Author/s:
Mobbs, JI;Di Paolo, A;Metcalfe, RD;Selig, E;Stapleton, DI;Griffin, MDW;Gooley, PR

Title:
Unravelling the Carbohydrate-Binding Preferences of the Carbohydrate-Binding Modules of AMP-Activated Protein Kinase

Date:
2018-02-02

Citation:
Mobbs, J. I., Di Paolo, A., Metcalfe, R. D., Selig, E., Stapleton, D. I., Griffin, M. D. W. & Gooley, P. R. (2018). Unravelling the Carbohydrate-Binding Preferences of the Carbohydrate-Binding Modules of AMP-Activated Protein Kinase. Chembiochem, 19 (3), pp.229-238. https://doi.org/10.1002/cbic.201700589.

Title: Unravelling the carbohydrate-binding preferences of the carbohydrate-binding modules of AMP-activated protein kinase.

Authors: Jesse Mobbs; Alex di Paolo; Riley Metcalfe; Emily Selig; David Stapleton; Michael Griffin; Paul Gooley

To be cited as: ChemBioChem 10.1002/cbic.201700589

Link to VoR: https://doi.org/10.1002/cbic.201700589

Unravelling the carbohydrate-binding preferences of the carbohydrate-binding modules of AMP-activated protein kinase.

Jesse I. Mobbs[1], Alex di Paolo[2], Riley D. Metcalfe, Emily Selig, David I. Stapleton, Michael D. W. Griffin and Paul R. Gooley[*]

[*]Department of Biochemistry and Molecular Biology, University of Melbourne, Parkville, Victoria 3010, Australia and Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria 3010, Australia.

Current address:

[1]Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia.

[2] New Technologies Development Department, Kaneka Eurogentec S.A. Biologics Division

14 Rue Bois Saint-Jean, 4102 Seraing, Belgium

* Correspondence to:

Paul R. Gooley

The Department of Biochemistry and Molecular Biology and Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Parkville, Victoria 3010, Australia

Ph: +61 3 9034 2273

Email: prg@unimelb.edu.au.

ABSTRACT

The β-subunit of AMP-activated protein kinase (AMPK), which exists as two isoforms (β1 and β2) in humans, has a Carbohydrate Binding Module (CBM) that interacts with glycogen. While the β1- and β2-CBMs are structurally similar, with strictly conserved ligand-contact residues, they show different carbohydrate affinities. β2-CBM shows the strongest affinity for both branched and unbranched oligosaccharides and we have recently shown that a Thr insertion in β2-CBM (Thr101) forms a pocket to accommodate branches. This insertion does not explain why β2-CBM binds all carbohydrate with stronger affinity. Here we show that residue 134 (Val for β2 and Thr for β1), which does not contact carbohydrate, appears to account for the affinity difference. Characterization by NMR, however, suggests that the mutant β2-Thr101Δ/Val134Thr differs from β1-CBM, and the mutant β1-Thr101ins/Thr134Val differs from β2-CBM. Furthermore, these mutants were less stable to chemical denaturation compared to wild type β-CBMs confounding the affinity analyses. To support the importance of Thr101 and Val134 we constructed the ancestral CBM. This CBM retained Thr101 and Val134 suggesting that the extant β1-CBM has a modest loss of function in carbohydrate binding. As the ancestor bound carbohydrate with equal affinity to β2-CBM we conclude that residue 134 does play an indirect role in carbohydrate binding.

Key words carbohydrate binding module (CBM); AMP-activated protein kinase (AMPK); protein thermodynamics; protein-ligand interaction; mutagenesis; ancient sequence reconstruction

INTRODUCTION

AMP-activated protein kinase (AMPK), comprised of three subunits (α-, β- and γ), is an essential, central metabolic enzyme involved in energy regulation of most eukaryotes [1]. Once activated, AMPK up-regulates catabolic pathways, including glucose uptake and fatty acid degradation, and inhibits anabolic pathways, such as protein and fatty acid synthesis [2-5]. Consequently, AMPK has been a target for metabolic diseases such as Type 2 diabetes [6]. The activation of AMPK is well-characterized. Phosphorylation of AMPK occurs by several up-stream kinases, and on a conserved Thr in the kinase domain of the α-subunit [7, 8]. AMPK is further activated allosterically by AMP and ATP binding to the γ-subunit [9]. Regulation by the β-subunit is more subtle, involving N-terminal myristoylation which localizes the enzyme to the membrane, but also may control phosphorylation of the α-subunit [10, 11]. The β-subunit also contains a carbohydrate-binding module (CBM) that may have roles in regulation by binding to glycogen [12-14] but also, through an interface with the α-subunit, provides a site for novel allosteric activation by small molecules [15-17]. Importantly, there are isoforms of each subunit. In humans there are two isoforms of the α-subunit (α1, α2), two of the β-subunit (β1, β2) and three of the γ-subunits (γ1, γ2, γ3). Therefore AMPK in humans can exist as one of 12 isoenzymes, which have been shown to be differentially expressed in tissues suggesting functional differences between these complexes and hence differences in ligand interactions [18].

Indeed, we have shown that the carbohydrate-binding module (CBM) of the β1 and β2 subunits have different affinities for carbohydrate [19] and for small molecules [15] either isolated or in the context of the intact enzyme. The β2-CBM binds all carbohydrates tested with greater affinity than the β1-CBM. The β2-CBM bind carbohydrates with a single α,1→6 branch with a stronger affinity than unbranched carbohydrates, whereas β1-CBM shows no difference. These trends are maintained in full-length protein where the AMPK heterotrimer,

α1β1γ1, bound maltoheptaose and glucosyl-maltoheptaose similarly (K_d of ~650 μM) but the heterotrimer, α1β2γ1, bound maltoheptaose with an affinity of 74 μM and glucosyl-maltoheptaose of 24 μM [20]. The structural reason for this preference by β2-CBM for single α,1→6 branched carbohydrate was explained by determining the crystal structure of the complex which showed that the insertion of Thr101 in β2-CBM forms a pocket that accommodates the branch [21]. The elucidation of the structures of the CBMs, however, does not explain the 4-fold tighter binding of unbranched carbohydrate to β2-CBM compared to β1-CBM, and consequently the aim of the study presented here was to attempt to clarify this difference. Furthermore, a feature that we have noted in the course of our characterization of the CBMs is that the β2-CBM shows microsecond to millisecond timescale structural dynamics, both free and bound to carbohydrate, that is absent in the β1-CBM [20]. Although the data were not definitive, this difference in dynamics may suggest a role in conformational selection for carbohydrate binding [21].

The CBMs are members of the carbohydrate-binding module-containing family 48, adopting a β-sandwich fold with a three-stranded antiparallel β-sheet below a five-stranded antiparallel β-sheet that forms the single carbohydrate binding site [21, 22]. The two CBMs show 81.7% sequence identity, thus differing by 15 residues, including the Thr101 insertion in β2-CBM. Importantly, all carbohydrate contact residues are strictly conserved. Given that the affinity of carbohydrate may involve subtle indirect interactions, we prepared mutants with two distinct approaches. Our first approach was to create individual point mutations that convert β2-CBM residues to their corresponding β1-CBM residue, and then determine if the mutation results in a loss of carbohydrate binding affinity. Mutations were chosen (F91Y, I92L, I103, D111N, V134T and I149V) due to their proximity to the carbohydrate binding site or protein motion observed in the protein dynamics studies and the already observed threonine deletion (ΔT101) was included [21].

The mutagenesis of residues based on differences in extant proteins ignores the evolution of the proteins and may not uncover epistatic mutations that indirectly contribute to binding and function. Two epistatic mutations that can be important are permissive and restrictive [23, 24]. Permissive mutations are those that are required for a function switching mutation to occur. On the other hand, restrictive mutations are those that are incompatible with function switching mutations, and stop evolution from returning to the ancestral state. Hence our second approach to mutating the CBMs was to recreate a likely ancestral protein. This approach, that has recently become possible and called ancestral sequence reconstruction (ASR), creates the ancestral sequences by a statistical method using extant protein sequences [25]. Using ancient sequence reconstruction we aimed to determine the ancestor sequence at which β1- and β2-CBM diverged, characterise the carbohydrate binding affinity of the ancient sequence and determine residues that are important for the differences in carbohydrate binding affinity.

RESULTS

Characterisation of mutants binding affinity via ITC

To determine which residues are important for the tighter carbohydrate affinity observed for β2-CBM, site specific mutations were created that convert a single β2-CBM residue into their corresponding β1-CBM residue. The mutations proposed were β2-F91Y, β2-I92L, β2-I103L, β2-D111N, β2-V134T, β2-I149V and the previously created β2-ΔT101 [19, 20]. To determine binding affinity for these mutants to unbranched β-cyclodextrin (BCD) and branched glucosyl-β-cyclodextrin (gBCD), ITC was employed. Figure 1 and Supplementary Figure 1 displays the dissociation constants (K_d) obtained for each CBM and Table 1 lists all thermodynamic parameters obtained from ITC. Similar to β2-CBM all mutant CBMs bind gBCD more tightly than they do BCD.

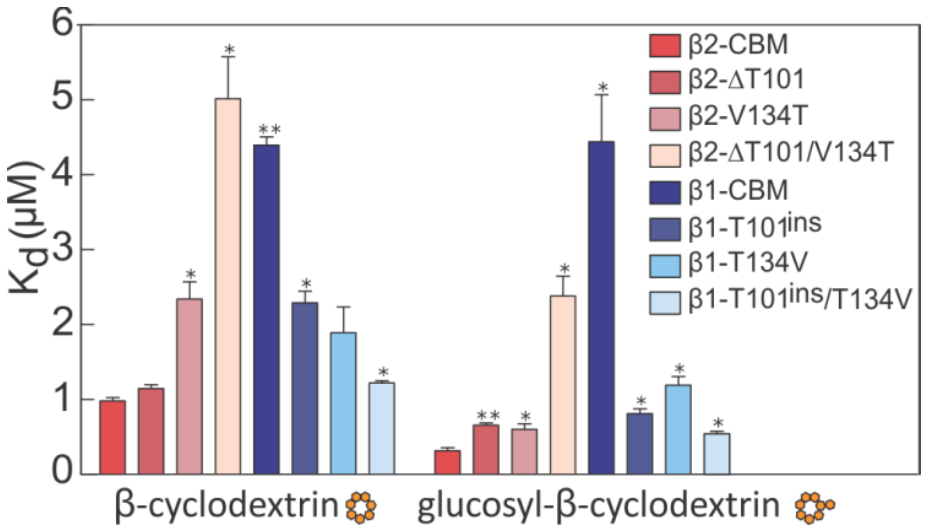

Figure 1. Bar chart of dissociation constants (K_d) of β2-CBM mutants and reversed β1-CBM mutants compared to wild type. β2-WT and mutations are shaded in red colours, while β1-WT and mutants are shaded in blue colours. Results obtained from binding to β-cyclodextrin are on the left and those from glucosyl-β-cyclodextrin are on the right. Values are reported as the mean and SEM of three experiments. Stars indicate significance from a two tailed t-test against β2-WT (* p<0.05, ** p<0.005).

For each of these mutations we compared dissociation constants to WT β2-CBM. If a residue was important for the enhanced affinity of WT β2-CBM, a weakening in affinity should be observed, and a shift towards that observed for WT β1-CBM. As expected, the threonine deletion (β2-ΔThr-101) showed a significant weakening in affinity for gBCD (0.66 ± 0.06 μM) and only a modest difference to BCD (1.15 ± 0.08 μM). On the other hand most other mutants (β2-F91Y, β2-I92L, β2-I103L, β2-D111N and β2-I149V) did not show a significant change in dissociation constant or thermodynamic parameters (Supplementary Figure 1). One mutant (β2-V134T), however did show significant change in dissociation constant and some of the thermodynamic properties. Binding to BCD, β2-V134T had a dissociation constant of 2.34 ± 0.32 μM, more than two-fold weaker than β2-CBM. This reduction in binding affinity was accompanied by a reduction in enthalpy change (ΔH = -6.9

± 0.1 kcal.mol^{-1}), when compared to β2-CBM WT (ΔH = -7.6 ± 0.4 kcal.mol^{-1}). The difference in binding affinity was not as significant when binding gBCD, with a dissociation constant of 0.60 ± 0.13 μM. The enthalpy change for gBCD was also similar to β2-CBM WT (-9.9 ± 0.5 kcal.mol^{-1} and -10.0 ± 0.7 kcal.mol^{-1} respectively).

Table 1. Dissociation constants (K_d) and thermodynamic parameters of wild-type (WT), single and double mutant and ancient sequence reconstructed (AS) CBMs.[a]

Protein	Carbohydrate[b]	K_d (μM)	ΔG (kcal.mol^{-1})	ΔH (kcal.mol^{-1})	TΔS (kcal.mol^{-1})
β1-WT	BCD	4.40 ± 0.19	-7.3 ± 0.1	-5.7 ± 0.3	1.4 ± 0.3
	gBCD	4.40 ± 1.10	-7.3 ± 0.1	-6.5 ± 0.4	0.8 ± 0.5
β2-WT	BCD	0.98 ± 0.07	-8.2 ± 0.1	-7.6 ± 0.4	0.5 ± 0.4
	gBCD	0.32 ± 0.07	-8.9 ± 0.1	-10.0 ± 0.7	-1.1 ± 0.8
β2-F91Y	BCD	1.01 ± 0.14	-8.2 ± 0.1	-7.9 ± 0.8	0.3 ± 0.8
	gBCD	0.21 ± 0.06	-9.1 ± 0.2	-10.2 ± 0.3	-1.1 ± 0.4
β2-I92L	BCD	0.84 ± 0.01	-8.2 ± 0.1	-7.8 ± 0.3	0.5 ± 0.3
	gBCD	0.36 ± 0.16	-8.8 ± 0.3	-10.0 ± 0.3	-1.2 ± 0.5
β2-ΔT101	BCD	1.15 ± 0.08	-8.1 ± 0.1	-8.3 ± 0.7	-0.3 ± 0.6
	gBCD	0.66 ± 0.06	-8.4 ± 0.1	-9.1 ± 0.5	-0.7 ± 0.6
β2-I103L	BCD	1.06 ± 0.16	-8.1 ± 0.1	-8.0 ± 0.1	0.1 ± 0.1
	gBCD	0.20 ± 0.04	-9.1 ± 0.1	-10.3 ± 0.5	-1.2 ± 0.6
β2-D111N	BCD	1.15 ± 0.17	-8.1 ± 0.1	-8.8 ± 1.2	-0.7 ± 1.2
	gBCD	0.29 ± 0.10	-8.9 ± 0.2	-11.3 ± 1.5	-2.33 ± 1.7
β2-V134T	BCD	2.34 ± 0.32	-7.7 ± 0.1	-6.9 ± 0.1	0.5 ± 0.5
	gBCD	0.60 ± 0.13	-8.5 ±0.1	-9.9 ± 0.5	-1.4 ± 0.5
β2-I149V	BCD	1.35 ±0.20	-8.0 ± 0.1	-8.2 ± 0.1	-0.2 ± 0.2
	gBCD	0.29 ± 0.04	-8.9 ± 0.1	-10.8 ± 0.1	-1.9 ± 0.2
β2-ΔT101/V134T	BCD	5.05 ± 0.98	-7.2 ± 0.1	-7.1 ± 0.1	0.1 ± 0.1
	gBCD	2.38 ± 0.46	-7.7 ± 0.1	-7.0 ± 0.2	0.7 ± 0.3
β1-T101^{ins}	BCD	2.29 ± 0.27	-7.7 ± 0.1	-6.7 ± 0.4	1.0 ± 0.4
	gBCD	0.81 ± 0.12	-8.3 ± 0.1	-8.4 ± 0.4	-0.1 ± 0.3
β1-T134V	BCD	1.89 ± 0.60	-7.8 ± 0.2	-7.3 ± 0.2	0.5 ± 0.5
	gBCD	1.19 ± 0.19	-8.1 ± 0.1	-7.4 ± 0.5	0.7 ± 0.6
β1-T101^{ins}/T134V	BCD	1.22 ± 0.06	-8.1 ± 0.1	-5.6 ± 0.3	2.4 ± 0.3
	gBCD	0.54 ± 0.06	-8.5 ± 0.1	-7.9 ± 0.7	0.6 ± 0.6
AS-CBM	BCD	1.20 ± 0.20	-8.1 ± 0.1	-8.2 ± 0.6	-0.5 ± 0.2
	gBCD	0.20 ± 0.02	-9.1 ± 0.1	-13.7 ± 0.8	-4.1 ± 0.8

[a]All values are reported as the mean ± 1 SD.

[b]BCD (β-cyclodextrin), gBCD (glucosyl-β-cyclodextrin).

From these results we decided to make a new suite of mutations to answer two questions: Firstly, is the effect on binding affinity additive, i.e. could we combine the Thr-101 deletion and the V134T substitution to further reduce binding affinity? And secondly, was the V134T substitution causing a specific change in binding affinity? To answer this second question we created the reverse mutations, substituting β1-CBM residues to their corresponding β2-CBM residues expecting strengthening of binding affinity. This suite of mutations therefore consisted of: β2-ΔT101/V134T, β1-T101^{ins}, β1-T134V and β1-T101ins/T134V. ITC was performed for each mutant with the unbranched carbohydrate BCD and the branched carbohydrate gBCD. Thermodynamic parameters and dissociation constants are reported in Table 1 and dissociation constants are plotted in Figure 1.

Insertion of threonine into β1-CBM (β1-Thr101^{ins}) increased binding affinity to both the unbranched BCD and branched gBCD consistent with previous observations [19]. The muation increased the enthalpy change for both BCD and gBCD (ΔH = -6.7 ± 0.4 kcal.mol^{-1} and -8.4 ± 0.4 kcal.mol^{-1} respectively) and decreased entropy for both carbohydrates (TΔS = 1.0 ± 0.4 kcal.mol^{-1} and -0.1 ± 0.3 kcal.mol^{-1}). The reverse of β2-V134T, i.e. β1-T134V, showed a significant increase in binding affinity to both the unbranched BCD (K_d = 1.89 ± 0.60 μM) and branched gBCD (K_d = 1.19 ± 0.19 μM). This mutation increased the enthalpy change to both BCD and gBCD (ΔH = -7.3 ± 0.2 kcal.mol^{-1} and -7.4 ± 0.5 kcal.mol^{-1} respectively) and decreased the entropy change for BCD (TΔS = 0.5 ± 0.5 kcal/mol). These data suggest that the mutations at position 134 may indeed cause a specific change to the binding affinity of the carbohydrates.

Combining the threonine deletion with the valine to threonine substitution (β2-ΔT101/V134T) resulted in an even greater weakening in binding affinity. The double mutant (β2-ΔT101/V134T) had a dissociation constant of 5.05 ± 0.98 μM for BCD, similar to the WT β1-CBM affinity for BCD of 4.4 ± 0.19 μM. For binding to gBCD the combination of

mutations (β2-ΔT101/V134T) also further reduced binding affinity (K_d = 2.38 ± 0.46 μM) compared to the individual single mutations. However, in this case the affinity for gBCD did not match that of WT β1-CBM (K_d = 4.4 ± 1.1 μM). The reverse double mutant (β1-T101^{ins}/T134V) had an increase in binding affinity to both BCD and gBCD (K_d = 1.22 ± 0.06 μM and 0.54 ± 0.06 μM respectively), approaching similar values to that of WT β2-CBM (K_d = 0.98 ± 0.07 μM and 0.32 ± 0.07 μM). The most notable changes in thermodynamics occurred for β2-ΔT101/V134T binding to gBCD, where enthalpy reduced (ΔH = -7.0 ± 0.2 kcal.mol^{-1}) and the entropy became favourable (TΔS = 0.7 ± 0.3 kcal.mol^{-1}), very similar to the WT β1-CBM (ΔH = -6.5 ± 0.4 kcal/mol and TΔS = 0.8 ± 0.5 kcal.mol^{-1}). This fits with the α,1→6 branch of gBCD being unable to fit in a defined manner, that is into a pocket, compared to WT β2-CBM [21].

Chemical shift difference of β2-ΔT101/V134T mutation

Previously we had hypothesised that differences in the molecular motion of the two CBMs may be related to their differences in affinities [20, 21]. Evidence of the role of structural dynamics included the observation that the addition of gBCD caused significant changes in chemical shift and line broadening of Trp133 in ^{15}N,^{1}H-HSQC NMR spectra of WT β2-CBM and not in WT β1-CBM [20]. We therefore performed similar ^{15}N,^{1}H-HSQC experiments of ^{15}N-labelled β2-Δ101/V134T in the free and gBCD bound states (Figure 2). As the double mutant (β2-ΔT101/V134) had similar binding affinities as β1-CBM, we expected upon carbohydrate binding it should show similar tryptophan chemical shift differences and line broadening. Although these mutations caused a decrease in chemical shift difference similar to β1-CBM, it retained the severe line broadening of the parent WT β2-CBM. These results indicate that although we have created a mutant with similar carbohydrate binding affinities to β1-CBM, the mutations have created a protein that is neither β1-CBM nor β2-CBM.

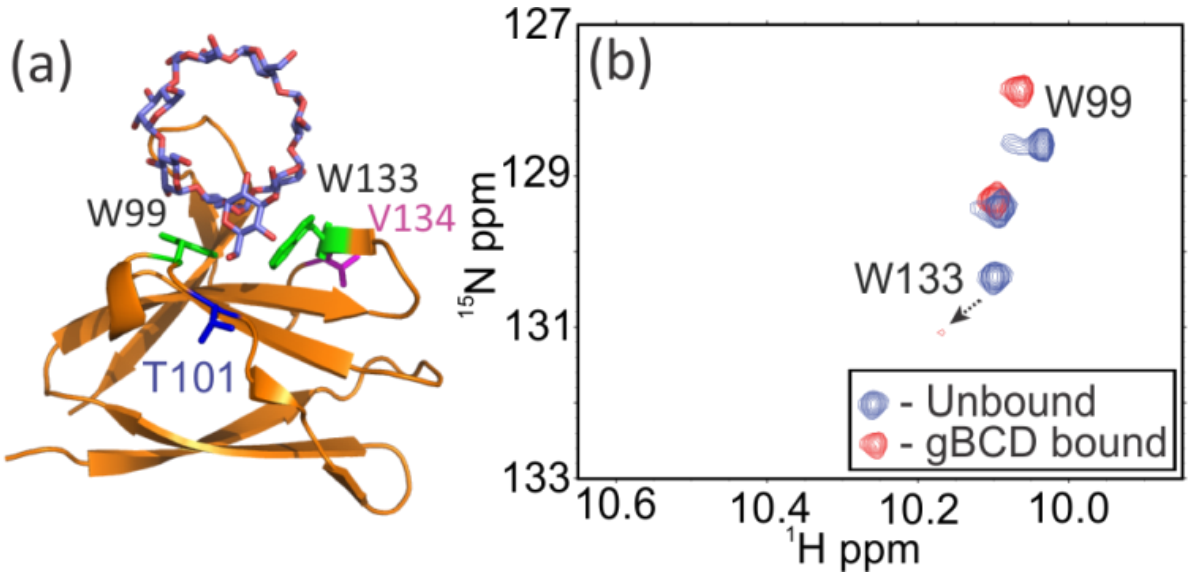

Figure 2. ^{15}N,^{1}H HSQC spectra of β2-ΔT101/V134T in the region of of the Trp side chain NH. (a) crystal structure of β2-CBM (pdb: 4YEE) in complex with gBCD [21]. The conserved Trp residues are highlighed in green and the two mutated residues T101 and V134 are shown in blue and pink respectively. (b) overlayed HSQC spectra of β2-ΔT101/V134T in the free and gBCD bound states. Trp133 has a small chemical shift difference but a large amount of line broadening.

Chemical stability of WT β1- and β2-CBM and mutants

To gain a global picture of the similarities of the WT β1- and β2-CBM, and the mutants β2-ΔT101, β2-V134T, β2-ΔT101/V134T, β1-T101^{ins}, β1-T134V and β1-T101^{ins}/T134V, chemical denaturation assays were performed. These experiments were performed for the unbound proteins and the gBCD-bound protein, from here on referred to as apo- and bound-state respectively. These experiments were performed by increasing the amount of denaturant, guanidine hydrochloride, and monitoring the maximum intrinsic fluorescence emission wavelength (λ_{max}) arising from Trp83, 99 and 133. This maximal intensity varied from ~342 nm for the apo-state to ~325 nm for the bound-states, possibly reflecting the change in hydrophobic environment for Trp99 and Trp133, which interact with the carbohydrate through carbohydrate/aromatic ring stacking [26].

The resulting data (Figure 3) were fitted to obtain the apparent Gibbs free energy of unfolding in the absence of denaturant ($\Delta G°(H_2O)_{N-U}$), the slope of the transition (m_{N-U}) and the denaturant midpoint (C_m) (Table 2). The stability curves of the apo-state of WT β1- and

Table 2 - Apparent thermodynamic parameter values $\Delta G°(H_2O)_{N-U}$, m_{N-U} and C_m for wild-type (WT), single and double mutants and ancient sequence reconstructed (AS) CBMs.[a]

	Apo			Bound		
	$\Delta G°(H_2O)_{N-U}$ (kcal.mol^{-1})	$-m_{N-U}$ (kcal.mol^{-1}M^{-1})	Cm [GdmCl M]	$\Delta G°(H_2O)_{N-U}$ (kcal.mol^{-1})	$- m_{N-U}$ (kcal.mol^{-1}M^{-1})	Cm [GdmCl M]
β1-CBM WT	4.17 ± 0.47	2.92 ± 0.27	1.42 ± 0.03	6.35 ± 0.27	3.73 ± 0.13	1.70 ± 0.04
β2-CBM WT	3.84 ± 0.12	2.72 ± 0.10	1.41 ± 0.02	8.52 ± 0.10	3.83 ± 0.07	2.22 ± 0.01
β2-ΔT101	5.32 ± 2.10	4.77 ± 1.74	1.06 ± 0.11	5.10 ± 0.28	2.88 ± 0.25	1.77 ± 0.08
β2-V134T	3.13 ± 0.26	2.42 ± 0.18	1.29 ± 0.02	7.66 ± 0.31	3.51 ± 0.09	2.17 ± 0.03
β2-ΔT101/V134T	2.56 ± 0.33	2.57 ± 0.29	0.99 ± 0.01	5.39 ± 0.29	3.61 ± 0.17	1.49 ± 0.01
β1-T101^{ins}	4.00 ± 0.10	3.41 ± 0.01	1.17 ± 0.02	6.77 ± 0.42	3.73 ± 0.20	1.81 ± 0.02
β1-T134V	3.24 ± 0.45	2.57 ± 0.27	1.24 ± 0.04	5.88 ± 0.30	3.13 ± 0.12	1.87 ± 0.03
β1-T101^{ins}/T134V	2.74 ± 0.26	2.26 ± 0.27	1.22 ± 0.04	5.06 ± 0.47	3.09 ± 0.19	1.63 ± 0.05
AS-CBM	5.08 ± 0.28	3.08 ± 0.18	1.65 ± 0.01	9.99 ± 0.40	3.61 ± 0.13	2.76 ± 0.01

[a]Values are reported as the mean ± SEM of three experiments.

β2-CBM overlap, indicating that they have similar apo-state stability and indeed give similar thermodynamic parameters (Table 2). In the bound-state there are clear differences in the stability of wild-type β1- and β2-CBM. Both CBMs are more stable in the bound-state compared to the apo state, where β2-CBM is more stable than β1-CBM. Bound β2-CBM has a $\Delta G°(H_2O)_{N-U}$ of 8.52 ± 0.10 kcal.mol^{-1}, while bound β1-CBM has a lower value of 6.35 ± 0.27 kcal.mol^{-1}. The higher stability of bound β2-CBM is expected since it has a stronger affinity for carbohydrate than β1-CBM, as observed by ITC, and additional interactions in the

form of hydrogen bonds and hydrophobic contacts observed in the X-ray crystal structures [21].

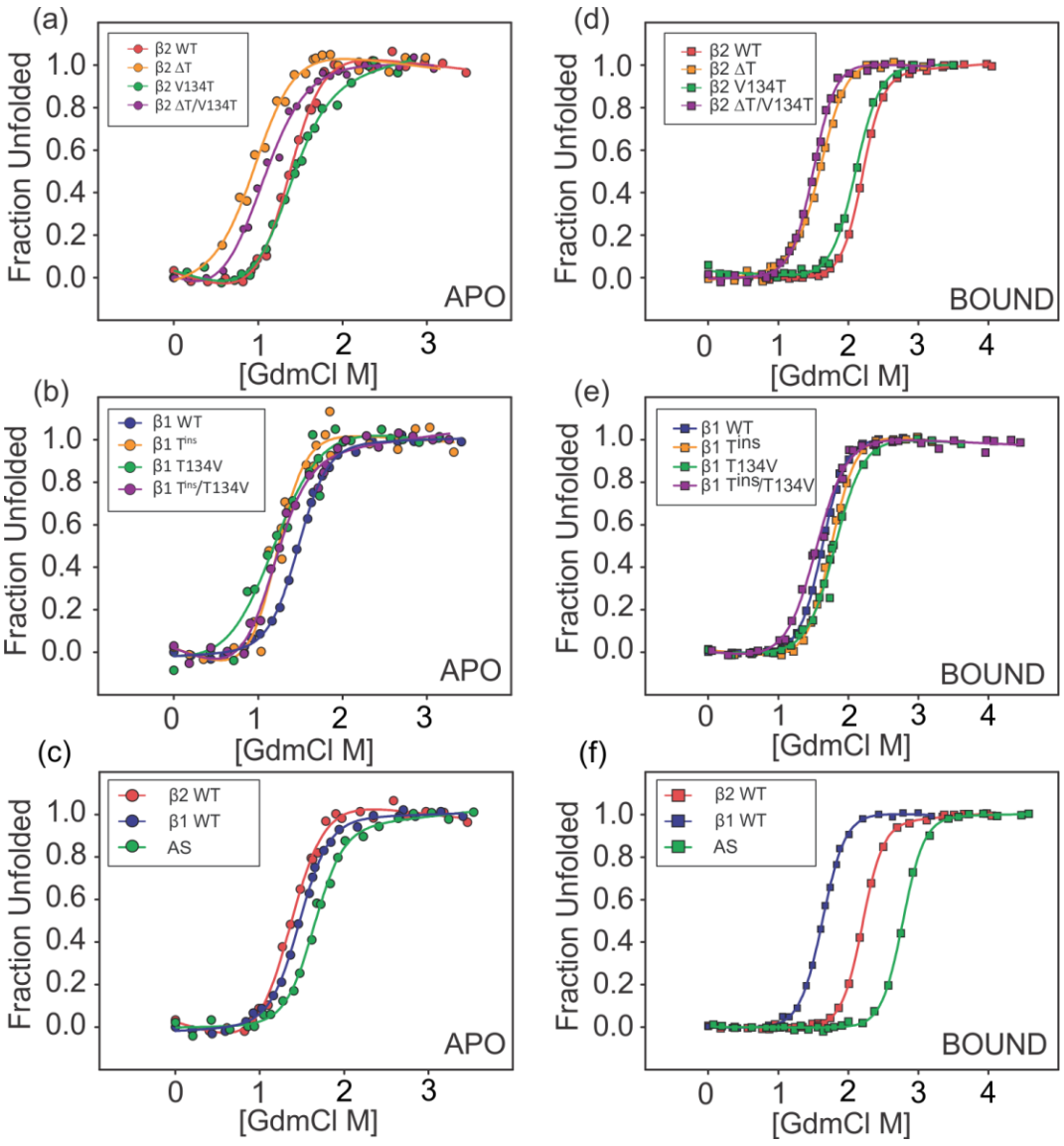

Figure 3. Chemical stability of CBM site specific mutants and ancient sequence reconstructed CBM (AS-CBM). Normalised unfolding curves as a function of guanidine hydrochloride (GdmCL) concentration and fitted to eq. 1. (a) Stability curves of β2-CBM WT and mutants in the apo-state. (b) β1-CBM WT and mutants in the apo-state. (c) β1-, β2-CBM WT and AS-CBM in the apo-state (d) β2-CBM WT and mutants in the glucosyl-β-cyclodextrin bound state. (e) β1-CBM WT and mutants in the bound state. (f) β1-, β2-CBM WT and AS-CBM in the bound-state.

As WT β2-CBM was much more stable than β1-CBM in the bound state, we expected mutants of β2-CBM to have a decrease in stability and mutants of β1-CBM to have an increase in stability when gBCD bound. Inserting the threonine into β1-CBM (β1-T101^{ins}) resulted in small changes to stability of the apo-state, indicating this mutation is well tolerated (Table 2). On the other hand, deletion of the threonine in β2-CBM (β2-ΔT101) gave denaturation curves that fit poorly to a two-state unfolding mechanism, suggesting that significant intermediates exist for this mutant. In the bound-state, β1-T101^{ins} is more stable ($\Delta G°(H_2O)_{N-U}$ = 6.77 ± 0.42 kcal.mol^{-1}) than wild type β1-CBM, but not significantly, but less than wild type β2-CBM). Bound β2-ΔT101 shows a significantly lower $\Delta G°(H_2O)_{N-U}$ (5.1 ± 0.28 kcal.mol^{-1}) than both wild types. While the fit to the denaturation curve was better for the bound state of β2-ΔT101 the parameter m_{N-U} (2.88 ± 0.25 kcal.mol^{-1}.M^{-1}) was much lower compared to the wild type proteins (3.73 ± 0.13 and 3.82 ± 0.07 kcal.mol^{-1}.M^{-1} for β1-CBM and β2-CBM respectively). Deviation from a two-state unfolding mechanism is expected to lower the m_{N-U} value [27], suggesting that this mutant may not obey a two-state mechanism in both the apo and bound states.

Residue 134, which we have shown is important for determining the affinity of the CBMs for carbohydrate, is exposed on the surface of the module but pointing away from the ligand binding site. For both mutations, β1-T134V and β2-V134T, the apparent thermodynamic parameters of unfolding trend to indicate decreased stability in the apo-state, although not significantly (Table 2). This pattern is maintained for the bound states of both these mutants. The combined double mutants, β1-T101^{ins}/T134V and β2-ΔT101/V134T, in the apo state show significantly lower $\Delta G°(H_2O)_{N-U}$ (2.74 ± 0.26 and 2.56 ± 0.33 kcal.mol^{-1} respectively) compared to the wild type proteins. The parameter m_{N-U} for β1-T101^{ins}/T134V is significantly lower (2.26 ± 0.27 kcal.mol^{-1}.M^{-1}) suggesting that this mutant is not consistent with a two-

state mechanism. On the other hand the apparent thermodynamics parameters for β2-ΔT101/V134T fit reasonably to a two-state mechanism. In the bound states the denaturation curves of both mutants fit well to a two-state mechanism, but $\Delta G°(H_2O)_{N\text{-}U}$ is significantly lower than the wild type modules (5.06 ± 0.47 kcal.mol^{-1} for β1-T101^{ins}/T134V and 2.56 ± 0.33 kcal.mol^{-1} for β2-ΔT101/V134T).

Ancient sequence reconstruction

While our single- and double-site mutants suggested position 134 played a significant indirect role in carbohydrate binding, the additional NMR and stability assays shows the mutants are dynamically distinct from the WT CBMs. For these reasons a vertical approach was made to study the mutations that evolved the two modern CBMs. To discover these mutations we decided to determine the sequence of the ancestral protein at the point of their divergence (AS-CBM). To do this, protein sequences of the β-subunit were obtained that had been used previously in a bioinformatics study [19] and a multiple sequence alignment was performed using ClustalW2 of the EMBL-EBI [28]. The sequence alignment was then used to create a phylogenetic tree using PhyML [29] (Supplementary Figure 3). The phylogenetic tree has three clear groups: the plants and fungi, the lancelets (small invertebrate marine-animal [30]) and the vertebrates. It is also clear that the β1- and β2-isoforms have evolved early in the vertebrate branch. Notably plants and fungi also have two isoforms of the β-subunit, yet these have occurred via a separate evolutionary event.

To determine the ancestral sequence of the β1- and β2-isoforms, ancient sequence reconstruction was performed using the FASTML webserver [31]. This method takes the multiple sequence alignment and phylogenetic tree and reconstructs the most likely sequence at each node using a maximum likelihood method. We first tested three different models of residue substitution (JTT, WAG and Dayhoff) [32-34] all giving the same ancient sequence,

therefore we continued with the default method (JTT). The ancient sequence was reconstructed with very high posterior probability across the whole sequence (Supplementary Figure 4), indicating a highly-likely sequence. Comparing the ancient sequence to β1- and β2-CBM we see that out of 15 residues that differ between β1- and β2-CBM the ancient sequence has 10 residues that match the β2-CBM sequence, suggesting the ancient sequence is more similar to β2-CBM than β1-CBM. The ancient sequence has both the threonine insertion (Thr101) and the valine (Val134) that modified carbohydrate binding affinity of β2-CBM upon mutagenesis. This analysis of the ancestral sequence suggests that before the divergence of the β1- and β2-CBM, the ancestor may have been a tight carbohydrate binder ,like β2-CBM, and that a gene duplication event has resulted in the evolution of a muscle-specific tight carbohydrate-binding β2-CBM and a more ubiquitous weak binding β1-CBM. Indeed, inspection of the multiple sequence alignment reveals the lancelet, which is often used to study the evolution of vertebrates [30], also has Thr101 and Val134 supporting these residues of the ancient sequence.

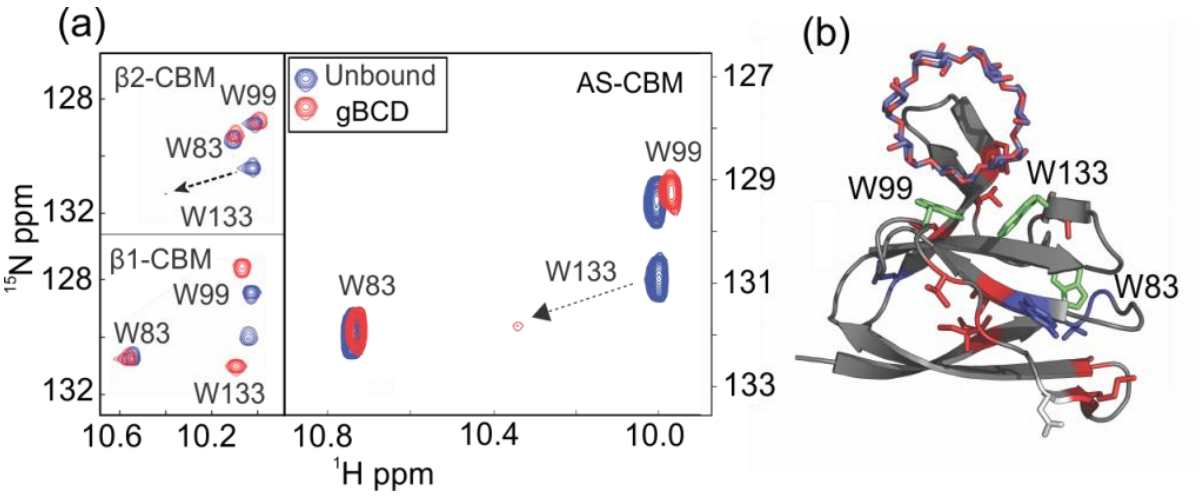

Figure 4. ^{15}N,^{1}H HSQC of AS-CBM in the region of the Trp side chain NH. (a) Titration of AS-CBM with glucosyl-β-cyclodextrin resulted in similar chemical shift and line broadening as β2-CBM for W133 and W99. W83 of AS-CBM has similar chemical shifts to β1-CBM and shows no change upon ligand binding. (b) Model of AS-CBM with tryptophan

residues highlighted in green. W99 and W133 are surrounded by β2-CBM residues (red) while W83 is surrounded by β1-CBM residues (blue).

Thermodynamics of Carbohydrate binding to AS-CBM

The AS-CBM was expressed and purified similar to the wild type proteins. ITC was performed to determine the carbohydrate binding affinity of AS-CBM and also to study the thermodynamics of the interactions. To compare to β1- and β2-CBM, binding affinities and thermodynamic parameters were obtained for the binding of BCD and gBCD (Table 1). AS-CBM has similar binding affinities for both carbohydrates as β2-CBM. For BCD the dissociation constant was 1.2 ± 0.2 μM, similar to the β2-CBM (0.98 ± 0.07 μM) and for gBCD the K_d was 0.20 ± 0.02 μM, approximately 60% tighter than β2-CBM (0.32 ± 0.07 μM). The Gibbs free energy (ΔG) of both interactions was similar to β2-CBM, representing the similar binding affinities. The change in enthalpy (ΔH) was more favourable in both cases (-8.2 ± 0.6 kcal.mol^{-1} and -13.7 ± 0.8 kcal.mol^{-1} for BCD and gBCD respectively), which was counteracted in both cases by a more negative contribution from entropy (TΔS) (-0.5 ± 0.2 kcal.mol^{-1} and -4.1 ± 0.8 kcal.mol^{-1} for BCD and gBCD respectively). These values agree with the sequence analysis suggesting that the modern CBMs diverged from an ancient β2-CBM-like protein, which could bind carbohydrate tightly.

Temperature dependence of the AS-CBM

To explore the temperature dependencies of carbohydrate binding to AS-CBM, ITC was performed at a series of temperatures (283.15, 288.15, 293.15, 298.15 and 303.15 K) for BCD and gBCD. The resulting thermodynamic parameters were fitted to a series of equations (Eq 2 to 4) to obtain the change in heat capacity (ΔCp) of each interaction (Supplementary Figure 5). Binding to the branched carbohydrate (gBCD) there was a slightly larger negative change in heat capacity (ΔCp = -130 ± 32 cal.mol^{-1}.K^{-1}) than binding to the unbranched

carbohydrate (BCD) ($\Delta Cp = -76 \pm 21$ cal.mol^{-1}.K^{-1}). However these are small negative changes in heat capacity and are similar to those obtained for WT β1- and β2-CBM [21] which ranged from -92 to -106 cal.mol^{-1}.K^{-1}. Just like the WT CBMs these small negative values are typical of carbohydrate binding and represent a balance of polar and hydrophobic interactions [35].

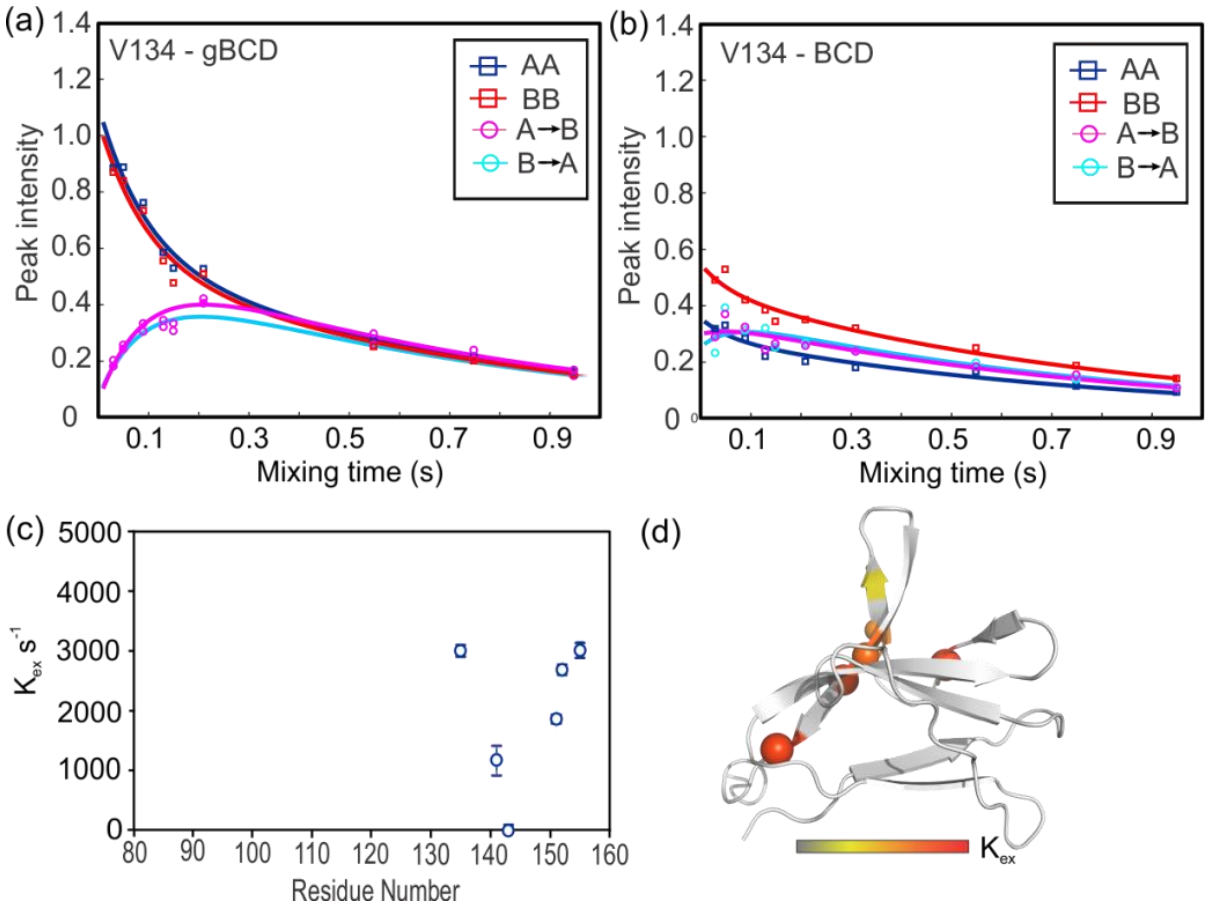

Figure 5. AS-CBM carbohydrate binding rates determined by ZZ-exchange and microsecond motion determined by CPMG relaxation dispersion. (a) Binding to branched gBCD and (b) unbranched BCD. ZZ-exchange experiments were performed at 10°C. Blue and red squares represent the apo and bound peaks, and the magenta and cyan circles represent the exchange peaks between apo and bound states. (c) AS-CBM showed microsecond motion in the region of the carbohydrate binding loop (shown in d) with similar K_{ex} values to WT β2-CBM.

Chemical stability of AS-CBM

To determine if the AS-CBM had similar stability to the WT β1- or β2-CBM, chemical denaturation studies were performed by monitoring the change in tryptophan fluorescence in the presence of increasing concentrations of guanidine hydrochloride. In comparison to the WT β1- and β2-CBM, the apo-state and gBCD bound-states of the AS-CBM should similar quality of fits to the denaturation curves, whereby the apparent thermodynamic parameters are consistent with a two-state mechanism of unfolding. In the apo-state AS-CBM had a $\Delta G°(H_2O)_{N-U}$ of 5.08 ± 0.28 kcal.mol^{-1}, indicating a small increase in stability compared to both WT CBMs (Figure 3 and Table 2). In the gBCD bound-state the $\Delta G°(H_2O)_{N-U}$ was 9.99 ± 0.40 kcal.mol^{-1}, which indicates a marked increase in stability for AS-CBM compared to both β1- and β2-CBM.

NMR properties, binding rates and microsecond motion of AS-CBM

^{15}N,^{1}H-HSQC spectra were recorded of AS-CBM in the apo and gBCD bound states. Viewing the overlayed spectra in the region of the tryptophan side chain NH groups we could compare the changes in chemical shift difference upon binding to what was observed for WT β1- and β2-CBM (Figure 4). WT β1- and β2-CBM show differences in chemical shifts and line broadening in this region which led us to the hypothesis that they have differences in protein dynamics [20]. For Trp99 and Trp133, AS-CBM has similar chemical shifts and line broadening to β2-CBM. These results concur since AS-CBM has a higher sequence similarity and similar binding affinity to β2-CBM. On inspection of the modelled structure of AS-CBM these tryptophan residues are indeed surrounded by β2-CBM residues (Supplementary Figure 4). Trp83 on the other hand has similar chemical shifts to β1-CBM which does not change upon carbohydrate binding. Inspection of the model reveals that Trp83 is surrounded by

multiple β1-CBM residues. These residues (Thr84 and Gly85) may be creating a similar chemical environment for Trp-83 in β1-CBM.

To determine the rates of carbohydrate binding of the AS-CBM, such that they could be compared to WT β1- and β2-CBM rates, ZZ-exchange NMR spectroscopy was performed at 10 °C in the presence of carbohydrate (BCD or gBCD) at an amount such that approximately 50% of the AS-CBM population was in the bound state. Volumes of the apo-state and bound-state peaks, and their corresponding exchange peaks were fitted (Figure 5) to a series of equations [36] to determine the off-rates (k_{off}) of carbohydrate binding (Table 3). These off-rates could then be used in conjunction with the ITC derived K_d to determine the association rates (k_{on}) of carbohydrate binding [21].

Table 3. Binding rates of wild-type and ancient sequence reconstructed (AS) CBMs to BCD and gBCD at 10 °C.

	BCD[a]			gBCD[a]		
	K_d (μM)	k_{off} (s^{-1})	k_{on} ($x10^7$ M^{-1} s^{-1})	K_d (μM)	k_{off} (s^{-1})	k_{on} ($x10^7$ M^{-1} s^{-1})
β2-CBM	0.88 ± 0.10	35.3 ± 11.9	4.0 ± 2.7	0.10 ± 0.06	3.2 ± 1.2	3.2 ± 1.4
β1-CBM	2.9 ± 0.50	42.8 ± 62.5	1.5 ± 2.2	2.16 ± 0.45	29.4 ± 26.3	1.4 ± 1.3
AS-CBM	0.73 ± 0.10	34.3 ± 27.1	4.7 ± 3.8	0.24 ± 0.04	3.2 ± 0.6	1.4 ± 0.4

[a] BCD (β-cyclodextrin), gBCD (glucosyl-β-cyclodextrin).

The off-rate of AS-CBM (k_{off} = 3.2 ± 0.6 s^{-1}) (Table 3) for branched carbohydrate (gBCD) was near identical to WT β2-CBM, but with a slower on-rate (k_{on} = 1.4 ± 0.4 $x10^7$ M^{-1} s^{-1} for AS-CBM), however the large errors on the β2-CBM values make it difficult to differentiate and would most likely suggest similar modes of binding. Like WT β2-CBM, k_{off} for gBCD was approximately 10-fold lower than β1-CBM. Binding to the unbranched carbohydrate (BCD), AS-CBM again has similar values to WT β2-CBM, with a k_{off} of 34.3 ±

27.1 s^{-1} and a k_{on} of 4.7 ± 3.8 x10^7 M^{-1} s^{-1}. The k_{on} for AS-CBM is faster than both β2-CBM and β1-CBM, however as the errors are quite high any differences are difficult to interpret.

As the AS-CBM had even greater stability than β2-CBM, we wanted to determine if AS-CBM had retained protein motion on the microsecond to millisecond timescale [20]. To do this we performed a ^{15}N CPMG relaxation dispersion experiment for AS-CBM in the apo-state. The AS-CBM showed microsecond motion in the carbohydrate binding β-hairpin (Figure 5) similar to WT β2-CBM whereas WT β1-CBM did not. The motion observed for AS-CBM was approximately 2-fold slower than WT β2-CBM (AS-BM k_{ex} = 2174 ± 94 s^{-1} and WT β2-CBM k_{ex} = 4936 ± 351 s^{-1}, 50% trimmed mean).

DISCUSSION

To probe for important residues of carbohydrate binding affinity, mutants in β2-CBM were created that converted individual residues to their corresponding β1-CBM residue, and similarly the reverse mutations were created into the background of β1-CBM. These mutants were chosen due to their proximity to the binding site or the binding-loop where they showed protein motion in β2-CBM [20, 21]. As expected, the threonine deletion (β2-ΔThr101) caused a significant decrease in affinity for the branched carbohydrate gBCD and only a modest decrease in affinity for the unbranched carbohydrate BCD. Whereas the valine to threonine mutation at position 134 (β2-V134T) caused a significant reduction in affinity for both carbohydrates. Insertion of Thr101 into β1-CBM and the mutation T134V caused a significant increase in affinity supporting the importance of the residue at position 134. These mutations (β2-ΔT101/V134T and β1-T101^{ins}/T134V) were indeed additive and had binding affinities that were close to the opposite isoforms for both carbohydrates. The residue in position 134 does not directly contact carbohydrate, but inspection of the available structures of β1-CBM (pdb: 4yef, 1z0m) suggests that the side-chain OH of Thr134 of β1-CBM can

form a hydrogen bond to the backbone carbonyl of residue 133 (Figure 6). This hydrogen bond may constrain the protein, reducing the dynamic freedom of Trp133 and restrict conformations that allow optimal carbohydrate contact, whereas no such hydrogen bond is formed for Val134 of β2-CBM.

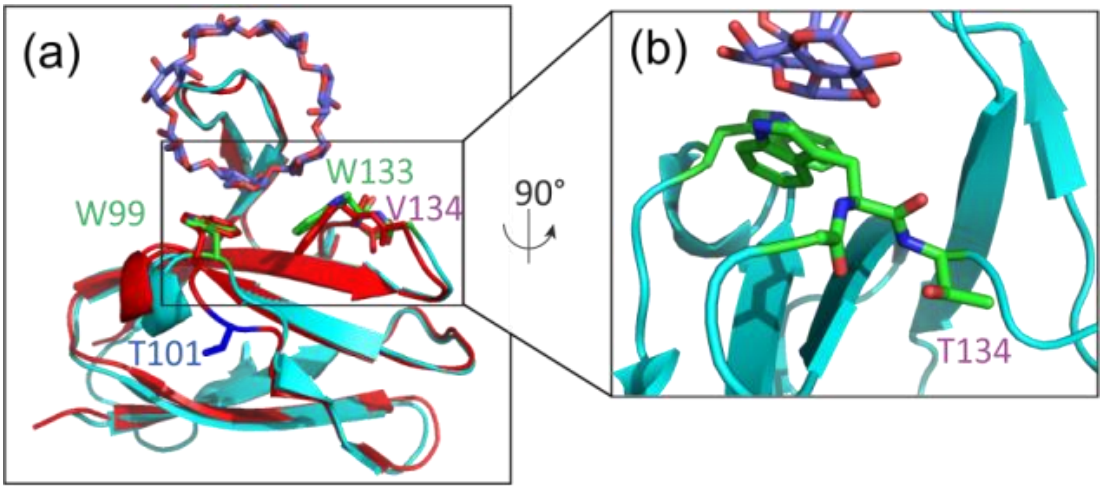

Figure 6 - Hydrogen bond of residue 134. (a) alignment of β1-CBM (cyan) and β2-CBM (red) with the residue at position 134 highlighted in pink. (b) zoom into β1-T134, which can form a hydrogen bond to the backbone carbonyl oxygen of residue 132, the residue adjacent to the carbohydrate contact residue W133.

Superficially, residues at positions 101 and 134 appear to distinguish β1- from β2-CBM. As the $^{15}N,^{1}H$ HSQC NMR spectra of β2-ΔThr101 mutant had reduced line broadening and chemical shift of Trp-133 upon carbohydrate binding [20], we expected the (β2-ΔT101/V134T) double mutant to have even less line broadening and chemical shift upon carbohydrate binding, and thus be similar to β1-CBM. Although it showed similar $^{1}H,^{15}N$ chemical shifts of Trp133 to β1-CBM, it displayed line broadening similar to β2-CBM, thus behaving in some hybrid manner. Comparison of the unfolding of the CBM mutants in chemical denaturant, either in the apo or bound states, showed all mutations, to some degree, caused a decrease in stability compared to the respective WT β-CBM. Notably, insertion of

threonine into β1-CBM (β1-T101^{ins}) had an insignificant effect on stability of the apo-state, whereas the deletion of threonine from β2-CBM (β2-ΔT101) resulted in a mutant that does not appear to unfold by a two-state mechanism. This is a common problem of the horizontal approach to understanding protein function by mutagenesis, where most mutations are out of context to the evolution of the protein, causing disruption of stabilising intramolecular interactions within the protein, and consequently decreasing protein stability [23, 37, 38]. In the bound state some of the mutations changed stability in the expected direction, for example the β2-ΔT101/V134T mutant, that had decreased binding affinity similar to β1-CBM, also had decreased stability in the bound state, but was significantly less stable than bound WT β1-CBM. These results raise a concern that although we could create mutants that mimic the opposite β-CBM isoform in regards to carbohydrate binding affinity, we have concomitantly created hybrid β-CBMs that have some properties dissimilar to both β-CBM isoforms. These results do not necessarily suggest that the residue at position 134 is not important for β-CBM binding affinity, but may require other complementary mutations.

To further investigate the differences of the two CBMs we recreated the sequence (AS-CBM) at which β1- and β2-CBMs diverged. The carbohydrate binding properties of AS-CBM were found to be similar to β2-CBM by ITC. Protein unfolding experiments revealed that the AS-CBM is moderately more stable in the apo state but had a significant increase in stability in the gBCD-bound state. Additional biophysical properties demonstrate the similarity of AS-CBM to β2-CBM. $^{15}N,^{1}H$ HSQC spectra of the AS-CBM revealed the side chain NH of the tryptophans in contact with carbohydrate have similar structural properties to β2-CBM. In both CBMs (AS- and β2-CBM) the chemical shift of Trp99 NH has a modest change, whereas Trp133 has a large chemical shift change and severe line broadening upon carbohydrate binding. ^{15}N CPMG relaxation dispersion experiments show that the AS-CBM has motion on the microsecond timescale similar to β2-CBM. While the rate of observed

motion was two-fold slower for AS-CBM, it is of interest that the increase in protein stability does not negate this motion.

The rates of carbohydrate binding were determined for AS-CBM by a combination of ZZ-exchange NMR spectroscopy experiments and ITC. On binding to both branched (gBCD) and unbranched carbohydrates (BCD) AS-CBM and WT β2-CBM showed similar k_{off} rates compared to WT β1-CBM. This suggests that for AS-CBM the α,1→6 branch binds into the pocket between the two tryptophan residues generated by Thr101 in the same way it does for β2-CBM. WT β1-CBM has very similar rates for both unbranched and branched carbohydrates, leading to the conclusion (along with other evidence) that it does not bind the α,1→6 branch with any specificity. Therefore, it is not reasonable to compare the rates of branched carbohydrate binding between β1-CBM and AS-CBM due to the different binding modes. On the other hand it is much more reasonable to compare binding rates for all three CBMs (β1-, β2- and AS-CBM) for unbranched carbohydrate (BCD). AS-CBM and β2-CBM both have similar k_{on} (4.7 ± 3.8 and $4.0 \pm 2.7 \times 10^7$ M^{-1} s^{-1}, respectively) and similar k_{off} (34.3 ± 27.1 and 35.3 ± 11.9, respectively) for BCD. The k_{off} for both CBMs (AS-CBM and β2-CBM) for unbranched carbohydrate is approximately ten-fold higher compared to branched carbohydrate, agreeing with the fact that the pocket formed by the insertion of Thr101 is important for mediating the added specificity for α,1→6 branched carbohydrate. Both β2-CBM and AS-CBM have on-rates that are more than two-fold faster than β1-CBM binding to BCD and although not statistically confirmed, the trend suggests that it may be the on-rate that mediates the differences in carbohydrate binding between β1- and β2-CBM. Potentially the μs-ms motion of the AS-CBM and β2-CBM play a role in this increased on-rate.

The biophysical properties of AS-CBM show that it behaves efficiently in carbohydrate binding. Importantly, the increase in stability of AS-CBM is expected as it positions the ancestral protein for further mutation, permissive and restrictive, for further modification of

function without loss of structure. The reconstruction process revealed that AS-CBM shared more similarities with residues that define the β2-CBM. In particular, the ancient sequence has the Thr101 insertion that gives β2-CBM its enhanced affinity for branched carbohydrates and has the Val-134 that we have found modifies CBM binding affinity. Analysis of the other extant sequences revealed that the lancelet, an organism often used to study the evolution of vertebrates [30], has both Thr101 and Val134. Further analysis also reveals that yeast and plants also have a residue at position 101, although only one of which has a threonine at this position. These results suggest that the ancient sequence of the β-CBMs was a "β2-like" CBM and that the modern day β1-CBM has evolved via a loss-of-function mutation. As β1-CBM binds carbohydrate with weaker affinity than β2-CBM [20] and has a more ubiquitous tissue distribution [39], β1-CBM may have evolved such that there is a more generalised CBM in tissues where tight carbohydrate binding is disadvantageous or not important. Indeed, it has been recently shown that autophosphorylation of Thr148 of the carbohydrate binding site significantly modulates glycogen binding suggesting molecular mechanisms to regulate AMPK localization and glycogen metabolism [13]. Futhermore, binding to carbohydrate may be of secondary importance to the modern AMPK CBMs where the module has evolved a gain-of-function, binding allosteric modulators, through its interface with the α-subunit kinase domain [15-17].

EXPERIMENTAL SECTION

Constructs and Site-directed Mutagenesis

The wild-type (WT) constructs for β1- and β2-CBM (Q75 – K156, β2-numbering) of the β-subunit of *Rattus norvegicus* AMPK were encoded into pGEX-6P-3 vectors which contain a glutathione S-transferase (GST) – tag and a rhinovirus 3C protease cleavage site. Site-specific point mutations were made of the CBM using a modified version of the site-directed

mutagenesis protocol described by Zheng et al. [40]. Primers were made with non-overlapping 3' ends and the mutation within a short overlapping region. PCR was performed using PrimeSTAR max DNA polymerase (Takara), with 20 - 100 ng of CBM template and 0.2 μM of each primer. The PCR products were treated with DpnI (BioLabs) at 37 °C for 2 hours to remove any WT template. The PCR products were then used to transform super competent DH5α cells by heat shock and incubated on LB agar (100 μg/ml Ampicillin) at 37 °C. Single colonies were transferred to 10 mL LB medium (100 μg/ml Ampicillin) and incubated overnight at 37 °C with shaking (180 rpm) for plasmid extraction by mini-prep. Mutants were then confirmed by Sanger sequencing, performed by the Molecular Diagnostics unit at the Centre for Translational Pathology of the University of Melbourne.

Protein Expression and Purification

Wild-type (WT) and mutant CBMs were expressed using the BL21 (DE3) strain of *Escherichia coli*. For experiments not requiring labelled protein, 1 L of 2xYT medium (100 μg/ml Ampicillin) was inoculated and grown to an optical density (600 nm) of 0.7 at 37 °C with shaking (180 rpm). Expression was induced with 1 mM IPTG and allowed to occur for 4 hours. After expression cells were harvested by centrifugation (4000 x g, 15 min, 4 °C) and cell pellets resuspended in phosphate buffered saline (1xPBS, pH 7.5), 1 mM EDTA, 1 mM DTT and a protease inhibitor tablet (Roche) before storage at -20 °C. Samples of protein for NMR experiments requiring ^{15}N-labelling were prepared as previously described [21, 41]. Protein was purified as previously described [21, 42].

Isothermal titration calorimetry (ITC)

Calorimetric measurements of WT and mutant CBMs binding to the oligosaccharides β-cyclodextrin (BCD) and glucosyl-β-cyclodextrin (gBCD) were performed using a MicroCal

iTC200. CBM (20 to 50 μM) was loaded into the cell and titrated with oligosaccharide (200 to 500 μM) with constant stirring of 1000 rpm. Each experiment was performed with an initial titration of 0.4 μL following by 16 titrations of 2.0 μL at 25 °C. To counter the effects of dilution, oligosaccharide was titrated into analysis buffer as a reference and subtracted from the experimental data. Analysis was performed using the MicroCal package of Origin 7 as follows. The amount of heat released was determined as the area under each peak and plotted against molar ratio ([ligand]/[protein]). Each data set was then fit to the single-site binding model to obtain binding affinity (K_a), stoichiometry (n) and enthalpy of binding (ΔH) [21].

Chemical denaturation stability assay

Stability of WT and mutant CBMs was studied by chemical denaturation, monitoring tryptophan fluorescence in the presence of increasing guanidine hydrochloride (GdmCl). UV transparent 96-well plates were set up with 0.05 mg/mL CBM with GdmCl concentrations between 0 to 4 M. Samples were excited at 280 nm and emission monitored from 300 to 400 nm using a plate reader at 25 °C. The resulting curves were fit to a 5-parameter Weibull distribution to obtain the wavelength at maximum intensity. Instead of monitoring the wavelength at maximum intensity it is preferable to use the maximum intensity itself to build the unfolding curve and fit these data. The wavelengths were then plotted against [GdmCl] and fit using the following function [43]:

$$\omega_{obs} = \frac{\left(\omega_f + m_f[GdmCl]\right) + (\omega_u + m_u[GdmCl])(exp - \left\{\frac{\Delta G°(H_2O)}{RT} - \frac{m[GdmCl]}{RT}\right\})}{1 + exp - \left\{\frac{\Delta G°(H_2O)}{RT} - \frac{m[GdmCl]}{RT}\right\}} \qquad \text{(Equation 1)}$$

Where ω_{obs} is the observed wavelength at maximum intensity rather the maximum intensity value, [GdmCl] is the concentration of guanidine hydrochloride, $\Delta G°(H_2O)$ is the standard Gibbs free energy of unfolding in the absence of denaturant, m the slope of the unfolding curve in the transition measuring the effect of guanidine hydrochloride on $\Delta G°$, T is the temperature, R is the gas constant and ω_f, m_f, ω_u and m_u describe the pre- and post-transition baselines. The effects of pre- and post-transition were subtracted from the curves by linear regression prior to fitting.

Ancient sequence reconstruction (ASR)

Sequences of the AMPK β-subunits and homologues were obtained from a previously reported β-subunit phylogenetic study [19]. A sequence alignment was generated using Clustal Omega of the EMBL-EBI webserver [44]. The sequence alignment was manually trimmed using JalView and a maximum likelihood tree recreated using PhyML [29]. The sequence alignment and phylogenetic tree were used to perform the maximum likelihood ancient sequence reconstruction using the FastML webserver [31] with the JTT model of substitution. An expression construct of the ancient sequence CBM (AS-CBM) was generated using site-directed mutagenesis of the β2-F91Y mutant. Recombinant AS-CBM was expressed and purified using the same methods as WT and mutant CBMs.

Affinity and temperature dependence of AS-CBM by ITC.

Carbohydrate-binding affinities of AS-CBM to BCD and gBCD via ITC were performed as described for WT and mutant CBMs. Temperature dependence of AS-CBM carbohydrate binding was also studied using ITC as reported previously [21]. In brief, ITC was performed for AS-CBM at multiple temperatures (10, 15, 20, 25 and 30 °C) to both carbohydrates

(gBCD and BCD). Results were fit to a single site binding model and the parameters fit to the following thermodynamic functions to determine ΔC_P:

$$\Delta H = \Delta H_0 + \Delta C_p(T - T_0) \quad \text{(Equation 2)}$$

$$\Delta S = \Delta S_0 + \Delta C_p \ln\left(\frac{T}{T_0}\right) \quad \text{(Equation 3)}$$

$$K_d = exp\left(\frac{-\Delta H + T\Delta S}{RT}\right) \quad \text{(Equation 4)}$$

AS-CBM dynamics studied by NMR spectroscopy

^{15}N CPMG relaxation dispersion experiments were performed as reported previously for WT-CBMs [21]. Briefly, a 500 µM sample of unbound AS-CBM was used and the experiment recorded of the unbound protein at 25 °C on a Bruker Avance III 600 MHz spectrometer with varying CPMG frequencies (0, 50, 2x75, 100, 2x150, 200, 300, 500, 2x700, 800, 900, 1000, 1500 Hz) and a constant delay of 80 ms. The data was processed using NMRPipe [45] and Sparky (Goddard, T.D. and Kneller, D.G., University of California, San Francisco) was used to assign peaks and determine peak intensities. All subsequent relaxation dispersion analysis was performed using NESSY [46].

ZZ-exchange spectroscopy experiments were also performed as previously reported for WT-CBMs [21]. 500 µM samples were used with the addition of BCD or gBCD at concentrations that resulted in resonances half in the free and half in the bound state (approximately 250 µM). Experiments were performed at 10 °C on a Bruker Avance IIIHD 700 MHz spectrometer with varying mixing times (0.03, 0.05, 0.09, 0.13, 0.15, 0.35, 0.55, 0.75 and 0.95 s). The data was processed using NMRPipe [45] and peak volumes determined using Sparky. MATLAB scripts generously donated by Dr Demers and Dr Mittermaier [47]

were used to calculate NMR-derived k_{off} values, and k_{on} was determined using ITC derived K_d values.

Acknowledgements

This work was supported by Australian Research Council (ARC) Discovery grant DP110103161 to P.R.G. and D.S.; equipment grants from the State of Victoria, ARC and Rowden White Foundation. M.D.W.G is the recipient of the C.R. Roper Fellowship and the Australian Research Council Future Fellowship (project number FT140100544).

References

[1] Hardie, D. G., *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 774-785.

[2] Bergeron, R., Russell, R. R., 3rd, Young, L. H., Ren, J. M., Marcucci, M., Lee, A., and Shulman, G. I., *Am. J. Physiol.* **1999**, *276*, E938-944.

[3] Chen, Z. P., McConell, G. K., Michell, B. J., Snow, R. J., Canny, B. J., and Kemp, B. E., *Am. J. Physiol. Endocrinol. Metab.* **2000**, *279*, E1202-1206.

[4] Davies, S. P., Carling, D., Munday, M. R., and Hardie, D. G., *Eur. J. Biochem.* **1992**, *203*, 615-623.

[5] Gwinn, D. M., Shackelford, D. B., Egan, D. F., Mihaylova, M. M., Mery, A., Vasquez, D. S., Turk, B. E., and Shaw, R. J., *Mol. Cell* **2008**, *30*, 214-226.

[6] Steinberg, G. R., and Kemp, B. E., *Physiol. Rev.* **2009**, *89*, 1025-1078.

[7] Hawley, S. A., Pan, D. A., Mustard, K. J., Ross, L., Bain, J., Edelman, A. M., Frenguelli, B. G., and Hardie, D. G., *Cell Metab.* **2005**, *2*, 9-19.

[8] Woods, A., Johnstone, S. R., Dickerson, K., Leiper, F. C., Fryer, L. G., Neumann, D., Schlattner, U., Wallimann, T., Carlson, M., and Carling, D., *Curr. Biol.* **2003**, *13*, 2004-2008.

[9] Oakhill, J. S., Steel, R., Chen, Z. P., Scott, J. W., Ling, N., Tam, S., and Kemp, B. E., *Science* **2011**, *332*, 1433-1435.

[10] Ali, N., Ling, N., Krishnamurthy, S., Oakhill, J. S., Scott, J. W., Stapleton, D. I., Kemp, B. E., Anand, G. S., and Gooley, P. R., *Sci. Rep.* **2016**, *6*, 39417.

[11] Oakhill, J. S., Chen, Z. P., Scott, J. W., Steel, R., Castelli, L. A., Ling, N., Macaulay, S. L., and Kemp, B. E., *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 19237-19241.

[12] Li, X., Wang, L., Zhou, X. E., Ke, J., de Waal, P. W., Gu, X., Tan, M. H., Wang, D., Wu, D., Xu, H. E., and Melcher, K., *Cell Res.* **2015**, *25*, 50-66.

[13] Oligschlaeger, Y., Miglianico, M., Chanda, D., Scholz, R., Thali, R. F., Tuerk, R., Stapleton, D. I., Gooley, P. R., and Neumann, D., *J. Biol. Chem.* **2015**, *290*, 11715-11728.

[14] Xu, H., Frankenberg, N. T., Lamb, G. D., Gooley, P. R., Stapleton, D. I., and Murphy, R. M., *Am. J. Physiol. Cell Physiol.* **2016**, *311*, C35-42.

[15] Scott, J. W., van Denderen, B. J., Jorgensen, S. B., Honeyman, J. E., Steinberg, G. R., Oakhill, J. S., Iseli, T. J., Koay, A., Gooley, P. R., Stapleton, D., and Kemp, B. E., *Chem. Biol.* **2008**, *15*, 1220-1230.

[16] Xiao, B., Sanders, M. J., Carmena, D., Bright, N. J., Haire, L. F., Underwood, E., Patel, B. R., Heath, R. B., Walker, P. A., Hallen, S., Giordanetto, F., Martin, S. R., Carling, D., and Gamblin, S. J., *Nat. Commun.* **2013**, *4*, 3017.

[17] Scott, J. W., Galic, S., Graham, K. L., Foitzik, R., Ling, N. X., Dite, T. A., Issa, S. M., Langendorf, C. G., Weng, Q. P., Thomas, H. E., Kay, T. W., Birnberg, N. C., Steinberg, G. R., Kemp, B. E., and Oakhill, J. S., *Chem. Biol.* **2015**, *22*, 705-711.

[18] Kemp, B. E., Stapleton, D., Campbell, D. J., Chen, Z. P., Murthy, S., Walter, M., Gupta, A., Adams, J. J., Katsis, F., van Denderen, B., Jennings, I. G., Iseli, T., Michell, B. J., and Witters, L. A., *Biochem. Soc. Trans.* **2003**, *31*, 162-168.

[19] Koay, A., Woodcroft, B., Petrie, E. J., Yue, H., Emanuelle, S., Bieri, M., Bailey, M. F., Hargreaves, M., Park, J. T., Park, K. H., Ralph, S., Neumann, D., Stapleton, D., and Gooley, P. R., *FEBS Lett.* **2010**, *584*, 3499-3503.

[20] Bieri, M., Mobbs, J. I., Koay, A., Louey, G., Mok, Y. F., Hatters, D. M., Park, J. T., Park, K. H., Neumann, D., Stapleton, D., and Gooley, P. R., *Biophys. J.* **2012**, *102*, 305-314.

[21] Mobbs, J. I., Koay, A., Di Paolo, A., Bieri, M., Petrie, E. J., Gorman, M. A., Doughty, L., Parker, M. W., Stapleton, D. I., Griffin, M. D., and Gooley, P. R., *Biochem. J.* **2015**, *468*, 245-257.

[22] Polekhina, G., Gupta, A., van Denderen, B. J., Feil, S. C., Kemp, B. E., Stapleton, D., and Parker, M. W., *Structure* **2005**, *13*, 1453-1462.

[23] Harms, M. J., and Thornton, J. W., *Curr. Opin. Struct. Biol.* **2010**, *20*, 360-366.

[24] Anderson, D. W., McKeown, A. N., and Thornton, J. W., *Elife* **2015**, *4*, e07864.

[25] Hall, B. G., *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5431-5436.
[26] Asensio, J. L., Arda, A., Canada, F. J., and Jimenez-Barbero, J., *Acc. Chem. Res.* **2013**, *46*, 946-954.
[27] Myers, J. K., Pace, C. N., and Scholtz, J. M., *Protein Sci.* **1995**, *4*, 2138-2148.
[28] McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P., and Lopez, R., *Nucleic Acids Res.* **2013**, *41*, W597-600.
[29] Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O., *Syst. Biol.* **2010**, *59*, 307-321.
[30] Putnam, N. H., Butts, T., Ferrier, D. E., Furlong, R. F., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J.-K. K., Benito-Gutiérrez, E. L., Dubchak, I., Garcia-Fernàndez, J., Gibson-Brown, J. J., Grigoriev, I. V., Horton, A. C., de Jong, P. J., Jurka, J., Kapitonov, V. V., Kohara, Y., Kuroki, Y., Lindquist, E., Lucas, S., Osoegawa, K., Pennacchio, L. A., Salamov, A. A., Satou, Y., Sauka-Spengler, T., Schmutz, J., Shin-I, T., Toyoda, A., Bronner-Fraser, M., Fujiyama, A., Holland, L. Z., Holland, P. W., Satoh, N., and Rokhsar, D. S., *Nature* **2008**, *453*, 1064-1071.
[31] Ashkenazy, H., Penn, O., Doron-Faigenboim, A., Cohen, O., Cannarozzi, G., Zomer, O., and Pupko, T., *Nucleic Acids Res.* **2012**, *40*, W580-584.
[32] Dayhoff, M. O., Schwatrz, R. M., and Orcutt, B. C., in *Atlas of Protein Sequence and Structure., Vol. 5* (Ed.: Dayhoff, M. O.), National Biomedical Research Foundation, Washington, DC, **1978**, pp. 345–352.
[33] Jones, D. T., Taylor, W. R., and Thornton, J. M., *Comput. Appl. Biosci.* **1992**, *8*, 275-282.
[34] Whelan, S., and Goldman, N., *Mol. Biol. Evol.* **2001**, *18*, 691-699.
[35] Williams, B. A., Chervenak, M. C., and Toone, E. J., *J. Biol. Chem.* **1992**, *267*, 22907-22911.
[36] Demers, J. P., and Mittermaier, A., *J. Am. Chem. Soc.* **2009**, *131*, 4355-4367.
[37] Bloom, J. D., Raval, A., and Wilke, C. O., *Genetics* **2007**, *175*, 255-266.
[38] Tokuriki, N., Stricher, F., Schymkowitz, J., and Serrano, L., *J. Mol. Biol.* **2007**, *369*. 1318-1332.
[39] Thornton, C., Snowden, M. A., and Carling, D., *J. Biol. Chem.* **1998**, *273*, 12443-12450.
[40] Zheng, L., Baumann, U., and Reymond, J. L., *Nucleic Acids Res.* **2004**, *32*, 5.
[41] Cai, M., Huang, Y., Sakaguchi, K., Clore, G. M., Gronenborn, A. M., and Craigie, R., *J. Biomol. NMR* **1998**, *11*, 97-102.
[42] Koay, A., Rimmer, K. A., Mertens, H. D., Gooley, P. R., and Stapleton, D., *FEBS Lett.* **2007**, *581*, 5055-5059.
[43] Pace, C. N., *Trends Biotechnol.* **1990**, *8*, 93-98.
[44] Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G., *Mol. Syst. Biol.* **2011**, *7*.
[45] Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A., *J. Biomol. NMR* **1995**, *6*, 277-293.
[46] Bieri, M., and Gooley, P. R., *BMC Bioinformatics* **2011**, *12*, 421.
[47] Demers, J.-P., and Mittermaier, A., *J. Am. Chem. Soc.* **2009**, *131*, 4355-4367.

TOC Graphic