CLAIMS

1. A method of identifying at least one of allosteric drug candidates, their binding sites, site-effector interactions, allosteric signalling to one or more target functional sites, and optimisation of site-effector pairs by adjusting or tuning of site-effector composition and interactions, the method comprising the following steps:

    i. obtaining, for a protein of interest, 3D structure information and information on a protein function to be a target of allosteric modulation;

    ii. identifying residues, of the protein of interest, that modulate one or more target functional sites of the protein of interest, by perturbing every residue or three-residue segment and quantifying allosteric signalling from each residue or three-residue segment to all other residues, at single-residue resolution in the protein of interest, based on the 3D structure information;

    iii. identifying residues that modulate one or more target functional sites, by perturbing a functional site and quantifying allosteric signalling to all residues, at single-residue resolution in the protein of interest, based on the 3D structure information;

    iv. identifying a set of binding patches associated with residues that are allosterically linked to at least one of the one or more target functional sites as identified in steps ii and iii, and selecting one or more binding patches of interest from the set of binding patches by quantifying allosteric signalling from each binding patch to a target functional site, geometrical and chemical properties of the respective patch;

    v. for each binding patch of interest, identifying one or more ligands and corresponding binding sites;

    vi. either:,

determining a set of allosteric site-effector pairs, each allosteric site-effector pair comprising a said ligand and a corresponding binding site as identified in step v, based on allosteric signalling from each pair to a said target functional site, the binding free energy, and physicochemical properties of the pair; or

if the one or more ligands do not possess a desired allosteric signalling and binding free energy, selecting a further one or more binding patches of interest and repeating from step v, and/or adjusting the compounds in step v, and repeating from step vi.

vii. outputting the set of site-effector pairs.

2. The method according to claim 1, wherein identifying one or more compounds comprises one or more of:

screening or designing a plurality of compounds for the binding patches, wherein the one or more compounds are selected from the plurality of compounds; and

identifying one or more compounds that possess a desired allosteric effect and binding free energy.

3. The method according to claim 2, wherein identifying one or more compounds comprises screening or designing a plurality of compounds for the binding patches, wherein the one or more compounds are selected from the plurality of compounds, step vi comprises ranking compounds in the plurality of compounds based on allosteric modulation and predicted binding free energy, and identifying the one or more compounds based on the ranking, and screening or designing compounds for binding patches, and ranking compounds based on allosteric modulation value and predicted binding affinity, are repeated for two or more rounds.

4. The method according to claim 1, wherein obtaining the 3D structure information comprises performing one or more of X-ray crystallography, nuclear magnetic resonance spectroscopy, and cryogenic electron microscopy, or performing computational structural prediction.

5. The method according to claim 1, wherein step ii involves simulating binding of a probe, using Structure-Based Statistical Mechanical Model of Allostery (SBSMMA), the model accounting for causality of allosteric communication and quantifying energetics of allosteric signalling and modulation, to each three-residue segment in a protein chain of the protein of interest, and/or, separately, simulating mutations at all residue positions, and calculating the allosteric modulation to all residues.

6. The method according to claim 1, wherein step iii involves Reverse Perturbation (RP) to identify allosterically-linked protein residues by simulated binding at one of the target functional sites.

7. The method according to claim 3, wherein, during each said round except for a final round, an analogue search and optimization process of obtained compounds is performed and an output of the process serves as inputs for a subsequent round.

8. The method according to claim 1, wherein steps ii, iii, iv, and vi involve use of Structure-Based Statistical Mechanical Model of Allostery (SBSMMA), the model accounting for causality of allosteric communication and quantifying energetics of allosteric signalling and modulation by analysing protein dynamics, evaluating per-residue allosteric potential, and computing the per-residue free energy associated with allosteric signalling based on a protein conformational ensemble of the protein of interest.

9. The method according to claim 8, wherein step vi comprises ranking the one or more site-effector pairs based on allosteric modulation, predicted binding free energy to the binding sites and physicochemical properties of site-effector pairs, and optionally comprises searching corresponding analogues and derivatives of the one or more compounds based on substructure matching or similarity and the ranking.

10. The method according to claim 3, wherein step iv comprises evaluating and prioritising the one or more binding patches based on allosteric modulation and a composition of residues.

11. The method according to claim 10, wherein:

the one or more binding patches are evaluated and prioritised based on properties of known allosteric sites in proteins and/or properties of known allosteric compounds for proteins; and/or

the one or more compounds are evaluated and prioritised based on molecular weight, shape and aromaticity, and/or properties of known allosteric sites in proteins and/or properties of known allosteric compounds for proteins.

12. The method according to claim 1, wherein for at least one of the one or more compounds, step vi comprises conducting a search of a chemical space surrounding the at least one compound, to identify further candidates of allosteric effectors based on substructure matching or similarity.

13. The method according to claim 1, wherein the one or more allosteric site-effector pairs are evaluated, prioritised and adjusted based on energetics of allosteric signalling and modulation at the one or more target functional sites, binding affinity and/or properties of site-effector compositions and interactions.

14. The method according to claim 13, wherein a simultaneous and mutual design of effector-site pairs with desirable properties comprises mutual adjustment of binding sites and effectors, via fine-tuning of the composition of allosteric site-effector pairs and intermolecular interactions of site-effector pairs, in respective adjustment, evaluation and prioritization steps across successive rounds, and optionally wherein evaluation, prioritisation and simultaneous and mutual adjustment of a site-effector pair continues in iterations until a site-effector pair with desired allosteric modulation and binding affinity is obtained.

15. A system for identifying at least one of allosteric drug candidates, their binding sites, site-effector interactions, allosteric signalling to one or more target functional sites, and optimisation of site-effector pairs by adjusting or tuning of site-effector composition and interactions, the system comprising:

a processor;

a memory device, accessible to the processor, the memory device comprising program code executable by the processor to:

i. obtain, for a protein of interest, 3D structure information and information on a protein function to be a target of allosteric modulation;

ii. identify residues, of the protein of interest, that modulate one or more target functional sites of the protein of interest, by perturbing every residue or three-residue segment and quantifying allosteric signalling from each residue or three-residue segment to all other residues, at single-residue resolution in the protein of interest, based on the 3D structure information;

iii. identify residues that modulate one or more target functional sites, by perturbing a functional site and quantifying allosteric signalling to all other residues, at single-residue resolution in the protein of interest, based on the 3D structure information;

iv. identify a set of binding patches associated with residues that are allosterically linked to at least one of the one or more target functional sites as identified in steps ii and iii, and select one or more binding patches of interest from the set of binding patches by quantifying allosteric signalling from each binding patch to a target functional site, geometrical and chemical properties of the respective patch;

v. for each binding patch of interest, identify one or more ligands and corresponding binding sites;

vi. either:
determine a set of allosteric site-effector pairs, each allosteric site-effector pair comprising a said ligand and a corresponding binding site as identified in step v,, based on allosteric signalling from each pair to a said target functional site, the binding free energy, and physicochemical properties of the pair; or
if the one or more ligands do not possess a desired allosteric signalling

and binding free energy, select a further one or more binding patches of interest and repeating from step v, and/or adjusting the compounds in step v, and repeat from step vi.

vii. output the set of site-effector pairs.