



Konferencija Srce DEI

# UNITY

Where **AI** meets **quantum chemistry** in drug discovery

*Pathway-centric, multi-target screening with AI-guided semi-empirical QM scoring*

**D. Tomić · B. Kolarek**

*Centre for Informatics and Computing*

*Ruđer Bošković Institute · Zagreb, Croatia*



**Srce DEI 2026**





## THE CLINICAL CHALLENGE

# Why combination therapy is essential — and hard

**2nd**

Most commonly diagnosed cancer in men worldwide (prostate)

**<10%**

5-year survival in KRAS-mutant pancreatic ductal adenocarcinoma

**~54,720**

Possible triple combinations from just 70 candidate drugs

**Castration-resistant prostate cancer** and **KRAS-mutant PDAC** share the same biology problem: *compensatory signalling* across interconnected pathways defeats single-target therapy. Combinations are essential — but the search space is too large for wet-lab triage and clinical trials are too expensive to run blind.



VINI · THE FOUNDATION

# A pathway-aware, multi-target in silico screening platform

## Pathway-aware

Built directly on the disease's KEGG cancer pathway — genes, proteins, and known interactions are first-class inputs.

## Context-specific

Cell-line mutation and gene expression data (CCLE / COSMIC) scale every interaction term to the biology being modelled.

## Multi-target by design

Scores combinations of drugs against all receptors in the pathway simultaneously — not one target at a time.

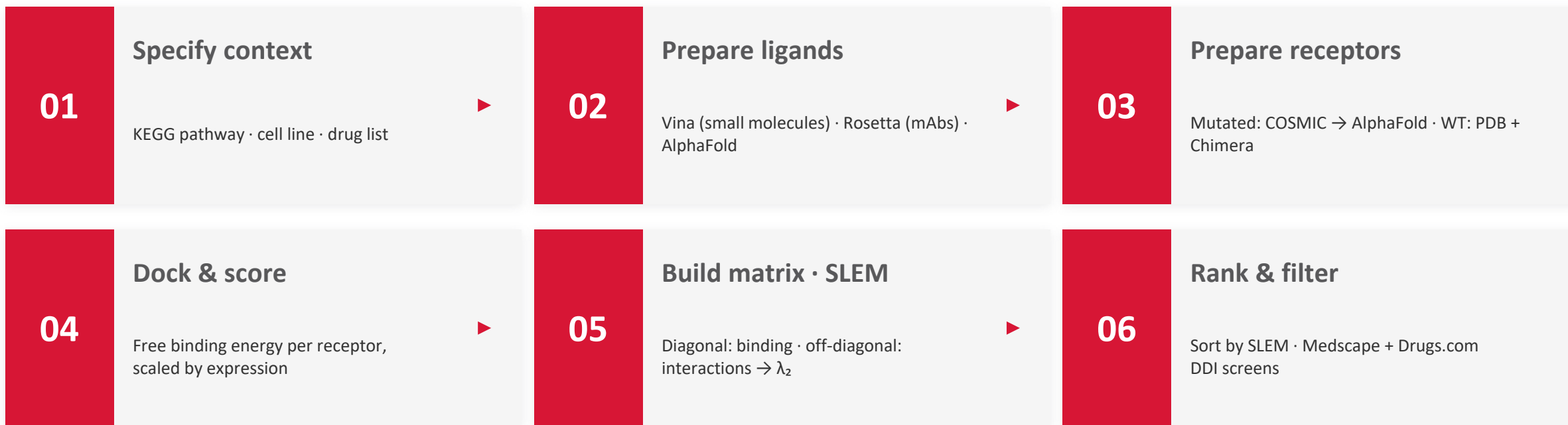
## Generalisable

Same engine, > 700 KEGG disease pathways. Applied across cancer types — and SARS-CoV-2 during the pandemic.



## VINI WORKFLOW

# Six steps from drug list to a ranked, DDI-clean shortlist



**Note** step 4 uses empirical scoring functions — *accurate enough to rank, not accurate enough to call winners. We return to this.*



## THE SCORING FUNCTION

# SLEM — second-largest eigenvalue in magnitude

## What it measures

Rate at which the cancer pathway transitions between states under a drug.

Higher SLEM  $\Rightarrow$  slower progression  $\Rightarrow$  better drug efficacy.

*Based on the Fastest-Mixing-Markov-Chain framework (Boyd et al., 2004).*

## What goes into the matrix

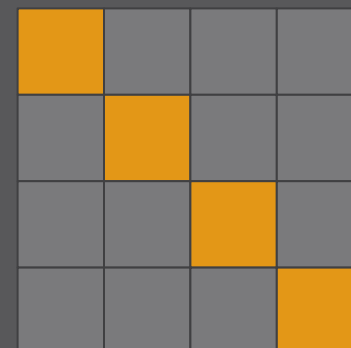
### Diagonal

Free binding energy between drug and each KEGG receptor, scaled by expression.

### Off-diagonal

Known receptor–receptor interactions from the KEGG pathway.

## MATRIX VIEW



- Orange square: Binding energies (drug  $\rightleftharpoons$  receptor)
- Gray square: Receptor–receptor interactions

$$\text{SLEM} = |\lambda_2(M)|$$

*One scalar per drug or combination · directly comparable*



VALIDATION

## Predicted SLEM vs. experimental GI50 across 28 NCI-60 cell lines



### KEY FINDINGS

|PCC|

**0.39 – 0.60**

*across all 28 cell lines, 9 KEGG pathways*

**100%**

agreement on DU-145 and PC-3 prostate lines

**79.3%**

agreement with in vivo outcomes across 16 cancer types



PROOF OF CONCEPT · CASTRATION-SENSITIVE PROSTATE

# From 54,720 candidates to 18 clinically viable triples

70	<b>Candidate drugs</b> <i>expert-curated small molecules</i>
22	<b>Top monotherapies</b> <i>highest individual SLEM</i>
1,540	<b>Triple combinations</b> <i>22 choose 3</i>
100	<b>Medscape DDI-clean</b> <i>top-ranked, no major DDI</i>
18	<b>Drugs.com DDI-clean</b> <i>clinically viable shortlist</i>

#	Combination	SLEM	Key targets
1	Temsirolimus · Abiraterone · Triptorelin	-50.01	mTOR · CYP17 · GnRH
2	Temsirolimus · Alectinib · Fludarabine	-44.61	mTOR · ALK · DNA
3	Temsirolimus · Alectinib · Venetoclax	-44.61	mTOR · ALK · BCL-2
4	Temsirolimus · Alectinib · Triptorelin	-44.58	mTOR · ALK · GnRH
5	Venetoclax · Fludarabine · Triptorelin	-43.86	BCL-2 · DNA · GnRH

Top 5 of 18 — recurring axes: mTOR · ALK · BCL-2 · DNA synthesis · androgen (AR / CYP17 / GnRH)

**The point** isn't just that VINI produces a clean shortlist — it's that the engine **works**. The remaining gap to clinical impact is **scoring accuracy**.



# Why empirical scoring functions cap our predictive accuracy



## Empirical scoring (Vina, Rosetta)

### How it works

Sums of weighted physicochemical terms — van der Waals, hydrogen bonds, hydrophobics, torsions — calibrated against a training set. Seconds per pose. **Fast.**

### What it misses

- Polarisation in the binding pocket
- Charge transfer across hydrogen bonds
- Accurate electrostatics and dispersion
- Solvent reorganisation thermodynamics

⇒ ranking is OK · absolute affinity is noisy · false-positive rate is high

## THE CONSEQUENCE — IN OUR OWN DATA

# 0.39 – 0.60

|PCC| of VINI vs. experimental GI50

**This is the empirical-scoring ceiling — not a model bug.**

Better pathway modelling can't push it higher.

More cell lines can't push it higher.

**Only better physics in the scoring step can.**



THE ACCURACY ↔ COMPUTE TRADE-OFF

## Three tiers of scoring — and where the sweet spot lives

### Empirical

*AutoDock Vina · Rosetta*

#### SPEED

seconds / pose

#### ACCURACY

moderate ranking · noisy affinities

**Fast — but capped**

*Initial triage at scale*

### Semi-empirical QM

*SQM2.20 · PM6/7 + corrections*

#### SPEED

minutes / pose

#### ACCURACY

near-QM affinities · fraction of cost

**The sweet spot**

*UNITY's scoring engine*

### Full QM

*DFT · ab initio · MP2 / CCSD(T)*

#### SPEED

hours – days / pose

#### ACCURACY

highest available accuracy

**Accurate — intractable**

*Reference / spot-checks*

**UNITY's bet:** semi-empirical QM gives most of the accuracy of full QM at a fraction of the cost — and AI makes it tractable at screening scale.



## Quantum effects retained · integrals parametrised



### Quantum physics — kept

- Electronic structure of the receptor–ligand system
- Polarisation in the binding pocket
- Charge transfer across hydrogen bonds
- Accurate description of non-covalent interactions



### Expensive integrals — approximated

- Two-electron integrals replaced by parametrised forms
- Core electrons folded into effective potentials
- Empirical corrections for dispersion, H-bonds, solvation
- Trained against high-level QM and experimental data

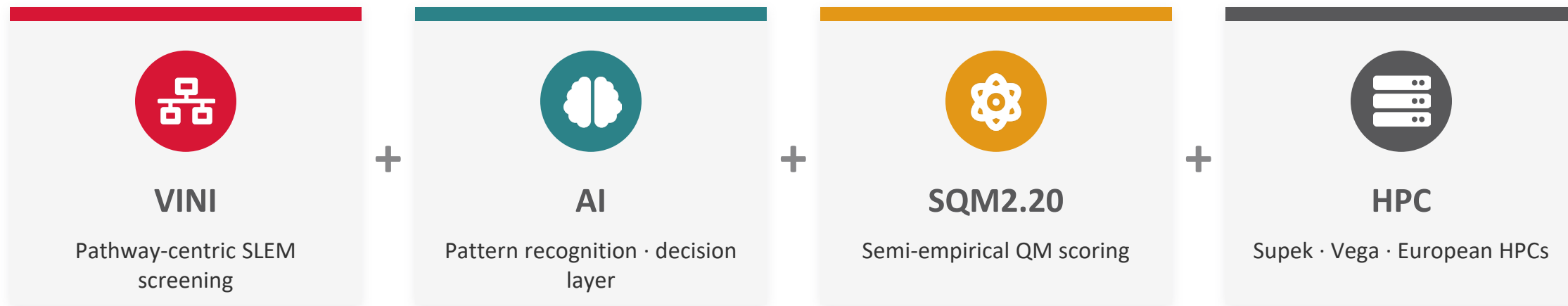
#### THE RESULT FOR PROTEIN–LIGAND BINDING AFFINITY

**Quantum-derived binding affinities** with realistic treatment of electrostatics and the hydrogen-bond network — without the cost of full QM.  
*Designed for non-covalent protein–ligand interactions, with empirical corrections trained against high-level reference data.*



## INTRODUCING UNITY

# VINI + AI-guided SQM scoring + HPC = UNITY



=

# UNITY

Quantum accuracy · AI-scale throughput · pathway-aware biology

*AI-guided semi-empirical QM virtual screening, on a pathway-centric backbone*



AI AS THE DECISION LAYER

## How AI makes minute-per-pose QM tractable for screening



### CONTINUOUS LEARNING LOOP

Every fresh SQM evaluation is added back to the reference database. The AI's similarity model retrains. Over time, fewer poses need expensive QM and average per-pose cost falls. The platform *gets faster the more it is used*.



## DEPLOYMENT

# HPC infrastructure, funding, and what comes next



## Supek

*Rudjer Bošković Institute · Croatia*

Local development & experiments



## Vega

*IZUM · Slovenia · EuroHPC JU*

Production screening (current VINI)



## EuroHPC

*European supercomputing network*

Scale-out for UNITY screening campaigns



## Funding

**Current** EU NextGenerationEU · EuroHPC JU · REG-2022R03-224

**Anticipated** Horizon Europe — UNITY platform deployment and validation across European HPC infrastructure.



## Beyond UNITY v1

**Personalised oncology** patient-level genomic imprint replaces cell-line proxy.

**Biologics & ADCs** extend SQM-AI scoring to mAbs and antibody–drug conjugates.



## TAKEAWAYS

# What to remember when you leave the room

1

### VINI works · the bottleneck is scoring

Pathway-aware multi-target screening is validated. The remaining gap to clinical impact lives in step 4: empirical scoring.

2

### Why semi-empirical QM

Quantum effects retained, expensive integrals approximated. Near-QM accuracy at minutes per pose — the only tractable middle tier.

3

### Why AI in the loop

Pattern recognition over a database of past SQM evaluations decides when fresh QM is needed. The platform gets faster as it learns.

4

### UNITY = VINI + AI + SQM2.20 + HPC

Quantum accuracy at AI-scale throughput, on a pathway-aware backbone — deployable on European supercomputing infrastructure.



THANK YOU

# We are seeking collaborators

*Academic · Clinical · Pharmaceutical · Funding · HPC infrastructure*

## CONTACT

**Dr. Draško Tomić**

Centre for Informatics and Computing  
Ruđer Bošković Institute, Zagreb, Croatia

**[drasko@irb.hr](mailto:drasko@irb.hr)**

## RECENT WORK

### **AACR 2026 Abstract #977**

Tomić D. (2026). VINI: A multimodal in silico platform for discovering rational drug combinations in KRAS-mutant pancreatic cancer.

*Cancer Res 2026;86(7 Suppl): Abstract nr 977*

### **Sci Rep 2024 · 14:18824**

Tomić D. et al. (2024). Exploring potential therapeutic combinations for castration-sensitive prostate cancer using supercomputers.

*doi.org/10.1038/s41598-024-69880-9*

**Code** [github.com/draskot/Vini](https://github.com/draskot/Vini)



**Funded by  
the European Union**  
NextGenerationEU



Ovo djelo je dano na korištenje pod licencom Creative Commons  
*Imenovanje* 4.0 međunarodna.

Srce politikom otvorenog pristupa široj javnosti osigurava  
dostupnost i korištenje svih rezultata rada Srca, a prvenstveno  
obrazovnih i stručnih informacija i sadržaja nastalih djelovanjem  
i radom Srca.

[www.srce.unizg.hr](http://www.srce.unizg.hr)

[creativecommons.org/licenses/by/4.0/deed](https://creativecommons.org/licenses/by/4.0/deed)

[www.srce.unizg.hr/otvoreni-pristup](http://www.srce.unizg.hr/otvoreni-pristup)

