections 4.3](#) and [4.4](#) the algorithms are tested in a more complex situation where two wet surfaces have to be coupled. The case in [subsection 4.3](#) seems simple, but the very different dynamics of the wet surfaces make it computationally challenging. The case in [subsection 4.4](#) proves the performance in a 3D biomechanically inspired geometry.

In the three cases, the CIQN algorithm is compared against the Aitken (atk) and a Broyden (brd) relaxation schemes. The algorithms are used to relax force and displacement separately. Also two different numbers of saved iterations are tested for the CIQN algorithm.

Continuity and stresses are imposed in the contact boundary (see [subsection 2.2](#)). No slip condition ($\mathbf{v} = \mathbf{0}$) is imposed everywhere in the fluid except for the output and the input. The output boundary condition is stress equal to zero weakly imposed. The values for the input fluid boundary condition are detailed in [Table 1](#) as well as the bulk properties. For the solid domain, the geometry is fixed everywhere except for the contact boundaries.

In all the cases, the tolerances for the non-linear solvers and the coupling where set to 1E-6. A typical convergence of the coupling for the different schemes is shown on [Figure 2](#). All the experiments were run in Marenstrum III supercomputer, based on Intel Sandy-Bridge E5-2670 processors, with 16 cores per node, 2 Gb/core of RAM and Infiniband interconnection.

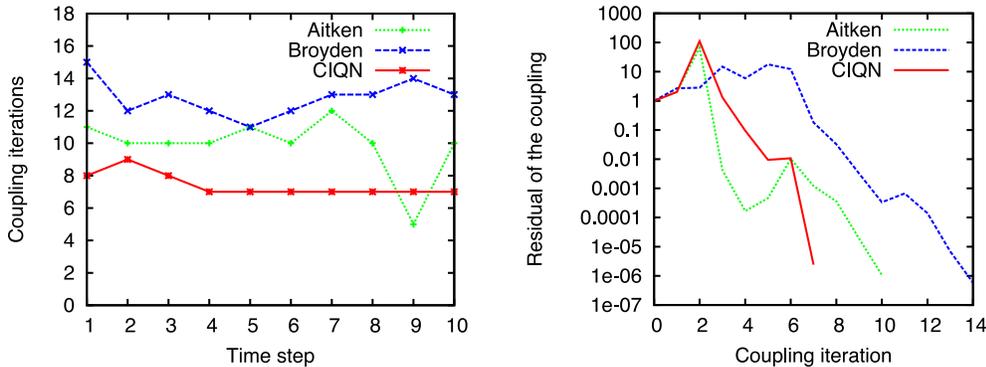


Fig. 2: Typical convergence for the coupling. Left: number of coupling iterations for the first 10 time steps. Right: residue of the coupling for the first time step.

For the experiments in [subsections 4.2](#) to [4.4](#), results are presented similarly. A table summarizes the results for all the possible combinations: Aitken, Broyden and

```

1 Initialize problem
2 Chose leader partition
   while not the last time step do                                     ▷ time loop
3    $x_\alpha^0 = x_\alpha^{ini}$ 
4    $\tilde{x}_\alpha^0 = H(x_\alpha^0)$ 
5    $r_\alpha^0 = \tilde{x}_\alpha^0 - x_\alpha^0$ 
6   update  $x_\alpha^1 = \tilde{x}_\alpha^1 + \omega_0 r_\alpha^0$ 
   while problem not converged do
7      $\tilde{x}_\alpha = H(x_\alpha)$ 
8      $r_\alpha = \tilde{x}_\alpha - x_\alpha$ 
9     build  $V_{\alpha i} = [\Delta r_\alpha^{I-1}, \dots, \Delta r_\alpha^0]$ ; with  $\Delta r_\alpha^I = r_\alpha^I - r_\alpha$ 
10    build  $W_{\alpha i} = [\Delta \tilde{x}_\alpha^{I-1}, \dots, \Delta \tilde{x}_\alpha^0]$ ; with  $\Delta \tilde{x}_\alpha^I = \tilde{x}_\alpha^I - \tilde{x}_\alpha$ 
11     ${}^1V_{\alpha i} = V_{\alpha i}$ 
    for  $j=1\dots q$  do                                               ▷ QR decomposition loop
12      if  $j > 1$  then
        |  ${}^{j+1}V_{\alpha i} = {}^1V_{\beta i} - 2u_\alpha u_\beta {}^1V_{\beta i}$  (as algorithm 5)
      end
        if I am the leader then
13       |  $v_\alpha = {}^jV_{\alpha i}$  with  $\alpha = j\dots p$  and  $i = j$  (cropped)
        else
14       |  $v_\alpha = {}^jV_{\alpha i}$  with  $\alpha = 1\dots p$  and  $i = j$  (not cropped)
        end
15       parallel compute  $\|v_\alpha\|$ 
        if I am the leader then
16       |  $n_\alpha = v_\alpha - \|v_\alpha\| {}^j e_\alpha$ 
        end
17       parallel compute  $\|n_\alpha\|$ 
18        $u_\alpha = n_\alpha / \|n_\alpha\|$ 
        end
19        $U_{ij} = (\delta_{i\gamma} - 2 {}^q u_i {}^q u_\gamma) \dots ({}^1V_{\alpha j} - 2 {}^1u_\alpha {}^1u_\beta {}^1V_{\beta j})$  (as algorithm 5)
20        $-Q_{\alpha i} r_\alpha = -(\delta_{i\gamma} - 2 {}^1u_i {}^1u_\gamma) \dots (r_\beta - 2 {}^q u_\beta {}^q u_\alpha r_\alpha)$  (as algorithm 6)
        if I am the leader then
21       | backsubstitute  $U_{ij} \lambda_j = -Q_{\alpha i} r_\alpha$ 
        end
22        $x_\alpha^{I+1} = \tilde{x}_\alpha + W_{\alpha i} \lambda_i$                                ▷ variable update
    end
  end
end

```

Algorithm 7: Compact Interface quasi-Newton algorithm.

CIQN, relaxed in displacement and force (atkd, atkf, brdd, brdf, ciqnd, ciqnf). For CIQN we also show the effect in increasing the number of past iterations used (rank 5 and rank 20). For each coupling strategy we show: the average coupling iterations; the computing time per time iteration (in [s]); the average subiterations per time step for the fluid (momentum and continuity) and the solid mechanics problems; and the average solver iteration per time step. For each case, five graphs refine the information in the tables. One of them show the accumulated solver iterations for the

Table 1: Domain dimensions, spatial and time discretization, bulk properties and boundary condition for the numerical experiments.

Property	Flexible bar	Tube expansion	two-field 2D	two-field 3D
fluid nod./elem.	2747/5245	836/3900	989/1838	20898/10187
solid nod./elem.	279/430	1320/6000	144/188	16949/76280
Time step [s]	0.1	4E-4	0.1	0.1
$\rho_f [g/cm^3]$	1	1	1	1
$\mu [Poise]$	0.04	0.03	0.01	0.03
$\rho_s [g/cm^3]$	1	1	10	1
$E [Baryes]$	3E5	3E7	2E5	1.5E4
$\nu [1]$	0.25	0.3	0.3	0.3
$v [cm/s]$	$v_x = 0.1$	$v_y = 30$	$v_y = -1.0$	$v_z = 1.0/\sin(2\pi t)$

best combination in each coupling algorithm family. Later, the iterations for all the tested schemes are detailed, aggregating the results by coupling scheme. Finally, a candle plot shows the distribution for the coupling iterations detailing the median and the interquartile range, a useful graph to evaluate the performance of the different schemes.

4.1. Method validation. The numerical method is validated with the benchmark **FSI3** proposed in section 4.3 of [22]. Results are represented in Figure 3. The obtained amplitude and frequency on the quasi-periodic period are $A_x = -2.60 \times 10^{-3} \pm 2.40 \times 10^{-3} [f = 10.8]$ and $A_y = 2.3 \times 10^{-3} \pm 33.7 \times 10^{-3} [f = 5.4]$ in concordance with the $A_x = -2.69 \times 10^{-3} \pm 2.53 \times 10^{-3} [f = 10.9]$ and $A_y = 1.48 \times 10^{-3} \pm 34.38 \times 10^{-3} [f = 5.3]$ obtained in the cited experiment¹. This validation is also presented in [31] using Alya as simulation tool but with a different coupling strategy, with similar results.

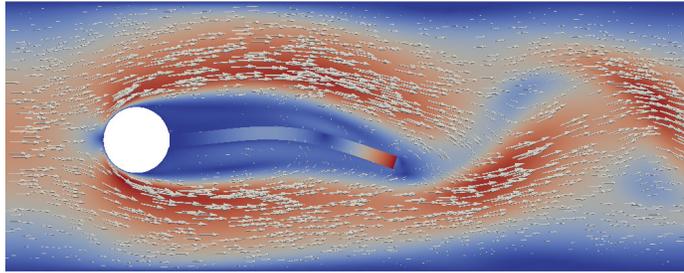


Fig. 3: Method Validation. Portion of the domain proposed by [22] in time $t = 3 [s]$. Deformation is represented on the bar and velocity field in the fluid domain.

4.2. Tube expansion. The problem setup is shown in Figure 4 and in Table 1. 100 time steps are computed for all the possible combinations in the relaxation schemes. The simulations ran using 3 cores for the fluid domain and 3 cores for the solid domain. Results are summarized in Table 2 and Figures 4 and 5.

¹The results are presented as in the original experiment: mean \pm amplitude[freq].

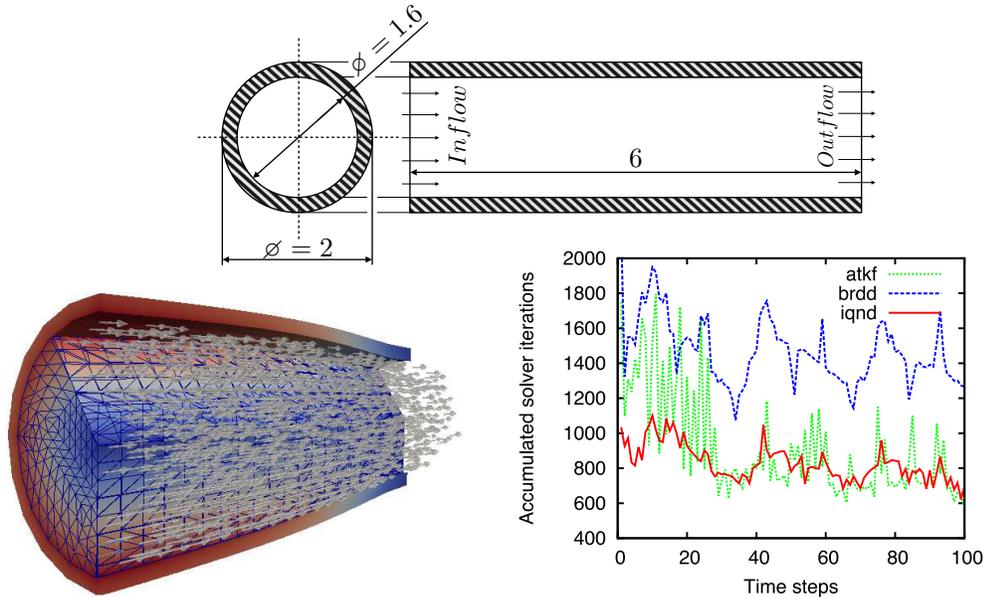


Fig. 4: Tube expansion experiment. Top: scheme of the problem. Left: Sectioned view of the tube expansion problem. Right: Accumulated solver iterations in each time step for the best variant in each algorithm family.

Table 2: Tube expansion experiment. Summary for the experiment of the distensible tube in subsection 4.2. † The Broyden case, relaxed in force fails at the 12th time step (see Figure 5).

coupling strategy	coupling iterations	time per iteration [s]	NSI		SLD	
			subit	sol ite	subit	sol ite
atk displ	10.23	6.8	16.60	782.09/127.06	10.40	1634.57
atk force	5.63	7.59	16.98	553.92/91.57	9.94	885.14
brd displ	10.29	7.63	15.94	919.92/147.09	10.32	1676.25
brd force†	—	—	—	—	—	—
iqnd rank 5	6.27	7.45	14.70	531.34/83.24	10.28	1011.08
iqnd rank 20	6.25	7.3	14.71	530.72/82.93	10.34	1010.13
iqnf rank 5	7.19	7.7	15.29	701.68/110.93	10.02	1121.49
iqnf rank 20	9.26	7.63	14.49	895.27/135.54	10.27	1498.03

In this type of problem, the Aitken algorithm performs better relaxed in force, while the Broyden algorithm diverges in the 12th time step for the force relaxation. If only the best set up is taken from each family (atkd, brdd and iqnd), CIQN requires 40 % less coupling iterations if compared with Broyden scheme and 2 % faster if compared with the Aitken strategy (see Table 2).

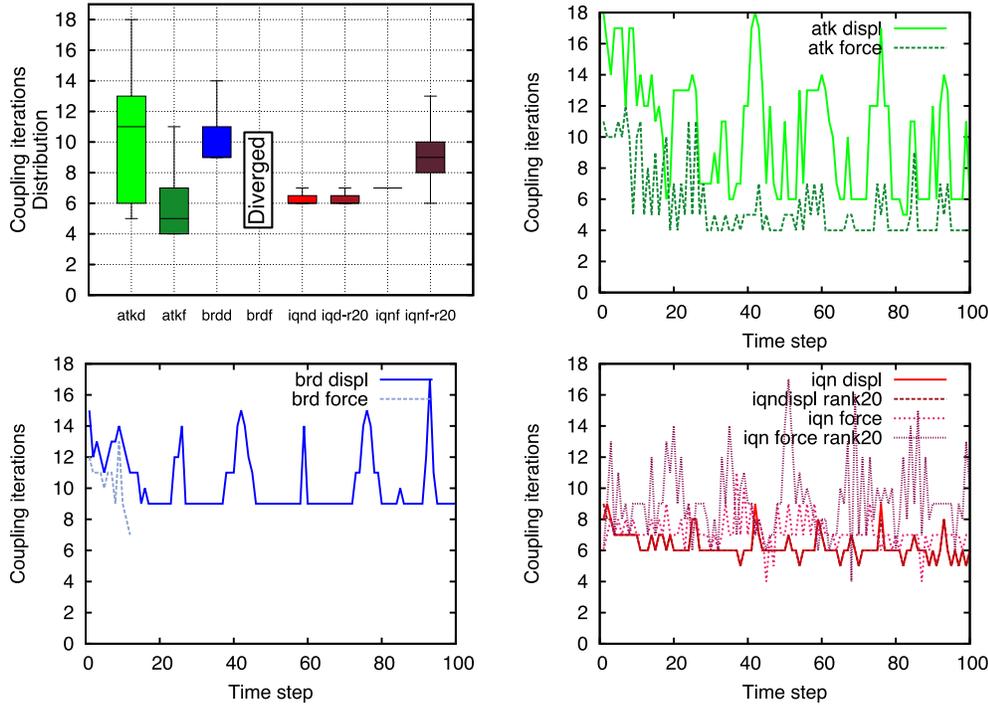


Fig. 5: Tube expansion experiment. Top left: coupling iterations distribution, Broyden relaxed in force (brdf) is absent because it diverges. Top right: both versions of the aitken relaxation are compared. Bottom left: both versions of Broyden relaxation are compared, note that the force version of the Broyden reached up to the 12th time step. Bottom right: the four versions of the interface quasi Newton relaxation are compared.

4.3. Two-field coupling 2D problem. The problem setup is described in Figure 6 and in Table 1. There are two contact surfaces to couple, an oscillating bar under a fluid flow and a deformable wall in the frontier of the fluid domain. In each case, 50 time steps are computed for all the possible combinations of relaxation schemes. The problem was solved using six cores, three for the solid problem, and three for the fluid problem. The results are summarized in Table 3 and Figures 6 and 7.

It is worth to note that in these cases, both solid problems are indirectly coupled by the fluid, on which the complexity of the n-field coupling lies. As both interfaces have very different dynamics, the Broyden and CIQN surpass the Aitken algorithm. Both, CIQN and Broyden requires 60 % less iterations than the Aitken (see Table 3). As the Aitken algorithm enforces the Jacobian to be a scalar for all the nodes in the contact surface, it has to increase the iteration number to reach convergence. The difference between Broyden scheme and CIQN scheme is not that notable. Despite CIQN requires 4 % more iterations, each iteration is computed 19 % faster, probably due to the smaller dispersion between iterations in the CIQN algorithm (see top left in Figure 7).

4.4. Two field 3D problem. The problem set-up is described in Figure 8 and in Table 1. A filled flexible tube is lying in a flexible surface which is in contact with

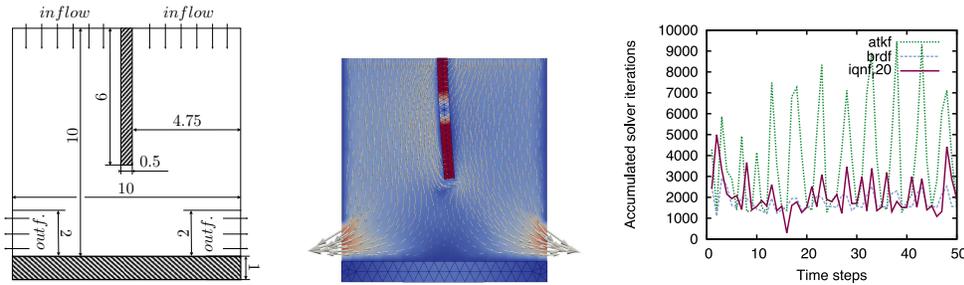


Fig. 6: Two field 2D problem. Left: scheme of the problem. Center: Snapshot for $t = 0.3$. As the range of deformation of the oscillating bar is bigger than the deformation of the bar in the bottom, the scale has been reduced to perceive the deformation in the bottom bar; effect that produces the colour scale saturation in the oscillating bar. Right: Accumulated solver iterations in each time step for the best variant in each algorithm family.

Table 3: Two field 2D problem. Summary for the experiment of the agitating bar in [subsection 4.3](#). The superiority of the Broyden and the CIQN schemes against the Aitken is clearly seen in the average coupling iterations.

coupling strategy	coupling iterations	time per iteration	NSI		SLD	
			subit	sol ite	subit	solite
atk displ	20.58	3.21	6.96	1098.64/359.02	12.22	2681.11
atk force	16.7	3.34	8.58	1127.05/382.15	11.48	1955.33
brd displ	10.76	3.20	7.16	620.10/205.04	12.20	1378.53
brd force	7.96	3.37	8.26	529.91/180.37	10.96	920.02
iqnd rank 5	9.20	3.07	7.68	575.75/193.1	12.88	1325.13
iqnd rank 20	9.24	2.76	7.68	576.44/193.31	12.96	1344.73
iqnf rank 5	12.76	2.68	14.64	1096.89/404.65	12.40	1695.81
iqnf rank 20	8.30	3.26	8.94	596.60/208.67	10.62	910.84

a big volume of another fluid. The reader can think about this academic example as a coronary artery close to the endocardium, that is in contact with the blood inside the ventricle. A constant velocity is imposed in the bigger fluid domain and an oscillating velocity in the distensible tube. 100 time steps were computed with all the possible relaxation combinations. The simulations ran using with 8 cores for the fluid domain and 8 cores for the solid domain. The results are summarized in [Table 4](#) and [Figures 8](#) and [9](#).

All the coupling schemes were capable to solve the system, with differences in the performance. For the best set up in each family, the CIQN algorithm required 8.5 % less iterations compared with the Aitken algorithm and 30 % less iterations if compared against the Broyden scheme. Talking about the processing time, CIQN performs 9 % and 6 % faster compared to Aitken and Broyden respectively (see [Table 4](#)).

Common behaviors for the analyzed cases. One of the variables in the CIQN scheme is the number of past iterations used for the approximation. Although one can expect that adding information from past iterations should improve the approximation, this does not correlate with the results. The reason for this, probably, is

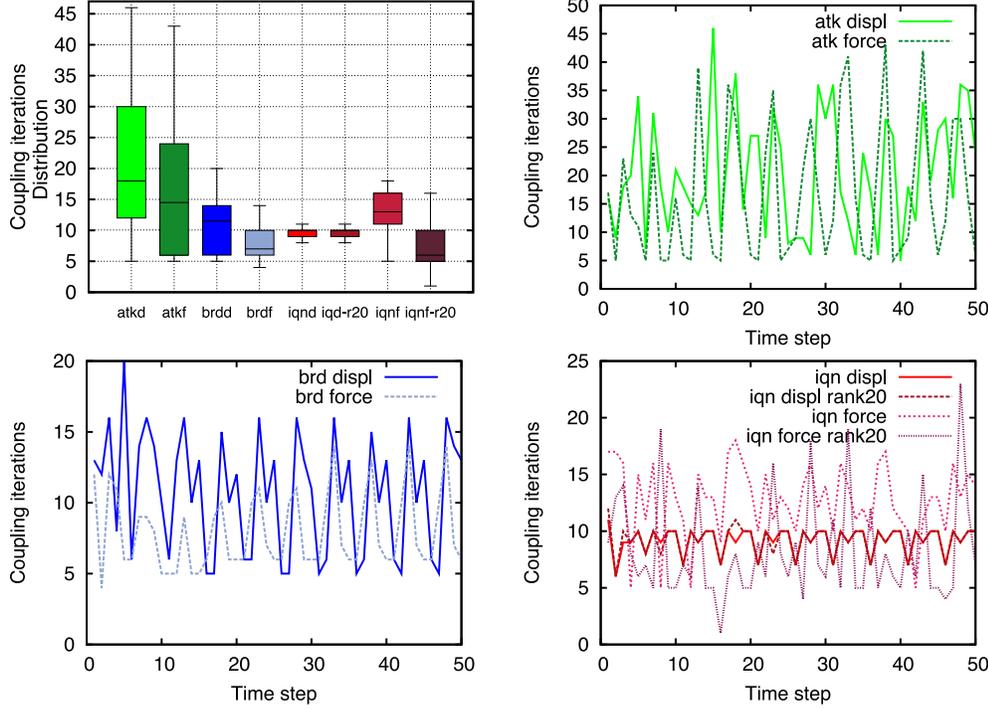


Fig. 7: Two field 2D problem. Top left: coupling iterations distribution for each scheme. Top right: Aitken family convergence. Bottom left: Broyden family convergence. Bottom right: CIQN family convergence.

Table 4: Two field 3D problem. Summary of the experiment in [subsection 4.4](#).

coupling strategy	coupling iterations	time per iteration	NSI		SLD	
			subit	sol ite	subit	sol ite
atk displ	9.36	16.70	18.70	2064.76/989.36	6.51	2513.36
atk force	9.52	16.27	20.88	2340.99/1078.36	5.99	2353.96
brd displ	12.16	16.31	18.50	2586.81/1271.28	6.53	3150.89
brd force	17.65	16.05	20.30	3565.25/1670.89	5.98	4488.76
iqnd rank 5	8.64	17.24	20.84	2040.46/961.70	6.72	2347.26
iqnd rank 20	8.66	17.37	20.81	2022.46/961.63	6.71	2375.82
iqnf rank 5	10.54	17.25	22.04	2715.19/1246.88	6.22	2724.44
iqnf rank 20	14.39	16.33	20.17	3605.16/1784.84	6.33	3496.33

that the older iterations are adding noise more than information. Another aspect to note is the narrow distribution obtained for the CIQN algorithm, where the interquartile range is generally much smaller compared with the Aitken and the Broyden algorithms. The number of coupling iterations for the CIQN remains almost constant in all the cases, while there is a big variability in the other coupling schemes.

4.4.1. Parallel performance analysis. As stated above, the aim of this paper is to design a high performance version of the interface quasi Newton algorithm. In order to show this, we present a performance analysis of the CIQN algorithm applied

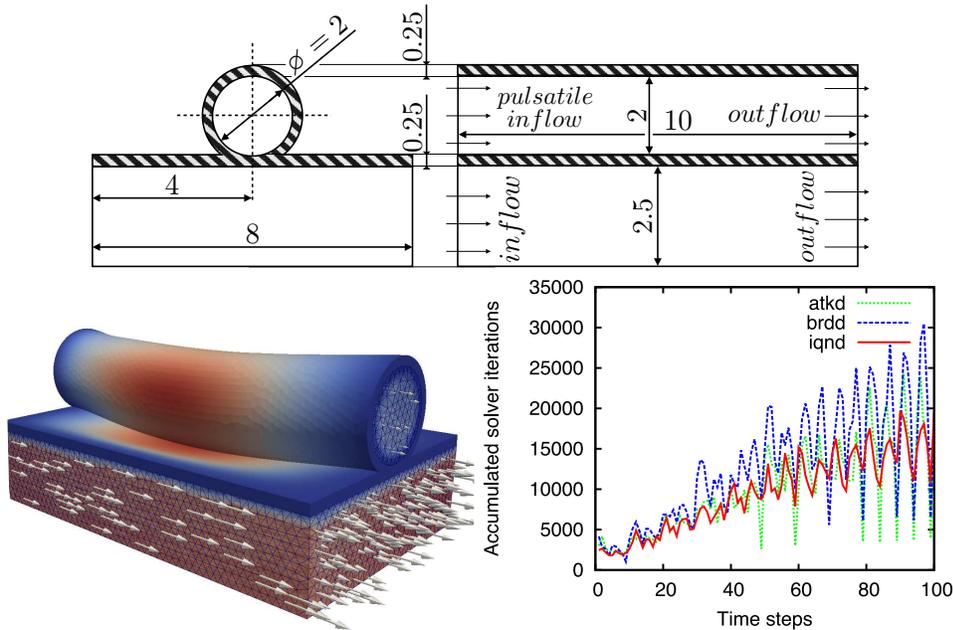


Fig. 8: Two field 3D problem. Top: scheme of the problem. Left: Result in time $t = 0.4$ [s] for the biomechanical inspired case. Right: Accumulated solver iterations in each time step for the best variant in each algorithm family.

to a modified version of case in [subsection 4.4](#). The original mesh was uniformly refined with the algorithm described in [18], obtaining a ~ 60 million elements mesh in the fluid and a ~ 40 million elements mesh in the solid.

First, the parallel performance of each parallel solver running independently was tested. An efficiency above 0.85 was obtained up to 2048 for both, the fluid mechanics and the solid mechanics solver (see [Table 5](#) and [Figure 10](#)).

Table 5: Parallel performance analysis. Speed-up and efficiency for both solvers.

Core count	Fluid mechanics		Solid mechanics	
	speed up	efficiency	speed up	efficiency
128	128.0	1.00	128.0	1.00
256	256.0	0.99	256.0	0.99
512	515.6	0.99	508.1	0.99
1024	1011.0	0.98	960.5	0.93
2048	1880.3	0.91	1793.0	0.87

The coupled problem has been solved using a Gauss-Seidel scheme, in which each parallel solver (fluid and structure) run one after the other. Due to this, one of the solvers is computing while the other is waiting for the results (see right side of [Figure 10](#)), reducing the efficiency of the coupled system.

To optimize the execution time, the number of cores assigned to each problem must be tuned. To obtain this optimum, a searching in core allocation is performed. First, the range of cores for each independent solver is chosen, in such a way that

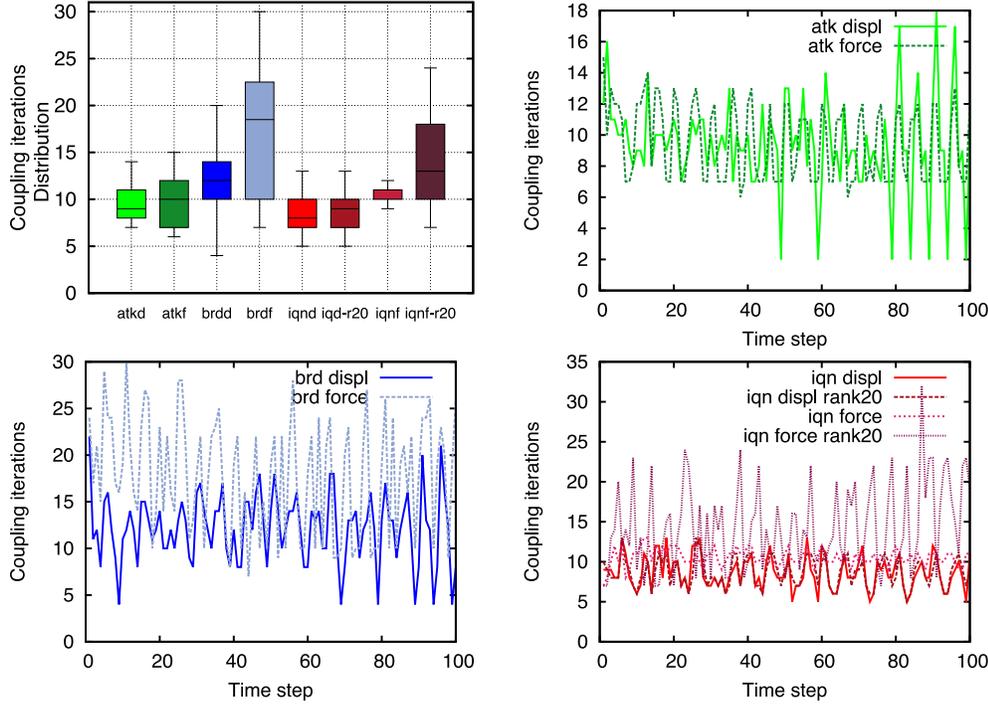


Fig. 9: Two field 3D problem. Top left: coupling iteration distribution. Top right: Aitken family convergence. Bottom left: Broyden family convergence. Bottom right: CIQN family convergence.

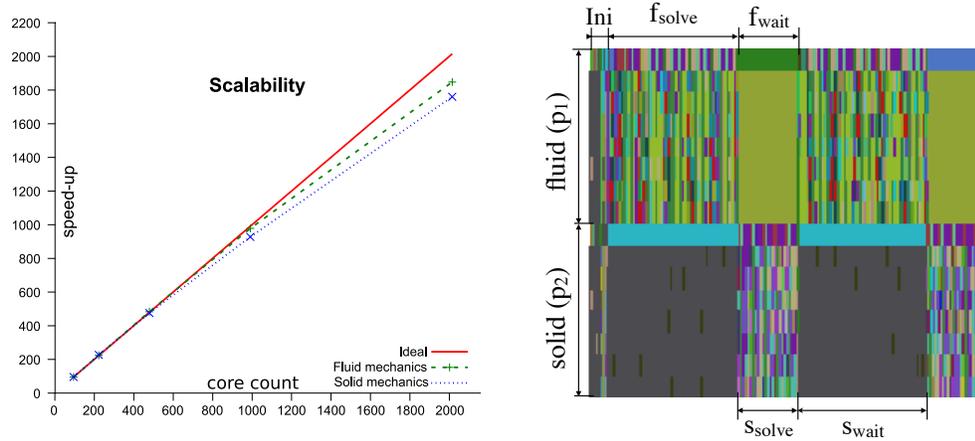


Fig. 10: Parallel performance analysis. Left: Speed-up for each one of the solvers (fluid mechanics and solid mechanics) uncoupled. Right: Trace example of the two field 3D problem for two FSI iterations. The Gauss-Seidel strategy forces one solver to be idle while the other is computing. $p = p_f + p_s$ processors are used to solve the coupled system. After the initialization stage (Ini); the fluid solver f_{solve} is executed by using $p_f = p_1$ processors. Later the solid solver s_{solve} uses $p_s = p_2$ processors.

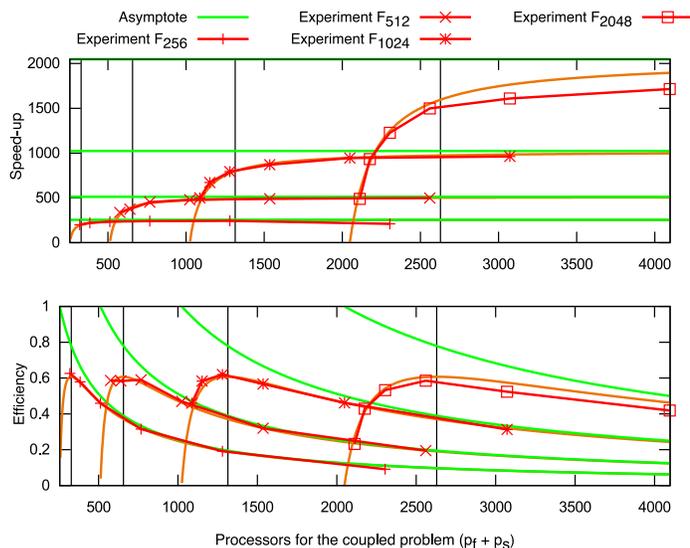


Fig. 11: Parallel performance analysis. Speed-up and Efficiency for a core allocation of $p = p_f + p_s$, where $p_f = \{256, 512, 1024, 2048\}$ in the core count for the fluid, and $p_s = \{64, 128, 256, 512, 1024, 2048\}$ for the solid. An optimal allocation p_{opt} which allows to achieve the maximum efficiency E_{p_f} of the coupled system can be found for each curve S_{p_f} .

each of them runs with the maximum efficiency E possible (see Figure 10). After this, in the coupled problem, the core count for one solver is fixed while ranging the number of cores for the other solver. Figure 11 shows speed-up and efficiency for four fixed values in the fluid solver core count p_f . In each case the core count for the solid mechanic solver p_s is ranged between 64 and 2048, with increments in power of two. This processes is performed for a core count of 256, 512, 1024, and 2048 in the fluid solver. For each speed-up curve, a maximum in the efficiency can be achieved. These maximum values of efficiency indicate the number of cores that should be assigned in order to attain an optimum execution time for a specific configuration of the Gauss-Seidel (algorithm 1).

In conclusion, the parallel implementation of the CIQN algorithm here presented allows to achieve the maximum values of efficiency possible for the CIQN algorithm in this fluid-structure interaction problem. This shows the feasibility of the CIQN algorithm to efficiently solve large-scale problems.

5. Conclusions. In this paper, we introduced a compact interface quasi-Newton (CIQN) coupling scheme, optimized for distributed memory architecture. The proposed algorithm avoids constructing dense matrices and reduces the number of operations to obtain a good performance in parallel computers. This algorithm also successfully dealt with added mass instability and two field coupling in all the tested cases.

We observe that the CIQN is robust in terms of coupling iteration counts because the CIQN scheme requires less solver iterations in the fluid and the solid modules to converge the problem. Also, in general, the CIQN requires less computing time per iteration.

As it has been said, there is no a silver bullet algorithm to tackle all the FSI

cases. The chosen algorithm must fit the features of the problem to solve, taking into account the dynamics and the possible numerical instabilities that may arise. In this work the superiority of the proposed high performance interface quasi Newton algorithm is shown, resulting in a cheaper and more stable scheme compared with the Aitken and the Broyden algorithms. Also, good performance in massively parallel computers has been proven.

REFERENCES

- [1] S. BADIA, A. QUAINI, AND A. QUARTERONI, *Modular vs. non-modular preconditioners for fluid-structure systems with large added-mass effect*, Computer Methods in Applied Mechanics and Engineering, 197 (2008), pp. 4216–4232, doi:10.1016/j.cma.2008.04.018.
- [2] C. G. BROYDEN, *A Class of Methods for Solving Nonlinear Simultaneous Equations*, Mathematics of Computation, 19 (1965), p. 577, doi:10.2307/2003941, <http://www.jstor.org/stable/2003941>.
- [3] H. J. BUNGARTZ, F. LINDNER, M. MEHL, AND B. UEKERMANN, *A plug-and-play coupling approach for parallel multi-field simulations*, Computational Mechanics, 55 (2015), pp. 1119–1129, doi:10.1007/s00466-014-1113-2.
- [4] J. C. CAJAS, G. HOUZEAUX, M. VÁZQUEZ, M. GARCIA, E. CASONI, H. CALMET, A. ARTIGUES, R. BORELL, O. LEHMKUHL, D. PASTRANA, Y. DANIEL, R. PONS, AND J. MARTORELL, *Fluid structure interaction based on hpc multi-code coupling*, (submitted), Computational mechanics, (2017).
- [5] J. C. CAJAS, M. ZAVALA, AND G. HOUZEAUX, *Fluid structure interaction in hpc multi-code coupling*, Civil-Comp Proceedings, 7 (2016), doi:10.4203/ccp.107.33.
- [6] R. CALDERER AND A. MASUD, *A multiscale stabilized ALE formulation for incompressible flows with moving boundaries*, Computational Mechanics, 46 (2010), pp. 185–197, doi:10.1007/s00466-010-0487-z.
- [7] E. CASONI, A. JÉRUSALEM, C. SAMANIEGO, B. EGUZKITZA, P. LAFORTUNE, D. D. TJAHAJANTO, X. SÁEZ, G. HOUZEAUX, AND M. VÁZQUEZ, *Alya: Computational Solid Mechanics for Supercomputers*, Archives of Computational Methods in Engineering, (2014), doi:10.1007/s11831-014-9126-8, <http://link.springer.com/10.1007/s11831-014-9126-8>.
- [8] E. CASONI, A. JÉRUSALEM, C. SAMANIEGO, B. EGUZKITZA, P. LAFORTUNE, D. D. TJAHAJANTO, X. SÁEZ, G. HOUZEAUX, AND M. VÁZQUEZ, *Alya: computational solid mechanics for supercomputers*, Archives of Computational Methods in Engineering, 22 (2015), pp. 557–576.
- [9] P. CAUSIN, J. F. GERBEAU, AND F. NOBILE, *Added-mass effect in the design of partitioned algorithms for fluid-structure problems*, Computer Methods in Applied Mechanics and Engineering, 194 (2005), pp. 4506–4527, doi:10.1016/j.cma.2004.12.005.
- [10] P. CROSETTO, P. REYMOND, S. DEPARIS, D. KONTAXAKIS, N. STERGIOPULOS, AND A. QUARTERONI, *Fluid-structure interaction simulation of aortic blood flow*, Computers and Fluids, 43 (2011), pp. 46–57, doi:10.1016/j.compfluid.2010.11.032, <http://dx.doi.org/10.1016/j.compfluid.2010.11.032>.
- [11] J. DEGROOTE, K. J. BATHE, AND J. VIERENDEELS, *Performance of a new partitioned procedure versus a monolithic procedure in fluid-structure interaction*, Computers and Structures, 87 (2009), pp. 793–801, doi:10.1016/j.compstruc.2008.11.013, <http://dx.doi.org/10.1016/j.compstruc.2008.11.013>.
- [12] C. FÖRSTER, W. A. WALL, AND E. RAMM, *Artificial added mass instabilities in sequential staggered coupling of nonlinear structures and incompressible viscous flows*, Computer Methods in Applied Mechanics and Engineering, 196 (2007), pp. 1278–1293, doi:10.1016/j.cma.2006.09.002.
- [13] M. W. GEE, U. KUTTLER, AND W. A. WALL, *Truly monolithic algebraicmultigrid for fluidstructure interaction*, International Journal for Numerical Methods in Biomedical Engineering, 85 (2010), pp. 987–1016, doi:10.1002/nme, arXiv:1010.1724.
- [14] J.-F. GERBEAU, M. VIDRASCU, AND P. FREY, *Fluidstructure interaction in blood flows on geometries based on medical imaging*, Computers & Structures, 83 (2005), pp. 155–165, doi:10.1016/j.compstruc.2004.03.083, <http://www.sciencedirect.com/science/article/pii/S0045794904003001>.
- [15] C. HABCHI, S. RUSSEIL, D. BOUGEARD, J. L. HARION, T. LEMENAND, A. GHANEM, D. D. VALLE, AND H. PEERHOSSAINI, *Partitioned solver for strongly coupled fluid-structure interaction*, Computers and Fluids, 71 (2013), pp. 306–319, doi:10.1016/j.compfluid.2012.11.004, <http://dx.doi.org/10.1016/j.compfluid.2012.11.004>.

- [16] G. HOUZEAUX, R. AUBRY, AND M. VÁZQUEZ, *Extension of fractional step techniques for incompressible flows : The preconditioned Orthomin (1) for the pressure Schur complement*, Computers and Fluids, 44 (2011), pp. 297–313, doi:10.1016/j.compfluid.2011.01.017, <http://dx.doi.org/10.1016/j.compfluid.2011.01.017>.
- [17] G. HOUZEAUX, J. C. CAJAS, M. DISCACCIATI, B. EGUZKITZA, A. GARGALLO-PEIRÓ, M. RIVERO, AND M. VÁZQUEZ, *Domain decomposition methods for domain composition purpose: Chimera, overset, gluing and sliding mesh methods*, Archives of Computational Methods in Engineering, (2016), pp. 1–38, doi:10.1007/s11831-016-9198-8, <http://dx.doi.org/10.1007/s11831-016-9198-8>.
- [18] G. HOUZEAUX, R. DE LA CRUZ, H. OWEN, AND M. VÁZQUEZ, *Parallel uniform mesh multiplication applied to a Navier-Stokes solver*, Computers and Fluids, 80 (2013), pp. 142–151, doi:10.1016/j.compfluid.2012.04.017, <http://dx.doi.org/10.1016/j.compfluid.2012.04.017>.
- [19] G. HOUZEAUX AND J. PRINCIPE, *A variational subgrid scale model for transient incompressible flows*, International Journal of Computational Fluid Dynamics, 22 (2008), pp. 135–152, doi:10.1080/10618560701816387, <http://www.tandfonline.com/doi/abs/10.1080/10618560701816387>.
- [20] G. HOUZEAUX, M. VÁZQUEZ, R. AUBRY, AND J. M. CELA, *A massively parallel fractional step solver for incompressible flows*, Journal of Computational Physics, 228 (2009), pp. 6316–6332, doi:10.1016/j.jcp.2009.05.019, <http://dx.doi.org/10.1016/j.jcp.2009.05.019>.
- [21] J. HRON AND S. TUREK, *A monolithic fem/multigrid solver for an ale formulation of fluid-structure interaction with applications in biomechanics*, Fluid-Structure Interaction, 53 (2006), pp. 146–170, doi:10.1007/3-540-34596-5-7.
- [22] J. HRON AND S. TUREK, *Proposal for numerical benchmarking of fluid-structure interaction between an elastic object and laminar incompressible flow*, Fluid-Structure Interaction, 53 (2006), pp. 371–385, doi:10.1007/3-540-34596-5-15, <http://link.springer.com/chapter/10.1007/3-540-34596-5-15>.
- [23] G. KARYPIS AND V. KUMAR, *A Fast and High Quality Multilevel Scheme for Partitioning Irregular Graphs*, SIAM Journal on Scientific Computing, 20 (1998), pp. 359–392, doi:10.1137/S1064827595287997.
- [24] U. KÜTTLER, M. GEE, C. FOSTER, A. COMERFORD, AND W. WALL, *Coupling strategies for biomedical fluid-structure interaction problems*, International Journal for Numerical Methods in Biomedical Engineering, 7 (2009), doi:10.1002/cnm.1281.
- [25] U. KÜTTLER AND W. A. WALL, *Fixed-point fluid-structure interaction solvers with dynamic relaxation*, Computational Mechanics, 43 (2008), pp. 61–72, doi:10.1007/s00466-008-0255-5.
- [26] R. LÖHNER, F. MUT, J. CEBRAL, R. AUBRY, AND G. HOUZEAUX, *Deflated preconditioned conjugate gradient solvers for the pressure-poisson equation: Extensions and improvements*, Int. J. Numer. Meth. Engn., 87 (2011), pp. 2–14.
- [27] H. G. MATTHIES AND J. STEINDORF, *Partitioned strong coupling algorithms for fluid-structure interaction*, Computers and Structures, 81 (2003), pp. 805–812, doi:10.1016/S0045-7949(02)00409-1.
- [28] M. MEHL, B. UEKERMANN, H. BIJL, D. BLOM, B. GATZHAMMER, AND A. VAN ZUIJLEN, *Parallel coupling numerics for partitioned fluid-structure interaction simulations*, Computers and Mathematics with Applications, 71 (2016), pp. 869–891, doi:10.1016/j.camwa.2015.12.025.
- [29] L. RADTKE, A. LARENA-AVELLANEDA, E. S. DEBUS, AND A. DÜSTER, *Convergence acceleration for partitioned simulations of the fluid-structure interaction in arteries*, Computational Mechanics, (2016), pp. 1–20, doi:10.1007/s00466-016-1268-0.
- [30] K. SCHEUFELE, *Robust Quasi-Newton Methods for Partitioned Fluid-Structure Simulations*, PhD thesis, University of Stuttgart, 2015.
- [31] B. UEKERMANN, J. C. CAJAS, B. GATZHAMMER, G. HOUZEAUX, M. MEHL, AND M. VÁZQUEZ, *Towards partitioned fluid-structure interaction on massively parallel systems*, Proceedings of WCCM XI/ECCM V/ECFD VI, Barcelona, (2014).
- [32] M. VÁZQUEZ, G. HOUZEAUX, S. KORIC, A. ARTIGUES, J. AGUADO-SIERRA, R. ARS, D. MIRA, H. CALMET, F. CUCCHIETTI, H. OWEN, A. TAHA, E. D. BURNES, J. M. CELA, AND M. VALERO, *Alya: Multiphysics engineering simulation toward exascale*, Journal of Computational Science, 14 (2016), pp. 15 – 27, doi:http://dx.doi.org/10.1016/j.jocs.2015.12.007, <http://www.sciencedirect.com/science/article/pii/S1877750315300521>. The Route to Exascale: Novel Mathematical Methods, Scalable Algorithms and Computational Science Skills.
- [33] F. VERDUGO AND W. A. WALL, *Unified framework for the efficient solution of n-field coupled problems with monolithic schemes*, Computer Methods in Applied Mechanics and Engineering, (2015), p. submitted, arXiv:1605.01522.

- [34] J. VIERENDEELS, L. LANOYE, J. DEGROOTE, AND P. VERDONCK, *Implicit coupling of partitioned fluid-structure interaction problems with reduced order models*, *Computers and Structures*, 85 (2007), pp. 970–976, doi:[10.1016/j.compstruc.2006.11.006](https://doi.org/10.1016/j.compstruc.2006.11.006).

Model for pressure drop and flow deflection in the numerical simulation of stents in aneurysms

S. Li*, J. Latt, B. Chopard

Department of Computer Science, University of Geneva, Carouge, 1227, Switzerland

SUMMARY

The numerical simulation of flow diverters like stents contributes to the development and improvement of endovascular stenting procedures, leading ultimately to an improved treatment of intracranial aneurysms. Due to the scale difference between the struts of flow diverters and the full artery, it is common to avoid fully resolved simulations at the level of the stent porosity. Instead, the effect of stents on the flow is represented by a heuristic continuum model. However, the commonly used porous media models describe the properties of flow diverters only partially, because they do not explicitly account for the deflection of the flow direction by the stent. We show that this deficiency can be circumvented by adopting the theoretical framework of screen models. The article first reviews existing screen models. It then proposes an explicit formula for the drag and the deflection coefficient, as predicted by each model, for both perpendicular and inclined angles. The results of 2D numerical simulations are used to formulate a generalization of these formulas, to achieve best results in the case of stent modeling. The obtained model is then validated, again through 2D numerical simulation. Copyright © 0000 John Wiley & Sons, Ltd.

Received ...

KEY WORDS: aneurysm; stent; screen model; CFD; lattice Boltzmann

1. INTRODUCTION

Intracranial aneurysm is a cerebrovascular disorder, consisting of a local dilation of the blood vessel, leading to an increased risk of rupture with possibly lethal consequences [1, 2]. A possible treatment of an unruptured intracranial aneurysm consists of the introduction of a flow diverter, often a stent, into the blood vessel. Stents are flexible, self-expanding porous tubular meshes. In order to improve the design of stents, and in this way improve the success rate of stenting procedures, numerical simulation has proven to be an invaluable tool [1, 3]. It allows fundamental investigation of hydrodynamic properties of stents, but also opens the window to patient-specific, personalized simulations. One of the possibilities is to study the alterations to the flow pattern before and after the placement of a stent, and then use this numerical data to determine the chances of success of the intervention [4].

*Correspondence to: Route de Drize 7, Carouge, 1227, Switzerland. E-mail: Sha.Li@unige.ch

Given the small-scale structure of flow diverters, it is however often prohibitively expensive to fully resolve the flow structure around the coils of a stent [3]. Indeed, the diameters of the struts can differ from the diameter of the artery by more than three orders of magnitude, and it can be very difficult to handle both scales with a unique fluid mesh. This problem is particularly severe for patient-specific simulations, in which the numerical results are expected in a short time frame. As an alternative, numerical simulations are often carried out on a coarse mesh, which cannot capture the coil structure of a stent. Instead, the effect of the stent on the flow is represented through a continuum model.

An approach advocated for example by [2, 5] proposes to model the stent like a porous media, using a term similar to Darcy's law. It is however questionable if the assumptions underlying porous media models remain valid in the case of stents. Such models assume for example that the underlying porous object has a finite thickness, while flow diverters in reality have very thin surfaces. As a consequence, porous media models do not explicitly account for the deflection of the flow through a thin surface.

In the present article, we take a more fundamental standpoint than [2, 5] and propose a theoretical framework to formulate, validate, and select continuum stent models. This framework is based on so-called screen models which, as opposed to porous media models, are explicitly built to reflect the physics of thin porous structures. We validate these models by running fully resolved 2D simulations across a stent. We show that while these models provide good results, and in particular are able to predict quite accurately both the pressure drop and the flow deflection across a stent, they also suffer from a major drawback, namely the assumption of a uniform, constant upstream flow in front of the screen. This assumption is not valid for stents deployed in aneurysms, in which case the flow pattern close to the stent surface can be quite complicated.

We therefore propose a correction to classical screen models in the form of a velocity reduction coefficient, which relates the presumed asymptotic upstream flow velocity to a local velocity, measured in the vicinity of different points on the stent surface. This model is incorporated into screen models to obtain a globally consistent model for flow diverters.

In a final part, we validate this flow diverter model by comparing fully resolved fluid flow simulations around a flow diverter to coarse-grained simulations using our flow diverter model. All simulations are carried out in 2D for the sake of simplicity, given that the main purpose of the article is to propose a fundamental framework for a new category of stent models. The authors intend to run 3D simulation and present them in a more application-oriented publication as part of their future research.

2. HYDRODYNAMIC MODELS OF FLOW DIVERTERS

2.1. General principles

Hydrodynamic flow diverter models represent the impact of the flow diverter on the flow pattern in an averaged manner, when the computational mesh is too coarse to fully resolve the structure of the flow diverter. State of the art models represent the impact of a flow diverter in terms of a pressure drop across its surface, and through an additional sink term in the momentum balance equation. Model constants for these two terms are adjusted empirically for specific flow diverters [2, 5]. While

this approach quite successfully reproduces some aspects of the flow structure, it fails to properly account for the flow deflection through the diverter, which can be particularly important in cases when the flow direction does not align with the normal of the diverter surface.

In this paper, we make the hypothesis that a flow diverter can be approximated as a screen, and we test different screen models for their validity, including their ability to predict flow deflection. The existing literature [7, 8] shows that screens are generally used as a kind of fluid distributor to achieve a more uniform velocity field, reduce the intensity of turbulence, or change the flow direction. The relevant hydrodynamic characteristics of screens are the pressure drop and the deviation of the flow direction. The pressure drop caused by the resistance of the screen depends on the properties of both screen and fluid, while the flow deviation depends on the inclination of screen. It is well acknowledged that screens refract the incident flow towards the locally normal direction to the screen surface [13, 14].

Considering an incident flow through a screen at an arbitrary angle as shown in Fig. 1, it is necessary to define three velocities:

- U_1 is the velocity of the incident flow far from the screen.
- U is the velocity at which the flow passes through the screen.
- U_2 is the downstream velocity far from the screen.

Here, θ_1, θ , and θ_2 are the angles between the direction of these velocities and the normal direction to the screen. Since the mass flux of fluid through the screen is conserved [11], the normal velocity components are consistent, i.e. $u_1 = u = u_2$. In the tangential direction, the screen refracts the incident flow toward the normal direction of the screen, leading to a corresponding decrease of the tangential velocity components, i.e. $(v_1 \geq v \geq v_2)$.

The resistance generated by the screen gives rise to a pressure difference between the two sides of the screen, and the pressure drop can be considered as a reactive force which drives the fluid through the screen. In hydrodynamics, the pressure drop is expressed by a non-dimensional drag coefficient which is generally defined as:

$$k_{\theta_1} = \frac{\Delta P}{\frac{1}{2}\rho U_1^2}, \tag{1}$$

where ΔP is the pressure drop and ρ the fluid density.

The drag coefficient k_{θ_1} is related to the screen's porosity and geometrical parameters, as well as the Reynolds number

$$\text{Re} = \frac{\rho U_{\text{ref}} L_{\text{ref}}}{\mu}. \tag{2}$$

Here, μ is the dynamic fluid viscosity, U_{ref} is a characteristic velocity, which in most screen models is taken to be the norm of the upstream velocity U_1 , and L_{ref} is a characteristic length. In the models of weaving screens, L_{ref} is often taken to be equal to the diameters of the screen wires, in which case the corresponding Reynolds number is labeled as Re_d .

Concerning the resistance in tangential directions, some authors define an analogous coefficient, named deflection force coefficient [7]:

$$F_{\theta_1} = \frac{\tau}{\frac{1}{2}\rho U_1^2}, \tag{3}$$

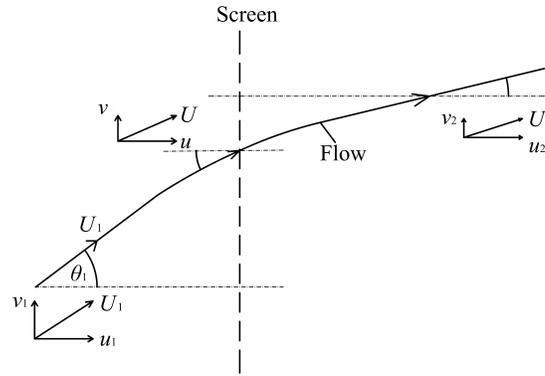


Figure 1. Flow through a screen at an incident angle

where τ is the tangential stress which causes a change of velocity in the tangential plane of the screen. When the upstream and downstream flows are far away from the screen, Taylor and Batchelor [8] propose that the tangential force could be obtained by measuring the change of direction of the stream passing through the screen, and the deflection force coefficient is expressed as:

$$F_{\theta_1} = 2 \cos \theta_1 \sec \theta_2 \sin(\theta_1 - \theta_2). \quad (4)$$

In order to express the deflection of velocity in a more direct way, another coefficient named deflection coefficient B [7] is proposed as:

$$B = 1 - v_2/v_1. \quad (5)$$

Combining Eqs. (4) and (5) yields a relationship between these two coefficients:

$$F_{\theta_1} = B \sin 2\theta_1 \quad (6)$$

From Eqs. (3) and (6), the tangential stress can be expressed directly from the deflection coefficient B and upstream velocities:

$$\tau = B u_1 v_1. \quad (7)$$

2.2. The drag coefficient

The computation of the drag coefficient for a screen that is inclined with respect of the incoming flow can be split into two steps. Firstly, the drag coefficient of a screen which is normal to the flow, k_0 , with $\theta_1 = 0$, is computed. In this case, the porosity and Reynolds number are the decisive factors. Secondly, the relationship between k_0 and k_{θ_1} , or in other words, the impact of the incident angle on the drag coefficient, is determined.

Kolodziej et al. [15] describe in detail the major theoretical approaches to model the pressure drop across a screen, and they summarize the most important models for k_0 .

Wieghardt (as cited in [15]) provides a correlation based on experimental data from the literature:

$$k_0 = 6 \frac{1 - \beta}{\beta^2} (\text{Re}_d / \beta)^{-1/3}, \text{Re}_d = \frac{\rho U_1 d}{\mu}, \quad (8)$$

where d is the diameter of wires in a weaving screen and β is the porosity of the screen, defined as:

$$\beta = A_{open}/A_{total}. \quad (9)$$

Here, A_{open} is the orthogonally projected open area of the screen and A_{total} is the total cross-sectional area.

Ehrhardt presents a correlation based on his measurements of oil, water and air flow through individual screens over a range $0.3 \leq Re_d \leq 300$ (as cited in [15]):

$$k_0 = \left(0.72 + \frac{49}{(Re_d/\beta)} \right) \frac{1 - \beta}{\beta^2}. \quad (10)$$

Wakeland and Keolian [17] propose their correlation by measuring the drag coefficient of more than 20 kinds of screens with different porosities and wire diameters for both low and intermediate Reynolds numbers.

$$k_0 = \frac{1 - \beta}{\beta^2} \left(\frac{17.0}{Re_d} + 0.55 \right) \quad (11)$$

Brundrett [16] proposes a correlation for screens with incident angles:

$$k_{\theta_1} = \cos^2 \theta \times \frac{1 - \beta^2}{\beta^2} \times \left[\frac{\sigma_M}{\sigma_{KE}} \times \frac{7.125}{Re_d \cos \theta} + \frac{0.75}{\log(Re_d \cos \theta + 1.25)} + 0.055 \log(Re_d \cos \theta) \right] \quad (12)$$

where σ_{KE} and σ_M are the correction factors for momentum and kinetic energy. In a uniform flow, the equality $\sigma_{KE} = \sigma_M = 1.0$ holds.

Therefore, according to Brundrett's model, k_0 should take the value

$$k_0 = \frac{1 - \beta^2}{\beta^2} \times \left[\frac{\sigma_M}{\sigma_{KE}} \times \frac{7.125}{Re_d} + \frac{0.75}{\log(Re_d + 1.25)} + 0.055 \log(Re_d) \right] \quad (13)$$

From the equations listed above, we can see that the equation for k_0 consists of three elements. The first is proportional to $1/Re_d$, and the second is a function of porosity, like $(1 - \beta)/\beta^2$ in Eq. 11, and the third element is some constants. All the models have a positive correlation with the term of $1/Re_d$, while their differences depend on the function of porosity and constant parameters.

Concerning the relationship between k_0 and k_{θ_1} , Schubauer and Spangenberg [9] plot $k_{\theta_1}/\cos^2 \theta_1$ against $Re_d \cos \theta_1$, over a range of angle $0^\circ \leq \theta \leq 45^\circ$. They show that there exists a unique relationship between the two items for all screens and Reynolds numbers. Hence,

$$\frac{k_{\theta_1}}{\cos^2 \theta_1} = f(Re_d \cos \theta_1). \quad (14)$$

In the case $\theta_1 = 0^\circ$, we find

$$k_0 = f(Re_d), \quad (15)$$

according to the definition of k_{θ_1} in Eq. 1,

$$\frac{k_{\theta_1}}{\cos^2 \theta_1} = \frac{\Delta P}{\frac{1}{2} \rho U_1^2 \cos^2 \theta_1} = \frac{\Delta P}{\frac{1}{2} \rho u^2}. \quad (16)$$

Combining Eq.14, Eq.15, and Eq.16 yields

$$\frac{k_{\theta_1}}{\cos^2\theta_1} = k_0(\text{Re}_d \cos \theta_1). \quad (17)$$

To explain more clearly the model of Schubauer and Spangenberg, in which the pressure drop is independent of the incident angle θ_1 , we define a new form of the drag coefficient and Reynolds number, which do not explicitly depend on θ_1 :

$$\begin{aligned} k^n &= \frac{k_{\theta_1}}{\cos^2\theta_1} = \frac{\Delta P}{\frac{1}{2}\rho u^2}, \\ \text{Re}_d^n &= \text{Re}_d \cos \theta_1 = \frac{\rho u d}{\mu}, \end{aligned} \quad (18)$$

where n stands for the normal direction. Equation 17 then becomes:

$$k^n = k_0(\text{Re}_d^n). \quad (19)$$

Equations 18 and 19 indicate that in case of a deviated flow, the pressure drop through a given screen only depends on the normal component of the velocity. This relation is valid for all screens, independently of the porosity and the incident angle. This equation is widely accepted for the computation of the drag coefficient at an incident angle. Among others, the derivation of Brundrett's model, Eq. 13, is based on this relationship.

As an alternative to Schubauer and Spangenberg, Reynolds [10] proposes a theoretical model for k_{θ_1} by introducing a local incident angle θ :

$$k^n = k_0(\text{Re}_d^n) \left[\frac{1 + \sec \theta}{2} \right]. \quad (20)$$

In the method of Reynolds, the pressure drop is determined not only by the normal velocity component, but also by the tangential velocity component, which is reflected by the local incident angle θ .

The models cited above can produce significantly different results, depending on the screen type and the range of Reynolds numbers. We therefore executed numerical tests in Section 4.1 for the various relations, to assess their validity.

2.3. The deflection coefficient

Various theoretical and semi-empirical equations are proposed to describe the relationship between deflection coefficient and drag coefficient. In early studies, a relationship between k_{θ_1} and the deflection force coefficient F_{θ_1} was introduced for small incident angles, like in the work of Schubauer and Spangenberg [9]:

$$B \approx \frac{F_{\theta_1}}{\theta_1} = \frac{2k_{\theta_1}}{8 + k_{\theta_1}}, \quad (21)$$

or in the model of Taylor and Batchelor [8]:

$$B \approx \frac{F_{\theta_1}}{\theta_1} = 1 - \frac{1.1}{1 + k_{\theta_1}}, \quad (22)$$

in which k^n is directly related with B . Elder [11] assumes that the flow past a screen is equivalent to a uniform flow past a row of equally spaced vortices of circulation. A correlation is deduced from this assumption:

$$B = 1 - \frac{1}{\sqrt{1 + \sqrt{k^n}}}. \tag{23}$$

Gibblings [12] proposes yet another equation, by assuming that the flow direction takes a value halfway between the upstream and downstream directions:

$$B = 1 + \frac{k^n}{4} - \left[\left(\frac{k^n}{4} \right)^2 + 1 \right]^{1/2}. \tag{24}$$

As pointed by Laws and Livesey in their review [7], the evaluation of the deflection coefficient presents considerable difficulties, because the screen is rippled. As a result, no entirely satisfactory form of this relationship has ever been established. The literature on this issue is scarce and has been produced more than thirty years ago. It is therefore necessary to study these models, and individually compare them with experimental data to investigate the relationship between the resistance and the deflection coefficient.

3. NUMERICAL APPROACH

3.1. Direct numerical simulation

There exist various types of screens, such as perforated plates, arrays of cylinders and wire-gauges. In the present work, a stent is represented by an array of 2D cylinders (the stent struts), and formally modelled as a screen. The drag coefficient and the deflection coefficient of the stent are measured through numerical simulation and compared with the predictions of the screen models in Section 2. Direct numerical simulation in 2D is launched with the help of the open-source fluid solver Palabos (<http://www.palabos.org>) based on lattice Boltzmann method (see for instance Chen and Doolen [18]).

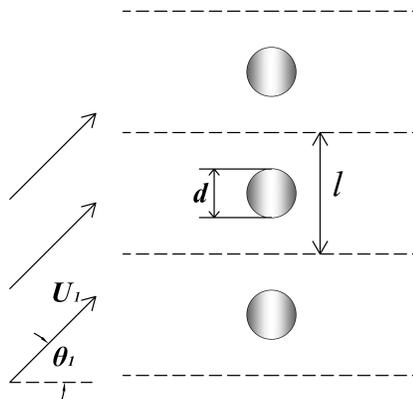


Figure 2. Periodic boundary condition

As shown in Fig. 1, the fluid is set up to flow through the stent, respectively screen, at an incident angle. It is assumed that the screen and flow field extend infinitely in the direction parallel to the stent, an effect which is achieved through periodic boundary condition on the top and bottom boundaries, as shown in Fig. 2. The fluid enters the numerical domain from the left at a uniform velocity and with an approaching angle which varies from 0° to 85° . A pressure boundary condition is imposed at the outlet, on the right. A no-slip boundary condition is applied on the surface of the struts. Both the inlet and outlet boundaries are far away from the stent, to guarantee that the upstream and downstream velocities have constant values at a sufficient distance from the stent. The density and viscosity of the fluid are similar to the ones of human blood: $\rho = 1000 \text{ kg/m}^3$ and $\mu = 3.3 \times 10^{-3} \text{ Pa} \cdot \text{s}$. The diameters d of the cylinders are taken to be $20 \mu\text{m}$, $30 \mu\text{m}$, and $40 \mu\text{m}$, corresponding respectively to a porosity of 0.8, 0.7, and 0.6. The distance l between is constant. The Reynolds number Re_d varies from 0.1 to 40, and is adjusted by changing the flow velocity at the inlet. The fluid mesh is uniform, and is resolved by 1000×200 nodes, with a discrete space step δx of $0.75 \mu\text{m}$.

The program measures directly both the normal and tangential components of the drag force on the cylinders, and the pressure drop ΔP and tangential component stress τ are derived from them according to the following relationship:

$$\begin{aligned}\Delta P &= \frac{F_n}{S}, \\ \tau &= \frac{F_t}{S},\end{aligned}\tag{25}$$

where F_n and F_t are the normal and tangential components of the total drag force exerted on the struts, and S represents the cross-sectional area of the stent. Since the upstream flow is uniform and the top and bottom boundaries are periodic, the value of ΔP and τ is the same for all the struts.

3.2. Coarse-grained force model

3.2.1. Force model In [2, 5], a coarse-grained model for the stents is proposed by means of a corrective source term to the momentum balance equations. In the present work, we follow the convention of screen models and use a body force term, which is correlated in a more sophisticated way to the local flow quantities, to reproduce both the expected pressure drop and velocity deviation. In fully resolved simulations, the body force is applied to all mesh nodes that are located inside stent struts. It is directly related to the upstream velocity, and to the parameters k^n and B predicted by the validated screen models, as follows:

$$\begin{aligned}f_n &= \frac{1}{\Delta L} \frac{1}{2} k^n \rho u_1^2, \\ f_t &= \frac{1}{\Delta L} \frac{1}{2} F_{\theta_1} \rho U_1^2 = \frac{1}{\Delta L} B u_1 v_1,\end{aligned}\tag{26}$$

where f_n and f_t are the normal and tangential components of the body force, and ΔL is the thickness of the volume over which the body force is applied. In the coarse grained case, we apply the force on one layer of nodes.

3.2.2. *Improved force model* Screen models, as they are presented in the literature, are only of limited use for our purpose, because they assume that the upstream velocity is known and is uniform, as shown in Fig. 1. For a stent placed in an artery, these assumptions are likely to be too restrictive. Since both the blood vessel and the stent are most often bent and follow a complex shape, the upstream velocity does not converge to a well-defined, asymptotic value. Furthermore, in an artery, just like in any other interior flow, the velocity exhibits a non-constant profile, varying from a zero velocity close to the walls to maximal velocity near the center.

It is therefore more adequate to relate the body force to a local value of the velocity $U(x, y)$, which is measured in the vicinity of the stent and can have different values along the surface of the stent, instead of a constant value U_1 . We therefore propose a novel phenomenological law in Section 4.3 which relates $U(x, y)$ to U_1 according to our numerical observations.

4. RESULTS

In this section, we investigate the drag coefficients k_0 and k^n , for angles in the range $0^\circ \leq \theta_1 < 90^\circ$, as well as the deflection coefficient B . The boundary conditions are periodic, to mimic a large system. The coefficients u_1 and v_1 , as defined in Section 3.1, are first measured in a fully resolved simulation of a stent with 2D cylindrical struts, and then compared to the various screen models.

4.1. Drag coefficient

4.1.1. *Drag coefficient at normal incidence k_0* The simulated drag coefficient at normal incidence k_0 is compared with the following four empirical correlations listed in Section 2.2: the model by Wieghardt (Eq. 8), as cited in [15], the model by Ehrhardt (Eq. 10), as cited in [15], the model by Wakeland and Keolian (Eq. 11) proposed in [17], and the model by Brundrett (Eq. 13) proposed in [16].

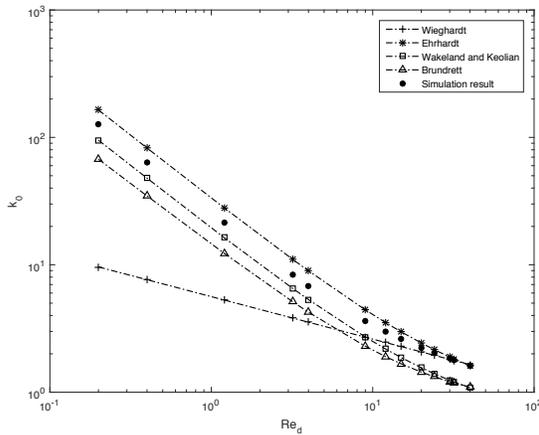


Figure 3. k_0 at a porosity of 0.6

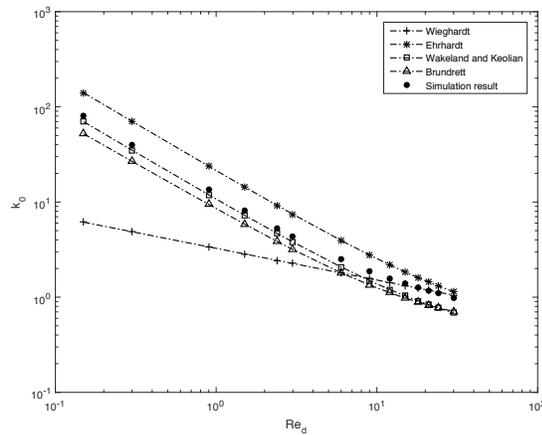
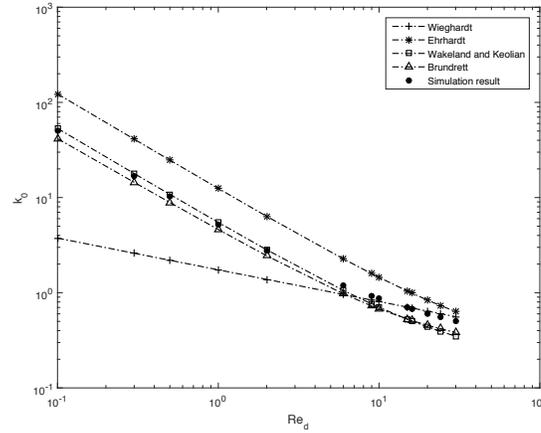


Figure 4. k_0 at a porosity of 0.7

Figure 3 to 5 compare the numerical data with model predictions at porosities of 0.6, 0.7 and 0.8. Figure 3 shows that the equation of Ehrhardt provides a best match for the drag coefficient at a screen porosity of 0.6, but Fig. 4 and Fig. 5 show that it does not work well at porosities of 0.7 and

Figure 5. k_0 at a porosity of 0.8

0.8. For porosities of 0.7 and 0.8, it can be seen that the model of Wakeland et al and the model of Brundrett are more consistent with the simulation results. However, while these two models provide good predictions for k_0 at low Reynolds number ($Re_d < 5$), they fail at higher Reynolds numbers ($5 < Re_d < 40$). In the latter regime, Wieghardt's model agrees much better with the numerical data. All in all, none of the models, taken by itself, provides satisfying results for the full range of Reynolds numbers and porosities of relevance to stents in aneurysms. But we notice that while the curve of Wieghardt's model does not exhibit the expected slope, the slope of the three other model curves is in accordance with the simulation results in a low Reynolds number region. This slope stems from the term proportional to $1/Re_d$, which is used in all the three models. To emphasize this similarity, we cast all three models into the following generic shape:

$$k_0(Re_d) = a/Re_d + b, \quad (27)$$

where a and b either depend on porosity or are constant parameters.

It should be pointed out that the discrepancies between the empirical models and our simulations are not surprising, as these models have themselves been obtained through correlations with different experimental setups, which are only partly comparable to the configuration of our numerical simulation. But given the fundamental importance of k_0 (the models for k^n and for B will be based on k_0), we now propose a new model, which yields a better match with our numerical simulation than the four models from the literature. It remains to be determined in future research if our own model is also appropriate for real, 3D stents in realistic arteries, or if one of the four models from the literature has a more general range of validity for actual stent simulations.

Based on Eq. 27, we propose a combination of terms depending on $1 - \beta$ and on $1/\beta$, inspired by the four models of the literature for the coefficients of a and b , which leads to a good match with our data. From a least-square fit, we obtain the following law:

$$k_0 = \left[12.25(1 - \beta) + 41.32 \left(\frac{1 - \beta}{\beta} \right)^2 \right] \frac{1}{Re_d} + \left[1.41(1 - \beta) + 1.11 \left(\frac{1 - \beta}{\beta} \right)^2 \right]. \quad (28)$$

The match between the numerical data points and our model, Eq. 28, is very good, as shown in Fig. 6, and we will therefore use this model as a starting point to derive more advanced screen-model based stent laws.

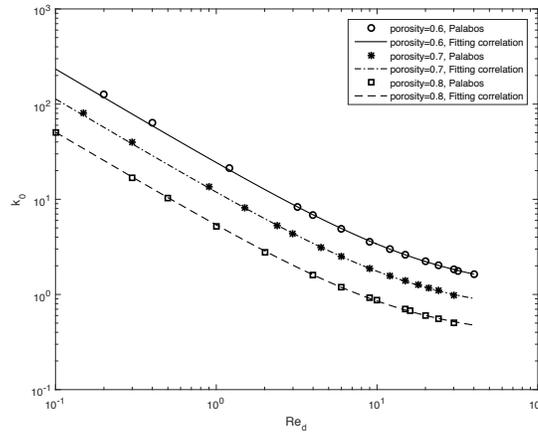


Figure 6. The fitting correlation for the drag coefficient k_0 at different porosities.

4.1.2. *Drag coefficient at an incident angle k^n* Now that the equation for k_0 is known, we turn to computing the drag coefficient of a more general screen, which a non-orthogonal alignment with respect to the flow. In Section 2.2, two models for the computation of k^n are presented. Both of them are now tested against numerical data, obtained with different parameters, including the angle θ_1 , the porosity β , and the Reynolds number Re_d . According to Eq. 18, k^n can be easily obtained from the pressure drop across the stent. The results are presented in Fig. 7. The black symbols represent the simulation results, while the blue lines and the blue crosses stand for the model of Schubauer *et al* and the model of Reynolds respectively.

It is observed that both models accurately predict k^n for small angles, such as 30° and 45° , deviate slightly at $\theta_1 = 60^\circ$, and show substantial deviations at large angles (75° and 85°).

Furthermore, both models perform well at low Reynolds number Re_d^n , when $Re_d^n < 1$, but some deviation appear when $Re_d^n > 1$. The discrepancies are particularly significant at angles of 75° and 85° .

As far as porosity is concerned, at intermediate Reynolds number region ($Re_d^n > 1$), the model of Schubauer *et al* is more accurate for $\beta = 0.6$, while the model of Reynolds is more accurate for $\beta = 0.8$, as it clearly stands out in the plots corresponding to angles of 75° and 85° .

In summary, both models are consistent with the simulation results at low incident angles for all the simulated Reynolds numbers. For large angles ($\theta_1 \geq 75^\circ$), the two models still perform well at low Reynolds numbers ($Re_d^n < 1$), but they can not accurately describe k^n in the intermediate Reynolds number region ($Re_d^n > 1$). The difference between the two models depends on the porosity. The model of Schubauer *et al* is better for the comparatively low value of $\beta = 0.6$ and the model of Reynolds is suitable for the comparatively high value of $\beta = 0.8$. The model of Schubauer *et al* can be considered a good candidate for the intermediate value $\beta = 0.7$, given its simplicity.

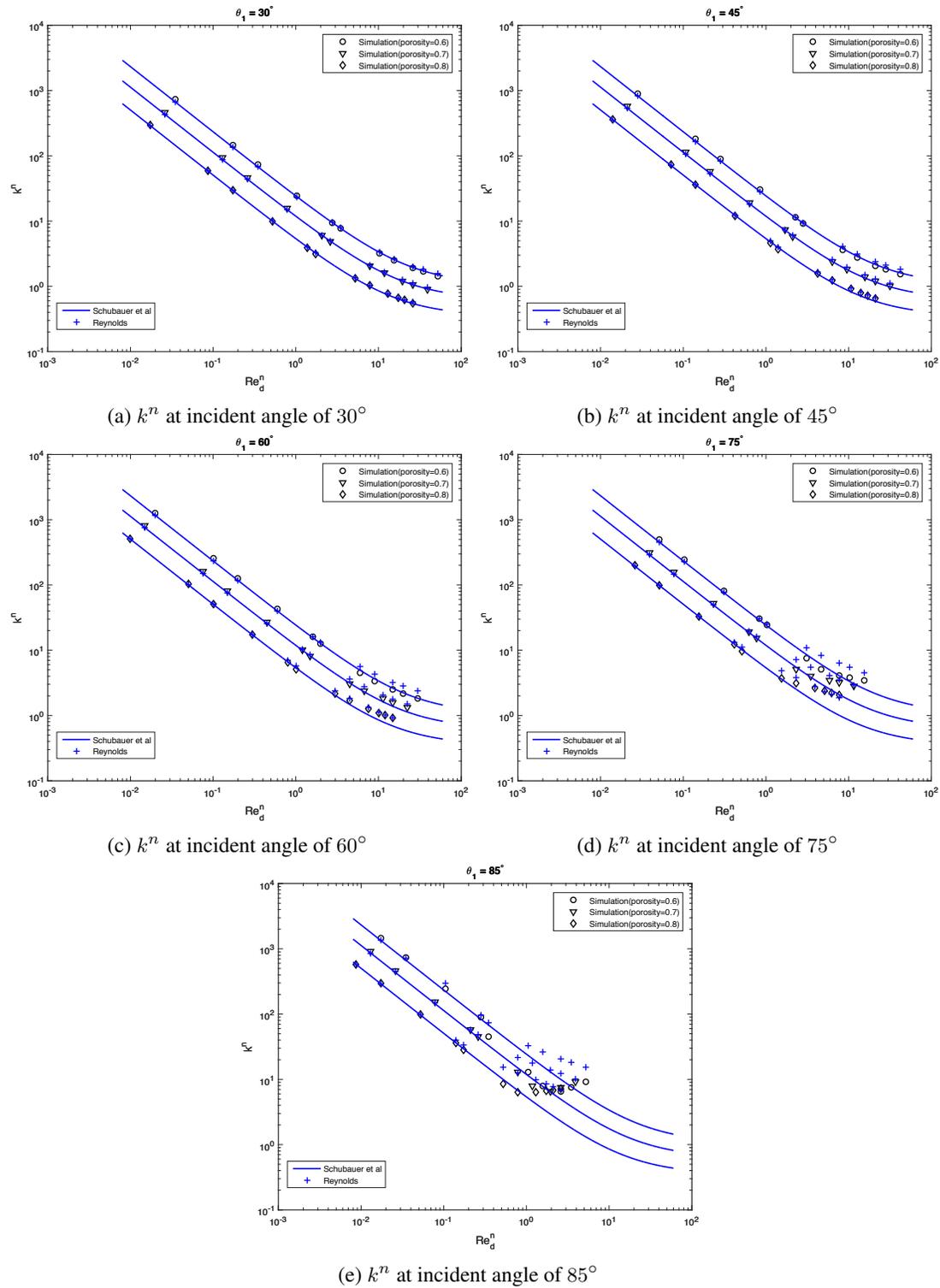


Figure 7. Compare simulated k^n with two models at different incident angles

4.2. Deflection

4.2.1. *The measuring of the deflection coefficient* The deflection coefficient B , defined by Eq. 5, represents the deviation of velocity caused by the screen. According to Eq. 5, B can be obtained

by measuring the downstream and upstream velocities. Another method is given by Eq. 7, which introduces a relationship between the deflection coefficient B and the tangential stress τ . The tangential component of the drag force on the cylinder struts is measured directly in the simulation, and τ is computed by dividing the tangential drag force by the cross sectional area of the stent. As a next step, B is computed from τ . For the sake of discussion, one may distinguish between the coefficient B in Eq. 5, as it is computed from the velocity (and which will therefore be called B_v), and the coefficient B in Eq. 7, which it is obtained from the drag force (and will therefore be called B_f). Equation 7, which is of central importance to our model, is only valid when B_f is equal to B_v . Figure 8 verifies the consistency of these two approaches.

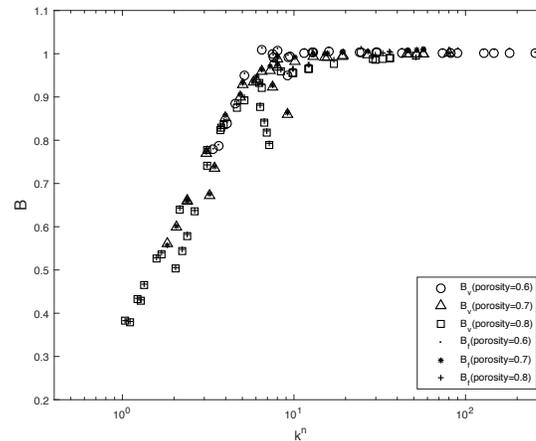


Figure 8. Two methods for computing the deflection coefficient B

In practice, these two methods can actually produce quite different results. It should be reminded that Eq. 7 is derived from Eq. 4, which assume the existence of an asymptotic upstream and downstream flow far from the screen. According to our numerical experiments, B_f is very sensitive to the distances separating the screen with the inlet and the outlet boundaries. At low Reynolds number in particular, a very long distance between the inlet and the screen is required for the model to be valid.

Moreover, Fig. 8 shows that the data points corresponding to a given porosity are strongly correlated, suggesting the existence of a simple relationship between k^n and B . At low angles (and using a logarithmic axis for k^n), the correlation is linear, while strong deviations from this trend are observed at large angles ($\theta_1 = 75^\circ$ and 85°). Consistently with previous observations, it is concluded that the hydrodynamic effect of screens obeys different laws at very large inclinations than at low and moderate inclinations.

4.2.2. Equations for B As described in Section 2.3, many attempts have been made to establish the relationship between the drag coefficient and deflection coefficient, and two main families of correlations are proposed: one correlates k_{θ_1} with B , as shown in the model of Schubauer et al (Eq. 21) and the model of Tayler et al (Eq. 22). The other relates k^n to B , as presented in the model of Elder (Eq. 23) and the model of Gibblings (Eq. 24).

Figure 9 compares the deflection coefficient measured in the simulation with the different model predictions. It is seen that none of the models provides a good match for B , and we therefore

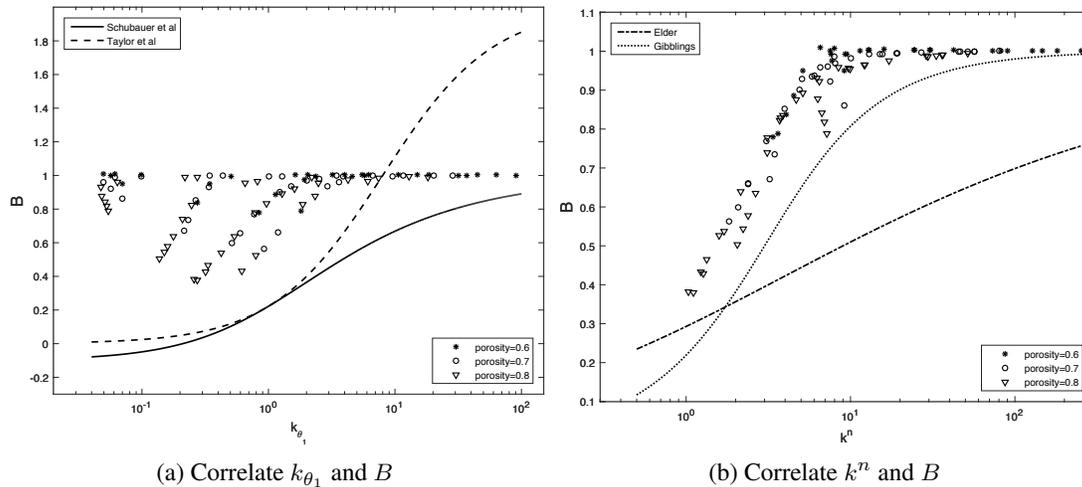


Figure 9. Compare the existing models of B with simulation results

need to formulate our own, more appropriate model. A qualitative comparison between Fig. 9a and Fig. 9b shows that the correlation between B and k_{θ_1} strongly depends on the porosity, while the correlation between B and k^n exhibits a more uniform trend, independent on porosity. Based on this observation, we therefore formulate a model of B depending on k^n .

It can further be pointed out that although none of the models directly matches the numerical data, the model of Gibblings follows most closely the general trends of the numerical curve. This observation is compatible with the findings of Laws and Livesy [7], who report that a best overall agreement is obtained with the Gibblings [12] model. It is therefore natural to use the general shape of this equation,

$$B = 1 + ak_{\theta_1} - [(ak_{\theta_1})^b + c]^{1/b}, \tag{29}$$

as a basis for the new model, and to reassign the constants to better fit the numerical data.

A least-square fit leads to the following model:

$$B = 1 + \frac{k^n}{3.4} - \left[\left(\frac{k^n}{3.4} \right)^3 + 0.797 \right]^{1/3} \tag{30}$$

Figure 10 shows that the fitted model of Eq. 30 matches the numerical results very well. Thus, with the constants adjusted to our specific 2D numerical setup, the model of Gibblings turns out to be of very high quality.

In summary, we have proposed in Sections 4.1 and 4.2 two model equations that allow to compute the values of k^n and B for our 2D stents. From there, Eq. 26 can be used to compute the body forces f_n and f_t , and apply them as a replacement for a real flow diverter in a coarse-grained fluid flow simulation.

4.3. Improve force model

In the previous sections, we have shown that screen models can be used to model a flow diverter through a body force term. We have however also discussed in Section 3.2.2 that in their original form, screen models are insufficient, as they are based on the assumption of a known constant

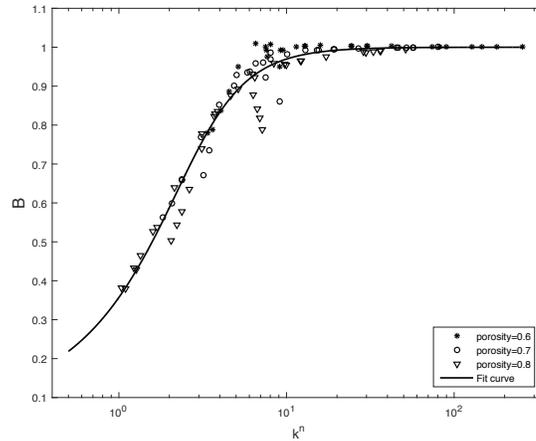


Figure 10. The fitting correlation of the deflection coefficient

upstream velocity, while the actual properties of a flow diverter may depend on a local, non-constant velocity $U(x, y)$. To relate these two concepts, we improved the original force model by introducing another coefficient: the velocity reduction coefficient

$$r = v/v_1, \tag{31}$$

which is the ratio between the local velocity and the upstream velocity in the tangential direction. For the normal velocity component, there is no need for a reduction coefficient, as the local normal velocity is equal to the upstream one due to flow conservation.

From numerical simulations of the fully resolved flow diverter, it is possible to plot the reduction coefficient r against Re_d^n , using the same porosities as the ones tested in the previous section.

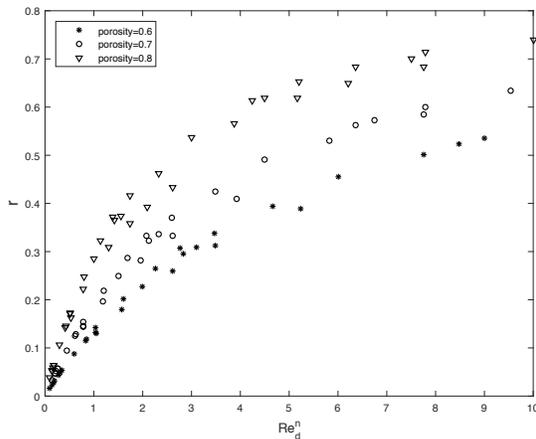


Figure 11. The velocity reduction coefficient versus Re_d^n

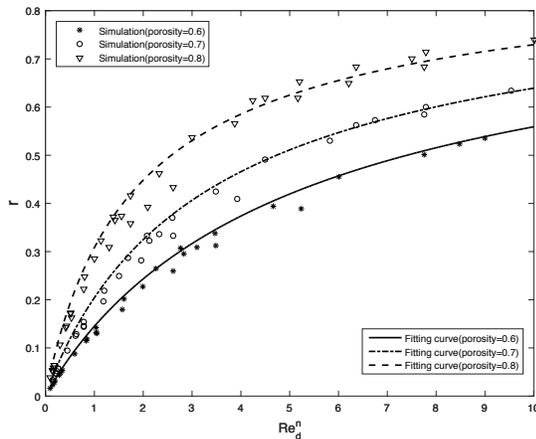


Figure 12. The fitting curve of the velocity reduction coefficient

Figure 11 shows that the reduction coefficient r depends not only Re_d^n , but also on the porosity. An appropriate model for the reduction coefficient could not be found in the existing literature. We

therefore propose a novel model based on a least-square fit of the data of Fig. 11.

$$\begin{cases} r &= 1 + a\text{Re}_d^n - [(a\text{Re}_d^n)^{\frac{3}{2}} + 1]^{\frac{2}{3}} \\ a &= 7.68\beta^2 - 8.842\beta + 2.734 \end{cases} \quad (32)$$

Eq. 32 adopts the same form as the model of the deflection coefficient, with a slight difference: the term Re_d^n is no longer multiplied by a constant, but by a function of the porosity. This form also guarantees that r is no greater than 1, because v should be smaller than v_1 .

Figure 12 shows that Eq. 32 offers a consistent fit with the simulation data, and that it is therefore appropriate to choose a model inspired by deflection coefficient, in spite of the different nature of these coefficients: The deflection coefficient B , Eq. 30, is related to k^n , while the velocity reduction coefficient r , Eq. 32, is related to Re_d^n .

The methodology to compute the body force through local velocities can be summarized by the following steps:

First, the pressure drop is computed from k_0 and k^n :

$$\begin{cases} \Delta P &= \frac{1}{2}k^n\rho u^2 \\ k^n &= k_0(\text{Re}_d^n) \text{ or } k^n = k_0(\text{Re}_d^n) \left[\frac{1+\sec\theta}{2} \right] \\ k_0 &= \left[12.25(1-\beta) + 41.32 \left(\frac{1-\beta}{\beta} \right)^2 \right] \frac{1}{\text{Re}_d} + \left[1.41(1-\beta) + 1.11 \left(\frac{1-\beta}{\beta} \right)^2 \right], \end{cases} \quad (33)$$

where the equation for k^n depends on the porosity of the screen and the Reynolds number of the flow.

Then, the tangential component τ of the external force is computed from the deflection coefficient B and from k^n :

$$\begin{cases} \tau &= Bw/r \\ B &= 1 + \frac{k^n}{3.4} - \left[\left(\frac{k^n}{3.4} \right)^3 + 0.797 \right]^{1/3}, \\ r &= 1 + a\text{Re}_d^n - [(a\text{Re}_d^n)^{\frac{3}{2}} + 1]^{\frac{2}{3}}, \end{cases} \quad (34)$$

where

$$a = 7.68\beta^2 - 8.842\beta + 2.734$$

Finally, the body force is computed with the help of Eq. 26:

$$\begin{cases} f_n &= \frac{1}{\Delta L} \frac{1}{2} k^n \rho u_1^2, \\ f_t &= \frac{1}{\Delta L} \frac{1}{2} F_{\theta_1} \rho U_1^2 = \frac{1}{\Delta L} B u_1 v_1. \end{cases}$$

This model can be used for simulating the influence of a flow diverter on the fluid using body forces.

4.4. Validation

4.4.1. An inclined stent in a straight channel In order to validate the model proposed in this article, we simulated fluid flows through an inclined stent placed in a channel.

In the direct numerical simulation, the inclined screen, represented by an array of 2D cylinders, is placed in a channel with a length $L = 23.5\text{ mm}$ and a diameter $D = 4.7\text{ mm}$. The stent with a porosity $\beta = 0.786$ and a cylinder diameter $d = 47.6\text{ }\mu\text{m}$ is deployed at a distance $L/4$ of the inlet with an incident angle of 45° . The full fluid mesh has a resolution of 4950×990 nodes. The density and viscosity of the fluid are $\rho = 1000\text{ kg/m}^3$ and $\mu = 3.3 \times 10^{-3}\text{ Pa}\cdot\text{s}$. The analytic velocity profile of a Poiseuille flow is applied as a boundary condition on the inlet. The outlet implements a pressure boundary condition. The Reynolds number is $\text{Re} = \frac{\rho u_{max} D}{\mu} = 600$, where u_{max} is the maximum velocity at the inlet boundary.

In the coarse-grained simulation using the improved screen model, body forces are applied throughout a plane with a one-cell thickness, superposed to the location of the stent nodes of the direct numerical simulation. Since the porosity is very close to 0.8, and the maximum $\text{Re}_d = \frac{\rho u_{max} d}{\mu} = 6.06$, we use the model of Reynolds (Eq.20) to compute k^n . Compared to the direct numerical simulation, the number of mesh nodes is reduced by a factor 9, with a resolution of 1650×330 nodes.

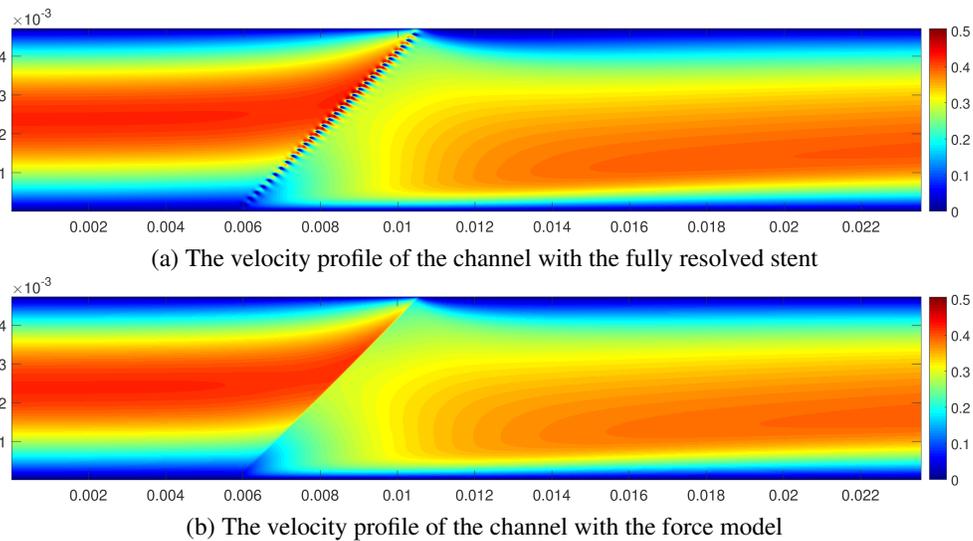


Figure 13. Velocity profile of the straight channel. The unit of the velocity is m/s , while the x - and y -coordinates are measured in m .

Figure 13 compares the velocity profiles of a fully-resolved simulation and of the improved force model. The force model proves capable to predict the velocity field very well.

the absolute error of the velocity of the force model from the fully resolved simulation is computed as follows:

$$\epsilon = \frac{1}{N} \sum_{i=1}^N |u_i^f - u_i^c| = 6.9698 \times 10^{-3}\text{ m/s} \quad (35)$$

where u_i^f and u_i^c are corresponding to the velocities of the fully resolved simulation (the fine grid) and the force model (the coarse grid). i is for the coarse grid.

The percent error equals to:

$$\delta = 100\% \times \frac{\epsilon}{u^f} = 2.3\% \quad (36)$$

The total pressure drop across the stent is also very well predicted, with a value of $\Delta P = 101.3 Pa$ for the improved stent model, and a value of $\Delta P = 106.3 Pa$ in the direct numerical simulation. Thus, the improved screen model introduces an error of just of 4.7% in the overall pressure drop, while the run time of the corresponding simulation is just 1/10 of the time of the fully-resolved one.

4.4.2. *A stent in aneurysm* After the validation of the stent model in a simple channel flow, we run a more demanding numerical experiment, in which a stent is placed into an idealized bent artery with an aneurysm.

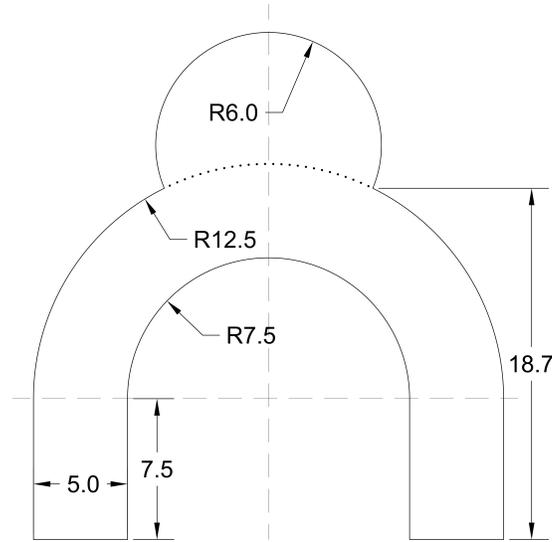


Figure 14. Representation of the bent artery with aneurysm used in the simulation, parametrized with the radius R . Units are in mm .

Figure 14 depicts the simulation setup for this test. The stent consists of struts with diameter of $40 \mu m$ and has a porosity of 0.8. The fluid enters at the left tube and leaves at the right tube. The diameter of the artery is $D = 5 mm$, and the inlet velocity is $0.396 m/s$. The Reynolds number is again equal to 600. The boundary conditions and the physical characteristics are the same as in the previous example in Section 4.4.1. The discrete space step δx of the fully resolved simulation is $1 \times 10^{-5} m$, while the coarse-grained simulation with force model uses half the resolution, with $\delta x = 2 \times 10^{-5} m$.

Figure 15 shows the velocity profiles obtained in a simulation without stent, with a fully resolved stent, and with the force model. The force model successfully prevents the flow from entering the aneurysm, just like the real stent. Figure 16 shows that without a stent, a large vortex is formed inside the aneurysm, while both the simulations of fully resolved stent and the force model prevent the formation of the vortex. The full run time with the force model is 1/3 of the run time of the fully-resolved simulation.

In order to quantitatively validate the force model, the RMS value of the velocity is computed inside the aneurysm as follows:

$$\bar{u} = \sqrt{\frac{1}{N\Delta s} \sum_{i=1}^N \|\vec{u}_i\|^2 \Delta s} \quad (37)$$

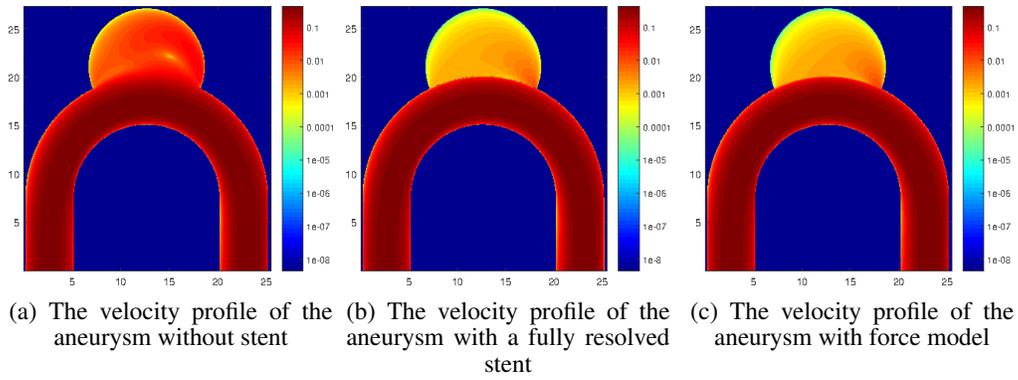


Figure 15. Comparison of velocity profiles. The unit of the velocity is m/s , while the x - and y - coordinates are measured in mm . The color palette is defined over a logarithmic scale to emphasize weak flows.

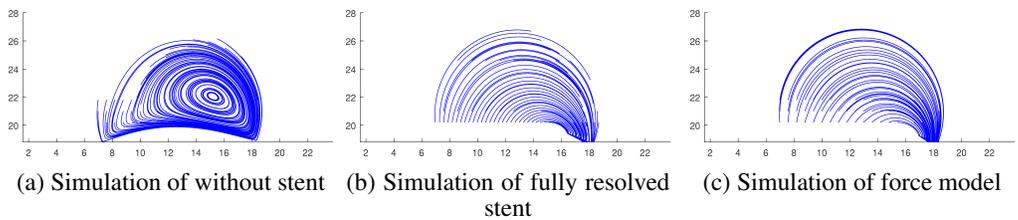


Figure 16. Streamlines in aneurysm

where N is the total number of mesh nodes contained in the aneurysm, and Δs is the area of per node.

	without stent	fully resolved stent	force model
Pressure drop (Pa)	28.21	29.17	29.11
Average velocity in aneurysm (m/s)	19.6×10^{-3}	2.08×10^{-3}	1.86×10^{-3}

Table I. The pressure drops and the average velocities in aneurysm of the three simulation cases

Table I shows the pressure drop between the inlet and the outlet and the average velocity inside the aneurysm for the three following cases: without stent, with fully resolved stent, and with force mode. It must be noted that the relative pressure drop caused by the stent is small, given the length of the full simulated artery section. Nevertheless, the value predicted with use of a force model is much better than the one without force model. The ability of the force model to predict the average velocity inside the aneurysm is even more impressive. Without force model, the average velocity is over predicted by almost a factor 10, while the force model introduces an error as small as 10%.

Besides, the local absolute error of the velocity inside the aneurysm is:

$$\epsilon = \frac{1}{N} \sum_{i=1}^N |u_i^f - u_i^c| = 1.7926 \times 10^{-4} m/s \tag{38}$$

where u_i^f and u_i^c are corresponding to the velocities of the fully resolved simulation and the force model inside the aneurysm. i is for the coarse grid.

The local percent error of the velocity equals to

$$\delta = 100\% \times \frac{\epsilon}{u^f} = 9.6\% \quad (39)$$

As a final test, the force model is applied in a pulsatile flow. The inlet velocity exhibits the following profile:

$$u = u_{min} + u_{min} |\sin(t/T)|. \quad (40)$$

Here, $u_{min} = 0.198 \text{ m/s}$, which is half of the inlet velocity of the steady flow, t is the time, with the units of s , and T is the length of the period, with a value of $T = 2 \text{ s}$. We chose this expression to create a steep time variation of the flow.

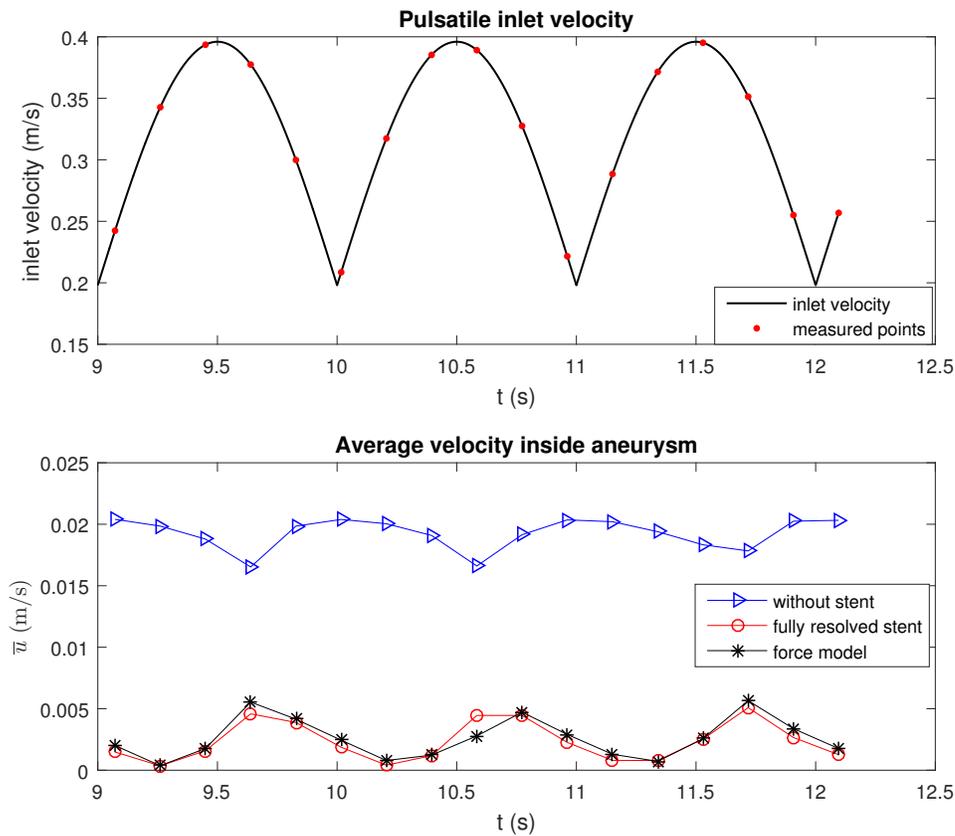


Figure 17. Pulsatile inlet flow

Figure 17 shows the numerical results of a pulsatile flow in a fully resolved simulation, and in a coarse-grained simulation with and without force model. The top figure presents the inlet velocity and the measured points. The bottom figure compares the RMS velocity inside the aneurysm \bar{u} in each of the three cases. To overcome the initial transient state, the results are compared starting at $t = 9 \text{ s}$. It can be seen in Figure 17 that the force model shows an excellent consistency with the fully resolved simulation, while the coarse-grained simulation without force model substantially overpredicts the flow.

The steady and the pulsatile test case above both show that the proposed force model is able to successfully replace the fully resolved flow diverter and decrease the computation cost of a simulation. It is a good model for both the pressure drop and the velocity deviation caused by the flow diverter, and it is available for general flow diverters with different porosities.

5. CONCLUSION

We propose a new modeling framework to represent a flow diverter through a continuum approach in coarse-grained blood-flow simulations of a human artery. While previous models available in the literature, such as [2], represent the flow diverter through a porous media model, we have instead decided to base our work on screen models, which are explicitly designed to represent the properties of thin surfaces. In particular, screen models fully consider the impact of a surface-like object on the incident flow, including both the pressure drop and the precise flow deviation.

We reviewed and validated the existing models for the drag coefficient k_0 at normal incidence, the drag coefficient k^n at an incident angle, and the deflection coefficient B , by comparing them with the result of numerical experiments around a fully resolved stent structure in 2D.

By means of a best fit through the numerical data set, we were able to adapt the model for the drag coefficient k_0 over the range $0.1 \leq Re_d \leq 40$ and thus obtain a more accurate model for our purposes, which we propose in Eq. 28. We then turned to the drag coefficient k_{θ_1} for incident angles, and determined the range of validity of two different models for this coefficient. Through numerical simulation, we confirmed the dominant role of the normal velocity component of the flow, as predicted by the screen model of Schubauer et al proposed in Eq. 19. At high porosity values however, we showed that the tangential velocity component, which is proposed by Reynolds in [10] through Eq. 20, but which often is neglected in the literature, must also be taken into account for better accuracy. These equation yield outstanding results, except for very large angles ($\theta > 75^\circ$), *i. e.* when the flow is close to parallel to the stent surface.

For the deflection coefficient finally, the data of the simulations appeared to differ most substantially from the models found in the literature. The model of Gibblings [12] however was able to predict the trends of our numerical data well and could be adjusted to achieve a very good fit after adjusting its model constants, as proposed in Eq. 30.

We argued however that the existing screen models are limited by the fact that they correlate the predicted values (drag and deflection coefficient) with an upstream velocity, which is assumed to be asymptotically constant. Given that this assumption does not hold in any complex geometry like a human artery, we introduced a corrective factor to the existing screen models in the form of a velocity reduction coefficient, in order to relate the predicted values to the local velocity, close to the stent.

In summary, we have performed a selection of screen models which we have combined and further improved to propose a novel modeling framework for stent-like objects. The obtained screen model is summarized in Section 4.3. To validate this model we carried out two numerical experiments, in which the actual stent geometry is replaced by the forcing term proposed by the model. In our examples, the total number of mesh nodes was reduced by a factor 9 and 4 respectively. The flow fields produced by the fully resolved and the coarse-grained simulations are very similar. These

two examples show that our proposed screen model provides a powerful theoretical base to the computational simulation of blood flow with stents, which extends the possibilities of classical models based on a porous media approach.

This research must be understood as a first-principle investigation of the physics of flow diverters, intended to provide a global theoretical framework for its modeling. The model presented in Section 4.3 has been tuned to yield an optimal match with the 2D model of this article. This model, in its exact current shape and with its constants, cannot be expected to provide, out of the box, an optimal prediction for all types of flow diverters. For example, while we found Eq. 28 to optimally predict the normal-incidence drag coefficient for our 2D stent representation, it might be that a given industrial flow diverter is better described by one of the other models reviewed in Section 4.1.1. Similarly, the model constants for the deflection coefficient in Eq. 30 will need to be adjusted to match the properties of specific industrial flow diverters. Besides, as discussed in Section 4.4.2, the application of the force model is also limited by its prerequisite conditions. Furthermore, in a real aneurysms, the bending of the flow diverter may cause an inhomogeneity of the porosity, which affects the performance of the stent, as investigated by [19]. Due to the fundamental nature of our work, we did not presently consider this problem. Actually, our model could deal with a local porosity. While not yet tested, we expect our approach to be able to address such an issue. We intent to conduct such an investigation of industrial 3D stents in realistic arteries in future research, which will be oriented towards high performance CFD and flow diverter specific parameter fitting.

ACKNOWLEDGEMENT

We acknowledge partial funding and access to high performance computing resources from the CADMOS center (<http://www.cadmos.org>) and partial funding from the European Union Horizon 2020 research and innovation programme for the CompBioMed project (<http://www.compbiomed.eu/>) under grant agreement 675451.

REFERENCES

1. Appanaboyina S, Mut F, Lhner R, Putman CM, Cebal JR. Simulation of intracranial aneurysm stenting: techniques and challenges. *Comput Methods Appl Mech Eng.* 2009; **198**(45-46):3567-3582.
2. Augsburg L, Reymond P, Rufenacht DA, Stergiopoulos N. Intracranial stents being modeled as a porous medium: flow simulation in stented cerebral aneurysms. *Ann Biomed Eng.* 2010; **39**(2):850-863.
3. Cebal JR, Lhner R. Efficient simulation of blood flow past complex endovascular devices using an adaptive embedding technique. *IEEE Trans Med Imaging.* 2005; **24**(4):468-476.
4. Anzai H, Falcone JL, Chopard B, Hayase T, Ohta M. Optimization of strut placement in flow diverter stent for four different aneurysm configurations. *J Biomech Eng.* 2014; **136**(6):0610061-0610067.
5. Raschi M, Mut F, Lhner R, Cebal JR. Strategy for modeling flow diverters in cerebral aneurysms as a porous medium. *Int J Numer Method Biomed Eng.* 2014; **30**(9):909-925.
6. N.M.P.Kakalis and A.P.Mitsos and J.V.Byrne and Y.Ventikos. The Haemodynamics of Endovascular Aneurysm Treatment: A Computational Modelling Approach for Estimating the Influence of Multiple Coil Deployment. *IEEE Trans. Med. Imaging.* 2008; **27**(6):814–824.
7. Laws EM, Livesey JL. Flow through screens. *Ann Rev Fluid Mech.* 1978; **10**:247-266.
8. Taylor GI, Bachelor GK. The effect of wire gauze on small disturbances in a uniform stream. *Quart Journ Mech and Applied Math.* 1949; **2**:1-29.
9. Schubauer GB, Spangenberg WG, Klebanoff PS. Aerodynamic characteristics of damping screens. *NACA, Technical Notes 2001.* 1950; Washington, USA.
10. Reynolds AJ. Flow deflection by gauze screens. *J Mech Eng Sci.* 1969; **11**(3):290-294.

11. Elder JW. Steady flow through non-uniform gauzes of arbitrary shape. *J Fluid Mech.* 1959; **5**(3):355-368.
12. Gibblings JC. The pyramid gauze diffuser. *Ing Arch.* 1973; **42**:225-233.
13. Mehta RD. Turbulent boundary layer perturbed by a screen. *AIAA Journal.* 1985; **23**(9):1335-1342.
14. Dunn W, Tavoularis S. The use of curved screens for generating uniform shear at low Reynolds numbers. *Exp Fluids.* 2007; **42**(2):281-290.
15. Kolodziej A, Jaroszynski M, Janus B, Kleszcz T, Lojewska J, Lojewski T. An experimental study of the pressure drop in fluid flows through wire gauzes. *Chem Eng Comm.* 2009; **196**(8):932-949.
16. Brundrett E. Prediction of pressure drop for incompressible flow through screens. *J Fluids Eng.* 1993; **115**(2):239-242.
17. Wakeland RS, Keolian RM. Measurements of resistance of individual square-mesh screens to oscillating flow at low and intermediate Reynolds numbers. *J Fluids Eng.* 2003; **125**(5):851-862.
18. Chen S, Doolen GD. Lattice Boltzmann method for fluid flows. *Annu Rev Fluid Mech.* 1998; **30**(1):329-364.
19. Shi Y, Tang GH, Tao WQ. Lattice Boltzmann study of non-Newtonian blood flow in mother and daughter aneurysm and a novel stent treatment. *Adv Appl Math Mech* 2013; **6**(2):165-178.