The circularly permuted globin domain of Androglobin

Brandon Reeder
University of Essex
Giuseppe Deganutti
Coventry University https://orcid.org/0000-0001-8780-2986
John Ukeri
University of Essex
Silvia Atanasio
University of Essex
Dimitri Svistunenko
University of Essex
Christopher Ronchetti
University of Essex
Juan Calos Mobarec
University of Essex
Marten Vos
Université Paris-Saclay https://orcid.org/0000-0003-0493-4831
Michael Wilson
University of Essex
Christopher Reynolds (ad5291@coventry.ac.uk)
Coventry University https://orcid.org/0000-0001-9267-5141

Article

Keywords: helix alignment, molecular dynamics, homology modelling, nitric oxide homeostasis, disulfide

Posted Date: May 24th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1685645/v1

The circularly permuted globin domain of Androglobin

Brandon J. Reeder[1]*, Giuseppe Deganutti[1,3], John Ukeri[1], Silvia Atanasio[1], Dimitri A. Svistunenko[1], Christopher Ronchetti[1], Juan Carlos Mobarec[1], Marten H. Vos[2], Michael T. Wilson[1], and Christopher A. Reynolds[1,3]*

[1]School of Life Sciences, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK.
[2]LOB, CNRS, INSERM, École Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau, France.
[3]Centre for Sport, Exercise and Life Sciences (CSELS), Alison Gingell Building, Coventry, CV1 5FB, UK.

*E-mail: reedb@essex.ac.uk (biochemical), and ad5291@coventry.ac.uk (modelling)

Supporting information for this article is available.

Abstract: Androglobin, is a recently discovered circularly permuted, multi-domain is hemoglobin. Using a remote homologue alignment method, coupled with molecular modelling and molecular dynamics, we identified the alignment to other hemoglobins. This guided the first stable recombinant expression of an androglobin domain and the first structural and biochemical characterization of the globin domain of androglobin, which is split by an IQ domain. Tyrosine is found in place of the highly conserved phenylalanine that resides in the highly conserved CD1 position, a structural feature unknown in eukaryotes but common in prokaryotic globins. As expressed, the heme iron is hexacoordinate in the ferrous form but partially pentacoordinate in the ferric form. Exceptional in the globin superfamily, but similar to other hemoproteins such as cytochrome c', the heme iron binds nitric oxide as a five coordinate complex. This work expands our knowledge of the fundamental chemistry of this hitherto elusive medically important protein.

Introduction

The hemoglobin (Hb) of the erythrocyte is one the most studied proteins in science. However, a decade after the discovery of androglobin (Adgb), a multi-domain Hb first identified in the testis of metazoans, there is still very little known about this novel atypical member of the globin superfamily. Adgb is unusual due to its long length of 1667 amino acids (human), with a centrally positioned globin domain.[1] This compares to the 140-190 amino acids of other human globins such as Hb,[2] myoglobin (Mb), neuroglobin (Ngb)[3] and cytoglobin (Cygb)[4] and is significantly larger than other multi-domain globins such as flavohemoglobin.[5] Initial sequence alignment analysis predicted the heme-binding globin domain of Adgb (Adgb-GD) to consist of the eight alpha helical structure (termed A to H) with the 3-on-3 alpha helical fold that encloses the heme moiety and is typical of most non-truncated globin architectures.[1] However, highly unusual in the globin family is that Adgb-GD is circularly permuted with a calmodulin binding domain situated between the H and A helices.[1] The N-terminal region of Adgb is reported to contain a calpain-like region and the C-terminal region is reported to contain sequences for a coiled-coil region, a nuclear localization signal (NLS) and an ER membrane endoplasmic reticulum retention signal.[1]

Since the discovery of Adgb by Hoogewijs et al. in 2012[1], very little has been elucidated about the properties of this highly unusual and newest member of the Hb superfamily, but its potential interaction with NO and calcium may underlie its relevance to spermatogenesis, ciliogenesis, as well as to potential diseases. A knockdown study in cells showed enhanced apoptosis and proliferation inhibition in glioma cell lines, relating to changes in the level of several proteins

involved in cell proliferation, survival or apoptosis, including STAT3 cleaved caspase-3 and Bcl-2.[6] Additionally, studies have shown that STXBP5 antisense proficient GR pancreatic cancer cell lines overexpress Adgb through ADGB promotor methylation, leading to drug resistance and inhibition of cell apoptosis. Recently, mRNA-Seq data from mammalian tissue has shown that Adgb is expressed in the lungs, brain and female reproductive tract. In each case, Adgb is specifically associated with cell types forming motile cilia.[7]

There remains a lack of data concerning the structure, properties or functions of this protein, due to the difficulty in generating such a large full-length hemoglobin by recombinant techniques and by instabilities of the heme-binding, circularly-permuted globin domain.[8]The alignment between Adgb and other proteins with regard to Adgb–GD.[1] was reassessed using a novel helix alignment method shown to work in the twilight zone of low sequence similarity.[9, 10, 11] Our unique identification of the globin domain yielded a structure by comparative modelling that was stable under 500 ns of molecular dynamics simulations, giving support for the potential expression of this protein despite initial negative results.[8] Thus, based on our new alternative alignment, we have expressed the circularly-permuted globin domain as a stable recombinant protein. An intramolecular disulfide bond linking the N and C terminal sections of the heme binding globin domain appears to stabilize the 'CD loop' heme pocket region and likely the whole globular domain.

With a stable form of the protein expressed, we have characterized the protein using optical, EPR, stopped-flow and femtosecond laser flash photolysis to show that Adgb binds nitric oxide (NO) as a five-coordinate heme iron. This is unlike other globins, but similar to that observed with other hemoproteins such as guanylate cyclase, cytochrome c prime (cyt c') and dissimilatory nitrate respiration regulator (DNR)[12, 13, 14]. Furthermore, the protein exhibits high nitrite reductase activity, which is influenced by the redox state of the disulfide bond and exhibits a high autoxidation similar to Ngb. Based on our findings we propose that the globin domain may serve as an NO storage or NO synthesis protein under hypoxic conditions. With a binding domain for a calcium-dependent calmodulin and a calcium dependent calpain-like sequence on the N terminal domain, it is very likely that Adgb plays a role in both NO homeostasis and calcium signaling, with repercussions on spermatogenesis, sperm maturation and cilia function.[15, 16, 17, 18]

Results and Discussion

Figure 1 (left) shows the scores for various alternative alignments of the six key helices of the Adgb-GD to their counterparts in alpha-Hb, beta-Hb, Cygb, Ngb, and Mb. Helix A (Adgb residues Val939-Glu951) is the first helix in alpha-Hb, beta- Hb, Cygb, Ngb and Mb, but in the naturally circular permuted Adgb it follows helix H (Adgb residues Phe866-Ser877) and the IQ domain (residues Val891-Thr933). The main peak lies at 0 for alpha-Hb, beta-Hb, Cygb, Ngb and Mb, which corresponds to the alignment in Figure 2; a small alternative peak lies at +4 for each globin bar Mb, which corresponds to shifting the Adgb helix A four residues to the right. For the alpha-Hb–Adgb-GD alignment, there were 17155 pairwise alignments (because there were 235 alpha-Hb sequences and 73 Adgb-GD sequences). When the Adgb sequences were moved between -25 and +25, the 0 alignment obtained 10664 votes and the +4 alignment obtained 3664 votes; when the alpha-Hb sequences were moved between -25 and +25, the zero alignment obtained 10001 votes and the +4 alignment obtained 774 votes. In this case, the mean of 10332 for alignment 0 is considerably higher than the next nearest mean of 2219 (alignment +4). A similar predominant score was obtained for all H1 alignments, and so when the scores for all five alignments were multiplied together, Figure 1 (right), there is an overwhelming preference for alignment 0, despite the low percentage identity for H1 of 15.3 %, 15.2 %, 16.6 %, 21.1 %, and 15.1 % for the Adgb-GD alignments to alpha-Hb, beta-Hb, Cygb, Ngb, and Mb respectively. The results for the other five helices are similar, with percentage identities ranging between 7.8 % (H6 of the Adgb-GD–Ngb alignment) and 25 %

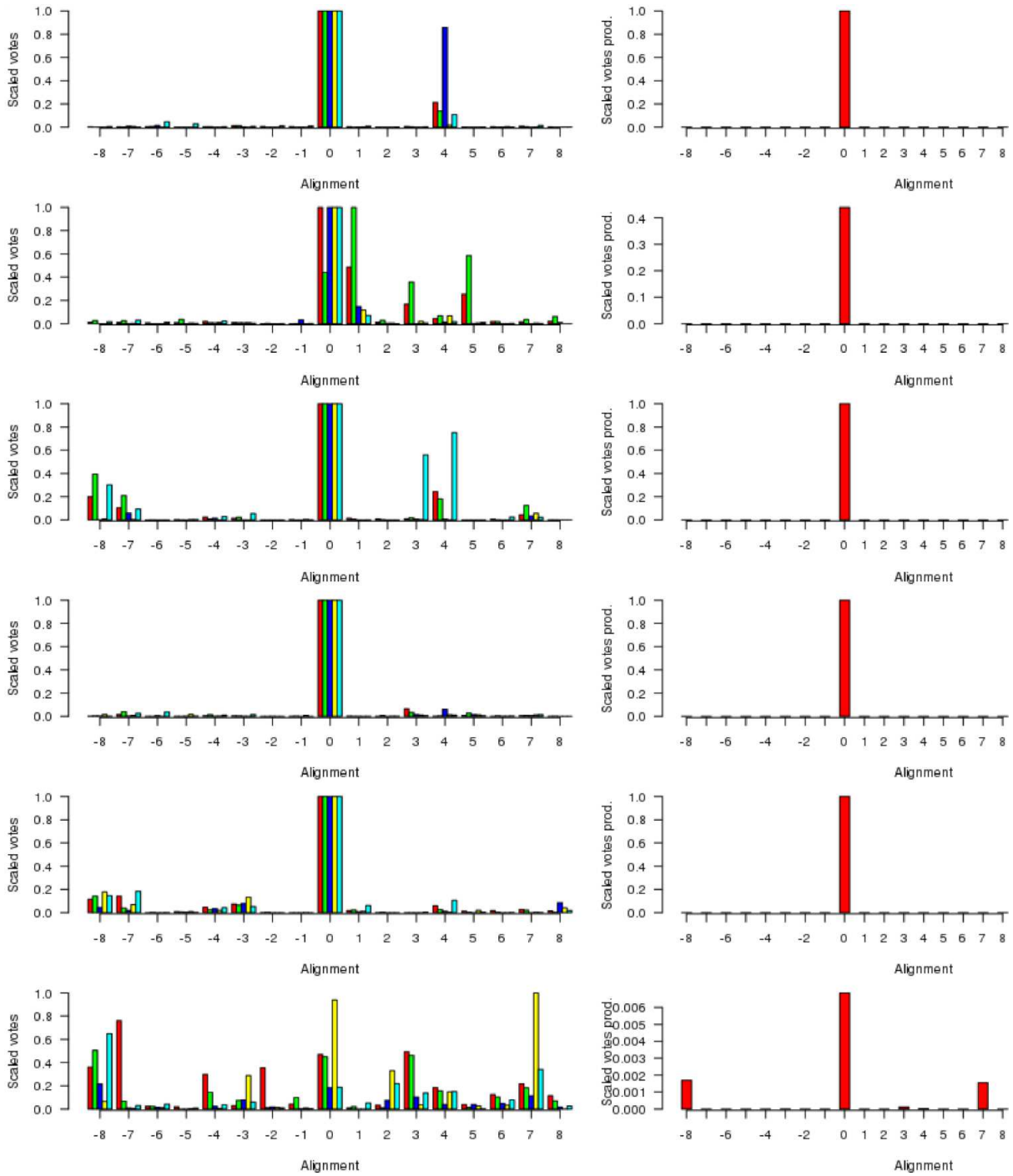

Figure 1. The alignment of the androglobin heme-binding globin domain helices. The left-hand column gives the individual alignments for helices A-B and helices E-H, one per row. The alpha-Hb (red), beta-Hb (green), Cygb (blue), Mb (yellow), and Ngb (cyan) alignments to Adgb are denoted by the bars representing the number of votes (scaled between 0 and 1). The right-hand column gives the consensus alignment where the votes for the individual alignments are multiplied together. The results point overwhelmingly to the 0 alignment given in Figure 2. An alignment of +1 would correspond to movement of the Adgb helix one residue to the right in Figure 2; an alignment of say -3 would correspond to movement of the Adgb helix three residues to the left in Figure 2.

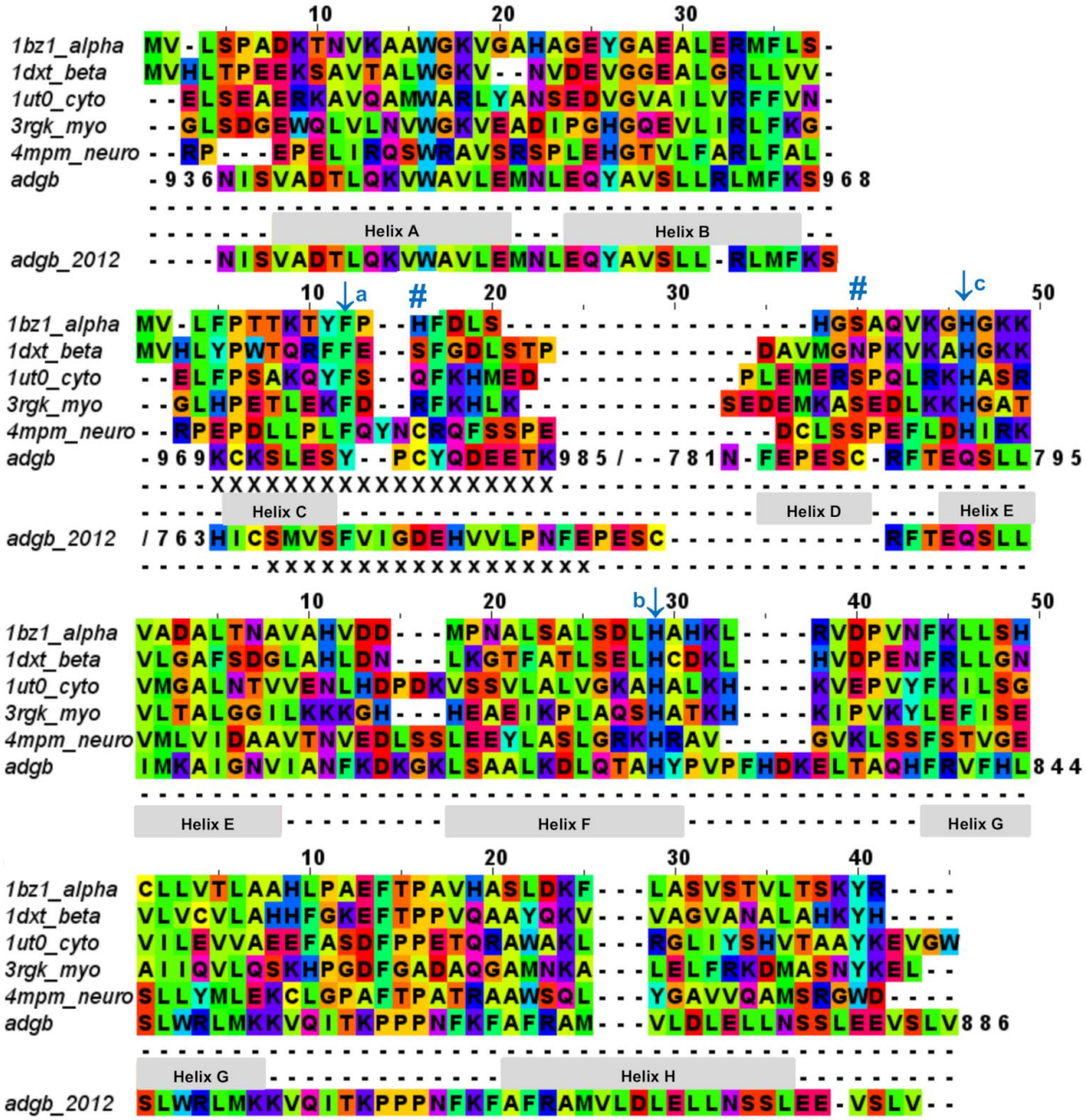

Figure 2. Androglobin heme-binding globin domain alignment. The alignment of the traditional globins was generated by structural alignment; the alignment of Adgb was determined using the in-house multi-template approach. The alignment denoted Adgb_2012, previously reported by Hoogewijs et al.[1], is shown where it differs from the 2022 alignment. The six alignment zones coincide (bar helices C and D) with globin helices denoted 'Helix A, 'Helix B' etc. as determined by inspection from the X-ray crystal structures. '/' denotes the chain break in the discontiguous Adgb sequence, 'XXX'… denotes the extra residues used in the Adgb construct (this work) while 'xxx…' denotes the extra residues in the globin domain of Adgb previously reported[1] and which here are identified by us as part of the preceding calpain C2-like domain (unpublished work). Key amino acids are denoted by labelled arrows, 'a' for the conserved CD1 loop aromatic residue, 'b' for proximal histidine and 'c' for the distal histidine/glutamine; the disulfide bond is denoted by '#'.

(H4 of the Adgb–Mb alignment), Table S1. Based on these percentages, Adgb-GD is marginally more similar to Mb (mean helical percentage identity 17.8 %) and less similar to beta-Hb (mean helical percentage identity 14.0 %). The full alignment is given in Figure 2 and the mean percentage identity for the pairwise alignments for each helix-helix profile alignment between androglobin and the other hemoglobins is given in Table S1.

The threefold Modeller structural alignment of both Adgb models (complex 1 and complex 2, see below) with the apo AlphaFold 2 model shows agreement in the sequence alignment except in two places, namely ^{828}PVPFHDKEL^{834} (where helix F of the AlphaFold 2 model fills the space normally occupied by the heme) and the distal half of helix H, which terminates prematurely to accommodate the IQ domain (^{872}DLWLLN^{877} is not helical).

The current list of Adgb sequences, greatly expanded since the original sequence analysis in 2012, together with the helix alignment analysis above, permits a re-assessment of the overall sequence identification of the globin domain of Adgb. From the previous assignment of the globin domain structure,[1] helices C to H are the first to be expressed on the N terminal side (His761 onwards). This is followed by C terminal helices A and B from Asp935 onwards. These sections are interconnected by a 33-amino acid section incorporating an IQ calmodulin binding domain (Figure S2a). This assignment took into account several key amino acids that are highly conserved throughout the hemoglobin superfamily. These comprised of: (i) The proximal histidine connecting the heme iron to the protein, situated on the F helix (F8); (ii) The E7 distal heme iron ligand, in Adgb this is a glutamine instead of the more common histidine residue;[1] (iii) The CD1 Phe residue functions as anchorage and binding of the heme group within the heme pocket.[19, 20] While our sequence alignment agrees with most of the original alignment, the identity of the C helix region and the CD1 residue differ. With the Phe770 amino acid assignment as the CD1 residue in human Adgb it becomes clear that in other species ~22% of such alignments places a Cys in the CD1 position (Table S2). Indeed, the previous alignment of Adgb globin domains shows three (9% of sequences) as Cys residues at the CD1 position, [1] but this was interpreted as mismatched alignments.

A Cys residue in the CD1 position is likely to have profound impact on the stabilization of heme binding as the substitution of the CD1 phenylalanine by cysteine removes an important contact with heme leaving a gap at the surface of the heme pocket which could result in instability. Mutations in human hemoglobin of CD1 Phe almost invariably result in instabilities in the protein, resulting in Heinz body formation, cyanosis and severe hemolytic anemia. A natural variation with a Cys residue in the CD1 position was found in the beta Hb chain of a Caucasian male infant (Hb Little Venice, β42[CD1] Phe→Cys).[21] At 2 years of age the infant showed severe chronic hemolytic anaemia, positive Heinz body formation, haptoglobin depletion and required a monthly regular transfusion regime used more commonly for cases of severe forms of thalassemia.

We conducted a reassessment of the C to D helix sequence assignment, disregarding the prerequisite of a Phe as the CD1 component of the structure, based on the globin domain helix alignments (Figures 1 and 2) . Our assessment places a different sequence as the C to D helix section thus placing a Tyr976 in the CD1 position of human Adgb globin domain (Figure S3, Figure S2B). Although essentially unique in eukaryotes, the presence of a Tyr in the CD1 position of prokaryotic Hbs is common and does not significantly affect heme binding, but can adversely affect oxygen binding affinity.[22] Sequence alignment comparison shows that the frequency of Tyr residues is ~82% with the majority of the remaining sequences being Phe residues (Table S2). Therefore, we propose the assignment of CD1 to Tyr976 instead of Phe770, resulting in a circular permutation where helices D to H are expressed on the N-terminal side, followed by helices A to C following the IQ calmodulin binding domain sequence.

The previous alignment sequence of the globin domain did not allow expression of a stable form of the globin domain to be generated recombinantly,[8] however, this new approach to sequence alignment resulted in expression of an stable globin domain (*vide infra*).

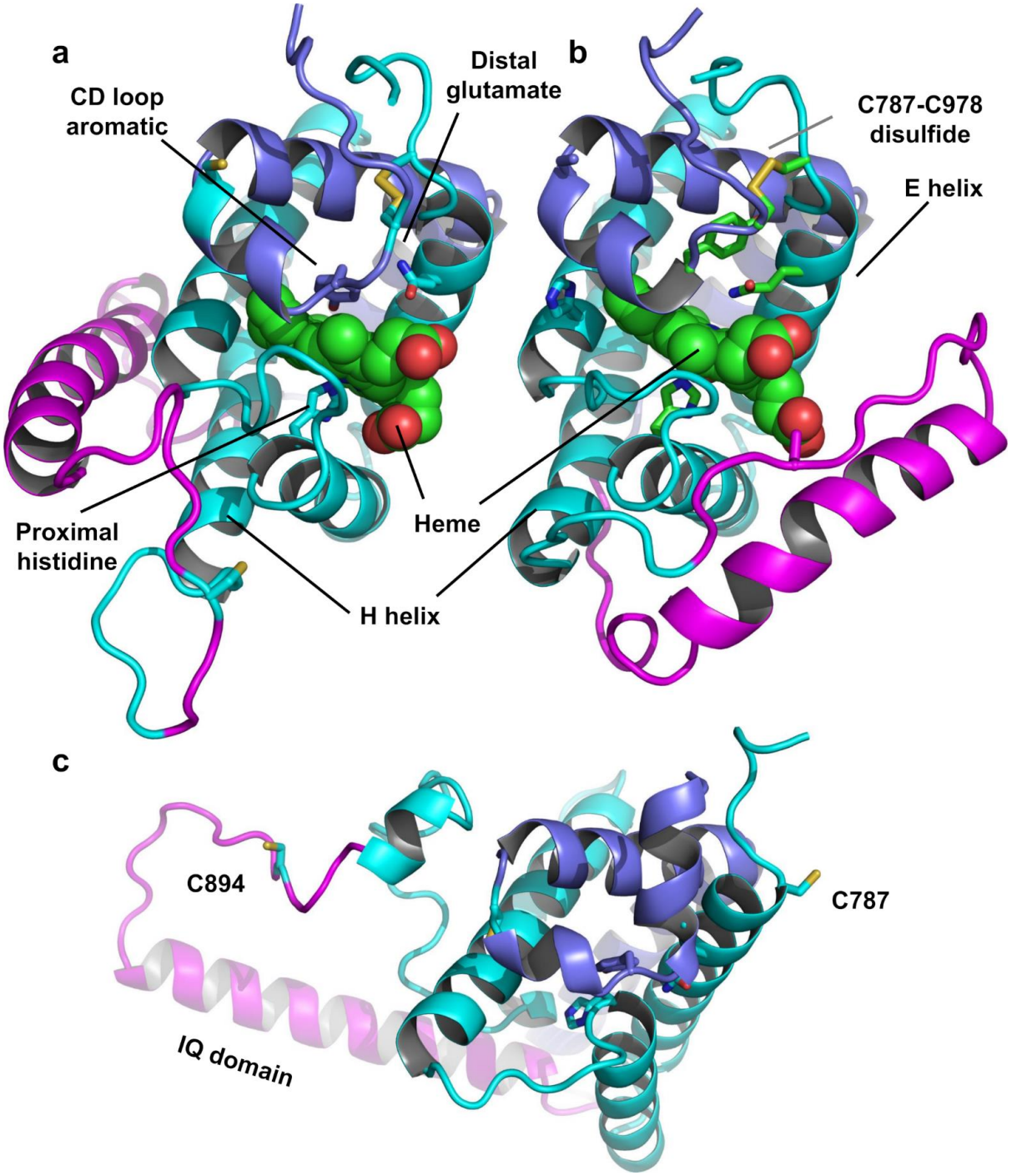

Figure 3. Structures of the heme-binding globin domain of androglobin. (a) The complex 1 structure determined using Modeller, (b) the complex 2 structure determined using Modeller and (c) The AlphaFold 2 structure. The structure is colored blue for the N-terminus (helices D-H), magenta for the IQ domain insertion and cyan for the C terminus (helices A - C).

Adgb-GD complexes 1 and 2 are shown in Figure 3, and align well with the five hemoglobin structural templates. The RMSDs over the common globin domain to myoglobin are 1.3 Å for complex 1, 1.1 Å for complex 2 and 2.3 Å for the AlphaFold 2 model, over the alpha helices, as determined using SSM.[23] The three models differ in the orientation of the IQ (calmodulin-binding) domain, which is not included in the RMSD calculations because it is absent in Mb and the other traditional Hbs. The orientation of the IQ domain differs in all three structures and is connected by a very flexible loop, and this flexibility is probably important for binding calmodulin. When superposed onto the full Adgb AlphaFold 2 structure, all three structures

present the IQ domain in an accessible orientation. The AlphaFold 2 and Modeller structures differ in two other aspects. Firstly, helix H is truncated in the AlphaFold 2 model. Secondly, helix E moves slightly into the space occupied by the heme; this structural topology may bear some resemblance to purified Adgb before the heme group is added back (see Methods). The globin domain contains four cysteine residues, none of which form a disulfide bond in the AlphaFold 2 model.

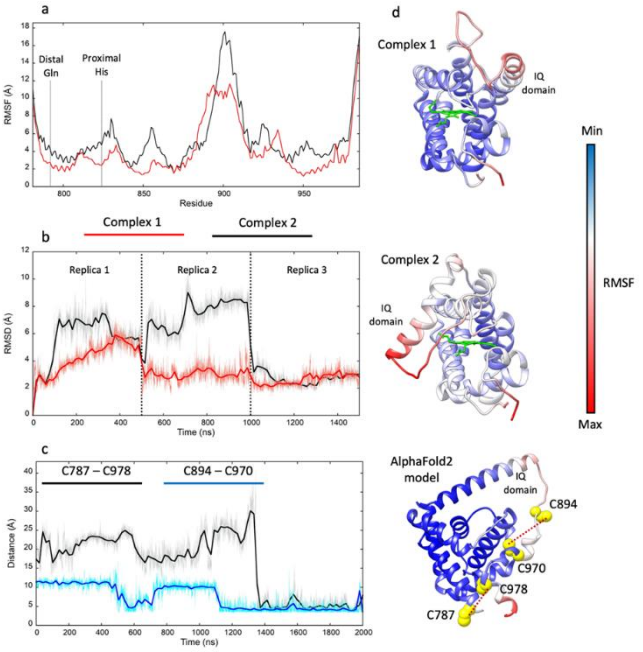

Figure 4. Molecular dynamics simulations of Adgb-GD complex 1, complex 2, and AlphaFold2 model. (a) RMSF comparison between complex 1 and 2. (b) Heme RMSD within complex 1 and 2 during the MD replicas. (c) Distances between pairs of cysteine residues during the MD simulation of the AlphaFold2 model; (d) RMSF plotted on complex 1, complex 2, and AlphaFold2 model (represented as ribbon, heme is represented as green stick); red ribbon color indicates high structural flexibility; the four cysteine of the domain and their distance are shown on the AlphaFold2 model.

Adgb-GD complexes 1 and 2 were interrogated through molecular (MD) dynamics simulations to assess stability over the time course. Three 500 ns MD replicas were produced for each model in complex with the heme (Figure 4). The emerging scenario suggests the overall higher stability of complex 1 as indicated by the root mean square fluctuation, RMSF, analysis (Figure 4a,d) and the root mean square deviation, RMSD, of the heme (Figure 4b, d). In both structures, the IQ domain and the loop connecting it to the H helix were the most dynamic parts (Supplementary videos 1 and 2), however, in complex 1 the whole structure underwent high thermal fluctuations; this affected the stability of the heme (Figure 4b). In both systems, the distal Gln792 remained in the proximity of the heme, while the CD Tyr976 was more flexible during the simulations, especially in complex 2.

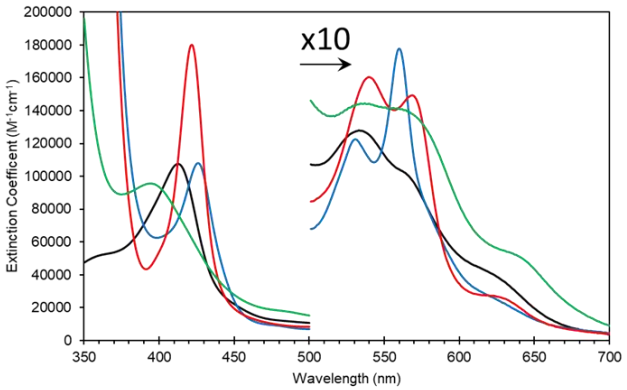

Figure 5. Optical characteristics of the androglobin globin domain. Ferric (black line), deoxy ferrous (blue line), ferrous-CO bound (red line) and ferrous-NO bound (green line).

As expressed, the optical properties of the Adgb-GD are shown in Figure 5 and calculated extinction coefficient and peak wavelengths in Table S3. The ferric protein has bands in the visible region at 534 and 566 nm regions, suggesting a hexacoordinate state of the heme iron like that observed for Cygb and Ngb. [3, 4] However, a small peak at ~630 nm suggests that the protein also has some pentacoordinate-like properties. Reduction to deoxy ferrous iron shows two prominent peaks at 531 and 560 nm indicative of hexacoordinate heme iron configuration. The CO-bound spectrum is typical for most globins, but the NO-bound spectrum exhibits a Soret peak at an unexpectedly

hypsochromic (blue) shifted peak at 395 nm. Typical wavelength maxima for hexacoordinate ferrous-NO bound protein are observed ~420-430 nm in other NO-bound globins such as Mb and Cygb.[24, 25] This suggests that the NO is bound in an unusual form in Adgb-GD.

The Electron Paramagnetic Resonance (EPR) spectrum of the ferric Adgb-GD at a slightly acidic pH (Figure 6a and 6b, pH 6) and 10 K shows a mixture of the high spin (S=5/2, HS) and low spin (S=1/2, LS) signals corresponding to penta- and hexa-coordinated heme iron, respectively. The HS signal with the perpendicular $g_x=g_y=5.95$ and the parallel $g_z=2.00$ components is typical of other globins in a pentacoordinate conformation. The g=2.95 and g=2.26 EPR signals are the g_x and g_y components of a LS signal, with the third g_z component likely to be too broad to be observed or off scale. The LS signal with these g-values has been classified as a low spin form in ferric Hb, with one of the axial ligands likely to be a histidine's nitrogen and the other not identified.[26] At pH 7 and 8, the HS signal is significantly smaller; however, this does not result in a noticeable increase of the low spin form. Instead, a prominent change in the line shape of the perpendicular components area, at g~6, is observed. Not only the g=6 signal becomes much wider, the new effective g-values become apparent – $g_x=6.5$ (or maybe lower – it is not straightforward to determine the value on a wing of the HS (S=5/2) EPR signal) and $g_y=5.3$. This makes the g_{12}-value, which is $g_{12}=(g_x+g_y)/2$ [27], less than g=6, which, in turn, is a strong indication of a presence of a quantum spin mixture of S=5/2 and S=3/2 states of the five 3d electrons in a ferric heme.[28] Thus, the increase of the pH of ferric Adgb-GD results in a deprotonation of a site, which has an effect on the heme geometry yielding, in turn, a change in the distribution of the energies between the $d(x^2-y^2)$ orbital, in the heme plane, and the degenerate d_{xy}, d_{xz} and d_{yz} orbitals.[28]

In Ngb, the CD loop contains a disulfide bond between Cys46 (helix C7) and Cys55 (helix D5), known to stabilize the distal histidine ligation and the redox thermodynamics of ferric Ngb.[29, 30, 31] With our reassignment of the C helix sequence, the CD section of Adgb-GD possesses a predicted disulfide bond from residues Cys787 and Cys978 (Figure S2). Absent in the original sequence alignment, this disulfide bond could stabilize this crucial juncture of the heme pocket, now predicted as the N and C terminal sections of the domain. With four cysteine residues in the globin domain sequence, we determined how many of those are surface exposed and free to bind dithiodipyridine. As shown in Figure 6c, there were two (1.9 ± 0.3) free sulfhydryl's per heme detected with TCEP reduced Adgb-GD, meaning that two of the four cysteines are surface exposed. In the protein as expressed in *E.coli* (without reduction by TCEP), no free cysteines were observed. Thus, as expressed, the globin domain possesses a single disulfide bond. From the predicted positions of the cysteines in complex 1 and 2, only two are predicted to be both surface exposed and close enough to form an intramolecular disulfide, that of Cys787 and 978 in the "CD loop" region of the protein. However, both these two (reduced) cysteine residues and Cys894, Cys970 moved sufficiently close to form a disulfide bond in the MD simulation of the Alphafold2 structure after 1.4 μs (Figure 4c), even though the SG atoms were originally 11.6 Å and 17.4 Å apart respectively (Figure 4d). Further studies will be required to confirm the position and the micro-environmental conditions favoring the formation of the disulfide bond within Adgb-GD. Adgb-GD binds to NO in the ferrous form to generate an unusual optical spectrum as observed in Figure 5, suggesting a five-coordinate NO heme iron. This is supported by EPR at 10 K (Figure 6d). The spectrum observed showed a characteristic three-line hyperfine signal, essentially identical to that reported for cyt c' from Shewanella frigidimarina and Alcaligenes xylosoxidans.[14] This is again consistent with five-coordinate NO binding to the ferrous heme.[14, 32]

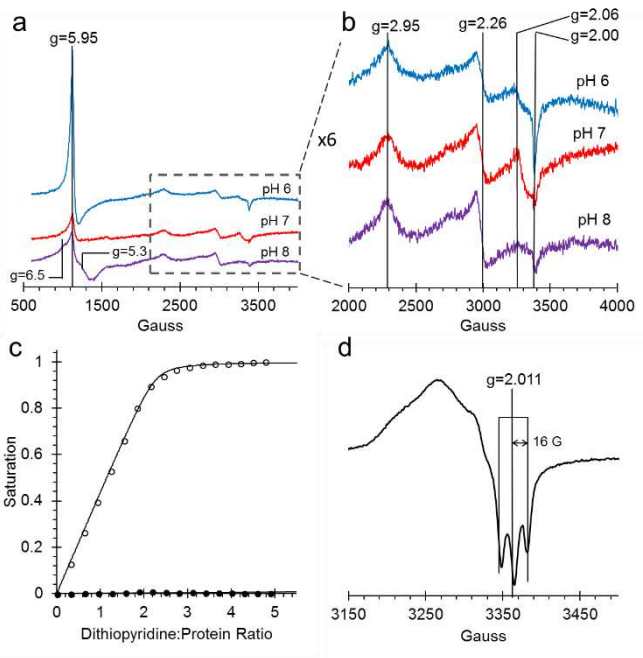

Figure 6. Androglobin heme iron ligation and cysteine oxidation states. (a), EPR spectra of 80 μM ferric androglobin globin domain showing both high spin and low spin Fe^{3+} EPR signals at various pH values. (b), Expanded view of (a) highlighting the low spin signals. (c), Measurement of surface exposed disulfide number and oxidation state. Protein (6.4 μM) was titrated with 4,4' dithiodipyridine and the optical changes followed at 324 nm. The fractional saturation of dithiodipyridine binding to free sulfhydryl as a function of dithiodipyridine concentration for TCEP reduced (○) and purified form (no disulfide reduction, ●) shows that one surface-exposed disulfide bond is present in the globin domain when expressed. (d), An EPR spectrum of ferrous NO bound Adgb-GD, exhibiting three lines, separated by 16 G, around g= 2.011, typical of other EPR spectra of five coordinate NO binding such as cyt c'.

The optical changes and kinetics of NO binding to deoxyferrous are shown in Figure S4. Increases at 390 nm are concurrent with decreases at 426 nm with a number of isosbestic points (e.g 407 nm, Figure 7a and 7b) consistent with a simple Fe^{2+} + NO → Fe^{2+}-NO binding mechanism. However, the time course follows a double exponential function (Figure S4c), suggesting either a heterogeneous population or an intermediate that is distinct from the NO-bound or deoxyferrous species. The former possibility is supported by a global fit to a 3-component serial mechanism showing that the putative intermediate is essentially in between the deoxyferrous species and ferrous-NO species (Figure S4d). A conformational change from a closed to open form of the protein may be required for NO binding. If in equilibrium, these two species would exhibit a fast phase for binding to the open conformation with a slower rate representing the change in equilibrium from the closed to open forms.[33] The oxidation state of the cysteines makes no significant difference in the rate of NO binding, either fast or slow phase, as seen in Figure S4e and S4f. Therefore, any effect of cysteine oxidation state on heme pocket dynamics does not hinder the entrance, and binding and dissociation of NO to the heme iron.

The five-coordinate NO-bound species found in this work are similar to those reported for cyt c', with the NO on the proximal side of the heme. The kinetics of NO binding to deoxyferrous Adgb-GD (Figure S4) with no observed hexacoordinate NO intermediate observed is consistent with the NO binding mechanism proposed for cyt c'.[14] Here NO can only bind to the distal iron location of cyt c' following a conformational change involving an occlusion in the heme pocket (by a Phe residue in the case of cyt c'). These two conformations of cyt c' are optically indistinguishable. Subsequent proximal histidine displacement, NO binding to the proximal side of the heme and NO dissociation from the distal side are rapid and not observed as separate events, but only as a single observed spectrokinetic event. This model may also be applied to the observations of NO binding to Adgb-GD. However, the kinetic trace for NO binding using stopped-flow is double exponential with Adgb but mono-exponential with cyt c'. A bi-molecular recombination of a ligand is often observed following laser flash photolysis due to an occluded/open conformational equilibrium model. In Adgb this equilibration appears slower, leading to the two distinct kinetics for NO binding observed in stopped flow on a slower

timescale. Hb has been reported to bind NO in a pentacoordinate form, but this is only observed at low NO to Hb ratios and only with the alpha chain.[34] At high ratios the NO is hexacoordinate when bound to the protein. This transition between penta- and hexa-coordinate was proposed to be sensitive to the quaternary structure conformation, such that the T-state promotes the breakage of the proximal histidine and the NO bound heme.[30, 35] Currently, we cannot discriminate between the NO bound to the proximal or distal side, but the proximal histidine-iron ligand in Adgb is more labile compared to other globins when NO is bound.

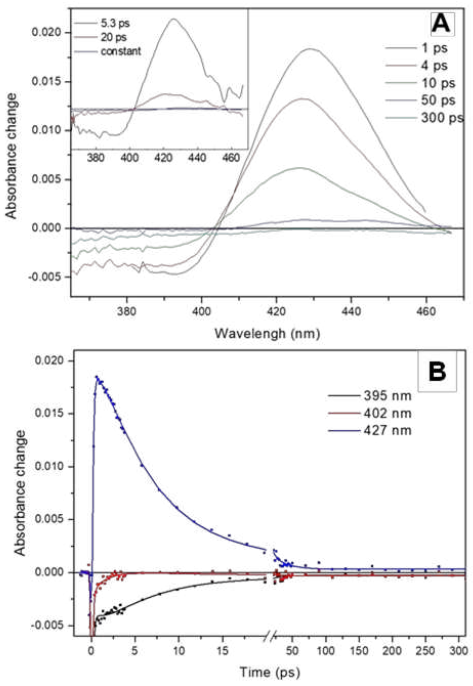

Figure 7. Ultrafast photo-dissociation and rebinding of NO from the ferrous Adgb-GD five coordinate NO complex. (a) Transient absorption spectra after different delay times upon excitation at 570 nm. Inset: Decay Associated Spectra corresponding to the NO geminate rebinding phases obtained from a global analysis (b) Dual-timescale kinetics and fits at selected wavelengths.

The transient absorption spectra observed following dissociation of the Adgb-GD:NO complex with a short light pulse is shown in Figure 7a. The spectra are characterized by a broad bleaching around 390 nm due to the disappearance of the 5-coordinate NO-bound state and a relatively strong induced absorption centered at 427 nm assigned to the 4-coordinate NO-dissociated state,[36, 37] supporting the EPR data (Figure 6d) that the NO bound state of the ferrous Adgb is pentacoordinate.

Apart from small relaxation signals with a time constant of ~1.5 ps, corresponding to a blue shift of the induced absorption band (Figure 7a inset) and assigned to vibrational cooling, the spectral evolution is characterized by a decay (associated spectra in the inset of Figure 7a) dominated by a 5.3 ps phase and a minor (~14 %) 20 ps phase (Figure 7b). The remaining spectrum after these phases corresponds to only ~1 % of the photo-dissociated NO, meaning that NO rebinding is almost completely geminate, and implying that dissociated NO stays within the confines of the heme pocket and only minor quantities of NO escape the heme pocket. Rebinding of NO to the heme iron from bulk solution outside the heme pocket would be expected to occur on the µs to ms timescale, as observed with Mb and Cygb for NO or other gases such as CO.[25, 38, 39] High-yield rebinding of NO to heme in a single 5-8 ps phase has been observed upon dissociation in all studied 5-coordinate heme-NO complexes in proteins thus far [40, 41, 42]). In Adgb, however a slower, 20 ps phase of NO binding is also present. This finding suggests a relaxation process competing with initial NO rebinding, allowing NO to explore a larger conformational space (rototranslational freedom), and indicating a less constrained heme pocket.

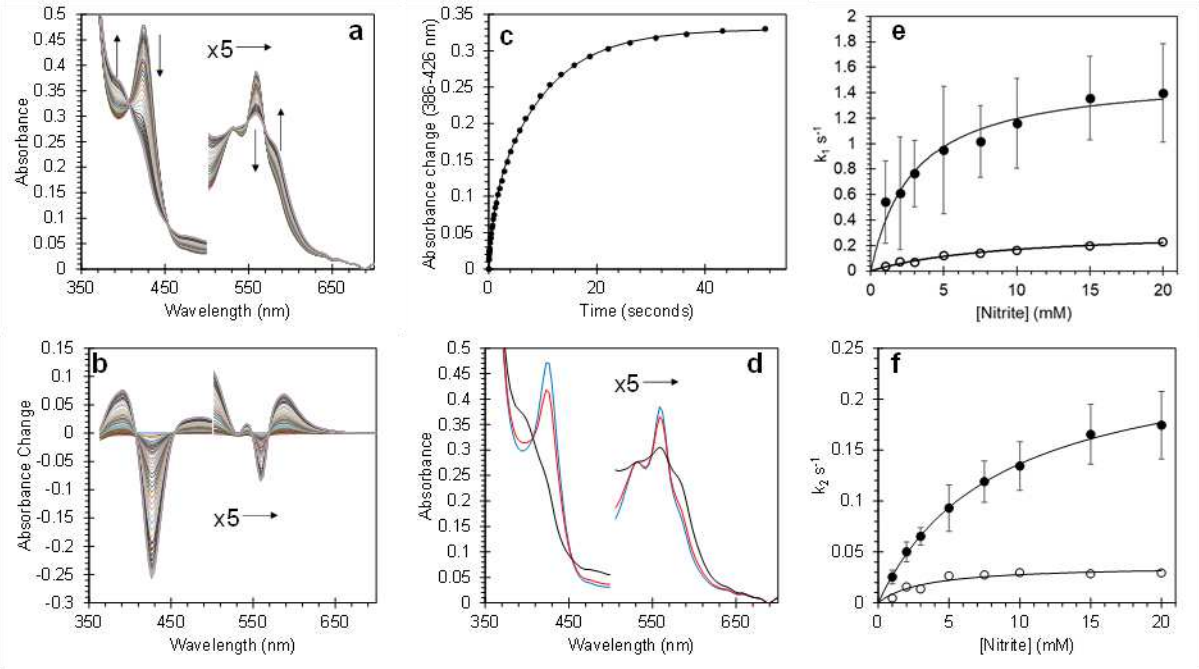

Figure 8. Nitrite reductase activity of Adgb-GD and effect of disulfide redox state on reductase activity. (a), Optical spectra of deoxyferrous Adgb-GD (5 μM) with sodium nitrite (5 mM). (b), Difference spectra with initial ferrous protein set to zero. (c), Time course of optical changes fitted to a double exponential function (k_1 = 9.25x10^{-1} s^{-1}, k_2= 1.02x10^{-1} s^{-1}. (d), Global fit of initial deoxyferrous protein (blue), intermediate (red) and final ferrous-NO bound spectrum (black). (e) and (f), The dependence of the rate constants for Adgb nitrite reductase activity. The observed rate constants of the fast (e) and slow (f) phase on the reaction in the presence (●) or absence (by reduction using TCEP, ○) of the CD disulfide bond.

Many heme proteins, including hemoglobins, are noted for their nitrite reductase (NiR) activity, giving potential functions in NO homeostasis depending on the oxygen concentration of their microenvironment.[25, 43, 44, 45] This reaction was assessed for Adgb-GD under anaerobic conditions (Figure 8).

The optical changes following this reaction are essentially identical to that of NO binding (Figure 8a and 8b; compare with Figure S4a and S4b) with a two stage kinetic trace (Figure 8c). The concentration dependence of the fast kinetics on nitrite (Figure 8e) exhibits a high error due to the small amplitude of the optical changes observed in the global fit (Figure 8d) but appears to follow a hyperbolic concentration dependence with an apparent K_D of 2.91 ± 0.57 mM NO_2^- and a maximum rate of 1.54 s^{-1} ± 0.01 s^{-1} (Figure 8e, ●). The slower rate representing the formation of the deoxyferrous-NO bound species also follows a hyperbola curve as a function of nitrite concentration with an apparent K_D of 8.42 ± 0.46 mM NO_2^- and a maximum rate of 2.54 x10^{-1} s^{-1} ± 6.1 x 10^{-4} s^{-1} (Figure 8f, ●). Unlike NO binding, the effect of reduction of the disulfide bond has a significant effect on the rate of NiR activity (Figure 8e and 8f) with TCEP reduced free sulfhydryl (○) exhibiting significantly decreased kinetics, both for the fast kinetics with an apparent K_D of 9.13 ± 1.38 mM NO_2^- and a maximum rate of 3.23 x10^{-1} s^{-1} ± 0.02 s^{-1} and slower NO binding kinetics of with an apparent K_D of 3.26 ± 1.12 mM $NO2^-$ and a maximum rate of 3.67 x10^{-2} s^{-1} ± 4.0 x 10^{-3} s^{-1}. This effect of sulfhydryl reduction on the NiR activity may arise from structural changes in the heme pocket, affecting exogenous ligand affinity, as observed for Ngb [31], or affecting the endogenous distal ligand off-rate, as observed for Ngb [46] and Cygb.[38]

The oxyform, which could not be generated without rapid autoxidation to ferric following dithionite removal by size-exclusion filtration, suggests that O_2 binding (indicated in the 2021 Uniprot entry) is not a physiological function of the protein except for the possibility of oxygen sensing. Similarly, NO dioxygenase activity is unlikely, as this requires a semi-stable ferrous-O_2 complex. The oxidation state of the iron *in vivo* remains unknown, like that of Ngb and Cygb. However, the oxygen tension in the testes is typically low (~10-15 mmHg), lower even than venal O_2 levels. [7, 47, 48] Consequently, the existence of the ferrous form of the heme iron *in vivo* cannot be discounted.

Recent studies have shown many globins appear to have properties relating to NO homeostasis, although the exact nature of this biochemistry *in vivo* is still under debate. Our results in Figure 8 confirm that Adgb-GD reacts with nitrite to generate NO under hypoxic/anoxic conditions and subsequently binds NO as a pentacoordinate species. The NiR rate constants of other globins are typically linear as a function of nitrite concentration.[25] Assuming that the rate of NO formation relates to the rate of ferrous-NO generation (Figure 8f), the initial slope of ~25 M^{-1} s^{-1} at low nitrite concentrations is higher than that for Ngb, Mb or Hb, and similar to that reported for globin X [43] and Cygb.[25, 49]

NO has important roles in the testes and hence NO homeostasis is important. Four Nitric Oxide Synthases (NOS) are present in the testes, endothelial NOS (eNOS), inducible NOS (iNOS), and neuronal NOS (nNOS) and a testis-specific nNOS (TnNOS).[15] NO has been proposed to play a unique role in modulating germ cell viability and development,[18] with high NO concentrations exhibiting a deleterious role in the mobility of spermatozoa [50] and thus acting on some aspects of male infertility.[15] This raises the possibility that Adgb could also play a role in NO homeostasis and hence germ cell viability. This ultimately depends on the functional properties of the other domains in Adgb and the interaction between the domains. However, the presence of a calpain-like motif on the N terminal side of the globin domain, as recorded in Uniprot, and a calmodulin binding domain directly dissecting the globin domain strongly suggest a role for calcium in the functional mechanism of this protein. The link between NO and calcium is well established.[51] Calcium channel blockers decrease intracellular calcium levels and increase the vasodilator efficacy of NO in smooth muscle.[52] Conversely, an increase in NO by vascular endothelial cells in the liver enhances calcium signaling in surrounding hepatocytes.[53] In the testes, calcium channel blockers to relieve hypertension causes reversible male infertility in mice.[54] Recently, the establishment of a role for Adgb in ciliogenesis has heightened the importance of Adgb. Overexpression studies show an Adgb-dependent increase in ciliated cells.[55] Expression of Adgb is in turn linked to Forkhead Box J1 (FOXJ1), a transcription factor involved in ciliogenesis, as overexpression of FOXJ1 directly led to increased Adgb mRNA levels through binding to the ADGB promotor.[55]

Conclusion

In summary, the helical alignment for Adgb-GD from our method, designed to work in the twilight zone, yielded an alternate helix alignment around the C and D helical region. This alignment is consistent with that obtained from the AlphaFold 2 model. Identification of a Tyr in the CD1 position (Tyr977 in the human Adgb) is to our knowledge unique for eukaryotic globins, but is common in prokaryotes such as that of truncated Hb from Mycobacterium tuberculosis. [22, 56]

Validation of our proposed helix alignment lies, at least in part, in the agreement with the AlphaFold-2 model and the generation of a stable recombinant form of the protein using the alternative globin domain sequence. This is supported from heme insertion and tight heme binding to generate spectra, optical and EPR, typical of pure authentic heme proteins. The femtosecond laser flash data indicates a true heme pocket from which ligand (NO) cannot escape and kinetics measured by stopped-flow that conform in general to known heme proteins. The presence of an intramolecular disulfide bond goes some way to explain the

stability of the recombinant protein, given that the N and C terminal regions of this circularly permuted globin domain are in the critical area of the heme-binding pocket.

A direct biochemical link between Adgb-GD and spermatogenesis or ciliogenesis is still to be determined; however, the evidence presented here illustrates the biochemical characteristics of the globin domain of Adgb, showing that the globin domain is capable of participating in NO sensing or regulation. This, together with the known calcium-linked structural aspects of calpain and calmodulin binding aspects Adgb, merit further investigation into the functional role of Adgb, given its potential medical significance.

Acknowledgements

BJR, DAS, MTW and CAR thank the Biotechnology and Biological Sciences Research Council (UK) (BB/T01802X/1) for financial support.

Keywords: helix alignment • molecular dynamics • homology modelling • nitric oxide homeostasis • disulfide

References

1. Hoogewijs D, Ebner B, Germani F, Hoffmann FG, Fabrizius A, Moens L*, et al.* Androglobin: a chimeric globin in metazoans that is preferentially expressed in Mammalian testes. *Mol Biol Evol* 2012, **29**(4)**:** 1105-1114.

2. Perutz MF. Structure of hemoglobin. *Brookhaven Symp Biol* 1960, **13:** 165-183.

3. Trent JT, 3rd, Watts RA, Hargrove MS. Human neuroglobin, a hexacoordinate hemoglobin that reversibly binds oxygen. *J Biol Chem* 2001, **276**(32)**:** 30106-30110.

4. Burmester T, Ebner B, Weich B, Hankeln T. Cytoglobin: a novel globin type ubiquitously expressed in vertebrate tissues. *Mol Biol Evol* 2002, **19**(4)**:** 416-421.

5. Zhu H, Riggs AF. Yeast flavohemoglobin is an ancient protein related to globins and a reductase family. *Proc Natl Acad Sci U S A* 1992, **89**(11)**:** 5015-5019.

6. Huang B, Lu YS, Li X, Zhu ZC, Li K, Liu JW*, et al.* Androglobin knockdown inhibits growth of glioma cell lines. *Int J Clin Exp Pathol* 2014, **7**(5)**:** 2179-2184.

7. Free MJ, Schluntz GA, Jaffe RA. Respiratory gas tensions in tissues and fluids of the male rat reproductive tract. *Biol Reprod* 1976, **14**(4)**:** 481-488.

8. Bracke A, Hoogewijs D, Dewilde S. Exploring three different expression systems for recombinant expression of globins: Escherichia coli, Pichia pastoris and Spodoptera frugiperda. *Anal Biochem* 2018, **543:** 62-70.

9. Lock A, Forfar R, Weston C, Bowsher L, Upton GJ, Reynolds CA*, et al.* One motif to bind them: A small-XXX-small motif affects transmembrane domain 1 oligomerization, function, localization, and cross-talk between two yeast GPCRs. *Biochim Biophys Acta* 2014, **1838**(12)**:** 3036-3051.

10. Taddese B, Upton GJ, Bailey GR, Jordan SR, Abdulla NY, Reeves PJ*, et al.* Do plants contain G protein-coupled receptors? *Plant Physiol* 2014, **164**(1)**:** 287-307.

11. Watkins HA, Chakravarthy M, Abhayawardana RS, Gingell JJ, Garelja M, Pardamwar M*, et al.* Receptor Activity-modifying Proteins 2 and 3 Generate

Adrenomedullin Receptor Subtypes with Distinct Molecular Properties. *J Biol Chem* 2016, **291**(22)**:** 11657-11675.

12. Cutruzzola F, Arcovito A, Giardina G, della Longa S, D'Angelo P, Rinaldo S. Distal-proximal crosstalk in the heme binding pocket of the NO sensor DNR. *Biometals* 2014, **27**(4)**:** 763-773.

13. Lawson DM, Stevenson CE, Andrew CR, Eady RR. Unprecedented proximal binding of nitric oxide to heme: implications for guanylate cyclase. *EMBO J* 2000, **19**(21)**:** 5661-5671.

14. Manole A, Kekilli D, Svistunenko DA, Wilson MT, Dobbin PS, Hough MA. Conformational control of the binding of diatomic gases to cytochrome c'. *J Biol Inorg Chem* 2015, **20**(4)**:** 675-686.

15. Lee NP, Cheng CY. Nitric oxide and cyclic nucleotides: their roles in junction dynamics and spermatogenesis. *Adv Exp Med Biol* 2008, **636:** 172-185.

16. Sataric MV, Zdravkovic S, Nemes T, Sataric BM. Calcium signaling modulates the dynamics of cilia and flagella. *Eur Biophys J* 2020, **49**(7)**:** 619-631.

17. Saternos HC, AbouAlaiwi WA. Signaling interplay between primary cilia and nitric oxide: A mini review. *Nitric Oxide* 2018, **80:** 108-112.

18. Zini A, O'Bryan MK, Magid MS, Schlegel PN. Immunohistochemical localization of endothelial nitric oxide synthase in human testis, epididymis, and vas deferens suggests a possible role for nitric oxide in spermatogenesis, sperm maturation, and programmed cell death. *Biol Reprod* 1996, **55**(5)**:** 935-941.

19. Hargrove MS, Krzywda S, Wilkinson AJ, Dou Y, Ikeda-Saito M, Olson JS. Stability of myoglobin: a model for the folding of heme proteins. *Biochemistry* 1994, **33**(39)**:** 11767-11775.

20. Roesner A, Fuchs C, Hankeln T, Burmester T. A globin gene of ancient evolutionary origin in lower vertebrates: evidence for two distinct globin families in animals. *Mol Biol Evol* 2005, **22**(1)**:** 12-20.

21. Henderson SJ, Timbs AT, McCarthy J, Gallienne AE, Proven M, Rugless MJ*, et al.* Ten Years of Routine alpha- and beta-Globin Gene Sequencing in UK Hemoglobinopathy Referrals Reveals 60 Novel Mutations. *Hemoglobin* 2016, **40**(2)**:** 75-84.

22. Ouellet H, Milani M, LaBarre M, Bolognesi M, Couture M, Guertin M. The roles of Tyr(CD1) and Trp(G8) in Mycobacterium tuberculosis truncated hemoglobin O in ligand binding and on the heme distal site architecture. *Biochemistry* 2007, **46**(41)**:** 11440-11450.

23. Mitchell EM, Artymiuk PJ, Rice DW, Willett P. Use of Techniques Derived from Graph-Theory to Compare Secondary Structure Motifs in Proteins. *Journal of Molecular Biology* 1990, **212**(1)**:** 151-166.

24. Li H, Hemann C, Abdelghany TM, El-Mahdy MA, Zweier JL. Characterization of the mechanism and magnitude of cytoglobin-mediated nitrite reduction and nitric oxide generation under anaerobic conditions. *J Biol Chem* 2012, **287**(43)**:** 36623-36633.

25. Reeder BJ, Ukeri J. Strong modulation of nitrite reductase activity of cytoglobin by disulfide bond oxidation: Implications for nitric oxide homeostasis. *Nitric Oxide* 2018, **72:** 16-23.

26. Blumberg WE, Peisach J. A unified theory for low spin forms of all ferric heme proteins as studied by EPR. In: Chance B, Yonetani T, Mildvan AS (eds). *Probes of structure and function of macromolecules and membranes*, vol. 2. Acedemic Press: New York, 1971, pp 215-229.

27. Droghetti E, Howes BD, Feis A, Dominici P, Fittipaldi M, Smulevich G. The quantum mechanically mixed-spin state in a non-symbiotic plant hemoglobin: the effect of distal mutation on AHb1 from Arabidopsis thaliana. *J Inorg Biochem* 2007, **101**(11-12)**:** 1812-1819.

28. Maltempo MM, Moss TH, Cusanovich MA. Magnetic studies on the changes in the iron environment in Chromatium ferricytochrome c'. *Biochim Biophys Acta* 1974, **342**(2)**:** 290-305.

29. Bellei M, Bortolotti CA, Di Rocco G, Borsari M, Lancellotti L, Ranieri A*, et al.* The influence of the Cys46/Cys55 disulfide bond on the redox and spectroscopic properties of human neuroglobin. *J Inorg Biochem* 2018, **178:** 70-86.

30. Hamdane D, Kiger L, Dewilde S, Green BN, Pesce A, Uzan J*, et al.* Coupling of the heme and an internal disulfide bond in human neuroglobin. *Micron* 2004, **35**(1-2)**:** 59-62.

31. Vinck E, Van Doorslaer S, Dewilde S, Moens L. Structural change of the heme pocket due to disulfide bridge formation is significantly larger for neuroglobin than for cytoglobin. *J Am Chem Soc* 2004, **126**(14)**:** 4516-4517.

32. Zhao Y, Hoganson C, Babcock GT, Marletta MA. Structural changes in the heme proximal pocket induced by nitric oxide binding to soluble guanylate cyclase. *Biochemistry* 1998, **37**(36)**:** 12458-12464.

33. Trent JT, 3rd, Hvitved AN, Hargrove MS. A model for ligand binding to hexacoordinate hemoglobins. *Biochemistry* 2001, **40**(20)**:** 6155-6163.

34. Fago A, Crumbliss AL, Peterson J, Pearce LL, Bonaventura C. The case of the missing NO-hemoglobin: spectral changes suggestive of heme redox reactions reflect changes in NO-heme geometry. *Proc Natl Acad Sci U S A* 2003, **100**(21)**:** 12087-12092.

35. Yonetani T, Tsuneshige A, Zhou Y, Chen X. Electron paramagnetic resonance and oxygen binding studies of alpha-Nitrosyl hemoglobin. A novel oxygen carrier having no-assisted allosteric functions. *J Biol Chem* 1998, **273**(32)**:** 20323-20333.

36. Negrerie M, Bouzhir L, Martin JL, Liebl U. Control of nitric oxide dynamics by guanylate cyclase in its activated state. *J Biol Chem* 2001, **276**(50)**:** 46815-46821.

37. Silkstone G, Kapetanaki SM, Husu I, Vos MH, Wilson MT. Nitric oxide binds to the proximal heme coordination site of the ferrocytochrome c/cardiolipin complex: formation mechanism and dynamics. *J Biol Chem* 2010, **285**(26)**:** 19785-19792.

38. Beckerson P, Reeder BJ, Wilson MT. Coupling of disulfide bond and distal histidine dissociation in human ferrous cytoglobin regulates ligand binding. *FEBS Lett* 2015, **589**(4)**:** 507-512.

39. Romberg RW, Kassner RJ. Nitric oxide and carbon monoxide equilibria of horse myoglobin and (N-methylimidazole)protoheme. Evidence for steric interaction with the distal residues. *Biochemistry* 1979, **18**(24)**:** 5387-5392.

40. Kruglik SG, Lambry JC, Cianetti S, Martin JL, Eady RR, Andrew CR*, et al.* Molecular basis for nitric oxide dynamics and affinity with Alcaligenes xylosoxidans cytochrome c. *J Biol Chem* 2007, **282**(7)**:** 5053-5062.

41. Liebl U, Lambry JC, Vos MH. Primary processes in heme-based sensor proteins. *Biochim Biophys Acta* 2013, **1834**(9)**:** 1684-1692.

42. Lobato L, Bouzhir-Sima L, Yamashita T, Wilson MT, Vos MH, Liebl U. Dynamics of the heme-binding bacterial gas-sensing dissimilative nitrate respiration regulator (DNR) and activation barriers for ligand binding and escape. *J Biol Chem* 2014, **289**(38)**:** 26514-26524.

43. Corti P, Xue J, Tejero J, Wajih N, Sun M, Stolz DB*, et al.* Globin X is a six-coordinate globin that reduces nitrite to nitric oxide in fish red blood cells. *Proc Natl Acad Sci U S A* 2016, **113**(30)**:** 8538-8543.

44. Gladwin MT, Kim-Shapiro DB. The functional nitrite reductase activity of the heme-globins. *Blood* 2008, **112**(7)**:** 2636-2647.

45. Jensen FB. The role of nitrite in nitric oxide homeostasis: a comparative perspective. *Biochim Biophys Acta* 2009, **1787**(7)**:** 841-848.

46. Tiso M, Tejero J, Basu S, Azarov I, Wang X, Simplaceanu V*, et al.* Human neuroglobin functions as a redox-regulated nitrite reductase. *J Biol Chem* 2011, **286**(20)**:** 18277-18289.

47. Massie ED, Gomes WR, VanDemark NL. Oxygen tension in testes of normal and cryptorchid rats. *J Reprod Fertil* 1969, **19**(3)**:** 559-561.

48. Reyes JG, Farias JG, Henriquez-Olavarrieta S, Madrid E, Parraga M, Zepeda AB*, et al.* The hypoxic testicle: physiology and pathophysiology. *Oxid Med Cell Longev* 2012, **2012:** 929285.

49. Corti P, Ieraci M, Tejero J. Characterization of zebrafish neuroglobin and cytoglobins 1 and 2: Zebrafish cytoglobins provide insights into the transition from six-coordinate to five-coordinate globins. *Nitric Oxide* 2016, **53:** 22-34.

50. Balercia G, Moretti S, Vignini A, Magagnini M, Mantero F, Boscaro M*, et al.* Role of nitric oxide concentrations on human sperm motility. *J Androl* 2004, **25**(2)**:** 245-249.

51. Muller U, Bicker G. Calcium-activated release of nitric oxide and cellular distribution of nitric oxide-synthesizing neurons in the nervous system of the locust. *J Neurosci* 1994, **14**(12)**:** 7521-7528.

52. Van Hove CE, Van der Donckt C, Herman AG, Bult H, Fransen P. Vasodilator efficacy of nitric oxide depends on mechanisms of intracellular calcium mobilization in mouse aortic smooth muscle cells. *Br J Pharmacol* 2009, **158**(3)**:** 920-930.

53. Charles A. Nitric oxide pumps up calcium signalling. *Nat Cell Biol* 1999, **1**(8)**:** E193-195.

54. Lee JH, Kim H, Kim DH, Gye MC. Effects of calcium channel blockers on the spermatogenesis and gene expression in peripubertal mouse testis. *Arch Androl* 2006, **52**(4)**:** 311-318.

55. Koay TW, Osterhof C, Orlando IMC, Keppner A, Andre D, Yousefian S*, et al.* Androglobin gene expression patterns and FOXJ1-dependent regulation indicate its functional association with ciliogenesis. *J Biol Chem* 2021, **296:** 100291.

56. Ouellet H, Juszczak L, Dantsker D, Samuni U, Ouellet YH, Savard PY*, et al.* Reactions of Mycobacterium tuberculosis truncated hemoglobin O with ligands reveal a novel ligand-inclusive hydrogen bond network. *Biochemistry* 2003, **42**(19)**:** 5764-5774.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupportingInformationv4.pdf
- adgbcomplex1MDmovie.mpg
- adgbcomplex2MDmvovie.mpg