## Supplementary Table 1: ROCK1 Actives

### Pyrazole

| Cpd # | Structure | ROCK1 $K_i$ (µM)* |
|-------|-----------|-------------------|
| 1     | ![Structure](image1.png) | 0.038 +/- 0.001 |
| 2     | ![Structure](image2.png) | 0.049 +/- 0.002 |
| 3     | ![Structure](image3.png) | 0.076 +/- 0.026 |
| 4     | ![Structure](image4.png) | 0.079 +/- 0.013 |
| 5     | ![Structure](image5.png) | 0.082 +/- 0.001 |
| 6     | ![Structure](image6.png) | 0.087 +/- 0.028 |
| 7     | ![Structure](image7.png) | 0.132 +/- 0.038 |
| 8     | ![Structure](image8.png) | 0.155 +/- 0.006 |
| 9     | ![Structure](image9.png) | 0.285 +/- 0.028 |
| 10    | ![Structure](image10.png) | 0.310 +/- 0.012 |
| 11    | ![Structure](image11.png) | 0.573 +/- 0.144 |
| 12    | ![Structure](image12.png) | 1.100 +/- 0.178 |
| 13    | ![Structure](image13.png) | 2.142 +/- 0.143 |
| 14    | ![Structure](image14.png) | 3.606 +/- 0.125 |
| 15    | ![Structure](image15.png) | 6.805 +/- 0.146 |

*Average +/- standard deviation of three measurements. Source data are provided as a Source Data file.

### Lactam/Pyridone

| Cpd # | Structure | ROCK1 $K_i$ (µM) |
|-------|-----------|------------------|
| 16    | ![Structure](image16.png) | 0.793 +/- 0.078 |
| 17    | ![Structure](image17.png) | 1.356 +/- 0.071 |
| 18    | ![Structure](image18.png) | 4.503 +/- 0.143 |
| 19    | ![Structure](image19.png) | 5.094 +/- 0.145 |
| 20    | ![Structure](image20.png) | 5.603 +/- 0.188 |
| 21    | ![Structure](image21.png) | 7.232 +/- 0.070 |

### Azaindole

| Cpd # | Structure | ROCK1 $K_i$ (µM) |
|-------|-----------|------------------|
| 22    | ![Structure](image22.png) | 0.097 +/- 0.004 |
| 23    | ![Structure](image23.png) | 2.202 +/- 0.030 |
| 24    | ![Structure](image24.png) | 4.393 +/- 0.131 |

### Indazole

| Cpd # | Structure | ROCK1 $K_i$ (µM) |
|-------|-----------|------------------|
| 25    | ![Structure](image25.png) | 1.826 +/- 0.066 |
| 26    | ![Structure](image26.png) | 6.639 +/- 0.512 |
| 27    | ![Structure](image27.png) | 6.709 +/- 1.592 |
| Compound ID | Enamine ID | 1H NMR | 13C NMR | HRMS (m/z) | Purity (HPLC area %) |
|-------------|------------|---------|---------|------------|---------------------|
| 1 | Z4540477201 | 5H NMR (400 MHz, DMSO-d6) δ 1.25 (t, J = 7.8 Hz, 3H), 3.52 (2H, 3H), 2.87 (2H, 3H), 2.78 (2H, 3H), 2.64 (2H, 3H), 2.63 (2H, 3H), 2.28 (2H, 3H), 2.26 (2H, 3H), 2.00 (2H, 3H), 1.84 (2H, 3H), 1.78 (2H, 3H), 1.30 (2H, 3H) | | | |
| 2 | Z4540473212 | 5H NMR (400 MHz, DMSO-d6) δ 1.32 (2H, 3H), 3.31 – 3.40 (m, 2H), 2.94 – 3.00 (m, 2H), 1.30 (3H) | | | |
| 3 | Z4540473133 | 5H NMR (400 MHz, DMSO-d6) δ 1.25 (t, J = 7.8 Hz, 3H), 3.52 (2H, 3H), 2.87 (2H, 3H), 2.78 (2H, 3H), 2.64 (2H, 3H), 2.63 (2H, 3H), 2.28 (2H, 3H), 2.26 (2H, 3H), 2.00 (2H, 3H), 1.84 (2H, 3H), 1.78 (2H, 3H), 1.30 (2H, 3H) | | | |
| 4 | Z4540473254 | 5H NMR (400 MHz, DMSO-d6) δ 1.32 (2H, 3H), 3.31 – 3.40 (m, 2H), 2.94 – 3.00 (m, 2H), 1.30 (3H) | | | |
| 5 | Z4540473219 | 5H NMR (400 MHz, DMSO-d6) δ 1.25 (t, J = 7.8 Hz, 3H), 3.52 (2H, 3H), 2.87 (2H, 3H), 2.78 (2H, 3H), 2.64 (2H, 3H), 2.63 (2H, 3H), 2.28 (2H, 3H), 2.26 (2H, 3H), 2.00 (2H, 3H), 1.84 (2H, 3H), 1.78 (2H, 3H), 1.30 (2H, 3H) | | | |
| 6 | Z4540473209 | 5H NMR (400 MHz, DMSO-d6) δ 1.25 (t, J = 7.8 Hz, 3H), 3.52 (2H, 3H), 2.87 (2H, 3H), 2.78 (2H, 3H), 2.64 (2H, 3H), 2.63 (2H, 3H), 2.28 (2H, 3H), 2.26 (2H, 3H), 2.00 (2H, 3H), 1.84 (2H, 3H), 1.78 (2H, 3H), 1.30 (2H, 3H) | | | |
| 7 | Z4540473457 | 5H NMR (400 MHz, DMSO-d6) δ 1.25 (t, J = 7.8 Hz, 3H), 3.52 (2H, 3H), 2.87 (2H, 3H), 2.78 (2H, 3H), 2.64 (2H, 3H), 2.63 (2H, 3H), 2.28 (2H, 3H), 2.26 (2H, 3H), 2.00 (2H, 3H), 1.84 (2H, 3H), 1.78 (2H, 3H), 1.30 (2H, 3H) | | | |
| 8 | Z4540473217 | 5H NMR (400 MHz, DMSO-d6) δ 1.25 (t, J = 7.8 Hz, 3H), 3.52 (2H, 3H), 2.87 (2H, 3H), 2.78 (2H, 3H), 2.64 (2H, 3H), 2.63 (2H, 3H), 2.28 (2H, 3H), 2.26 (2H, 3H), 2.00 (2H, 3H), 1.84 (2H, 3H), 1.78 (2H, 3H), 1.30 (2H, 3H) | | | |
| 9 | Z4540473502 | 5H NMR (400 MHz, DMSO-d6) δ 1.25 (t, J = 7.8 Hz, 3H), 3.52 (2H, 3H), 2.87 (2H, 3H), 2.78 (2H, 3H), 2.64 (2H, 3H), 2.63 (2H, 3H), 2.28 (2H, 3H), 2.26 (2H, 3H), 2.00 (2H, 3H), 1.84 (2H, 3H), 1.78 (2H, 3H), 1.30 (2H, 3H) | | | |
| 10 | Z4540473564 | 5H NMR (400 MHz, DMSO-d6) δ 1.25 (t, J = 7.8 Hz, 3H), 3.52 (2H, 3H), 2.87 (2H, 3H), 2.78 (2H, 3H), 2.64 (2H, 3H), 2.63 (2H, 3H), 2.28 (2H, 3H), 2.26 (2H, 3H), 2.00 (2H, 3H), 1.84 (2H, 3H), 1.78 (2H, 3H), 1.30 (2H, 3H) | | | |
| 11 | Z4540473531 | 5H NMR (400 MHz, DMSO-d6) δ 1.25 (t, J = 7.8 Hz, 3H), 3.52 (2H, 3H), 2.87 (2H, 3H), 2.78 (2H, 3H), 2.64 (2H, 3H), 2.63 (2H, 3H), 2.28 (2H, 3H), 2.26 (2H, 3H), 2.00 (2H, 3H), 1.84 (2H, 3H), 1.78 (2H, 3H), 1.30 (2H, 3H) | | | |
| 12 | Z4540473508 | 5H NMR (400 MHz, DMSO-d6) δ 1.25 (t, J = 7.8 Hz, 3H), 3.52 (2H, 3H), 2.87 (2H, 3H), 2.78 (2H, 3H), 2.64 (2H, 3H), 2.63 (2H, 3H), 2.28 (2H, 3H), 2.26 (2H, 3H), 2.00 (2H, 3H), 1.84 (2H, 3H), 1.78 (2H, 3H), 1.30 (2H, 3H) | | | |
| 13 | Z4540473720 | 5H NMR (400 MHz, DMSO-d6) δ 1.25 (t, J = 7.8 Hz, 3H), 3.52 (2H, 3H), 2.87 (2H, 3H), 2.78 (2H, 3H), 2.64 (2H, 3H), 2.63 (2H, 3H), 2.28 (2H, 3H), 2.26 (2H, 3H), 2.00 (2H, 3H), 1.84 (2H, 3H), 1.78 (2H, 3H), 1.30 (2H, 3H) | | | |

Supplementary Table A: Chemical Characterization of ROCK1 Actives
| Compound # | Enzyme ID | 1H NMR (400 MHz, DMSO-d6) δ | 13C NMR | HRMS (m/z) | Purity (HPLC area % at 254 nm) |
|------------|-----------|-----------------------------|---------|-------------|----------------------------------|
| 15 | Z3504273136 | 1H NMR (400 MHz, DMSO-d6) δ: 13.02 (s, 1H), 8.83 (d, J = 5.9 Hz, 1H), 8.29 (d, J = 8.6 Hz, 1H), 7.92 – 7.83 (m, 2H), 7.74 – 7.65 (m, 2H), 5.78 (q, J = 2.2 Hz, 1H), 1.62 (d, J = 2.1 Hz, 3H), 1.44 (q, J = 5.8 Hz, 2H), 3.84 (d, J = 7.2 Hz, 2H), 2.07 (m, 13.47, 6.8 Hz, 1H), 0.83 (q, J = 6.7 Hz, 6H) | 13C NMR (101 MHz, DMSO) δ: 128.34, 126.06, 125.00, 120.89, 109.80, 80.80, 40.62, 40.41, 40.20, 40.00, 39.79, 39.58, 39.37, 37.37, 29.55, 20.17 | Predicted m/z=324.1851 | Matched m/z=324.1851 | 100.00 |
| 16 | Z3504071831 | 1H NMR (500 MHz, DMSO-d6) δ: 10.41 – 10.39 (m, 1H), 7.80 (d, J = 8.8 Hz, 1H), 7.8 (s, 1H), 7.46 (m, J = 7.0 Hz, 2H), 7.31 – 7.24 (m, 2H), 7.24 – 7.18 (m, 2H), 7.17 (s, 1H), 6.98 (m, J = 8.8, 2.3 Hz, 1H), 6.60 (d, J = 5.5 Hz, 1H), 4.86 (s, J = 4.8 Hz, 1H), 2.48 (s, J = 4.8 Hz, 1H) | 13C NMR (126 MHz, DMSO) δ: 152.14, 150.38, 139.02, 138.15, 129.54, 128.93, 128.48, 128.21, 128.10, 119.34, 116.31, 114.89, 104.94, 104.02, 55.21, 40.64, 40.40, 40.36, 40.27, 40.24, 39.83, 39.68, 39.31, 37.53, 0.9 | Predicted m/z=324.1850 | Matched m/z=324.1850 | 100.00 |
| 17 | Z3504071832 | 1H NMR (400 MHz, DMSO-d6) δ: 9.10 (d, J = 5.9 Hz, 1H), 8.75 (s, 1H), 8.08 (m, 1H), 7.83 (d, J = 15.8, 1H), 7.80 – 7.72 (m, 2H), 7.67 (q, J = 1.4 Hz, 1H), 6.5 (446 m, 2H), 4.6 (446 m, 2H) | 13C NMR (101 MHz, DMSO) δ: 129.05, 126.02, 124.80, 124.72, 124.05, 124.59, 124.48, 124.39, 124.07, 42.49, 42.42, 42.40, 42.00, 40.39, 39.79, 38.58, 39.51, 39.37, 38.15 | Predicted m/z=324.1856 | Matched m/z=324.1856 | 100.00 |
Supplementary Note 1: Selection of Protein Structure for Docking

Eighteen ROCK1 protein structures were exported from the protein kinase database within MOE (Chemical Computing Group), which contains curated, pre-aligned kinase structures. The protein databank codes for the structures were: 2ESM, 2ETK, 2ETR, 2V55, 3D9V, 3NCZ, 3NDM, 3TV7, 3TWJ, 3V8S, 4L6Q, 4W7P, 4YVC, 4YVE, 5BML, 5HVU, 5KKS, and 5KKT. Those were further examined in detail using SeeSAR (version 9.2, BioSolveIT GmbH, Sankt Augustin, Germany, 2020, www.biosolveit.de/SeeSAR) with the goal to find one structure suitable to re-dock all 18 co-crystallized reference compounds into it with reasonable accuracy and associated affinity estimates. This kind of validation is crucial as throughout the process, the number of compounds to be docked reaches several millions. So, for practical reasons they can only be docked into one structure; but also, for statistical reasons, the best possible structure for scoring has to be chosen, as selecting just 100 from millions requires a close to perfect enrichment ratio of actives over inactives chosen.

To find the most suitable structure, the binding modes of the co-crystallized ligands in all 18 crystal structures were visually inspected. All of them share an H-bond interaction with the backbone NH of Met323. Several ligands interact additionally with the neighbouring backbone carbonyl of Glu321, forming a solid hinge binding motif. Further H-bonds are made with the gatekeeper or DFG-loop amino acids, further away from the hinge, which makes sense due to the drug-like size of the crystalized molecules. As in the first step of the Chemical Space Docking only fragment-sized molecules will be docked, placement constraints need to focus on one region only, as otherwise, the fragments may not be able to span the distance between them. Therefore, multiple alternative pharmacophoric constraints were defined, but all only in the hinge region:

- Essential H-Bond-Acceptor opposite of the backbone nitrogen of Met323
- H-Bond-Donor opposite of the backbone carbonyl of Glu321 or H-Bond-Acceptor opposite of the backbone nitrogen of Met323, but at least one of them needed to be fulfilled
- No constraints.

To find a good docking setup, the 18 co-crystallized ligands were docked into all 18 crystal structures applying the three described pharmacophore constraints. Binding sites on the crystal structures were determined using the binding site mode in SeeSAR (version 9.2) by copying all 18 ligands into that mode and picking the amino acids within 6.5 Å distance around the union of all ligand atoms.

Best results were obtained with the second setup in the crystal structures 2ETR, 2V55, 4YVC, 5KTT. The superposition of those was visually inspected and 2V55 was dismissed as it shows big differences to the other structures and hence did not seem to be a good consensus structure. As the remaining three are very similar the electron density was checked using EDIA. Results are shown in Supplementary Figure 1. EDIA represents the electron density by different colors, areas which show high electron density are depicted in blue, whilst areas with low density are highlighted in red. The three remaining crystal structures are all well resolved in the hinge area, but overall 2ETR shows the
biggest blue areas, hence is the most trustworthy structure. Due to the good docking results in structure 2ETR (res. 2.6 Å) under application of pharmacophore constraint 2, this setup is used for the chemical space docking.

Supplementary Figure 1: Electron densities of ROCK-1 crystal structures: PDB ID 2ETR (left), 4YVC (middle), 5KKT (right). Areas with high electron density are depicted in blue, areas with a low density in red.
### Supplementary Table 3: Data Collection and Refinement Statistics

|                        | ROCK1-Cmp 1 (7S25) | ROCK1-Cmp 22 (7S26) |
|------------------------|--------------------|--------------------|
| **Data collection**    |                    |                    |
| Space group            | P 2₁               | P 2₁               |
| Cell dimensions        |                    |                    |
| a, b, c (Å)            | 57.03, 100.96, 176.22 | 57.14, 101.60, 182.66 |
| α, β, γ (°)            | 90, 90.84, 90      | 90, 91.98, 90      |
| Resolution (Å)         | 2.34 (2.38-2.34)*  | 2.74 (2.79-2.74)   |
| R_sym or R_merge       | 0.039 (1.076)      | 0.053 (0.952)      |
| I / σI                 | 16.2 (1.4)         | 14.4 (1.3)         |
| Completeness (%)       | 98.3 (99.6)        | 99.9 (99.9)        |
| Redundancy             | 4.3 (4.3)          | 4.3 (4.0)          |
| **Refinement**         |                    |                    |
| Resolution (Å)         | 2.34               | 2.74               |
| No. reflections        | 82006/1238         | 54015/733          |
| R_work / R_free        | 21.7/25.6          | 24.3/27.6          |
| No. atoms              |                    |                    |
| Protein                | 11226              | 11231              |
| Ligand                 | 72                 | 78                 |
| Water                  | 143                | 37                 |
| B-factors              | (Ask for input)    |                    |
| Protein                | 45.3               | 54.8               |
| Ligand/ion             | 75.1               | 104.4              |
| Water                  | 71.6               | 101.0              |
| R.m.s. deviations      |                    |                    |
| Bond lengths (Å)       | 0.017              | 0.014              |
| Bond angles (°)        | 1.95               | 1.74               |

*Values in parentheses are for highest-resolution shell.
Supplementary Figure 2: Polder omit map for 1 calculated with Phenix by omitting all copies of the ligand and local bulk solvent masks. The mFo-DFc omit difference density is shown at 3 sigma contour level.

Supplementary Figure 3: Polder omit map for 22 calculated with Phenix by omitting all copies of the ligand and local bulk solvent masks. The mFo-DFc omit difference density is shown at 3 sigma contour level.
Supplementary Note 3: Comparison with an alternative ligand-based virtual screening approach

When structure-based virtual screening is considered too expensive, multi-step filter cascades are often times considered the only viable alternative. Fast methods are then used to trim down the massive amounts of data and slower methods are used towards the end of the funnel. 2D fingerprint similarity is an extremely fast approach for such filtering. In order to test the performance of this methodology, we collected a set of highly active ROCK1 inhibitors from the Chembl database. We first assessed the similarity of these compounds compared to our hit-set, using ECFP4 fingerprints and the SpaceLight program.² The respective highest similarity values of a hit compared to any of the Chembl actives cover a significant range from about 0.3 to 0.6. We then used each one of the Chembl actives in turn to retrieve all compound from the version of REAL Space used here throughout, with a 2D similarity >0.25. These searches resulted in a total of 6 mio compounds. Supplementary Figure 4 shows the distribution of 2D similarities in this set, as well as the similarity to each one of the hits presented herein.

Supplementary Figure 4: distribution of similarities in REAL Space compared to Chembl ROCK1 actives together with the respective highest similarity compared to the hit molecules presented here. Source data are provided as a Source Data file.

Several observations are noteworthy. (1) the number of molecules retrieved this way is on the same order of magnitude as the number of molecules docked in the Chemical Space Docking process, so the computational effort is not significantly lower and the approach is certainly not trivial. (2) the range of similarities is quite wide and the distribution of similarities to the novel hits is quite even. So there is no particular bias
towards higher similarities in the hit-set, which supports the hypothesis that docking helps avoid such bias. (3) the number and range of similarities in the set of Chembl ROCK1 actives is quite large as well. So the alternative approach might be viable here but most certainly not on targets without such a rich body of data, as is available for ROCK1.

Supplementary Note 4: Comparison to Fully Enumerated Docking

In the Chemical Space Docking approach, the search space is significantly pruned by selecting building blocks that dock well to the binding site. The chemical subspace associated with each building block is fully enumerated through its associated reaction and all complementary building block partners. But how do the results compare with those obtained by brute-force, fully enumerated docking?

A docking campaign that enumerates the full library would be extremely time-consuming and expensive, but we conducted two types of experiments to compare our results to exhaustive docking. In the first approach, we randomly sampled 1% of the compounds from all sub-libraries, giving a sample size that is 1% of the total enumerated set and evenly distributed across all reactions in the chemical space. The total number of docked molecules in Chemical Space Docking and the random sample of 1% of the entire space is roughly the same, about 8 million.

We compared the distribution of docking scores from the random sample to those obtained from Chemical Space Docking (as described in the manuscript). Supplementary Figure 5 shows the comparison of the distribution of docking scores of both samples. The vast majority of the compounds sampled from the full library are poor dockers, while Chemical Space Docking significantly enriches for good dockers. This demonstrates that Chemical Space Docking is much more efficient at finding compounds that have high docking scores.

In the second approach, we selected two chemical subspaces (reactions and associated building blocks) from which experimentally validated hits were obtained and whose size reasonably allowed fully enumerated docking: subspaces s38 (~3.8 million compounds) and s270302 (~838 thousand compounds). We enumerated and docked all compounds in these subspaces and compared the distribution of docking scores to those selected by Chemical Space Docking.

Supplementary Figures 6 and 7 show the comparisons of the Chemical Space Docking results to the full enumeration of the compounds in subspaces s38 and s270302. For subspace s38 (Supplementary Figure 6) the chemical space docking approach does a very good job retrieving the best dockers from the full-enumeration docking, and its distribution of scores almost matches a very sparsely populated tail of the full distribution. A striking amount of unproductive effort is avoided by limiting the very poor dockers (score <1) to only a few hundred.

Chemical subspace s270302 (Supplementary Figure 7) looks a little different. Again we see a significant waste of effort in the > 95% of the fully enumerated library that fails to
achieve a reasonable docking score. A comparison of the distributions of the high scoring compounds shows that, for this subspace, Chemical Space Docking does indeed miss some of the compounds that were identified by docking the fully enumerated library. This results not from a failure to identify an initial building block as a good docking component, but rather from the filtering using the six exclusion criteria, as detailed in the main manuscript. This criteria clearly excluded initial building blocks that led to complete products that docked well, but whose instantiation was never realized because the building blocks were filtered out beforehand.

Supplementary Figure 5: Comparison of the distribution of docking scores from chemical space docking with that obtained from a random sample of 1% of the full chemical space. Docking scores are binned in 1 unit increments of estimated pIC50 values and reported as a histogram. Bars are labeled by the percentage of the compounds in the sampled chemical space. Source data are provided as a Source Data file.
Supplementary Figure 6: Comparison of the docking score distributions for chemical subspace s38. Docking scores are binned in 1 unit increments of estimated pIC50 values and reported as a histogram. Bars are labeled by the percentage of the compounds in the sampled chemical space. Source data are provided as a Source Data file.

Supplementary Figure 7: Comparison of the docking score distributions for chemical subspace s270302. Docking scores are binned in 1 unit increments of estimated pIC50 values and reported as a histogram. Bars are labeled by the percentage of the
compounds in the sampled chemical space. Source data are provided as a Source Data file.

Supplementary References

(1) Meyder, A., Nittinger, E., Lange, G., Klein, R., and Rarey, M. (2017) Estimating Electron Density Support for Individual Atoms and Molecular Fragments in X-ray Structures. *J Chem Inf Model* 57, 2437–2447.

(2) Bellmann, L., Penner, P., and Rarey, M. (2021) Topological Similarity Search in Large Combinatorial Fragment Spaces. *J Chem Inf Model* 61, 238–251.