SimDock Pro: Web Application Architecture & Tiered Feature Breakdown

Executive Overview

SimDock Pro is a cloud-native molecular docking platform designed to democratize access to structure-based drug design. Unlike legacy desktop software that requires complex installation and command-line knowledge, SimDock Pro runs entirely in the browser.

The platform is architected around a **single-engine, Multi-Tier** philosophy. The core backend (High-Throughput Docking, Python/FastAPI) remains the same for all users, but the **Feature Gates** and **Infrastructure Limits** scale according to the user tier:

- **Student Tier (The Sandbox):** Focused on education, visualisation, and learning the "Why" behind the science.
- **Researcher Tier (The Workhorse):** Focused on high-throughput screening, ADMET safety, advanced chemical modification, and budgeting.
- **Enterprise Tier (The Infrastructure):** Focused on security, local deployment, compliance, and governance.

Tier 1: Student Edition ("The Learning Sandbox")

Target User: Undergraduates, High School AP Chemistry students, Beginners.

Core Philosophy: "Educational Friction." While it should be easy to use, it must visually demonstrate what is happening (e.g., cleaning a protein) so students learn the concepts.

Features & Limitations

Feature Category	Feature Description	Novelty/Experience
Input Limit	Max 5 Ligands + 1 Protein.	Allows comparing a small series (e.g., Aspirin vs. Ibuprofen) without overwhelming the server.
Docking Engine	AutoDock Vina (Standard) only.	CPU-based execution. Limited to short, standard-precision runs.

Visualization	"Arcade Mode" Viewer.	A gamified 3D view. Highlights bonds in glowing colors and uses simple tooltips (e.g., "Good Fit!" vs. "Clash!").
Data Prep	Educational "Clean-Up" Animation.	Instead of fixing errors silently, the UI shows a quick animation: "Stripping Water Molecules..." -> "Adding Polar Hydrogens..." -> "Merging Non-Polar..." so the student understands the prep workflow.
Results	Top 9 Poses.	Shows a broader range of binding modes (the "Top 9" grid) to discuss alternative docking possibilities in class.
Export	Screenshot & PyMOL Session.	Students can download a `.pse` file to open in PyMOL for homework assignments.

Technical Constraints

- **Compute:** Runs on shared, low-priority server instances.
- **Storage:** Non-persistent. Projects vanish after the session closes unless saved locally.

Tier 2: Researcher Edition ("The Scientific Workhorse")

Target User: PhD Candidates, Postdocs, Academic Labs, Boutique Biotechs.

Core Philosophy: "The Galaxy View & Safety Net." Shifting from docking one molecule to managing campaigns, with AI assisting in safety and cost.

Exclusive Novelty Features

A. The "Glass Panel" Dashboard

- **Description:** A semi-transparent overlay UI that floats above the 3D molecular view.
- **Function:** Displays real-time logs ("Minimizing structure..."), convergence plots, and RMSD charts without obscuring the protein structure.

B. "Smart Ligand Prep" (Bulk Ingest)

- **Description:** Drag-and-drop 100+ files (SDF, PDBQT, CSV) at once.

- **Novelty:** The **"Processing Ring"** animation. Files visually dissolve into particles and fly into a ring, which spins as it processes them (adding charges, generating conformers).

C. Active Site Auto-Detection ("Glowing Spheres")

- **Description:** Auto-detection of binding pockets.
- **Novelty:** The protein renders in a "Ghostly" transparent mode. Deep binding pockets appear as **pulsing, glowing spheres**. Clicking a sphere snaps the grid box to it.

D. The "Galaxy" Visualization

- **Description:** A massive 3D scatter plot representing thousands of docking results.
- **Novelty:** Points are clustered by **Chemical Similarity** (Tanimoto). Users can "lasso" clusters of similar compounds.

E. "Sparkle" Interactive Modification

- **Description:** A holographic wireframe editor for lead optimization.
- **Novelty:** Users click a "Sparkle" button on a docked ligand. Ghostly wireframes appear, suggesting modifications (e.g., "Add Hydroxyl here"). Clicking it instantly modifies and re-docks.

F. One-Click Mutation Testing

- **Description:** Rapid resistance testing.
- **Function:** Right-click any protein residue -> "Mutate to [Amino Acid]". The system repacks side chains and re-runs the dock to see if the drug still binds (useful for studying drug resistance).

G. "Safety Net" (Real-Time ADMET)

- **Description:** Background AI inference for safety.
- **Function:** A "Shield Icon" glows Green or Red while you design.
    - **Red Alert:** "Warning: Modification makes molecule insoluble (LogP > 5)."
    - **Yellow Alert:** "Predicted Toxicity issue."
- **Goal:** Prevents researchers from wasting time on impossible drugs.

H. Generative AI Copilot ("Local AI")

- **Description:** A sandboxed AI assistant.
- **Function:**
    - *Query:* "Summarize binding differences between Ligand A and B."
    - *Generative:* "Suggest a modification to fill this hydrophobic pocket."

I. Compute Cost Estimator

- **Description:** Budget management for cloud runs.
- **Function:** Before submitting a 50,000-ligand screen, the system calculates:

"Estimated Cost: $450 (Cloud Compute) | Time: 4 Hours."

- **Control:** Managers can set budget caps per project.

J. Real-Time "Multiplayer" Collaboration

- **Description:** Live session URLs (e.g., simdock.io/live/xyz-123).
- **Function:** Two researchers see the same protein. Includes **3D "Sticky Notes"**—users can attach comments to specific atoms (e.g., "This bond looks unstable").

K. Auto-Methods Writer

- **Description:** One-click generation of publication materials.
- **Output:** A PDF containing exact methods text and high-res, white-background rendered images.

Tier 3: Enterprise Edition ("The Corporate Infrastructure")

Target User: Big Pharma (Pfizer, Merck, etc.), CROs, Government Labs.

Core Philosophy: "Control, Governance & Security." Since they run locally (On-Prem), cloud costs don't apply, but compliance and security are paramount.

Infrastructure & Security Features

A. Self-Hosted Docker Deployment ("Air-Gapped")

- **Architecture:** The entire web app (Frontend React + Backend FastAPI + Redis + Postgres) runs inside the client's private network.
- **Security: Zero-Data Egress**. The application never "phones home." All "Buy Button" features switch to offline databases of internal inventory.

B. Single Sign-On (SSO) & RBAC

- **Integration:** Full support for SAML 2.0/OIDC (Okta, Microsoft Entra ID).
- **RBAC Roles:**
    - *Admin:* Global settings.
    - *Chemist:* Full access.
    - *Viewer:* Read-only.
    - *Guest:* Temporary 48-hour links for external partners.

C. Immutable Audit Logging

- **Compliance:** Mandatory for FDA 21 CFR Part 11.
- **Function:** Every action is logged to a write-only database with a SHA-256 hash: [User

ID] [Date/Time] [Action] [Input Hash].

D. Governance & Intent Tags

- **Description:** Mandatory metadata for tracking *why* science is being done.
- **Function:** Before running a job, a popup asks: "Why are you running this?"
- **Tags:** "Scaffold Hopping", "Side-chain check", "Random Screen".
- **Goal:** Helps managers track R&D strategy across the organization.

E. License Key Management (Offline)

- **Novelty:** Uses cryptographically signed license files that validate offline, supporting strict firewalled environments.

F. High-Scale Job Queue System

- **Architecture:** Integration with **Kubernetes** or internal HPC clusters.
- **Scale:** Optimizes job distribution across internal hardware resources.

Feature Comparison Matrix

Feature	Student Tier	Researcher Tier	Enterprise Tier
Deployment	Public Web (SaaS)	Public Web (SaaS)	Private Cloud / On-Prem
Max Ligands	5 per run	500+ per run	Unlimited (1M+)
Docking Engine	Vina (Basic)	Vina, AutoDock-GPU	All + Proprietary Engines
Visuals	Arcade + Edu-Prep	Galaxy + Glass Panel	Galaxy + Glass Panel
Data Prep	Visual Step-by-Step	Smart Prep Ring	Smart Prep Ring

Safety/ADMET	None	"Safety Net" AI	"Safety Net" AI
Collaboration	None	Multiplayer + Sticky Notes	Multiplayer + RBAC
Modifications	None	Sparkle + Mutate	Sparkle + Mutate
Budgeting	None	Cost Estimator	Internal Resource Allocation
Security	Standard HTTPS	Standard HTTPS	SSO, Audit Logs, Air-Gap
Reporting	Screenshot	Auto-Methods PDF	FDA Compliance Report