

# Is allostery a fuzzy concept?

Veronica Morea<sup>1</sup> , Francesco Angelucci<sup>2</sup> and Andrea Bellelli<sup>3</sup> 

1 Institute of Molecular Biology and Pathology, CNR, Rome, Italy

2 Department of Life, Health, and Environmental Sciences, University of L'Aquila, Italy

3 Department of Biochemical Sciences "A. Rossi Fanelli", Sapienza University of Rome, Italy

## Keywords

allostery; arginine repressor; KNF model; MWC model; phosphoglycerate dehydrogenase; uracil phosphoribosyltransferase

## Correspondence

A. Bellelli, Department of Biochemical Sciences "A. Rossi Fanelli", Sapienza University of Rome, Piazzale Aldo Moro 5, 00185 Rome, Italy  
E-mail: [andrea.bellelli@uniroma1.it](mailto:andrea.bellelli@uniroma1.it)

(Received 17 November 2023, revised 30 January 2024, accepted 11 March 2024)

doi:10.1002/2211-5463.13794

Allostery is an important property of biological macromolecules which regulates diverse biological functions such as catalysis, signal transduction, transport, and molecular recognition. However, the concept was expressed using two different definitions by J. Monod and, over time, more have been added by different authors, making it fuzzy. Here, we reviewed the different meanings of allostery in the current literature and found that it has been used to indicate that the function of a protein is regulated by heterotropic ligands, and/or that the binding of ligands and substrates presents homotropic positive or negative cooperativity, whatever the hypothesized or demonstrated reaction mechanism might be. Thus, proteins defined to be allosteric include not only those that obey the two-state concerted model, but also those that obey different reaction mechanisms such as ligand-induced fit, possibly coupled to sequential structure changes, and ligand-linked dissociation-association. Since each reaction mechanism requires its own mathematical description and is defined by it, there are many possible 'allosteries'. This lack of clarity is made even fuzzier by the fact that the reaction mechanism is often assigned imprecisely and/or implicitly in the absence of the necessary experimental evidence. In this review, we examine a list of proteins that have been defined to be allosteric and attempt to assign a reaction mechanism to as many as possible.

## The origins of the concept of allostery

The concept and term of allostery was originally introduced by J. Monod and co-workers in 1963 to describe the behavior of enzymes whose catalytic activity is regulated by ligands that present a chemical structure different from that of the substrate and bind to a site other than that for substrate, the interaction between the effector and the substrate being mediated by structural changes of the protein [1]. Monod and J.P. Changeux at the time were studying the catalytic properties of biosynthetic threonine deaminase, and its inhibition by isoleucine [1]. They remarked that threonine deaminase catalyzes the initial reaction of a

biosynthetic pathway, and its inhibitor is the final product of the same pathway, thus realizing an elegant negative feedback regulation mechanism. This concept was born as an intuition of general applicability, but without a precise mathematical formulation; we refer to it as 'model-independent allostery'.

A refined structural interpretation, coupled with a rigorous mathematical formulation was proposed 2 years later by Monod, Wyman and Changeux [2], in a widely cited paper titled 'On the nature of allosteric transitions: a plausible model'. The concerted allosteric model of Monod, Wyman and Changeux (often nicknamed

## Abbreviations

Hb, hemoglobin; KNF, the sequential model of ligand binding of Koshland, Néméthý and Filmer (for the appropriate references, see text); MWC, the two-state, concerted model of ligand binding of Monod, Wyman and Changeux (for the appropriate references, see text); PEP, phosphoenolpyruvate; PFK, phosphofructokinase; PGDH, D-3-phosphoglycerate dehydrogenase.

MWC) postulates that allosteric proteins are symmetric oligomers that are stable in two different quaternary conformations, in free equilibrium with each other. The two conformations are envisaged as true thermodynamic states, named T (tense) and R (relaxed), which have low and high affinity for ligands, respectively. Provided that the low affinity state (T) is more populated in the absence of ligands, ligation biases the allosteric equilibrium in favor of the high affinity state (R) and causes the R state to overcome the T state. The MWC model describes positive homotropic cooperativity (i.e., cooperativity among identical ligand binding sites) and negative and/or positive heterotropic regulation (i.e., linkage between binding sites for different ligands). The ligand partition function of the MWC model of a  $n$ -sites oligomer is as follows:

$$\Xi = (1 + [X]/K_R)^n + L_0 (1 + [X]/K_T)^n, \quad (1)$$

where  $K_R$  represents the dissociation constant of ligand X from the R state protein;  $K_T$  the dissociation constant of ligand X from the T state protein (by definition  $K_R \ll K_T$ ); and  $L_0$ , the most characteristic parameter of the model, the equilibrium constant of the interconversion of the two states in the absence of ligand, i.e.:

$$L_0 = [T_0]/[R_0], \quad (2)$$

where  $[T_0]$  and  $[R_0]$  represent the concentrations of the protein in the unliganded T and R state, respectively, the suffix 0 indicating absence of bound ligands.

Regulatory effects occur because of changes of the allosteric constant L: the effect of ligand X is to bias L in favor of the R state:

$$L_i = L_0 (K_R/K_T)^i, \quad (3)$$

where  $i$  represents the number of bound ligands and varies between 0 and  $n$ .

Positive homotropic cooperativity occurs when  $L_0 \gg 1$  (in the absence of ligand the T state predominates), and  $L_n \ll 1$  (in the fully liganded protein the R state predominates). Notice that the T and R states in themselves are non-cooperative; cooperativity occurs because of the population switch induced by ligation.

The effect of regulatory ligands different from X, and binding to other sites in the macromolecule, is to bias  $L_0$  in favor of the T state (for negative regulation) or in favor of the R state (for positive regulation), as follows:

$$L_0' = L_0 (1 + [Y]/{}^T K_Y)^m / (1 + [Y]/{}^R K_Y)^m \quad (4)$$

where  $L_0'$  is the allosteric constant in the presence of the effector Y,  ${}^R K_Y$  and  ${}^T K_Y$  are the dissociation

equilibrium constants of Y from the R and T states respectively, and  $m$  is the number of binding sites for ligand Y, which may or may not equal  $n$ , the number of binding sites for ligand X. Eqns 3 and 4 mathematically express the concept of ‘population shift’ or ‘conformational selection’. This reaction mechanism we would like to define ‘allostery proper’, being the definitive formulation adopted by Monod and co-workers.

The MWC model has two extremely noteworthy properties: (a) it can explain positive, but not negative, homotropic cooperativity, except under very special experimental conditions; and (b) it is concerted, i.e., it requires that all ligation intermediates, as well as the fully liganded and fully unliganded species, both for the T and R states, are symmetric, and that all subunits have the same tertiary structure and ligand affinity. Indeed, the only effect of ligands Y and X is to bias the allosteric equilibrium, in principle without changing the ligand binding properties of the T and R states. Actually, it has been experimentally observed in hemoglobin and other proteins obeying this reaction mechanism that binding of allosteric effectors often modifies  $K_T$  as well as  $L_0$ .

The concept of allostery was thus born with an original sin: two definitions, both by Monod and co-workers, one phenomenological, mostly focused on heterotropic regulation, as exemplified by the inhibition of Thr deaminase by Ile [1]; the other mechanistic, mostly focused on positive homotropic cooperativity [2], as exemplified by oxygen binding to hemoglobin. The second definition is more rigorous, as it includes a full mathematical description of the allosteric model, and more comprehensive, as it considers homotropic cooperativity as well as heterotropic regulation.

Since the formulation of the MWC model, several other models were proposed, usually to explain the results of specific experiments carried out on single proteins, most often hemoglobin, and, therefore, they usually lack the general applicability of the original MWC model [3–9]. These models maintain the fundamental hypothesis of two different symmetric quaternary structural arrangements/states, whose equilibrium is governed by an allosteric constant L in all states of ligation. These models belong to a single family of ‘two-state models’, and in what follows we shall consider them as equivalent, for reasons that will become clear in the course of the analysis below.

### Other reaction mechanisms involving cooperativity and heterotropic regulation

Unfortunately, the phenomena described by model-independent allostery may occur via different

reaction mechanisms, conformation selection being only one among many. Alternative to the MWC model, and historically more ancient, is the sequential model initially formalized by Pauling [10], and refined by Koshland, Neméthy and Filmer (i.e., the KNF model) [11]. Importantly, sequential models were devised to explain homotropic cooperativity, and the effect of heterotropic regulation was conspicuously absent in the original formulations; indeed, Koshland never referred to his model as ‘allosteric’. Sequential models postulate that the structural changes responsible for the regulation of ligand affinity occur at the tertiary level because of a ligand-induced fit mechanism, and as a result of pairwise interactions between contacting subunits of the oligomer. Therefore, sequential models differ from two-state models in that: (a) they deny a ligand-independent equilibrium between different quaternary structural conformations of the protein, even though they may admit a ligand-independent conformational equilibrium between different tertiary structures of each subunit; (b) they postulate that partially liganded derivatives are asymmetric, liganded and unliganded subunits having different tertiary structures; and (c) they can explain both positive and negative homotropic cooperativity, depending on the nature of the intersubunit contacts. The ligand partition function of the KNF model varies depending on the hypothesized functional geometry of the oligomeric protein under consideration, i.e., on whether each subunit transmits the conformational change to one, two, or more other subunits. Thus, the ligand partition function of the KNF model should be derived by considering the number of subunits in the oligomer and their reciprocal arrangement. As an example, for a tetramer in which each subunit interacts with all other three subunits, the ligand partition function is as follows:

$$\Xi = 1 + 4 K_X K_{01}^3 [X] + 6 K_X^2 K_{01}^4 K_{11} [X]^2 + 4 K_X^3 K_{01}^3 K_{11}^3 [X]^3 + K_X^4 K_{11}^6 [X]^4, \quad (5)$$

where  $K_X$  is the ‘intrinsic’ association constant of ligand X and  $K_{01}$  and  $K_{11}$  are called interaction parameters, and describe the modulation of the affinity constant due to the interaction between a liganded and an unliganded subunit ( $K_{01}$ ) and between two liganded subunits ( $K_{11}$ ). A  $K_{00}$  interaction parameter for pairs of unliganded subunits is not required because interaction parameters are conceived as relative terms and thus one of them (i.e.,  $K_{00}$ ) can be assigned an unitary value. The difference between Eqns 5 and 1 is obvious and does not require specific comments. In Eqns 1 through 5, we respected the original formulation used

by the respective authors; therefore, in Eqns 1 through 4 we used dissociation constants, whereas in Eqn 5 we used association constants.

As for the MWC model, several variants of the KNF model can be envisaged. These variants are encompassed in the first definition of the model and usually imply considering different functional geometries of the macromolecule and adding further interaction parameters, to take into account heterotropic effects. The KNF model has been less widely used than the MWC model, essentially because of the difficulty to identify the functional geometry of the macromolecule, i.e., which intersubunit contacts transmit information on the ligation state and how, which makes the definition of functional geometry somewhat arbitrary. Indeed Pauling himself remarked that the oxygen binding isotherms of human hemoglobin could be fitted equally well by assuming a tetrahedral functional geometry (each subunit interacts with all the other three) or a square functional geometry (each subunit interacts with two others).

The KNF model can describe the same phenomena described by the MWC model, thus the two models, in spite of being different in the reaction mechanisms they postulate, present overlap in their applicability. As we remark below, all two-state models and all sequential models have been tested on the oxygen binding isotherms of hemoglobin and on other proteins, and have usually been proved to be capable of describing cooperative ligand binding isotherms with great accuracy. This demonstrates that the ability of fitting ligand binding data is not proof of validity of a model and that more refined experiments are necessary to assign a reaction mechanism, and thus a model, to an allosteric protein.

Ligand-linked association-dissociation is another mechanism able to explain homotropic cooperativity in ligand binding and heterotropic regulation, for reasons analogous to those considered for the MWC model. Indeed, in this case, the protein has two states, represented by different degrees of oligomerization and, if the different oligomers have different ligand affinity, positive homotropic cooperativity will ensue [12]. Moreover, heterotropic ligands may affect the association-dissociation constant. The difference between ligand-linked association-dissociation and the MWC model is that the association constant of the monomers in the former model, which plays the same role as the allosteric constant  $L_0$  in the latter, is not dimensionless; however, ligand-linked association-dissociation is a kind of population selection.

Strictly speaking, intermediate reaction mechanisms and models do not exist. Either a protein obeys a

sequential mechanism or it does not; either it presents an allosteric equilibrium between two states or it does not; etc. These statements do not refer to the actual value of the parameters involved: e.g. in a protein that obeys a two-state reaction mechanism the value of  $L_0$  may be very large, implying minimal population of the  $R_0$  state, but this does not make it to approximate a sequential reaction mechanism, since if it presents homotropic or heterotropic effects at some degree of ligation the allosteric constant, be it  $L_0$  or  $L_i$ , must approximate unity to allow the quaternary structure switch.

Mixed reaction mechanisms are possible if the same protein presents two or even all three mechanisms together, i.e., not an intermediate mechanism but the sum of two or three. Wyman called this effect ‘nesting’ to describe a cooperative substructure embedded into a larger more cooperative superstructure. Nesting has been invoked for some  $O_2$  carriers [5,6], in which the T state, besides being in equilibrium with the R state, might present sequential cooperativity. Another possibility is the association of the allosteric T–R quaternary structure equilibrium with ligand-linked dissociation, which has been observed, for example, in hemoglobin.

The problem that we face in scientific literature is that many proteins that are classified as ‘allosteric’, because they present homotropic cooperativity and/or heterotropic regulation, obey reaction mechanisms different from the two-state model (see Table 1 below). In other words, there is not one single mathematical definition corresponding to allostery, but many. However, since this situation is rarely acknowledged, allostery has become an imprecisely defined concept. Once this condition is acknowledged, it follows that it would be advisable to define which type of allostery each protein possesses.

### Assignment of the reaction mechanism to a list of allosteric proteins

Since allostery is an appealing concept, it has been invoked for a vast number of proteins, and it is commonly used to refer to homotropic cooperativity, heterotropic regulation, or both. Moreover, as discussed above, this concept has been associated in the literature with several reaction mechanisms, in addition to the MWC model. The problem we face is thus to classify the reaction mechanisms capable of producing homotropic and heterotropic regulation and to define the criteria for assigning the appropriate mechanism.

### Identification of the evidence required for mechanism assignment

Identification of the reaction mechanism underlying allostery cannot be based on ligand binding experiments and crystallographic structures of the fully unliganded and fully liganded states alone, but requires subtle and refined experiments designed to test the following specific features: (a) whether the structure of the ligation intermediates is symmetric or asymmetric, and equal to either the fully liganded or the fully unliganded derivative; (b) whether there is evidence for ligand-independent quaternary structure change; (c) whether the structures of the ‘unfavorable’ states hypothesized by the MWC model (i.e., liganded T and unliganded R) can be determined, together with those of the more favorable states (i.e., unliganded T and liganded R); (d) whether cooperativity and regulation depend on protein concentration (which suggests ligand-linked association-dissociation); and (e) whether the protein presents negative homotropic cooperativity, which is incompatible with population selection, and is thus proof of a sequential model (some very uncommon exceptions may occur). Thus, the reaction mechanism of a ligand binding protein can be identified as MWC-like if it obeys any of the following criteria: symmetry of ligation intermediates; evidence of the ligand-independent (quaternary) structure change; evidence for the existence of the unfavorable states. A ligand binding protein can be assumed to obey a KNF reaction scheme if it presents significant intersubunit asymmetry, especially in the incompletely saturated ligation intermediates or negative homotropic cooperativity. Finally, a protein can be assigned a ligand-linked association-dissociation mechanism if it presents evidence of concentration-dependent changes in ligand affinity. The information required for assignment of the reaction mechanism may not be always available, in which case the assignment should be considered presumptive or provisional.

In addition to the experimental criteria listed above, we also examined a structural parameter recently identified by ourselves based on the analysis of the 3D structures of the liganded and unliganded state of some allosteric proteins whose reaction mechanism could be assigned with certainty [13]. Following this analysis, we found that, as expected, the MWC-like proteins present small ligand-dependent tertiary structure changes coupled to large quaternary structure changes, whereas the opposite occurs for KNF-like proteins [13]. We measured the RMSD

**Table 1.** List of the allosteric proteins analyzed by Daily and Gray [14], with some additions (\*), ordered according to the (presumed) thermodynamic mechanism. List of 'allosteric' proteins, integrated with our analysis of the reaction mechanism. Since neither the MWC nor the KNF models consider monomeric proteins, these are not assigned a specific reaction mechanism and are satisfactorily described by Wyman's linked functions. In some cases, the reaction mechanism is presumptive, because the available information is insufficient for a certain attribution. The proteins we added to the list by Daily and Gray [14] are marked with an asterisk, and were considered allosteric by the authors who studied them, but not necessarily because of heterotropic regulation; e.g. Furukawa *et al.* [29] consider D-lactate dehydrogenases allosteric because they present homotropic cooperativity, although via a sequential reaction mechanism. The references quoted in the table usually report and describe the structure indicated by the pertinent PDB code(s); in some cases, further references were added that describe the functional properties of the protein. The criteria for the assignment of the reaction mechanism are described in the text. The parameter R1 cannot be determined in the cases of monomeric proteins, or ligand-linked association-dissociation. 'Undefined' indicates that the available information does not allow the identification of the reaction mechanism; 'Not applicable' indicates that the monomeric state of the protein is not compatible with the criteria defined in Section 'Identification of the evidence required for mechanism assignment'. Covalent modification may cause structural and functional changes, but is not compatible with either the MWC or KNF reaction mechanism, which presume rapid and reversible chemical equilibria; however, each state a covalently modifiable protein (e.g. because of phosphorylation) may obey a KNF or MWC model, as it occurs, for example, in the case of glycogen phosphorylase.

Protein and PDB codes; references in square brackets	Thermodynamic mechanism	Notes
Hemoglobin* [13,15] 2DN3; 2DN2	Concerted, MWC-like, and ligand-linked oligomerization	Heterotetramer; see Section 'Hemoglobin'. R1 = 0.28
Asp-transcarbamylase [13] 6AT1; 8ATC	Concerted, MWC-like	12-mer; demonstrated ligand-independent T-R structure change. R1 = 0.28
Phosphofructokinase [13] 6PFK; 4PFK	Concerted, MWC-like	Homotetramer; ligand-independent T-R structure change is induced by effectors. R1 = 0.60
Glycogen phosphorylase [13] 1GPB; 7GPB	Probably concerted, MWC-like; admits covalent modification (phosphorylation)	Homodimers reversibly associating to homotetramers. R1 = 0.51
Glucosamine-6-P deaminase [16,17] 1CD5; 1HOT	Probably concerted, MWC-like	Symmetric homo-6-mer; R1 = 0.25
Chorismate mutase [18] 2CSM; 1CSM	Possibly concerted, MWC-like	Homodimer, activated by Trp and inhibited by Tyr. R1 = 0.68
Anthrilate synthase [19] 1I7S; 1I7Q	Possibly concerted	Biosynthesis of Trp; inhibited by Trp. A <sub>2</sub> B <sub>2</sub> heterotetramer. R1 = 0.49
Human mitochondrial malic enzyme [20–22] 1QR6; 1PJ2	Possibly concerted	Symmetric homotetramer; oxidative decarboxylation of malate to pyruvate; positive cooperativity for malate; activation by fumarate. Stoichiometric ratio fumarate: malate = 1 : 1. R1 = 0.58
<i>L. casei</i> L-lactate dehydrogenase [23] 2ZQY; 2ZQZ	Concerted	Cooperative homotetramer, inhibited by fructose 1,6 bisphosphate. allosteric structure change demonstrated by its pH dependence. R1 = 0.26
Transcriptional activator DctD [24] 1L5Z; 1L5Y	Probably concerted	Symmetric homodimer. R1 = 0.27
3-phosphoglycerate dehydrogenase [25,26] 1PSD; 1YBA	Sequential, atypical because of significant quaternary structure change	Homotetramer; Ser biosynthesis pathway; negative homotropic cooperativity for inhibitor L-Ser. R1 = 0.56
Arg repressor (ArgR) [27,28] 1XXC; 1XXA	Sequential	Homo 6-mer; negative homotropic cooperativity for Arg. R1 = 0.81
<i>Pseudomonas aeruginosa</i> D-lactate dehydrogenase* [13,29] 6ABJ; 5Z20	Sequential	Homotetramer; negative homotropic cooperativity. R1 = 0.73
Uracil-phosphoribosyl transferase from <i>S. solfataricus</i> [30,31] 1XTU; 1XTT; 1XTV	Sequential	Homotetramer. Inhibited by CTP; negative homotropic cooperativity for UMP. Significant intramolecular asymmetry for 1XTV. R1 ~ 0.6
3-deoxy-D-arabino- heptulose-7-phosphate (DAHP) synthase [32–34] 1KFL; 1GG1	Possibly sequential.	Biosynthesis of aromatic aa. Symmetric (D2) homotetramer. non-cooperative; non-competitively Inhibited by Phe. R1 = 0.45
<i>E. coli</i> purine nucleotide phosphorylase (PNP)* [13] and references therein; 1ECP; 4TTA	Sequential	Homo 6-mer, dimer of trimers; negative homotropic cooperativity; strong intersubunit asymmetry. R1 = 1.05

**Table 1.** (Continued).

Protein and PDB codes; references in square brackets	Thermodynamic mechanism	Notes
<i>E. coli</i> Met repressor [35] 1CMB; 1CMA	Possibly sequential	Homodimer; ligand- induced asymmetry
dTMP synthase* [36] 1CI7	Sequential	Homodimer; negative homotropic cooperativity; no known effectors
Purine repressor PurR [37,38] 1DBQ; 1WET	Undefined	Homodimer; symmetric when liganded to guanine and DNA; minor asymmetry when ligand-free. R1 = 0.87
Tetracycline resistance repressor (TetR) [39] 2TRT; 2XB5; 1QPI	Undefined	Symmetric homodimer. Binds either to tetracycline or to DNA; Tc binding site at the monomer-monomer interface. R1 = 1
Fructose bisphosphatase (FBPase-1) [40] 1EYJ; 1EYI	Undefined; insufficient information	Homotetramer. Significant subunit asymmetry. R1 > 1
ADP-ribosylation factor 1 (arf1) [41] 1HUR	Undefined; insufficient information	Symmetric homodimer
GTP cyclohydrolase I [42] 1WPL; 1IS7	Undefined, complex: inhibited by the biopterin- binding regulatory protein GFRP. Multiple aggregation states	Biosynthesis of biopterin. GTP cyclohydrolase I is a 10-membered ring; the regulatory protein GFRP is a 5-membered ring; the complex of the two involves one cyclohydrolase and two GFRP forming a hetero 20-mer
Heat labile enterotoxin [43,44] 1LTT; 1LTR	Undefined	Binds lactose; hetero 7-mer composed by a 5-membered ring of identical subunits plus two other subunits
ATP sulfurylase [45] 1M8P; 1I2D	Undefined	Symmetric homo 6-mer, inhibited by 3'-phosphoadenosine-5'-phosphosulfate; R1 = 0.88
<i>E. coli</i> arabinose binding protein (AraC) [46,47] 1XJA; 2ARA; 2ARC	Ligand-linked association-dissociation; forms two different homodimers, depending on the presence of arabinose	Possibly higher assembly states in the absence of arabinose. Open structure in spite of isologous interfaces
G-protein rab11 [48] 1OIV; 1OIW	Ligand-linked oligomerization?	Quasi-symmetric homodimer in the inactive GDP-bound state; monomer when the GTP-bound
Kinase domain of insulin-like GF receptor IGRK [49,50] 1P4O; 1K3A	Covalent modification (phosphorylation); ligand-linked oligomerization?	Monomer when unliganded; heterodimer when liganded?
<i>E. coli</i> repressor of the biotin operon (BirA; BioR) [51,52] 1BIA; 1HXD	Ligand-linked oligomerization	Monomer when free, dimer when biotin- and DNA-bound
Tyr kinase domain of human insulin receptor (IRK) [53] 1IRK; 1IR3	Ligand-linked oligomerization; phosphorylation	
ATP-phosphoribosyl transferase [54] 1NH8; 1NH7	Ligand-linked association-dissociation; possibly concerted	His biosynthesis; inhibited by His. Dimer when His-free, hexamer when His-bound
Lac repressor (LacR) [55,56] 1TLF; 1EFA	Ligand-linked oligomerization	Homotetramer when bound to isopropyl-beta-D-thiogalactoside; homodimer when bound to DNA
Nitrogen fixation enzyme fixJ [57,58] 1DBW; 1D5W	Ligand-linked oligomerization; covalent modification (phosphorylation)	Monomer/dimer equilibrium
<i>E. coli</i> H <sub>2</sub> O <sub>2</sub> sensor (oxyR) [59] 1I69; 1I6A	Covalent modification; redox-dependent oligomerization (intrasubunit disulfide bridge formation)	Homodimer in the reduced state, monomer in the oxidized state
<i>E. coli</i> chemotaxis protein CheY [60] 3CHY; 1FQW	Covalent modification (phosphorylation)	Monomer
Protein kinase B (PKB) [61] 1GZK; 1O6K	Covalent modification (phosphorylation)	Monomer; suggested conformational equilibrium
Anti sigma factor antagonist SpoIIAA [62] 1H4Y; 1H4X	Covalent modification (phosphorylation)	Monomer
Caspase [63] 1SHJ; 1F1J	Covalent modification (disulfide formation)	Homodimer

**Table 1.** (Continued).

Protein and PDB codes; references in square brackets	Thermodynamic mechanism	Notes
MAP kinase <a href="#">ERK2</a> [64,65] <a href="#">SUMO</a> (previously <a href="#">1ERK</a> ); <a href="#">2ERK</a>	Covalent modification (phosphorylation)	Monomer
GTP binding protein Ran [66] <a href="#">11BR</a>	Not applicable	Monomer
Protooncogene RAS [67] <a href="#">4Q21</a> ; <a href="#">6Q21</a>	Not applicable	Monomer; activated by GTP; inhibited by GDP
GTPase <a href="#">cdc42</a> [68] <a href="#">1AN0</a> ; <a href="#">1NF3</a>	Not applicable	Monomer
<a href="#">rac1</a> [69] <a href="#">1HH4</a> ; <a href="#">1MH1</a>	Not applicable	Monomer
Protein Tyr phosphatase 1 B (PTP1B) [70,71] <a href="#">1T48</a> ; <a href="#">1PTY</a>	Not applicable	Monomer
GTP binding protein <a href="#">sec4</a> (Rab family) [72] <a href="#">1G16</a> ; <a href="#">1G17</a>	Not applicable; possible Ligand-linked oligomerization	Monomer
GTPase <a href="#">rheb</a> [73] <a href="#">1XTQ</a> ; <a href="#">1XTS</a>	Not applicable	Monomer
Elongation factor <a href="#">Eftu</a> [74,75] <a href="#">1TUI</a> ; <a href="#">1EFT</a>	Not applicable	Monomer; GTP-dependent binding to aminoacyl-tRNA
GTPase <a href="#">ypt7p</a> [76] <a href="#">1KY3</a> ; <a href="#">1KY2</a>	Not applicable	Monomer
G-protein <a href="#">rap2a</a> [77] <a href="#">1KAO</a> ; <a href="#">2RAP</a>	Not applicable	Monomer
GTP binding protein <a href="#">YsxC</a> [78] <a href="#">1SVI</a> ; <a href="#">1SVW</a>	Not applicable	Monomer
GTP binding protein <a href="#">arf6</a> [79] <a href="#">1E0S</a> ; <a href="#">2J5X</a> (formerly <a href="#">1HFV</a> )	Not applicable	Monomer
Transducin alpha subunit [80,81] <a href="#">1TAG</a> ; <a href="#">1TND</a>	Not applicable	Alpha subunit of the heterotrimeric GTP binding protein transducin
<a href="#">rab7</a> [82] <a href="#">1VG1</a> ; <a href="#">1VG8</a>	Not applicable	Monomer; binding of GTP/GDP regulates affinity for REP1 protein
GTPase <a href="#">rhoA</a> [83,84] <a href="#">1FTN</a> ; <a href="#">1A2B</a>	Not applicable	Monomer; binds GTP/GDP
GTP hydrolase <a href="#">Gi</a> [85,86] <a href="#">1GDD</a> ; <a href="#">1GIA</a>	Not applicable	Alpha subunit of heterotrimeric GTP binding protein

values between the  $\alpha$ -carbons of liganded and unliganded structures and called R1 the ratio between the RMSD values measured for the isolated subunits and the RMSD values measured for the whole oligomers. We found that R1 values are  $<0.3$  in typical two-state, MWC-like proteins and  $>0.7$  in typical sequential, KNF-like proteins (see Table 1). In the case of strong but ordered intersubunit asymmetry, R1 may exceed unity. Intermediate values are observed when the protein obeys mixed reaction mechanisms (see above) or when the protein follows a sequential, KNF-like model and the ligand-induced structural changes are very small at both the tertiary and quaternary level. The R1 parameter is not a clear-cut proof of the reaction mechanism, but it may be taken as an indication when more refined information is lacking. Unfortunately, the parameter R1 cannot be determined for monomeric proteins, or for proteins that undergo ligand-linked association-dissociation equilibria, and it is not meaningful in the case of proteins that undergo covalent modification.

### Analysis of a list of 'allosteric' proteins

Table 1 lists the proteins considered allosteric by Daily and Gray [14], with the addition of some others (marked by an asterisk), considered allosteric by the authors of the papers cited in the Table. It is crucial for the scope of this review that we use a list of allosteric proteins compiled by other authors, otherwise our analysis would be biased by our views on allostery. However, we felt free to implement the original list by Daily and Gray, chosen because it is extensive, with other allosteric proteins, so defined by the authors of the pertinent papers, because Daily and Gray apparently considered allostery a synonymous of 'presenting a heterotropically regulated structure change', thus excluding pure homotropic cooperativity. While Daily and Gray did not try to analyze reaction mechanisms, in this work we carried out an extensive literature search for each of the proteins in the list to assign to each of them the presumed or demonstrated reaction mechanism. In some cases, we confirmed what was reported in the literature, but in others we added

a missing assignment or corrected the original one. The criteria to assign the reaction mechanism are summarized above and more extensively explained in Table 1 of ref. [13].

The Table is by no means exhaustive, but is, in our opinion, representative of what researchers may define an ‘allosteric protein’, and this is what matters for our analysis. Entries are ordered according to the reaction mechanism, with the caution that any single protein may present more than one reaction mechanism: e.g., ligand-linked oligomerization may coexist with concerted cooperativity, or with covalent modification. Clearly, allosteric proteins share neither the reaction mechanism nor the aggregation state, and the only property the entries of Table 1 have in common is that they bind ligands and present homotropic and/or heterotropic regulation of ligand affinity (i.e. all of them present model-independent allostery); moreover they present ligand-dependent structure changes, but we suspect that no protein can bind a ligand without experiencing some structure change.

Inspection of Table 1 shows that, for each of the well-established reaction mechanisms between proteins and ligands described in the previous paragraph, pertinent examples can be quoted; moreover, some unexpected additions are present in the Table, namely covalently regulated proteins and monomeric proteins. In the case of many entries of the Table the reaction mechanism was correctly assigned, in some cases no specific reaction mechanism was identified, or a reaction mechanism had been assigned in the absence of sufficient evidence, or was wrongly assigned. This is often the case for the two-state MWC model that, due to its popularity, has been associated with a large number of allosteric proteins, even when insufficient information was available. As an example, whenever the fully liganded and fully unliganded structures of a ligand binding protein have been determined, the former has been assumed to pertain to the R state, and the latter to the T state. This language is misleading, as it conveys the idea that in the protein a ligand-independent equilibrium exists between the two quaternary structures, when in fact none has been demonstrated to occur (or not to occur).

Table 1 lists 55 proteins, tentatively grouped according to their reaction mechanism as follows (to some proteins more than a single mechanism is assigned). Ten entries can be assigned to the class that obeys a concerted reaction mechanism based on population selection, hence the MWC model or any of its more recent variants. Eight entries obey a ligand-linked oligomerization mechanism; this class is most probably

underrepresented, but there is some overlap with the class of proteins that require covalent modification. As already stated, ligand-linked oligomerization is a kind of population selection.

Eight entries of Table 1 could be more or less confidently assigned to the sequential reaction mechanism proposed by Koshland *et al.* [11]. Two of these, D-lactate dehydrogenase and dTMP synthase, were not present in the original list by Daily and Gray, probably because they have no known heterotropic effectors; we added them because the authors of the relevant papers, explicitly call them allosteric as they present homotropic regulation of substrate affinity; thus they serve to further illustrate how subjective the definition of allostery has become.

Seven entries of Table 1 are labeled as ‘undefined’. These proteins obey either a MWC-like or a KNF-like reaction mechanism, but the available information is insufficient to assign the mechanism more precisely, or they present features that may be compatible with both mechanisms (e.g., a small quaternary structure change, evident as a value of RI close to unity, coupled to minimal asymmetry in both the liganded and unliganded states).

Six entries of Table 1 refer to proteins in which a structure change is caused by covalent binding of the heterotropic effector (e.g., phosphorylation). Covalent modification, unless rapid and reversible, does not fit with either the MWC or KNF reaction mechanism, which postulate rapidly reversible equilibria, thus these proteins do not conform to Eqns 1–5. However, the two states of the native and covalently modified protein should be considered as two different entities, and the possibility arises that either of them obeys a MWC or KNF mechanism or presents ligand-linked dissociation-association. An example is glycogen phosphorylase, whose phosphorylated state seems to obey a MWC-like reaction mechanism.

Finally, 16 entries of Table 1 refer to proteins that are monomers in solution or have been studied as monomeric components of assemblies made of different subunits. These obviously lack homotropic interactions and are allosteric only in the sense that they bind (at least) two ligands to different sites, and the affinity of either influences that of the other(s); thus, their functional behavior can be described using Wyman’s linkage functions [87–89]. Monomeric proteins evade our criteria for assigning a reaction mechanism and are discussed under a separate heading (see below).

Some proteins fall in two categories (e.g., monomers and covalent modification). We may add that in some cases of covalent modification, or other types of



structure changes, the protein may undergo a slow rearrangement that may mimic equilibrium cooperativity. These cases may be difficult to detect in the absence of specific kinetic information and may contaminate the mechanism assigned to some of the proteins in Table 1.

Some reaction mechanisms are probably underrepresented in Table 1. In particular, certain positive criteria for the sequential reaction mechanism are very restrictive (e.g. negative cooperativity essentially proves the mechanism, but positive cooperativity does not disprove it); thus it is likely that several ‘undefined’ entries might in fact be cases of KNF-like reaction mechanism. However, for the purposes of the present work, an unequal representation of the possible reaction mechanisms that may be grouped under the umbrella of allostery does not invalidate the finding that allostery is a fuzzy concept, in need of some revision, or at the very least of the recognition that it groups different entities.

### Stoichiometry of ligands and effectors

The mechanism of action of heterotropic effectors in proteins that present population selection is to bias the allosteric (or oligomerization) constant. Since these proteins are oligomers whose function is regulated at the level of the quaternary structure, the binding site of the heterotropic effector is not related to the number of protomers: e.g. the O<sub>2</sub> affinity of hemoglobin is regulated by CO<sub>2</sub>, whose stoichiometric ratio to oxygen is 1 : 1 (one site per protomer), by bezafibrate, with a ratio to oxygen of 1 : 2 (two sites per tetramer), and by glycerate bisphosphate with a ratio to oxygen of 1 : 4 (one site per tetramer). By contrast, in the proteins that obey the sequential KNF reaction scheme heterotropic regulation usually occurs at the tertiary structure level, and in all cases considered in Table 1 they possess one binding site *per* subunit.

An interesting observation is that in some proteins that obey a KNF reaction mechanism heterotropic effectors present two different stoichiometric ratios between subunits and effectors; this is observed, for example in PGDH and in the Arg repressor [13]. In these cases, the true stoichiometry of the effector is 1 *per* subunit, as expected for KNF proteins, but the functional regulation is fully achieved at lower stoichiometric ratios (0.5 *per* subunit or 2 *per* tetramer in the case of PGDH). It is also common that in these cases the binding of the effector presents negative cooperativity (see below). In these cases, the stoichiometry one calculates from enzyme activity differs from the one observed by directly measuring effector

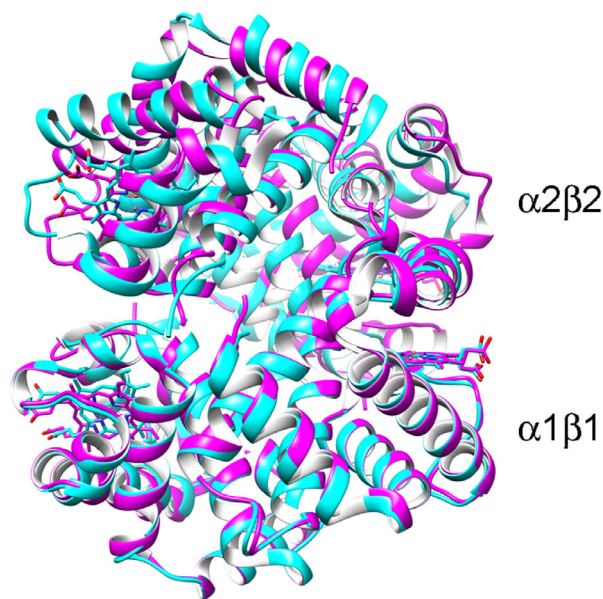
binding or by determining the number of bound effectors by X-ray crystallography.

### Some examples

The principal aim of the present review is the assignment of reaction mechanisms to the allosteric proteins listed in Table 1. Since this process is quite intricate and may not be obvious to all readers, we provide here some examples of MWC-like and a KNF-like cases, to practically illustrate how the criteria defined under Section ‘Identification of the evidence required for mechanism assignment’ are used. We do not provide examples of the assignment of the ligand-linked association-dissociation mechanism, because it is straightforward and not open to doubt.

### Hemoglobin

Hemoglobin is the prototype of a protein that obeys a two-state MWC-like reaction mechanism (see Fig. 1). The rich spectroscopic properties of the heme, the ease of preparation and crystallization, and the availability



**Fig. 1.** Allosteric quaternary structure transition in human hemoglobin. The quaternary structure change of human hemoglobin is described as a rotation and translation of one  $\alpha 1\beta 1$ -type heterodimer with respect to the other. The  $\alpha 1\beta 1$  heterodimers form T state (purple) and R state HbA (light blue) are superimposed, to demonstrate their different spatial relationship with the  $\alpha 2\beta 2$  heterodimers form T state and R state HbA. The tertiary structure changes are comparatively small, as one may notice in the superposition of the  $\alpha 1\beta 1$  heterodimers.

of different derivatives have allowed researchers to conduct experiments that are impossible on other proteins. Since we reviewed the allosteric properties of hemoglobin elsewhere [15], we shall not go in detail in this work, but shall list only the key pertinent features that allow the unequivocal assignment of Hb to the concerted allosteric mechanism. We also remark that several variants of the MWC model were devised to obtain a more precise description of refined experiments conducted on this protein; but in this work, we are not concerned with subtleties that would require identification criteria that cannot be met for any other protein among those listed in Table 1.

The oxygen binding isotherms of hemoglobin were very satisfactorily described using all sequential models [10,11], and all concerted models [2–9], thus demonstrating that ability to describe the ligand binding curve is not proof of the applicability of the model. The assignment of the reaction mechanism of Hb is based on the following experimental evidence:

- 1 The less populated states of Hb could be crystallized and solved, together with the most populated ones: i.e. we know the structures of  $T\text{Hb}$ ,  $T\text{Hb}(\text{O}_2)_4$ , and  $R\text{Hb}(\text{O}_2)_4$  [15] and references therein. All structures are symmetric or almost so. The structure of  $R\text{Hb}$  is known via its mimics, e.g.,  $R\text{Hb}$  reacted with BME. Thus, the ligand biases the allosteric equilibrium constant but does not determine the quaternary structure of the protein, coherently with the MWC model and its variants.
- 2 The allosteric structure change could be detected in real time, via the associated spectroscopic signals, in the following experiments: the modulated photo-excitation of  $\text{Hb}(\text{CO})_4\text{-Hb}(\text{CO})_3$ ; the titration of the Root effect  $\text{HbIV}(\text{CO})_4$  from trout with pH; the rearrangement of Hb after photolysis [90]. Thus, not only the ligand-independent quaternary structure equilibrium has been demonstrated, but the allosteric constants have been directly measured, at least for some derivatives.
- 3 The structure of doubly liganded intermediates could be solved thanks to the replacement of two iron atoms in the tetramer with Mn or Zn which form stable five-coordinate complexes mimicking the unliganded Fe. The other two Fe atoms were CO-liganded [15] and references therein. Doubly liganded intermediates crystallize as symmetric T state tetramers; no intermediate states were found. Actually, intermediate quaternary structures could only be observed by restricting the motion of the subunits by cross-linking.

Thus, in the case of Hb all three criteria of MWC-like reaction mechanism are fulfilled, and there can be no reasonable doubt on the assignment.

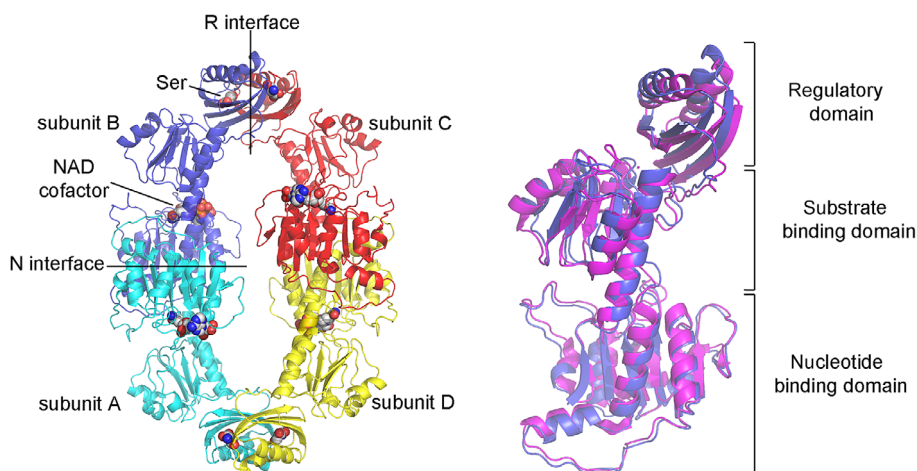
### Phosphofructokinase

Phosphofructokinase presents several features compatible with the concerted, MWC-like reaction mechanism [91]: (a) it presents positive homotropic cooperativity for fructose 6-phosphate (F6P), and heterotropic inhibition by phosphoenolpyruvate (PEP); (b) it is a homotetramer presenting isologous interfaces and may be described as a dimer of dimers; (c) the homotetramer is symmetric both in the F6P-liganded state and the PEP-inhibited state, and the two quaternary structures differ because of a rotation of one dimer with respect to the other, much like hemoglobin; and (d) the ligand-induced quaternary structure change is much larger than the tertiary structure change of each subunit. Unfortunately, direct evidence of the allosteric transition at constant degree of ligation with F6P is not available; but the enzyme from *Escherichia coli* in the absence of F6P and PEP crystallizes in the putative R state, a quite surprising finding, and crystallization in the T state is only obtained in the presence of PEP [91]. One might explain this finding as evidence of an allosteric equilibrium between two states, biased by the inhibitor. Thus, it seems very plausible that PFK qualifies for a concerted, MWC-like reaction mechanism.

### Phosphoglycerate dehydrogenase from *E. coli*

D-3-phosphoglycerate dehydrogenase (PGDH) is the enzyme that catalyzes the first step of L-Ser biosynthesis pathway. The *E. coli* enzyme is a homotetramer that is inhibited by L-Ser, whose binding stoichiometry is one molecule per subunit (see Fig. 2).

PGDH is assumed by some authors to obey the two-state MWC model [25], but in our opinion this assignment should be revised, because the enzyme presents negative homotropic cooperativity with respect to both the phosphoglycerate substrate [92], and the inhibitor L-Ser [93], a feature that cannot be accounted for by the MWC model or its variants. The structural parameters of *E. coli* PGDHs are quite compatible with the hypothesis of KNF-like reaction mechanism: the enzyme is a homotetramer; the quaternary structure obtained in the presence of the inhibitor L-Ser is symmetric, whereas that obtained in the presence of the substrates  $\text{NAD}^+$  and alpha-ketoglutarate presents significant asymmetry. Thus, we feel confident in assigning PGDH to the KNF-like group.



**Fig. 2.** Structure of *E. coli* PGDH. Left panel: tetrameric structure of PGDH; the subunits are shown in different colors; the interfaces between regulatory domains (R interfaces) are at the top and bottom of the structure, those between NADH binding domains on the left and right corners of the rhomboidal structure (N interfaces). The substrate (NAD<sup>+</sup>) and inhibitor (Ser) are shown as spheres. Right panel: superposition of subunits B from Ser-free and Ser-bound PGDH, showing the tertiary structural changes and reorientation of the regulatory domain with respect to the nucleotide binding domain.

### D-Lactate dehydrogenase

Bacterial lactate dehydrogenases can be distinguished according to their stereospecificity into D- and L-LDHs. D- and L-LDHs are evolutionarily unrelated and present wide divergence even within their respective families. In Table 1 there are separate entries for D- and L-LDHs and it is noteworthy that, though both present homotropic cooperativity, D-LDHs obey a KNF-like reaction mechanism and some of them present negative homotropic cooperativity for their substrate, whereas most L-LDHs are supposed to obey a MWC-like mechanism and present positive homotropic cooperativity and allosteric regulation by fructose 1,6 bisphosphate. Furukawa *et al.* characterized the functional and structural properties of three D-LDHs from Gram negative bacteria and correctly assigned their reaction mechanism as sequential, KNF-like, because of their negative homotropic cooperativity [29]. However, these authors claimed both families of LDHs to be allosteric, because of their homotropic cooperativity; indeed, heterotropic regulation is not reported for D-LDHs. Thus these authors completely dissociate the concept of allostery from the reaction mechanism. The structure of D-LDHs is that of a homotetramer, and in the substrate-bound form, intramolecular asymmetries are evident, consistent with the assigned reaction mechanism. While we applaud the attention given by Furukawa *et al.* to the reaction mechanism, we remark that the definition of allostery they adopt is peculiar: since D-LDHs do not present

(known) heterotropic regulation, they do not qualify for the definition given by Monod *et al.* [1]; and since they do not obey the MWC reaction mechanism, they do neither qualify for the definition given by Monod *et al.* [2]. D-LDHs qualify for the sequential KNF reaction mechanism, but Koshland never used the term allostery to refer to his model [11]. The ‘sequential allosteric transitions’ referred to in the very title of the paper are events that neither belong to the MWC nor to the KNF model and testify that the term ‘allostery’ can be stretched in whatever direction the author feels appropriate.

### Allostery in monomeric proteins

Table 1 includes several monomeric proteins that present functional regulation and structure changes operated by heterotropic effectors, which have been described as instances of model-independent allostery [94]. Monomeric proteins, however, are not described by Eqns 1–5, thus application of the concept of allostery to them requires further elaboration. Moreover, in the case of monomeric proteins the criteria we elaborated to assign the reaction mechanism are not applicable, as they rely mainly on the observation of homotropic cooperativity, and of intramolecular symmetry.

We begin our analysis with a word of caution: as Hans Frauenfelder showed many years ago [95], proteins are dynamic objects undergoing small scale structural fluctuations, which may be as subtle as

movements of the side chains of single amino acids. This realizes a continuum of conformational isomers that is better described as a statistical distribution around a thermodynamic minimum, rather than as a tertiary structure change governed by the allosteric constant  $L$ . The vast majority of authors rightly excludes this phenomenon from the concept of allostery, and indeed this type of conformational fluctuations would occur in monomeric as well as in multimeric proteins, and in both the T and R states (if they exist in the protein under consideration) or in all the liganded and unliganded states.

Chi *et al.* [96] analyzed the reaction mechanism of a monomeric PDZ domain and considered the problem in some detail. They state: ‘Conformational selection and induced fit are two well-known mechanisms of allosteric protein-ligand interaction.’ Thus, as in the case of oligomeric proteins, the term allostery may cover at least two different reaction mechanisms, each with its own mathematical formulation. Unfortunately establishing the reaction mechanism of heterotropically regulated monomeric proteins is exceedingly difficult and has been possible only in a handful of cases. The criteria we underlined in Section ‘Assignment of the reaction mechanism to a list of allosteric proteins’ are not applicable to monomeric proteins, and one must rely on kinetic evidence; however kinetic evidence is conclusive only if the kinetic constants present some special relationships among each other [96,97].

From the data available in the literature, we are not able to assign a reaction mechanism to the monomeric proteins listed in Table 1, thus in the present work we do not attempt their classification. Suffice it to say that if one wants to extend the concept of allostery to monomeric proteins, one finds there the same general problem we encountered in oligomeric proteins, i.e. allostery is an umbrella that covers more than a single reaction mechanism and cannot be defined by a single mathematical expression.

## Concluding remarks

Regulation of cell functions is key to physiology, and dysregulation is almost always a cause of disease. An important mechanism of regulation of protein (and, consequently, cell) function is allostery, which Monod called ‘the second secret of life’. The concept of allostery was initially formulated by J. Monod to describe heterotropic regulation of enzymes and other ligand binding proteins [1], and later in more stringent mechanistic terms [2]. Unfortunately, Monod’s mechanism is not the only reaction mechanism capable of producing heterotropic regulation, and thus several different

mathematical formulations are able to describe allostery; as a consequence many different types of allostery exist. This condition, which is *per se* undesirable, is further complicated by the fact that it is essentially unrecognized. Our analysis demonstrates that the concept of allostery has been extended to cover also cases unforeseen by classical thermodynamic models (e.g. monomeric proteins presenting Wyman’s heterotropic linkage [87,88], or proteins whose function is regulated by covalent chemical modification, which is thermodynamically irreversible, even though it can be undone enzymatically). We do not pretend to re-define allostery; however, our work demonstrates that in its current usage this term includes many different reaction mechanisms, and does not have any precise and universally accepted meaning; at the very least, we should try to classify the many possible types of allostery.

It is unlikely that imprecise definitions are of help for science, e.g. if one tries to define structural parameters of ‘allosteric’ proteins, one ends up with very heterogeneous results; but we could obtain a reasonable description of some structural properties of allosteric proteins by grouping them according to their reaction mechanism [13]. The scope of this review is to warn researchers that the significance and usage of the term has become so broad that an effort is required to specify which type of allostery one is referring to. Some authors tried to address the point we analyze in this work, but unfortunately their conclusions disagree. Changeux and Edelstein [98] proposed that the vast majority of cases of allostery obey a population selection mechanism. Cui and Karplus [99] conducted a review analysis similar to the one presented here, but their examples are fewer and selected to illustrate some different possible cases, rather than to provide criteria for assigning the reaction mechanisms; their analysis includes proteins that obey a sequential reaction mechanism. Hilser *et al.* [100] proposed a comprehensive model capable of including conformation selection and sequential cooperativity, an approach opposite to the one we followed in this work: they try to unify, we try to separate and classify. Nussinov and co-workers proposed the provocative, but imprecise idea that ‘allostery is an intrinsic property of all dynamic proteins’ [101], which is tantamount to say that allostery is the common property of ligand binding proteins that share none. Indeed, if all dynamic proteins were allosteric, we could dispose of the term allostery altogether. We suspect that this interpretation of allostery confuses two different concepts, whose boundary may indeed be blurry: the existence of two (or more) discrete thermodynamic states, as envisaged by the MWC model, and

the existence of conformational ensembles first described by Frauenfelder and co-workers. The latter phenomenon is indeed general, but is extended also to proteins whose function is not ligand binding, and to every state of truly allosteric proteins that may either obey the MWC or the KNF reaction mechanism. Clarity, once again, is provided by the mathematical definition, since the concept for Frauenfelder's conformational ensembles is described as the statistical distribution of minimally different conformers around an energy minimum, rather than via equations similar to Eqns 1–4. A protein that obeys a two-state reaction mechanism would have two widely separated minima (for the T and R state, respectively), with a distribution of minimally different conformers around each of them, but no intermediates in between.

An unexpected finding of the present analysis is that proteins that bind their ligands according to a sequential reaction mechanism and at the same time present heterotropic regulation are relatively uncommon, even though some cases listed as 'undefined' in Table 1 may qualify. We speculate that heterotropic regulation exerted at a tertiary structure level in an oligomeric cooperative enzyme is scarcely appealing from the viewpoint of physiology, because it is bound to produce incomplete inhibition or activation unless the concentration of the effector is high enough to saturate all the binding sites in the oligomer. Indeed, the cases of PGDH and the Arg repressor seem to confirm this speculation since in these cases effector binding is negatively cooperative but full inhibition does not require full saturation of the effector binding sites [13]; this however implies a reaction mechanism that involves a symmetric quaternary structural transition, albeit not requiring the free equilibrium between two different conformations. The KNF model does not forbid, but neither requires, quaternary structure changes, thus we consider this property atypical but not incompatible with respect to the premises of that model. The negative cooperativity of effector binding suggests that the quaternary structure change does not reflect an equilibrium between two conformations but is a case of ligand-induced fit. We may summarize the preliminary results obtained from the comparative study of proteins that obey the KNF reaction mechanism and present heterotropic regulation as follows: (a) combination of the sequential reaction mechanism and heterotropic regulation is relatively uncommon; (b) a stoichiometry of one effector binding site per subunit is observed; (c) negative cooperativity in effector binding is frequent, as are subtle tertiary and quaternary structure changes; (d) often (but not necessarily always) the binding site of the effector lies in proximity of the

intersubunit interfaces. Further work is required to validate these results.

## Acknowledgements

We gratefully acknowledge many long and fruitful discussions with Prof Maurizio Brunori on the subjects covered in this work. Funding was received from: PRIN 2017, project 2017483NH8\_05 to VM, and Progetto POR FESR Lazio 2014–2020 PROT. N. A0375-2020-36575 to AB.

## Conflict of interest

The authors declare no conflict of interest.

## Author contributions

AB designed the study and wrote the manuscript. VM and FA carried out the structural analyses.

## References

- 1 Monod J, Changeux JP and Jacob F (1963) Allosteric proteins and cellular control systems. *J Mol Biol* **6**, 306–329.
- 2 Monod J, Wyman J and Changeux JP (1965) On the nature of allosteric transitions: a plausible model. *J Mol Biol* **12**, 88–118.
- 3 Minton AP and Imai K (1974) The three-state model: a minimal allosteric description of homotropic and heterotropic effects in the binding of ligands to hemoglobin. *Proc Natl Acad Sci USA* **71**, 1418–1421.
- 4 Szabo A and Karplus M (1972) A mathematical model for structure-function relations in hemoglobin. *J Mol Biol* **72**, 163–197.
- 5 Di Cera E, Robert CH and Gill SJ (1987) Allosteric interpretation of the oxygen-binding reaction of human hemoglobin tetramers. *Biochemistry* **26**, 4003–4008.
- 6 Gill SJ, Robert CH, Coletta M, Di Cera E and Brunori M (1986) Cooperative free energies for nested allosteric models as applied to human hemoglobin. *Biophys J* **50**, 747–752.
- 7 Ackers GK, Doyle ML, Myers D and Daugherty MA (1992) Molecular code for cooperativity in hemoglobin. *Science* **255**, 54–63.
- 8 Viappiani C, Abbruzzetti S, Ronda L, Bettati S, Henry ER, Mozzarelli A and Eaton WA (2014) Experimental basis for a new allosteric model for multisubunit proteins. *Proc Natl Acad Sci USA* **111**, 12758–12763.
- 9 Yonetani T, Park SI, Tsuneshige A, Imai K and Kanaori K (2002) Global allostery model of hemoglobin. Modulation of O<sub>2</sub> affinity, cooperativity,

- and Bohr effect by heterotropic allosteric effectors. *J Biol Chem* **277**, 34508–34520.
- 10 Pauling L (1935) The oxygen equilibrium of hemoglobin and its structural interpretation. *Proc Natl Acad Sci USA* **21**, 186–191.
  - 11 Koshland DE Jr, Némethy G and Filmer D (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry* **5**, 365–385.
  - 12 Saccoccia F and Bellelli A (2020) Ligand-linked association-dissociation in transport proteins and hormone receptors. *Curr Protein Pept Sci* **21**, 993–1010.
  - 13 Morea V, Angelucci F, Tame JRH, Di Cera E and Bellelli A (2022) Structural basis of sequential and concerted cooperativity. *Biomolecules* **12**, 1651.
  - 14 Daily MD and Gray JJ (2009) Allosteric communication occurs via networks of tertiary and quaternary motions in proteins. *PLoS Comput Biol* **5**, e1000293.
  - 15 Bellelli A (2010) Hemoglobin and cooperativity: experiments and theories. *Curr Protein Pept Sci* **11**, 2–36.
  - 16 Horjales E, Altamirano MM, Calcagno ML, Garratt RC and Oliva G (1999) The allosteric transition of glucosamine-6-phosphate deaminase: the structure of the T state at 2.3 Å resolution. *Structure* **7**, 527–537.
  - 17 Oliva G, Fontes MR, Garratt RC, Altamirano MM, Calcagno ML and Horjales E (1995) Structure and catalytic mechanism of glucosamine 6-phosphate deaminase from *Escherichia coli* at 2.1 Å resolution. *Structure* **3**, 1323–1332.
  - 18 Strater N, Hakansson K, Schnappauf G, Braus G and Lipscomb WN (1996) Crystal structure of the T state of allosteric yeast chorismate mutase and comparison with the R state. *Proc Natl Acad Sci USA* **93**, 3330–3334.
  - 19 Spraggon G, Kim C, Nguyen-Huu X, Yee MC, Yanofsky C and Mills SE (2001) The structures of anthranilate synthase of *Serratia marcescens* crystallized in the presence of (i) its substrates, chorismate and glutamine, and a product, glutamate, and (ii) its end-product inhibitor, L-tryptophan. *Proc Natl Acad Sci USA* **98**, 6021–6026.
  - 20 Tao X, Yang Z and Tong L (2003) Crystal structures of substrate complexes of malic enzyme and insights into the catalytic mechanism. *Structure* **11**, 1141–1150.
  - 21 Sauer LA (1973) Mitochondrial NAD-dependent malic enzyme: a new regulatory enzyme. *FEBS Lett* **33**, 251–255.
  - 22 Xu Y, Bhargava G, Wu H, Loeber G and Tong L (1999) Crystal structure of human mitochondrial NAD(P)<sup>+</sup>-dependent malic enzyme: a new class of oxidative decarboxylases. *Structure* **7**, 877–889.
  - 23 Arai K, Ishimitsu T, Fushinobu S, Uchikoba H, Matsuzawa H and Taguchi H (2010) Active and inactive state structures of unliganded *Lactobacillus casei* allosteric L-lactate dehydrogenase. *Proteins* **78**, 681–694.
  - 24 Park S, Meyer M, Jones AD, Yennawar HP, Yennawar NH and Nixon BT (2002) Two-component signaling in the AAA + ATPase DctD: binding Mg<sup>2+</sup> and BeF<sup>3-</sup> selects between alternate dimeric states of the receiver domain. *FASEB J* **16**, 1964–1966.
  - 25 Schuller DJ, Grant GA and Banaszak LJ (1995) The allosteric ligand site in the V<sub>max</sub>-type cooperative enzyme phosphoglycerate dehydrogenase. *Nat Struct Biol* **2**, 69–76.
  - 26 Thompson JR, Bell JK, Bratt J, Grant GA and Banaszak LJ (2005) V<sub>max</sub> regulation through domain and subunit changes. The active form of phosphoglycerate dehydrogenase. *Biochemistry* **44**, 5763–5773.
  - 27 Van Duyne GD, Ghosh G, Maas WK and Sigler PB (1996) Structure of the oligomerization and L-arginine binding domain of the arginine repressor of *Escherichia coli*. *J Mol Biol* **256**, 377–391.
  - 28 Jin L, Xue WF, Fukayama JW, Yetter J, Pickering M and Carey J (2005) Asymmetric allosteric activation of the symmetric ArgR hexamer. *J Mol Biol* **346**, 43–56.
  - 29 Furukawa N, Miyanaga A, Nakajima M and Taguchi H (2018) Structural basis of sequential allosteric transitions in tetrameric d-lactate dehydrogenases from three gram-negative bacteria. *Biochemistry* **57**, 5388–5406.
  - 30 Jensen KF, Arent S, Larsen S and Schack L (2005) Allosteric properties of the GTP activated and CTP inhibited uracil phosphoribosyltransferase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *FEBS J* **272**, 1440–1453.
  - 31 Arent S, Harris P, Jensen KF and Larsen S (2005) Allosteric regulation and communication between subunits in uracil phosphoribosyltransferase from *Sulfolobus solfataricus*. *Biochemistry* **44**, 883–892.
  - 32 Staub M and Dénes G (1969) Purification and properties of the 3-deoxy-D-arabino- heptulosonate-7-phosphate synthase (phenylalanine sensitive) of *Escherichia coli* K12. I. Purification of enzyme and some of its catalytic properties. *Biochim Biophys Acta* **178**, 588–598.
  - 33 Staub M and Dénes G (1969) Purification and properties of the 3-deoxy-D-arabino- heptulosonate-7-phosphate synthase (phenylalanine sensitive) of *Escherichia coli* K12. II. Inhibition of activity of the enzyme with phenylalanine and functional group-specific reagents. *Biochim Biophys Acta* **178**, 599–608.
  - 34 Shumilin IA, Zhao C, Bauerle R and Kretsinger RH (2002) Allosteric inhibition of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase alters the coordination of both substrates. *J Mol Biol* **320**, 1147–1156.

- 35 Rafferty JB, Somers WS, Saint-Girons I and Phillips SE (1989) Three-dimensional crystal structures of *Escherichia coli* met repressor with and without corepressor. *Nature* **341**, 705–710.
- 36 Anderson AC, O'Neil RH, DeLano WL and Stroud RM (1999) The structural mechanism for half-the-sites reactivity in an enzyme, thymidylate synthase, involves a relay of changes between subunits. *Biochemistry* **38**, 13829–13836.
- 37 Schumacher MA, Choi KY, Lu F, Zalkin H and Brennan RG (1995) Mechanism of corepressor-mediated specific DNA binding by the purine repressor. *Cell* **83**, 147–155.
- 38 Schumacher MA, Glasfeld A, Zalkin H and Brennan RG (1997) The X-ray structure of the PurR-guanine-purF operator complex reveals the contributions of complementary electrostatic surfaces and a water-mediated hydrogen bond to corepressor specificity and binding affinity. *J Biol Chem* **272**, 22648–22653.
- 39 Hinrichs W, Kisker C, Düvel M, Müller A, Tovar K, Hillen W and Saenger W (1994) Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance. *Science* **264**, 418–420.
- 40 Choe JY, Fromm HJ and Honzatko RB (2000) Crystal structures of fructose 1,6-bisphosphatase: mechanism of catalysis and allosteric inhibition revealed in product complexes. *Biochemistry* **39**, 8565–8574.
- 41 Amor JC, Harrison DH, Kahn RA and Ringe D (1994) Structure of the human ADP-ribosylation factor 1 complexed with GDP. *Nature* **372**, 704–708.
- 42 Maita N, Hatakeyama K, Okada K and Hakoshima T (2004) Structural basis of bioprotein-induced inhibition of GTP cyclohydrolase I by GFRP, its feedback regulatory protein. *J Biol Chem* **279**, 51534–51540.
- 43 Sixma TK, Pronk SE, Kalk KH, van Zanten BA, Berghuis AM and Hol WG (1992) Lactose binding to heat-labile enterotoxin revealed by X-ray crystallography. *Nature* **355**, 561–564.
- 44 Matković-Calogović D, Loregian A, D'Acunto MR, Battistutta R, Tossi A, Palù G and Zanotti G (1999) Crystal structure of the B subunit of *Escherichia coli* heat-labile enterotoxin carrying peptides with anti-herpes simplex virus type 1 activity. *J Biol Chem* **274**, 8764–8769.
- 45 MacRae IJ, Segel IH and Fisher AJ (2002) Allosteric inhibition via R-state destabilization in ATP sulfurylase from *Penicillium chrysogenum*. *Nat Struct Biol* **9**, 945–949.
- 46 Soisson SM, MacDougall-Shackleton B, Schleif R and Wolberger C (1997) Structural basis for ligand-regulated oligomerization of AraC. *Science* **276**, 421–425.
- 47 Weldon JE, Rodgers ME, Larkin C and Schleif RF (2007) Structure and properties of a truly apo form of AraC dimerization domain. *Proteins* **66**, 646–654.
- 48 Pasqualato S, Senic-Matuglia F, Renault L, Goud B, Salamero J and Cherfils J (2004) The structural GDP/GTP cycle of Rab11 reveals a novel interface involved in the dynamics of recycling endosomes. *J Biol Chem* **279**, 11480–11488.
- 49 Munshi S, Hall DL, Kornienko M, Darke PL and Kuo LC (2003) Structure of apo, unactivated insulin-like growth factor-1 receptor kinase at 1.5 Å resolution. *Acta Crystallogr D Biol Crystallogr* **59**, 1725–1730.
- 50 Favelyukis S, Till JH, Hubbard SR and Miller WT (2001) Structure and autoregulation of the insulin-like growth factor 1 receptor kinase. *Nat Struct Biol* **8**, 1058–1063.
- 51 Weaver LH, Kwon K, Beckett D and Matthews BW (2001) Corepressor-induced organization and assembly of the biotin repressor: a model for allosteric activation of a transcriptional regulator. *Proc Natl Acad Sci USA* **98**, 6045–6050.
- 52 Wilson KP, Shewchuk LM, Brennan RG, Otsuka AJ and Matthews BW (1992) *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. *Proc Natl Acad Sci USA* **89**, 9257–9261.
- 53 Hubbard SR, Wei L, Ellis L and Hendrickson WA (1994) Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* **372**, 746–754.
- 54 Cho Y, Sharma V and Sacchettini JC (2003) Crystal structure of ATP phosphoribosyltransferase from *Mycobacterium tuberculosis*. *J Biol Chem* **278**, 8333–8339.
- 55 Friedman AM, Fischmann TO and Steitz TA (1995) Crystal structure of lac repressor core tetramer and its implications for DNA looping. *Science* **268**, 1721–1727.
- 56 Bell CE and Lewis M (2000) A closer view of the conformation of the lac repressor bound to operator. *Nat Struct Biol* **7**, 209–214.
- 57 Birck C, Mourey L, Gouet P, Fabry B, Schumacher J, Rousseau P, Kahn D and Samama JP (1999) Conformational changes induced by phosphorylation of the FixJ receiver domain. *Structure* **7**, 1505–1515.
- 58 Gouet P, Fabry B, Guillet V, Birck C, Mourey L, Kahn D and Samama JP (1999) Structural transitions in the FixJ receiver domain. *Structure* **7**, 1517–1526.
- 59 Choi H, Kim S, Mukhopadhyay P, Cho S, Woo J, Storz G and Ryu SE (2001) Structural basis of the redox switch in the OxyR transcription factor. *Cell* **105**, 103–113.
- 60 Volz K and Matsumura P (1991) Crystal structure of *Escherichia coli* CheY refined at 1.7-Å resolution. *J Biol Chem* **266**, 15511–15519.
- 61 Yang J, Cron P, Thompson V, Good VM, Hess D, Hemmings BA and Barford D (2002) Molecular mechanism for the regulation of protein kinase B/Akt by hydrophobic motif phosphorylation. *Mol Cell* **9**, 1227–1240.

- 62 Seavers PR, Lewis RJ, Brannigan JA, Verschueren KH, Murshudov GN and Wilkinson AJ (2001) Structure of the *Bacillus* cell fate determinant SpoIIAA in phosphorylated and unphosphorylated forms. *Structure* **9**, 605–614.
- 63 Hardy JA, Lam J, Nguyen JT, O'Brien T and Wells JA (2004) Discovery of an allosteric site in the caspases. *Proc Natl Acad Sci USA* **101**, 12461–12466.
- 64 Zhang F, Strand A, Robbins D, Cobb MH and Goldsmith EJ (1994) Atomic structure of the MAP kinase ERK2 at 2.3 Å resolution. *Nature* **367**, 704–711.
- 65 Canagarajah BJ, Khokhlatchev A, Cobb MH and Goldsmith EJ (1997) Activation mechanism of the MAP kinase ERK2 by dual phosphorylation. *Cell* **90**, 859–869.
- 66 Vetter IR, Arndt A, Kutay U, Görlich D and Wittinghofer A (1999) Structural view of the ran-importin beta interaction at 2.3 Å resolution. *Cell* **97**, 635–646.
- 67 Milburn MV, Tong L, deVos AM, Brünger A, Yamaizumi Z, Nishimura S and Kim SH (1990) Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. *Science* **247**, 939–945.
- 68 Garrard SM, Capaldo CT, Gao L, Rosen MK, Macara IG and Tomchick DR (2003) Structure of Cdc42 in a complex with the GTPase-binding domain of the cell polarity protein, Par6. *EMBO J* **22**, 1125–1133.
- 69 Grizot S, Fauré J, Fieschi F, Vignais PV, Dagher MC and Pebay-Peyroula E (2001) Crystal structure of the Rac1-RhoGDI complex involved in NADPH oxidase activation. *Biochemistry* **40**, 10007–10013.
- 70 Puius YA, Zhao Y, Sullivan M, Lawrence DS, Almo SC and Zhang ZY (1997) Identification of a second aryl phosphate-binding site in protein-tyrosine phosphatase 1B: a paradigm for inhibitor design. *Proc Natl Acad Sci USA* **94**, 13420–13425.
- 71 Wiesmann C, Barr KJ, Kung J, Zhu J, Erlanson DA, Shen W, Fahr BJ, Zhong M, Taylor L, Randal M *et al.* (2004) Allosteric inhibition of protein tyrosine phosphatase 1B. *Nat Struct Mol Biol* **11**, 730–737.
- 72 Stroupe C and Brunger AT (2000) Crystal structures of a Rab protein in its inactive and active conformations. *J Mol Biol* **304**, 585–598.
- 73 Yu Y, Li S, Xu X, Li Y, Guan K, Arnold E and Ding J (2005) Structural basis for the unique biological function of small GTPase RHEB. *J Biol Chem* **280**, 17093–17100.
- 74 Kjeldgaard M, Nissen P, Thirup S and Nyborg J (1993) The crystal structure of elongation factor EF-Tu from *Thermus aquaticus* in the GTP conformation. *Structure* **1**, 35–50.
- 75 Polekhina G, Thirup S, Kjeldgaard M, Nissen P, Lippmann C and Nyborg J (1996) Helix unwinding in the effector region of elongation factor EF-Tu-GDP. *Structure* **4**, 1141–1151.
- 76 Constantinescu AT, Rak A, Alexandrov K, Esters H, Goody RS and Scheidig AJ (2002) Rab-subfamily-specific regions of Ypt7p are structurally different from other RabGTPases. *Structure* **10**, 569–579.
- 77 Cherfils J, Ménétrey J, Le Bras G, Janoueix-Lerosey I, de Gunzburg J, Garel JR and Auzat I (1997) Crystal structures of the small G protein Rap2A in complex with its substrate GTP, with GDP and with GTPgammaS. *EMBO J* **16**, 5582–5591.
- 78 Ruzheinikov SN, Das SK, Sedelnikova SE, Baker PJ, Artymiuk PJ, García-Lara J, Foster SJ and Rice DW (2004) Analysis of the open and closed conformations of the GTP-binding protein YsxC from *Bacillus subtilis*. *J Mol Biol* **339**, 265–278.
- 79 Pasqualato S, Ménétrey J, Franco M and Cherfils J (2001) The structural GDP/GTP cycle of human Arf6. *EMBO Rep* **2**, 234–238.
- 80 Lambright DG, Noel JP, Hamm HE and Sigler PB (1994) Structural determinants for activation of the alpha-subunit of a heterotrimeric G protein. *Nature* **369**, 621–628.
- 81 Noel JP, Hamm HE and Sigler PB (1993) The 2.2 Å crystal structure of transducin-alpha complexed with GTP gamma S. *Nature* **366**, 654–663.
- 82 Rak A, Pylypenko O, Niculae A, Pyatkov K, Goody RS and Alexandrov K (2004) Structure of the Rab7:REP-1 complex: insights into the mechanism of Rab prenylation and choroideremia disease. *Cell* **117**, 749–760.
- 83 Wei Y, Zhang Y, Derewenda U, Liu X, Minor W, Nakamoto RK, Somlyo AV, Somlyo AP and Derewenda ZS (1997) Crystal structure of RhoA-GDP and its functional implications. *Nat Struct Biol* **4**, 699–703.
- 84 Ihara K, Muraguchi S, Kato M, Shimizu T, Shirakawa M, Kuroda S, Kaibuchi K and Hakoshima T (1998) Crystal structure of human RhoA in a dominantly active form complexed with a GTP analogue. *J Biol Chem* **273**, 9656–9666.
- 85 Coleman DE, Lee E, Mixon MB, Linder ME, Berghuis AM, Gilman AG and Sprang SR (1994) Crystallization and preliminary crystallographic studies of Gi alpha 1 and mutants of Gi alpha 1 in the GTP and GDP-bound states. *J Mol Biol* **238**, 630–634.
- 86 Mixon MB, Lee E, Coleman DE, Berghuis AM, Gilman AG and Sprang SR (1995) Tertiary and quaternary structural changes in Gi alpha 1 induced by GTP hydrolysis. *Science* **270**, 954–960.
- 87 Wyman J (1948) Heme proteins. *Adv Protein Chem* **4**, 407–531.
- 88 Wyman J (1964) Linked functions and reciprocal effects in hemoglobin: a second look. *Adv Protein Chem* **19**, 223–286.
- 89 Bellelli A and Carey J (2018) *Reversible Ligand Binding: Theory and Experiment*. Wiley, Oxford, UK.



- 90 Bellelli A and Brunori M (1994) Optical measurements of quaternary structural changes in hemoglobin. *Methods Enzymol* **232**, 56–71.
- 91 Schirmer T and Evans PR (1990) Structural basis of the allosteric behaviour of phosphofructokinase. *Nature* **343**, 140–145.
- 92 Winicov I and Pizer LI (1974) The mechanism of end product inhibition of serine biosynthesis. IV. Subunit structure of phosphoglycerate dehydrogenase and steady state kinetic studies of phosphoglycerate oxidation. *J Biol Chem* **249**, 1348–1355.
- 93 Grant GA, Xu XL, Hu Z and Purvis AR (1999) Phosphate ion partially relieves the cooperativity of effector binding in D-3-phosphoglycerate dehydrogenase without altering the cooperativity of inhibition. *Biochemistry* **38**, 16548–16552.
- 94 Changeux JP (2012) Allostery and the Monod-Wyman-Changeux model after 50 years. *Annu Rev Biophys* **41**, 103–133.
- 95 Ansari A, Berendzen J, Bowne SF, Frauenfelder H, Iben IE, Sauke TB, Shyamsunder E and Young RD (1985) Protein states and proteinquakes. *Proc Natl Acad Sci USA* **82**, 5000–5004.
- 96 Chi CN, Bach A, Engstrom A, Wang H, Strømgaard K, Gianni S and Jemth P (2009) A sequential binding mechanism in a PDZ domain. *Biochemistry* **48**, 7089–7097.
- 97 Hammes GG, Chang YC and Oas TG (2009) Conformational selection or induced fit: a flux description of reaction mechanism. *Proc Natl Acad Sci USA* **106**, 13737–13741.
- 98 Changeux JP and Edelstein SJ (2011) Conformational selection or induced fit? 50 years of debate resolved. *F1000 Biol Rep* **3**, 19.
- 99 Cui Q and Karplus M (2008) Allostery and cooperativity revised. *Protein Sci* **17**, 1295–1307.
- 100 Hilser VJ, Wrabl OJ and Motlagh HN (2012) Structural and energetic basis of Allostery. *Annu Rev Biophys* **41**, 585–609.
- 101 Gunasekaran K, Ma B and Nussinov R (2004) Allostery is an intrinsic property of all dynamic proteins. *Proteins* **57**, 433–443.