Published in final edited form as:
Nature. 2007 July 19; 448(7151): 325–329. doi:10.1038/nature05959.

Conformational entropy in molecular recognition by proteins: A first example from the calmodulin system

Kendra King Frederick, **Michael S. Marlow**, **Kathleen G. Valentine**, and **A. Joshua Wand**
Johnson Research Foundation and Department of Biochemistry & Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania, 19104 USA

Abstract

Molecular recognition by proteins is fundamental to almost every biological process, particularly the protein associations underlying cellular signal transduction. Understanding the basis for protein-protein interactions requires the full characterization of the thermodynamics of their association. Historically it has been virtually impossible to experimentally estimate changes in residual protein entropy, a potentially important component of the change in the free energy of protein association. However, solution NMR spectroscopy has recently emerged as a powerful tool for characterizing the dynamics of proteins and has thereby gained access to their residual entropy[1]. Here we show that the change in internal dynamics of the protein calmodulin varies significantly upon binding a variety of target domains. Surprisingly, the apparent change in the corresponding residual entropy is linearly related to the change in the overall binding entropy. These results indicate that changes in residual protein conformational entropy can contribute significantly to the free energy of protein-ligand association. It seems evident that protein entropy can be exploited in the maturation of high affinity interactions either by biological evolution or by human intervention such as in the design of protein-targeted pharmaceuticals.

Numerous structural studies have revealed that protein-protein interfaces often involve dozens of amino acids and thousands of Å^2 of contact area [2]. It has also become apparent that a non-uniform contribution of individual residues to the free energy of binding can exist and that static structural analyses can mask important factors underlying the high affinity interactions between proteins[3]. Of particular interest here is the role of the residual protein conformational entropy in modulating the free energy of the association of a protein with a ligand, i.e.

$$\begin{aligned}\Delta G_{binding} &= \Delta H_{binding} - T\Delta S_{binding} \\ &= \Delta H_{binding} - T\left(\Delta S^{protein}_{binding} + \Delta S^{ligand}_{binding} + \Delta S^{solvent}_{binding}\right)\end{aligned} \quad (1)$$

The trivial decomposition made in Equation 1 emphasizes the fact that the entropy of binding ($\Delta S_{binding}$), obtainable by calorimetric methods, is comprised of contributions

Correspondence and requests for materials should be addressed to A.J.W. (wand@mail.med.upenn.edu).
The authors declare that they have no competing financial interests.
Supplementary Information is linked to the online version of the paper at www.nature.com/nature

associated with the ligand, solvent and the protein. It is well-established that gross disorder-order transitions of a ligand can profoundly influence the entropy of macromolecular associations[4]. It is also well-established that burial of hydrophobic surface area and the consequent release of hydration waters to the bulk solvent can also contribute significantly to the thermodynamics of binding [5]. What is less understood is the potential entropic contributions from a structured protein ($\Delta S_{binding}^{protein}$), which includes changes in its internal residual conformational entropy ($\Delta S_{conf}^{protein}$) as well as changes in rotational and translational entropy[6–8]. Here we focus on $\Delta S_{conf}^{protein}$. As is obvious from Equation 1, the measurement of total system thermodynamic parameters does not resolve contributions from internal protein conformational entropy. The estimation of changes in conformational entropy due to protein-protein association from molecular dynamics simulations remains a considerable challenge[9]. Direct measurement of the residual conformational entropy of the protein in its free and complexed states is therefore required. Recent developments in NMR relaxation methods and analysis now make this feasible.

We have employed calmodulin as a model system to investigate the role for changes in protein conformational entropy in the high affinity association of proteins. Calmodulin is a central participant in the calcium-mediated signal transduction pathways of eukaryotes[10]. It interacts with and regulates the activity of approximately three hundred proteins[11]. In principle, the characterization of calmodulin's internal dynamics should facilitate measurement of conformational entropy changes through a "counting of states" implicit in molecular motion[1]. Previously, using NMR relaxation methods, we have shown that calcium-saturated calmodulin itself is an unusually dynamic protein, which is characterized by a broad, non-uniform multi-modal distribution of the amplitude of fast side chain dynamics[12]. Binding of a target domain to calcium-saturated calmodulin causes a significant redistribution of the fast side chain dynamics in calmodulin[12]. This raises the possibility that CaM employs its internal residual entropy to "tune" its affinity for ligands.

Here we explore this question by using NMR methods to determine the dynamical response of CaM to the binding of six target domains. Peptides representing the calmodulin-binding domains of the smooth muscle myosin light chain kinase (smMLCKp)[13], the neuronal and endothelial nitric oxide synthases (nNOSp and eNOSp)[14], the calmodulin kinase kinase (CaMKKαp)[15], the calmodulin kinase I (CaMKIp)[16], and the phosphodiesterase (PDEp)[17] have been employed. All have a basic amphiphillic character and form alpha-helical structure when bound to calmodulin (Figure 1a). Four of the peptides have been found previously by isothermal titration calorimetry to have roughly the same affinity for calmodulin but with widely different thermodynamic parameters defining the free energy of association[18]. We have repeated the ITC measurements at a temperature (35 °C) that is more optimal for solution NMR spectroscopy and have characterized the binding of the two additional domains (Figure 1b). In the case of the CaMKKαp and smMLCKp domains, binding is driven by a large favourable change in system enthalpy overcoming a large unfavourable change in system entropy. At the other extreme, nNOSp binding is driven by a large favourable change in system enthalpy accompanied by a small favourable change in system entropy. The PDEp, CaMKIp, and eNOSp calmodulin-binding domains represent

intermediate cases. The entropy of binding of these domains varies by 90 kJ/mole and changes sign.

Titration of CaM with each of the peptides reveals that all six of the resulting complexes have a 1:1 stoichiometry and are in slow exchange with their dissociated components on the NMR ^{1}H chemical shift time scale (not shown). The CaM-smMLCKp, CaM-PDEp and CaM-CaMKKαp complexes show very little conformational heterogeneity, as judged by ^{15}N- and ^{13}C-HSQC spectra, while the CaM-nNOSp, CaM-eNOSp and CaM-CaMKIp complexes show limited heterogeneity at a small number of locations in the calmodulin molecule. This was found to largely arise from populations of minor rotameric orientations of methyl bearing side chains. These results indicate a range of localized conformational heterogeneity in calmodulin across the six calmodulin complexes. This heterogeneity represents classical conformational entropy.

The sub-ns dynamics of the polypeptide backbone of calmodulin in the six complexes were probed using ^{15}N dipolar relaxation by its attached hydrogen[19] and by transverse cross-correlated relaxation between ^{13}CO chemical shift anisotropy and the ^{13}CO-^{13}Cα dipolar interactions [20]. The sub-ns dynamics of methyl groups of calmodulin side chains were characterized using ^{2}H spin relaxation methods[21]. The degree of spatial restriction of each motional probe was assessed via the squared model-free generalized order parameter (O^2)[22]. Values of the O^2 parameter can vary from 0 to 1, corresponding to isotropic disorder and a fixed orientation in the molecular frame, respectively. We denote the O^2 parameters for motion of the amide N-H bond, C_α-C′ bond and methyl symmetry axis as O^2_{NH}, $O^2_{C\alpha CO}$ and O^2_{axis}, respectively.

The dynamics of the backbone of calmodulin are invariant across the complexes, as indicated by the average O^2_{NH} and $O^2_{C\alpha CO}$ parameters (Supplementary Information). In contrast, the motion of methyl-bearing amino acid side chains varies significantly with the nature of the target domain. There are 56 methyl-bearing amino acids providing 80 methyl groups as probes distributed across the primary sequence of calmodulin and include nine methionines that line the target domain binding sites formed in the various complexes.

To project the change in internal protein dynamics in terms of residual protein entropy, we follow Karplus and co-workers and describe the protein as a disjoint multidimensional harmonic well[8]:

$$S_{conf}^{protein} = \sum_{i=1}^{N} p_i S_i^h - k_B \sum_{i=1}^{N} p_i \ln p_i \quad (2)$$

where S_i^h represents the entropy manifested by fast intra-well motion and the second term corresponds to the classical conformational entropy arising from the N distinct conformations. Here S_i^h is obtained from interpretation of local order parameters, which is model-dependent. We choose a simple harmonic oscillator treatment[23]. *Differences* in entropy calculated from changes in $O^2$are fairly insensitive to the model used[24]. Calcium-saturated calmodulin is used as the reference state for obtaining $\Delta S_{conf}^{protein}$. The second term

of Equation 2 represents classical entropy from the local heterogeneity of side chain conformers. This can be manifested on a range of time scales. Some methyl sites exhibited slowly interconverting conformational heterogeneity on the chemical shift time scale. This was interpreted as classical entropy with the population of each state (p_i) estimated from the intensity of cross peaks. This contributed less than 2% of the estimated change in conformational entropy due to binding. It has also been shown that fast motion *between* rotamer wells contributes significantly to low O^2_{axis} parameters[24]. This also represents conformational entropy and was estimated using a previously described model[24]. This resulted in a roughly constant 15% of the total conformational entropy.

Notwithstanding the limitations imposed by correlated motion[24], the changes in the residual entropy of calmodulin upon binding to the six peptides, obtained by simple summation of the individual local entropies, shows a linear correlation ($R^2 = 0.78$) with the corresponding entropy of binding (Figure 2). Taken at face value, half of the binding entropy is reflected in the motion of the methyl-bearing amino acid side chains. There is no *a priori* reason for such a correlation. However, the linearity of the correlation implies that either the change in the residual entropy of calmodulin upon binding a target domain is a major contribution to the binding entropy or that the various sources of entropy change in concert (see Equation 1). Regardless, it appears that the residual entropy of calmodulin can vary sufficiently to impact the free energy changes arising from high affinity protein associations. Therefore the commonly held view that high affinity interactions are necessarily energetically dominated by specific structural (enthalpic) interactions must be relaxed to include the structural dynamics and heterogeneity that contributes to residual protein entropy.

The binding of target domains to CaM results in a distribution of O^2_{axis} parameters that is remarkable for its distinct clustering into three apparent classes of motion (Figure 3). The relative populations of these motional classes in calmodulin vary considerably across the six complexes. Although the distinctive grouping of order parameters, seen across all six complexes studied here, is often obscured in other proteins[25], the motional origin of these classes is clear. In the case of calmodulin, two fundamental types of motion occurring on the sub-ns time scale are involved: motion within a rotamer well and motion between rotamer wells of side chain torsion angles[24]. It has been shown that the class of motion centred on a O^2_{axis} value of ~0.35 generally involves a significant contribution from rotameric interconversion on the nanosecond or faster time scale as it leads to a significant averaging of scalar coupling (J) constants[24]. More recent experimental results[26] and theoretical simulations[27] suggests this to be general. The distribution of motion at the other extreme is centred on a O^2_{axis} value of ~0.8, which represents highly restricted motion within a rotamer well. The class of motion centred on a O^2_{axis} value of ~0.6 involves little detectable rotamer interconversion and is restricted to motion within a single rotamer well. The precise value reflects intra-well motion and the effects of superposition of similar motion about connected torsion angles. We have termed these groupings the J-, ω- and α-classes of motion, respectively[1]. The distributions of O^2_{axis} parameters in each of the six physiologically-relevant calmodulin complexes are satisfactorily described by a sum of three Gaussians (Figure 3 and Supplementary Information).

There exists a surprising correlation between the populations of the three classes of motion and the change in total system entropy for binding (Figure 3). The population of the J-class is negatively correlated with the entropic contribution (−TΔS) to the free energy of binding. The populations of the ω- and α-classes are positively correlated. The correlations are remarkably linear for all three classes. This is a direct, relatively model-insensitive indication that the residual entropy of calmodulin changes in concert with the change in the entropy of binding and that this variation can, in part, be identified with the motional class of the involved side chains.

In summary, we have employed a battery of NMR methods to characterize the dynamical response of calmodulin to the binding of six target regulatory domains. This view has been interpreted in terms of the changes in conformational entropy of calmodulin upon binding. The behaviour of the six physiologically relevant interactions indicates that the residual entropy of structured proteins can enter very significantly into high affinity interactions between proteins. It will therefore be of great interest to determine how prevalent this is across the universe of protein-protein interactions in biology. Finally, the results presented here suggest that residual protein entropy can indeed play a significant role in more complex protein functions such as allostery[7].

METHODS

Sample preparation & isothermal titration calorimetry

Calmodulin and synthetic peptides and complexes were prepared as described previously[28]. All solutions were prepared in 20 mM imidazole (pH 6.5), 100 mM KCl, 6 mM $CaCl_2$ and 0.02 % (w/v) NaN_3. For isothermal titration calorimetry, calcium-saturated calmodulin (200 μM) was used to titrate dilute solutions of peptide (10–20 μM) in order to avoid artefacts arising from peptide aggregation. ITC measurements were performed with a VP-ITC (Microcal) thermostatically regulated at 35 °C. Data were analyzed with Origin (v.5) software. NMR samples were slightly (~10%) over-titrated with peptide to ensure full complex formation.

NMR methods

O^2_{axis} parameters were determined from T_1 and $T_{1\rho}$ deuterium relaxation[21] measured at two magnetic fields. Rotational correlation times and O^2_{NH} were determined by ^{15}N relaxation measurements[19] obtained at two magnetic fields. $O^2_{CO-C\alpha}$ parameters were determined by transverse cross-correlated relaxation between ^{13}CO chemical shift anisotropy and the ^{13}CO-$^{13}C\alpha$ dipolar interactions as described elsewhere[20]. All measurements were made at 35 °C. Model free parameters[22] were determined using a comprehensive grid search approach[29]. The average error of O^2_{axis}, O^2_{NH} and $O^2_{CO-C\alpha}$ parameters across all complexes were estimated by Monte Carlo sampling to be 0.016, 0.011 and 0.024, respectively. The squared generalized order parameters and effective correlation times have been deposited in the Biological Magnetic Resonance Data Bank (http://www.bmrb.wisc.edu/).

Data interpretation

The summed distribution of O^2_{axis} parameters of all six complexes was fitted to a random distribution and to one, two and three Gaussian distribution models. The bin size for this analysis was determined from a well-established formula for optimal bin width[30] and was found to be 0.05. Only the three Gaussian model could satisfactorily describe the data ($p < 0.0001$). The fitting procedure is described in the Supplementary Information. The summed distribution yielded fitted 3-Gaussian distributions centred on O^2_{axis} values of 0.35 (J-class), 0.58 (α-class) and 0.78 (ω-class). These centres were fixed in subsequent fitting of the O^2_{axis} distributions of the individual complexes from which the relative populations of each motional class were obtained. Uncertainties in the fitted populations were estimated by varying the O^2_{axis} parameters by two standard deviations. This results in asymmetric error bars. The total change in conformational entropy of calmodulin upon binding a target domain was calculated as the sum of three terms:

$\Delta S^{protein}_{conf} = \Delta S^{harm} + \Delta S^{fast}_{rotamer} + \Delta S^{slow}_{rotamer}$. Changes in conformational entropy (ΔS^{harm}) expressed as changes in motion within a rotameric well on the fast time scale (sub-ns) were obtained from the experimentally determined O^2_{axis} parameters using a simple harmonic oscillator model.[23] To calculate changes in entropy derived from motion of the same oscillator, site-to-site comparison to free calcium-saturated calmodulin was used as the reference state. The change in the entropy reflected by the change in the motion of each methyl symmetry axis was estimated using $\Delta S = -18 \times \Delta O^2_{axis}\ \mathrm{J\,mol^{-1}\,K^{-1}}$. See Li et al.[23] for further details of the model. A classical entropy term ($\Delta S^{fast}_{rotamer}$) was added to represent minor conformers that are sampled due to fast rotameric interconversion that also contributes to the generalized order parameter[24]. For the small number of methyl sites having multiple conformations in slow exchange on the NMR chemical shift time scale, an additional classical entropy contribution ($\Delta S^{slow}_{rotamer}$) was calculated based on the measured intensities. See Supplementary Information for additional details and results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a grant from the National Institutes of Health. We are grateful to Professor S. W. Englander for helpful discussion and to Professor Mark I. Greene for access to isothermal titration calorimetry instrumentation.

References

1. Igumenova TI, Frederick KK, Wand AJ. Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution. Chem Rev. 2006; 106:1672–1699. [PubMed: 16683749]
2. Wodak SJ, Janin J. Structural basis of macromolecular recognition. Adv Prot Chem. 2002; 61:9–73.
3. Clackson T, Wells JA. A hot spot of binding energy in a hormone-receptor interface. Science. 1995; 267:383–386. [PubMed: 7529940]

4. Spolar RS, Record MT Jr. Coupling of local folding to site-specific binding of proteins to DNA. Science. 1994; 263:777–784. [PubMed: 8303294]
5. Sturtevant JM. Heat capacity and entropy changes in processes involving proteins. Proceedings of the National Academy of Sciences of the United States of America. 1977; 74:2236–2240. [PubMed: 196283]
6. Steinberg IZ, Scheraga HA. Entropy changes accompanying association reactions of proteins. J Biol Chem. 1963; 238:172–181. [PubMed: 13983721]
7. Cooper A, Dryden DTF. Allostery without conformational change -a plausible model. Eur Biophys J Biophys Lett. 1984; 11:103–109.
8. Karplus M, Ichiye T, Pettitt BM. Configurational entropy of native proteins. Biophys J. 1987; 52:1083–1085. [PubMed: 3427197]
9. Grunberg R, Nilges M, Leckner J. Flexibility and conformational entropy in protein-protein binding. Structure. 2006; 14:683–693. [PubMed: 16615910]
10. Kahl CR, Means AR. Regulation of cell cycle progression by calcium/calmodulin-dependent pathways. Endocr Rev. 2003; 24:719–736. [PubMed: 14671000]
11. Yap KL, et al. Calmodulin target database. J Struct Funct Genom. 2000; 1:8–14.
12. Lee AL, Kinnear SA, Wand AJ. Redistribution and loss of side chain entropy upon formation of a calmodulin-peptide complex. Nat Struct Biol. 2000; 7:72–77. [PubMed: 10625431]
13. Lukas TJ, et al. Calmodulin binding domains: characterization of a phosphorylation and calmodulin binding site from myosin light chain kinase. Biochemistry. 1986; 25:1458–1464. [PubMed: 3754463]
14. Zhang M, Vogel HJ. Characterization of the calmodulin-binding domain of rat cerebellar nitric oxide synthase. J Biol Chem. 1994; 269:981–985. [PubMed: 7507114]
15. Tokumitsu H, et al. Calcium/calmodulin-dependent protein kinase kinase: identification of regulatory domains. Biochemistry. 1997; 36:12823–12827. [PubMed: 9335539]
16. Goldberg J, Nairn AC, Kuriyan J. Structural basis for the autoinhibition of calcium/calmodulin-dependent protein kinase I. Cell. 1996; 84:875–887. [PubMed: 8601311]
17. Charbonneau H, et al. Evidence for domain organization within the 61-kDa calmodulin-dependent cyclic nucleotide phosphodiesterase from bovine brain. Biochemistry. 1991; 30:7931–7940. [PubMed: 1651111]
18. Wintrode PL, Privalov PL. Energetics of target peptide recognition by calmodulin: a calorimetric study. J Mol Biol. 1997; 266:1050–1062. [PubMed: 9086281] Brokx RD, et al. Energetics of target peptide binding by calmodulin reveals different modes of binding. J Biol Chem. 2001; 276:14083–14091. [PubMed: 11278815]
19. Farrow NA, et al. Backbone dynamics of a free and a phosphopeptide-complexed Src homology-2 domain studied by N-15 NMR relaxation. Biochemistry. 1994; 33:5984–6003. [PubMed: 7514039]
20. Wang T, Cai S, Zuiderweg ER. Temperature dependence of anisotropic protein backbone dynamics. J Am Chem Soc. 2003; 125:8639–8643. [PubMed: 12848571]
21. Muhandiram DR, et al. Measurement of H-2 T-1 and T-1p relaxation-times in uniformly C-13-Labeled and fractionally H-2-labeled proteins in solution. J Am Chem Soc. 1995; 117:11536–11544.
22. Lipari G, Szabo A. Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules. 1 Theory and Range of Validity. J Am Chem Soc. 1982; 104:4546–4559.
23. Li Z, Raychaudhuri S, Wand AJ. Insights into the local residual entropy of proteins provided by NMR relaxation. Prot Sci. 1996; 5:2647–2650.
24. Lee AL, et al. Temperature dependence of the internal dynamics of a calmodulin-peptide complex. Biochemistry. 2002; 41:13814–13825. [PubMed: 12427045]
25. Best RB, Clarke J, Karplus M. The origin of protein sidechain order parameter distributions. J Am Chem Soc. 2004; 126:7734–7735. [PubMed: 15212494]
26. Chou JJ, Case DA, Bax A. Insights into the mobility of methyl-bearing side chains in proteins from (3)J(CC) and (3)J(CN) couplings. J Am Chem Soc. 2003; 125:8959–8966. [PubMed: 12862493]

27. Best RB, Clarke J, Karplus M. What contributions to protein side-chain dynamics are probed by NMR experiments? A molecular dynamics simulation analysis. J Mol Biol. 2005; 349:185–203. [PubMed: 15876377]

28. Kranz JK, et al. A direct test of the reductionist approach to structural studies of calmodulin activity: relevance of peptide models of target proteins. J Biol Chem. 2002; 277:16351–16354. [PubMed: 11904288]

29. Dellwo MJ, Wand AJ. Model-independent and model-dependent analysis of the global and internal dynamics of cyclosporine-A. J Am Chem Soc. 1989; 111:4571–4578.

30. Scott D. On optimal and data-based histograms. Biometrika. 1979; 10:605–610.

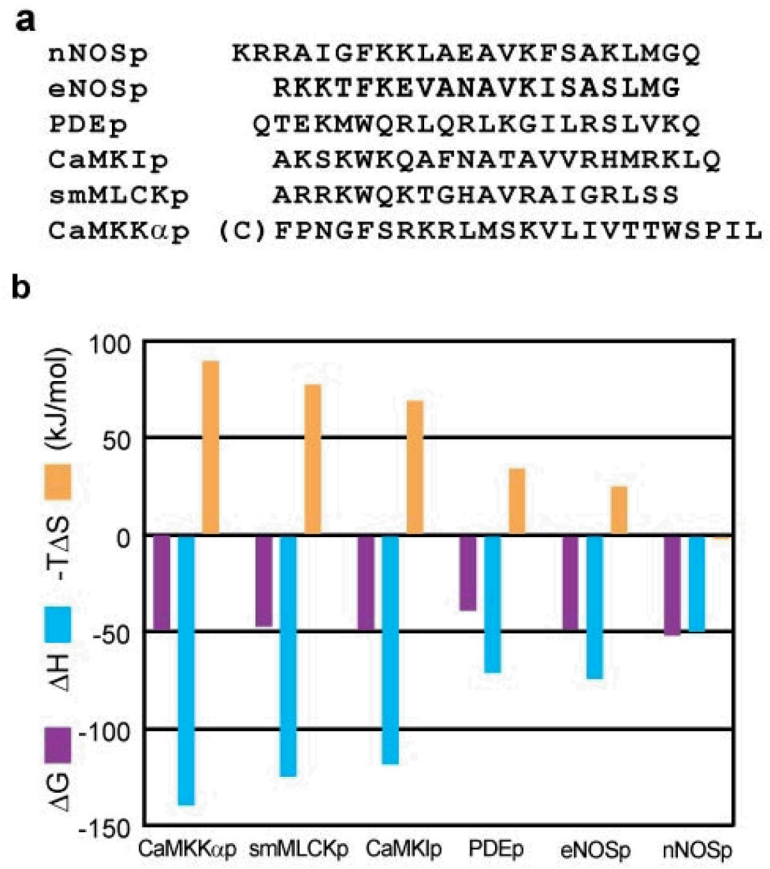

Figure 1. Calmodulin-binding peptides

a, Sequence alignment of the six calmodulin-binding peptides. Note that CaMKKαp binds in opposite orientation to CaM relative to the other five peptides. For PDEp, a C15S mutation has been used to avoid complications with oxidation. **b**, Thermodynamic parameters. The Gibbs free energy (ΔG), enthalpy (ΔH) and entropy (−TΔS) for the formation of the six calcium-saturated CaM-peptide complexes at 35 °C. Values are tabulated in the Supplementary Information.

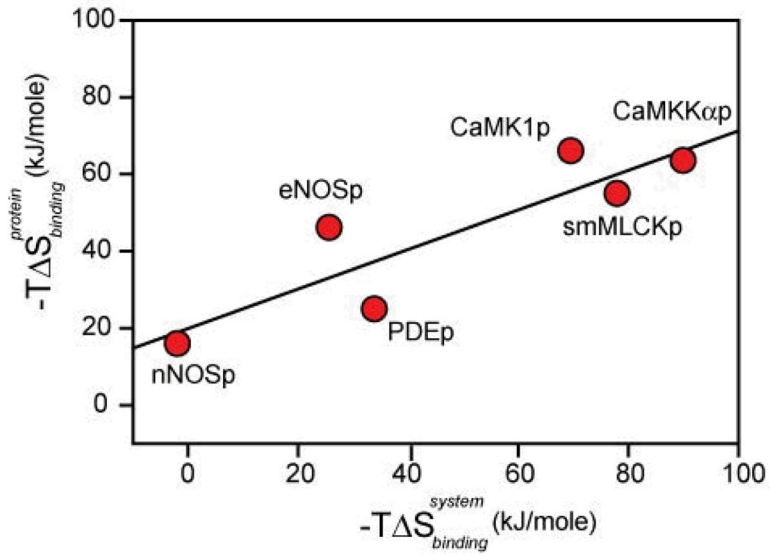

Figure 2. Correlation of the change in residual conformational entropy of calmodulin with the change in the total entropy of binding of a target domain

The change in conformational entropy was estimated using Equation 2 as described in the Methods and Supplementary Information. Propagation of measurement error in fitted order parameters results in uncertainties in conformational entropy less than the size of the symbols used. The fitted linear correlation coefficient (R^2) of conformational entropy versus the entropy of binding is 0.78 with a slope of 0.51.

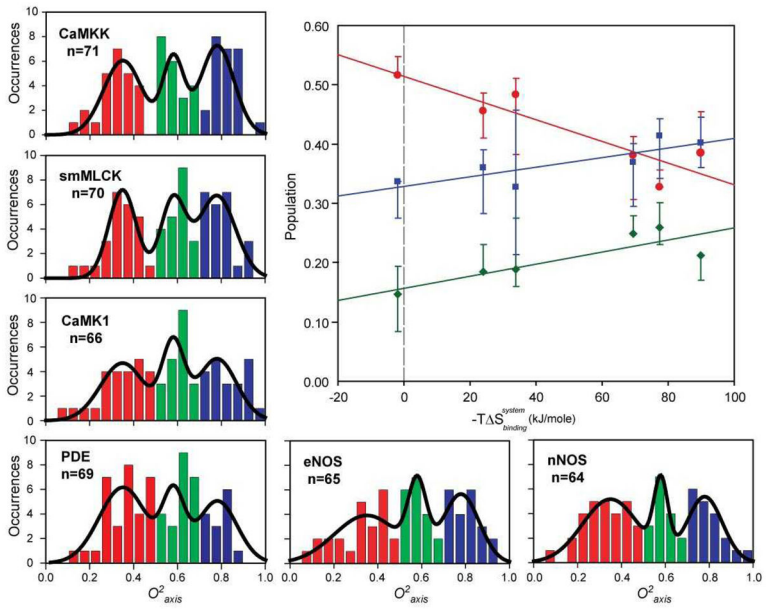

Figure 3. Distribution of the amplitude of methyl-bearing side chain motion of calmodulin in complex with target domains and correlation with the change in total entropy of binding

Shown are histograms of the squared generalized order parameters for calmodulin methyl group symmetry axes obtained at 35 °C. The solid lines represent fitted 3-Gaussian distributions centred on O^2_{axis} values of 0.35 (J-class, red), 0.58 (α-class, green) and 0.78 (ω-class, blue). The relative populations of each class were derived from the fitted 3-Gaussian distributions for each complex. The change in population of the J, α and ω classes with the $-T\Delta S^{system}_{binding}$ have fitted linear correlation coefficients (R^2) of −0.83, +0.74 and +0.70, respectively. Correlation of the number of sites assigned to each class by simple binning, as colour-coded, yielded similar results (see Supplementary Information). Error bars reflect the variation of the population of each motional class that results from an increase or decrease in the measured O^2_{axis} values by two standard deviations (see Figure 2).