2013

Structural and Functional Characterization of a Permuted Hemoglobin

Michael P. (Michael Patrick) Murphy
Western Washington University

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF A PERMUTED HEMOGLOBIN

By

Michael Patrick Murphy

Accepted in Partial Completion

Of the Requirement for the Degree

Master of Science

Kathleen L. Kitto, Dean of the Graduate School

ADVISORY COMMITTEE

Chair, Dr. Spencer Anthony-Cahill

Dr. P. Clint Spiegel

Dr. Serge Smirnov

MASTER'S THESIS

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Western Washington University, I grant to Western Washington University the non-exclusive royalty-free right to archive, reproduce, distribute, and display the thesis in any and all forms, including electronic format, via any digital library mechanisms maintained by WWU.

I represent and warrant this is my original work, and does not infringe or violate any rights of others. I warrant that I have obtained written permissions from the owner of any third party copyrighted material included in these files.

I acknowledge that I retain ownership rights to the copyright of this work, including but not limited to the right to use all or part of this work in future works, such as articles or books.

Library users are granted permission for individual, research and non-commercial reproduction of this work for educational purposes only. Any further digital posting of this document requires specific permission from the author.

Any copying or publication of this thesis for commercial purposes, or for financial gain, is not allowed without my written permission.

Michael Patrick Murphy

November 21, 2013

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF A PERMUTED HEMOGLOBIN

A Thesis
Presented to
The Faculty of
Western Washington University

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

by
Michael Patrick Murphy
November 2013

Abstract:

The tetrameric composition of human hemoglobin complicates protein engineering efforts that are required to improve its potential as an oxygen-carrying therapeutic. In our research to design a single-chain version of the hemoglobin molecule (scHb), we have co-expressed a circularly permuted human β-globin (cp-β) with human α-globin. At micromolar concentrations, the purified recombinant globins appear to associate to form an α-cpβ heterodimer in solution rather than the expected α_2-cpβ_2 heterotetramer. Compared to recombinant human hemoglobin, the α-cpβ heterodimer exhibits a stronger ligand binding affinity. Knowledge of the intermolecular interactions favoring formation of the α-cpβ heterodimer will be instrumental in understanding the global structural consequences of the cpβ mutation, and in directing future protein engineering efforts to optimize the function of permuted hemoglobins. X-ray diffraction of α-cpβ crystals has been employed to determine the molecular structure of this protein complex at near-atomic resolution (2.7 Å). Examination of the structure shows that, in the crystal, the subunits associate to form a heterotetramer similar to that of wild-type hemoglobin in the high affinity "R-state". The structure also confirms the incorporation of the β-globin His146 (wild-type numbering) into the cpβ linker, which removes an important ionic interaction that stabilizes the low oxygen affinity (T state) conformation. Structural information obtained as a result of this work will guide future protein engineering efforts to enhance cooperativity of oxygen binding and the T-state stability of future hemoglobin constructs.

ACKNOWLEDGEMENTS

I would like to express my deep and abiding gratitude to:

My advisor, Professor Spencer Anthony-Cahill for your excellent mentorship, humor, and patience;

Professor P. Clint Spiegel for guiding me through crystallography and structure refinement;

Justin Walters for his tutelage in crystallography techniques;

The Fred Hutchison Cancer Research Center for use of their facilities;

Professor John Olson at Rice University for providing the rHb0.0 clone, allowing us access to his laboratory for the ligand binding studies, and for the advice on the data analysis and interpretation;

All of the past and present members of Team Spencer, especially Ngoc-An Huyn, Jen Liddle, and Matt Lockett;

Erin Macri for her help with the mass spectrometry;

And the Murphy clan for their unwavering love, encouragement, and patience, especially my parents Kevin and Lynette, my brother Shaun, my grandparents Don, Carol, Martha and Stu, my uncle Brent, my uncle Bruce, and my godparents Tim and Stacy.

This work was supported by an AREA award from the National Institute of Health (#R15 HL081068-01).

TABLE OF CONTENTS

Abstract	iv
Acknowledgements	v
List of Figures	vii
List of Tables	ix
Introduction	1
Materials and Methods	23
Results and Discussion	33
Conclusions	69
References	71
Appendix	

LIST OF FIGURES

		Page
Figure 1.	Whole blood and red blood cell transfusions and donations from 1989-2004	1
Figure 2.	Adverse reactions from transfusions 2006-2008 in the US	2
Figure 3.	Hemoglobin-based oxygen carriers	3
Figure 4.	Oxygen dissociation curves for various HBOCs	4
Figure 5.	Cartoon diagram of the deoxy- "T" and oxy- "R" states of hemoglobin	5
Figure 6.	Somatogen hemoglobin permutant development process	6
Figure 7.	Change in peritoneal hemoglobin concentration over time for Somatogen permutants	7
Figure 8.	Change in mean arterial pressure in conscious rats following injection of permutants	7
Figure 9.	Illustration of the $\alpha_1\beta_1$ interface and the $\alpha_1\beta_2$ interface	8
Figure 10.	Illustration of the distances between the termini within hemoglobin	9
Figure 11.	Illustration of the process of circular permutation of the β-globin	10
Figure 12.	Illustration of the external loops of the α_1- & β_1-globins and their relative proximity	11
Figure 13.	Comparison of the sequences of wtHb and α-cpβ	11
Figure 14.	Scheme for the development of a single-chain hemoglobin	12
Figure 15.	Proposed steps for fusion of the four globin subunits into a single-chain hemoglobin	12
Figure 16.	Oxygen binding equilibrium curves for HbA and α-cpβ	14
Figure 17.	Oxygen binding affinity profiles for myoglobin and hemoglobin	14
Figure 18.	HPSEC data showing the relative proportion of hemoglobin species for α-cpβ	15
Figure 19.	Illustration of important interactions that stabilize the T-state of hemoglobin	17
Figure 20.	Protein engineering schemes to increase the stability of T-state hemoglobin	18
Figure 21.	Proton NMR spectra of deoxy form hemoglobin permutants	19
Figure 22.	Proton NMR spectra of carbonmonoxy form hemoglobins	19
Figure 23.	Oxygen binding equilibrium curves for HbA and engineered permutants	20
Figure 24.	Structure determination by the X-ray crystallography process	21
Figure 25.	SDS-PAGE of α-cpβ proteins	37
Figure 26.	ESI-MS of α-cpβ at high purity	38
Figure 27.	ESI-MS of α-cpβ-hug at high purity	38

Figure 28. SDS-PAGE of cyano-met isoform hemoglobins 39
Figure 29. SDS-PAGE gel of α-cpβ and diα-cpβ at normal and high loadings 40
Figure 30. Image of stacked 24-well crystallization trays 41
Figure 31. General layout of a gradient tray 43
Figure 32. Tray 11 crystal results (100x light microscope) for α-cpβ 44
Figure 33. Tray 12 crystal results (100x light microscope) for α-cpβ 44
Figure 34. Tray 2x crystal results (100x light microscope) for α-cpβ 45
Figure 35. Tray 7x crystal results (100x light microscope) for α-cpβ 45
Figure 36. Tray 9x crystal results (100x light microscope) for α-cpβ 46
Figure 37. Tray 10x crystal results (100x light microscope) for α-cpβ 46
Figure 38. Tray 11x crystal results (100x light microscope) for α-cpβ 47
Figure 39. Tray 12x crystal results (100x light microscope) for α-cpβ 47
Figure 40. Example of birefringence of the polystyrene trays 48
Figure 41. Examples of α-cpβ crystals under normal light and polarized light 49
Figure 42. Examples of diα-cpβ crystals under normal light and polarized light 49
Figure 43. Diffraction patterns of α-cpβ crystals at lower resolution 53
Figure 44. Diffraction pattern image from the highest resolution diffraction data set for α-cpβ 54
Figure 45. Histograms of relevant validation parameters for the final α-cpβ model 59
Figure 46. Cartoon model of the final α-cpβ structural model 63
Figure 47. Overlaid cartoon models of the final α-cpβ structure and wtHb 63
Figure 48. Electron density and stick model of the final α-cpβ structure of the cpβ linker 64
Figure 49. Structure refinement view of one cpβ-globin heme-binding pocket 65
Figure 50. Structure refinement view of the cpβ-globin termini 65
Figure 51. Structure refinement view of the α–globin termini 66
Figure 52. Structural homology of cpβ-globin to the oxy and deoxy wtHb structures 67
Figure 53. Structural homology of α-cpβ α-globins to the oxy and deoxy wtHb globin structures 68
Figure 54. Final α-cpβ structure with future fusion sites between the α_1 and $cp\beta_1$ globins 68

LIST OF TABLES

		Page
Table 1.	Sedimentation equilibrium study results for wtHb and α-cpβ	16
Table 2.	IMAC wash steps	25
Table 3.	IEC purification protocol	26
Table 4.	SEC purification protocol	27
Table 5.	Growth optimization data for hemoglobin mutants	34
Table 6.	Expected and observed MWs for α-cpβ calculated from ESI-MS m/z ratios	38
Table 7.	Crystallization conditions reported for hemoglobin structures in the PDB	42
Table 8.	Gradient crystallization conditions extrapolated from sparse matrix hits for α-cpβ	43
Table 9.	Tray origin of α-cpβ crystals and cane position for X-ray diffraction	51
Table 10.	Crystal diffraction information for α-cpβ crystals	52
Table 11.	Progress of α-cpβ model refinement	60
Table 12.	X-ray data collection and refinement statistics for the final α-cpβ model	61
Table 13.	Example structures collected from the PDB of different R and T state configurations	64

Introduction:

The American Association of Blood Banks has predicted an acute shortage of packed red blood cells (RBCs) by 2030 as demand for blood is expected to surpass the rate at which it can be supplied [1] if trends continue as illustrated in Figure 1. This will increase the cost of medical procedures and prevent patients from receiving life-saving transfusions. In an attempt to mitigate the detrimental effects of a blood shortage, and reduce instances of adverse reactions to transfusions (Figure 2), efforts are underway in many labs to develop hemoglobin-based blood substitutes to augment the supply of donor-derived blood.

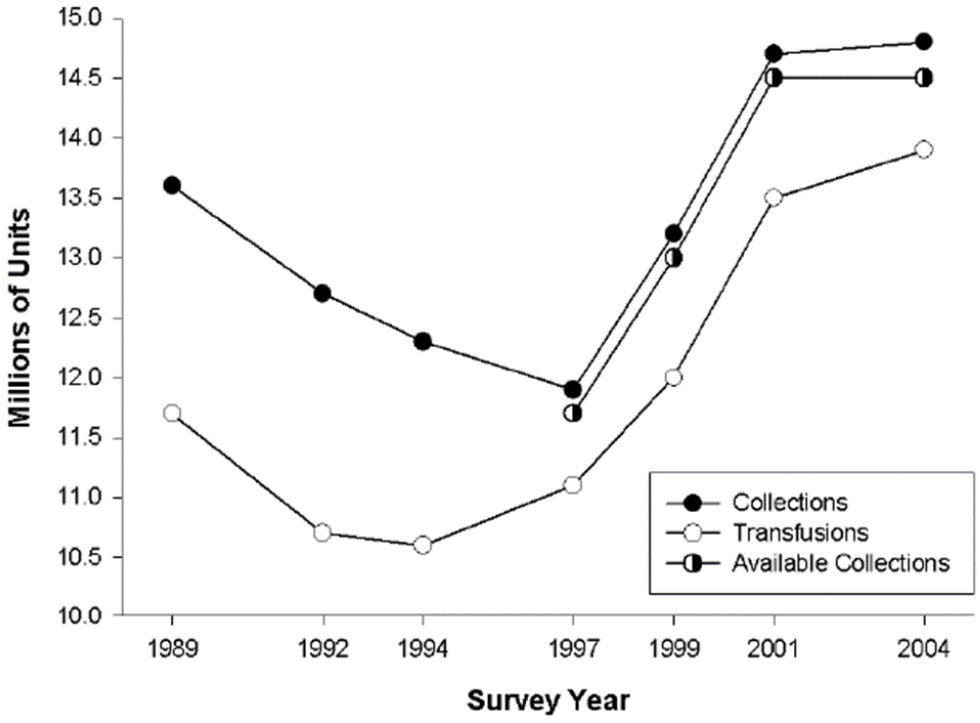

Figure 1. Whole blood and red blood cell transfusions and donations from 1989-2004. Available collections represent those units that have been approved for administration [1].

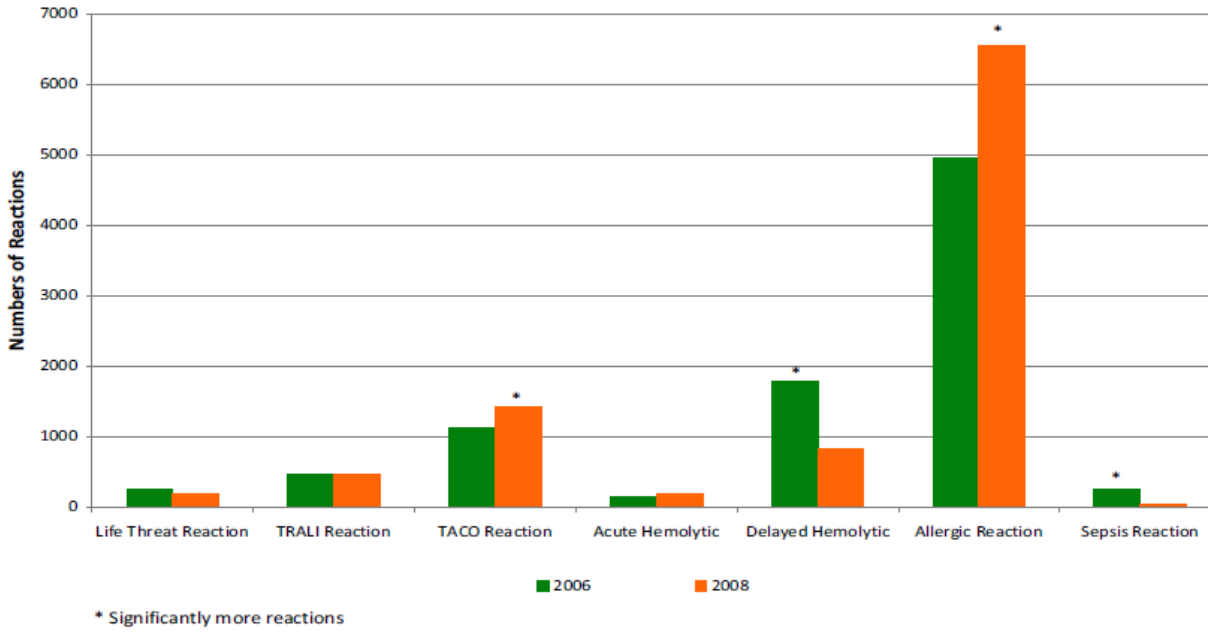

Figure 2. Adverse reactions from transfusions 2006-2008 in the US [2].

Many approaches have been taken to develop a blood replacement, unfortunately none of these emerging therapeutics has been proven sufficiently safe or effective to supplement the donor-derived blood supply. We believe that our approach to developing a hemoglobin-based blood substitute will aid efforts to develop a safe, effective, and economically viable blood replacement therapy. A long standing goal of many in private and public enterprise has been the development of a cell-free hemoglobin-based oxygen carrier (HBOC) that can be readily manufactured in large quantities, is stable, and has functional properties equal to or exceeding those of red blood cells. HBOCs are excellent candidates to fill the gap between the supply and demand for blood because they are derived from a naturally-occurring oxygen transporter. Representative HBOCs are illustrated in Figure 3.

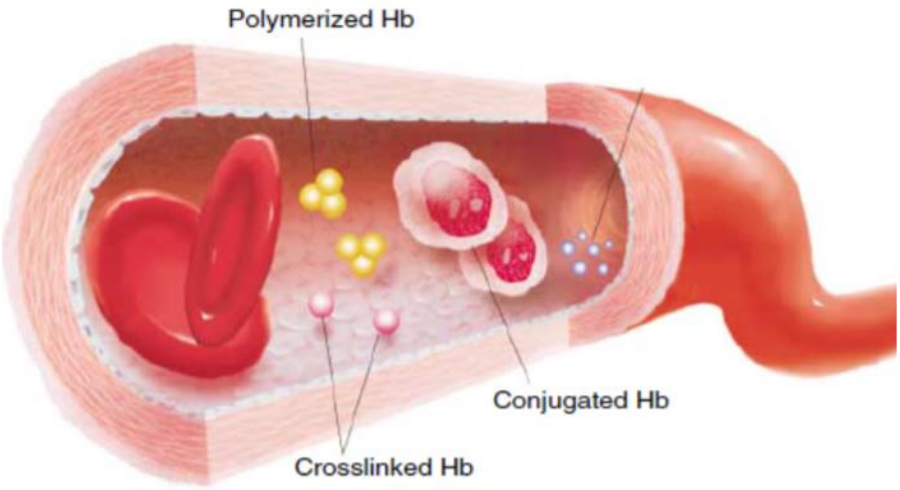

Figure 3. Hemoglobin-based oxygen carriers [3]

Native human hemoglobin is a soluble globular heme-protein composed of four subunits: two α- and two β-globins. Heme is a tetrapyrrole porphyrin macrocycle that coordinates an iron ion and enables each globin subunit to bind an oxygen ligand. The globins are composed primarily of α-helices, with 8 helices per β-globin (denoted A-H) and 7 homologous helices per α-globin (denoted A-C and E-H). These globin subunits spontaneously associate into αβ-heterodimers, through the formation of inter-subunit hydrogen bonds and salt-bridges. Two αβ–heterodimers associate to form the $\alpha_2\beta_2$ heterotetramer (i.e., native hemoglobin is a dimer of heterodimers). Interactions between the two associated heterodimers enable hemoglobin to bind oxygen cooperatively, a function that is essential for efficient oxygen transport. Figure 4 illustrates the oxygen binding affinity profiles for a variety of HBOCs; αα-crosslinked Hb is the only HBOC with a profile similar to that of human red blood cells [4].

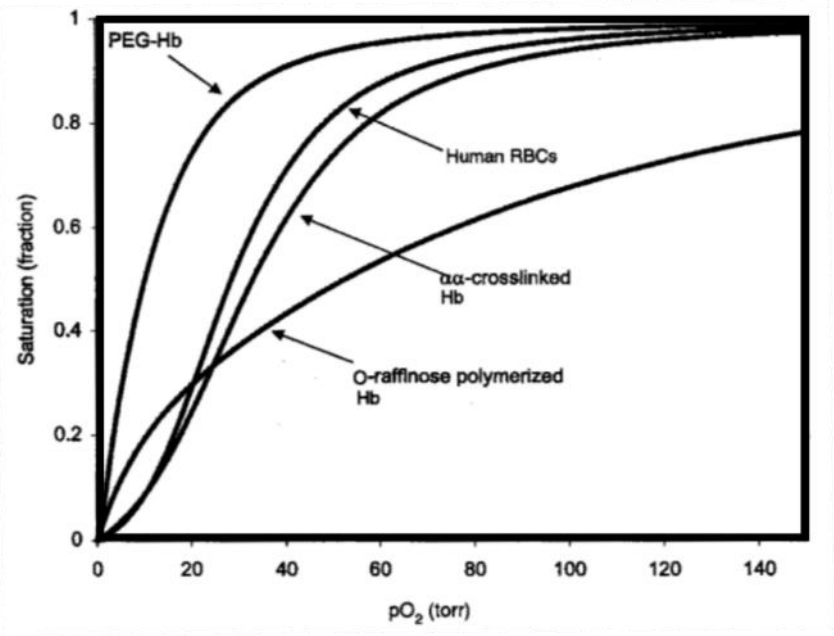

Figure 4. Oxygen dissociation curves for various HBOCs. Curves are shown for human red blood cells, an αα cross-linked Hb (a covalently cross-linked HBOC generated by recombinant DNA technology), and other oxygen carrying therapeutics based on chemical modification of native hemoglobin [4].

Hemoglobin's oxygen binding cooperativity is mediated by the allosteric shift that occurs upon binding an oxygen ligand to the low oxygen affinity, or deoxy-Hb, conformation. Oxygen binding to the heme iron induces strain, which is relieved by movement of the F-helix and the associated remodeling of the interface between the α/β heterodimers. These structural shifts induce strain in adjacent globin subunits, which favors a global change in the conformation of the hemoglobin tetramer into the high oxygen binding affinity, or R state, conformation. When oxygen concentrations are lowered, these structural shifts are reversed, oxygen ligands are released, and all subunits cooperatively return to the lower affinity state. This mechanism enables efficient oxygen transport between areas of high and low oxygen concentration. The R and T states, illustrated in Figure 5, are characterized by changes in the relative orientations between of the two αβ-heterodimers.

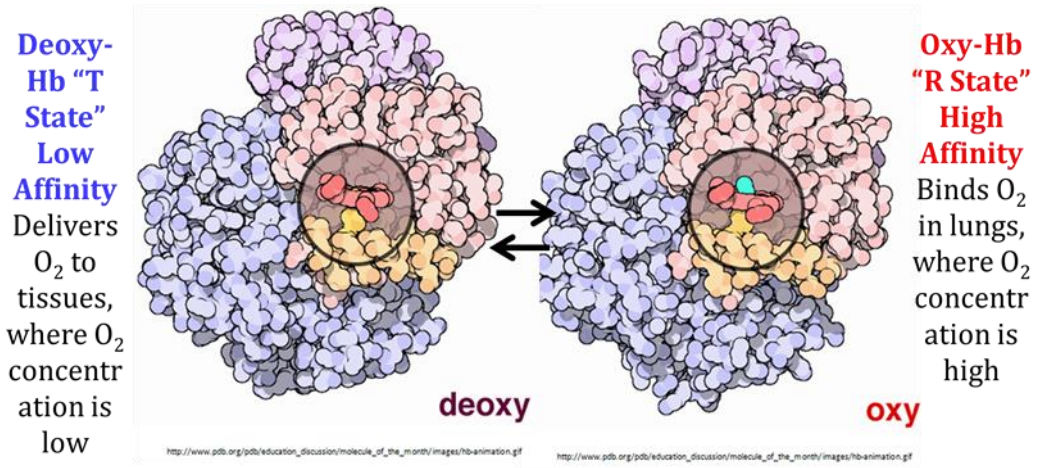

Figure 5. Cartoon diagram of the deoxy- "T" and oxy- "R" states of hemoglobin. The ligand binding heme subunit is denoted in red and encircled by a black ring at the center of each diagram. An oxygen molecule can be seen bound to the heme ligand in the oxy-hemoglobin diagram. From [5], used without permission.

Direct administration of wild type hemoglobin (wtHb) as a therapeutic would be the simplest method to restore oxygen-carrying capacity, unfortunately wtHb cannot be used directly as a cell-free therapeutic due to toxic side effects. Cell-free Hb exhibits significantly lower stability compared to Hb contained within erythrocytes (red blood cells, or RBCs). Outside of RBCs the Hb tetramer dissociates into αβ-heterodimers, which are small enough to be filtered by the kidneys, where the subsequent release of heme in unnaturally high amounts [6] causes renal failure. Additionally, cell-free oxy-hemoglobin scavenges nitric oxide (NO), a potent vasodilator, resulting in increased in blood pressure through arterial smooth muscle contraction.

Renal toxicity can be ameliorated by preventing the dissociation of the 64 kDa heterotetramer into 32 kDa heterodimers. The rHb1.1 molecule, illustrated in Figure 6, showed no renal toxicity in Phase I and II safety trials, and is no larger than wt Hb [7], [8]. It

has also been reported that larger oligomeric Hbs exhibit reduced blood pressure elevation ("vasoactivity" or "pressor response"), which is a significant undesirable side effect [8] (see Figures 7 and 8). Work in our lab is aimed at addressing the vasoactivity of cell-free Hb through protein engineering. Specifically, we have covalently linked the four globin subunits to make a hemoglobin composed of a single polypeptide chain (single-chain hemoglobin or scHb). We plan to generate polymeric hemoglobins of defined molecular weight using the scHb as a protomer. We hypothesize that such engineered HBOCs will exhibit a reduction in pressor response and enhancement in therapeutic value as was observed for 130 kDa and 260 kDa recombinant hemoglobins developed by Somatogen [8].

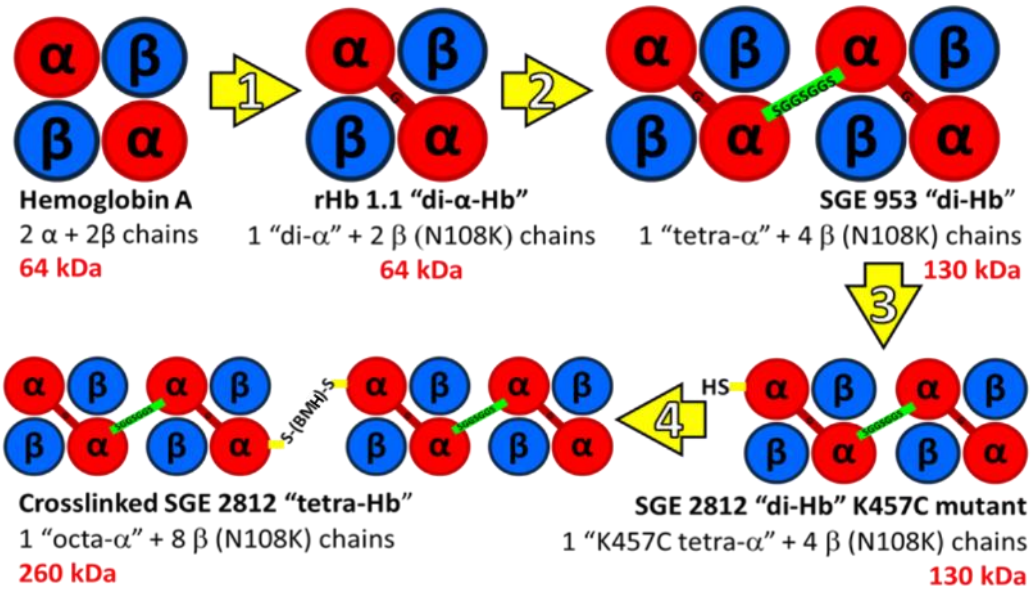

Figure 6. Previous work performed at Somatogen to increase the size of hemoglobin through covalent linkages [8]. Red circles with 'α' symbol represent α-globins, blue circles with 'β' represent β-globins, colored lines connecting circles represent covalent linkages, and 'HS-' represents the addition of a reactive thiol moiety by mutation of lysine 457 to cysteine.

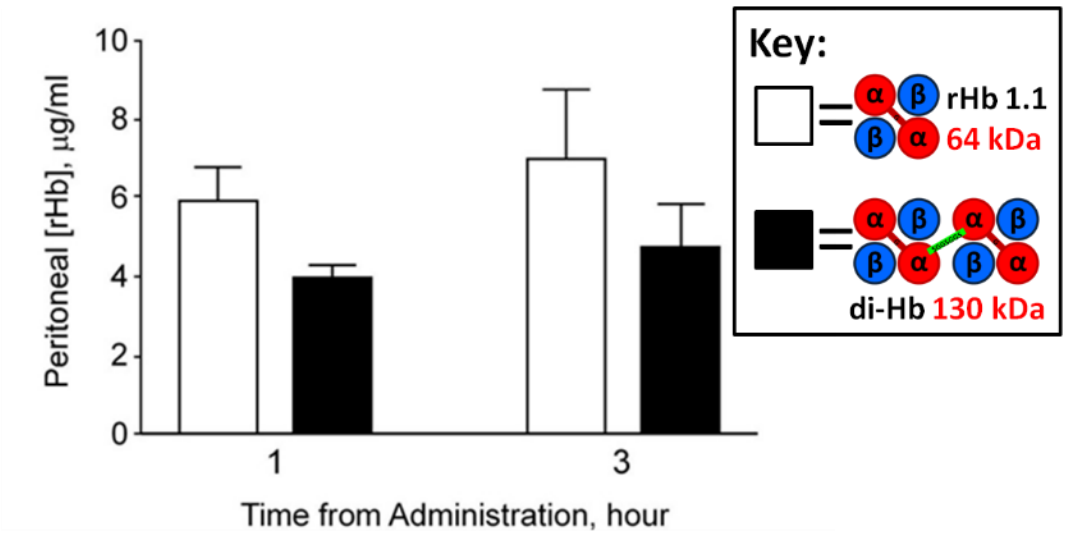

Figure 7. Change in peritoneal hemoglobin concentration as a function of time and size of hemoglobin molecule. Open bars are for a 64 kDa hemoglobin and the filled bars are for a 130 kDa polymeric hemoglobin [8].

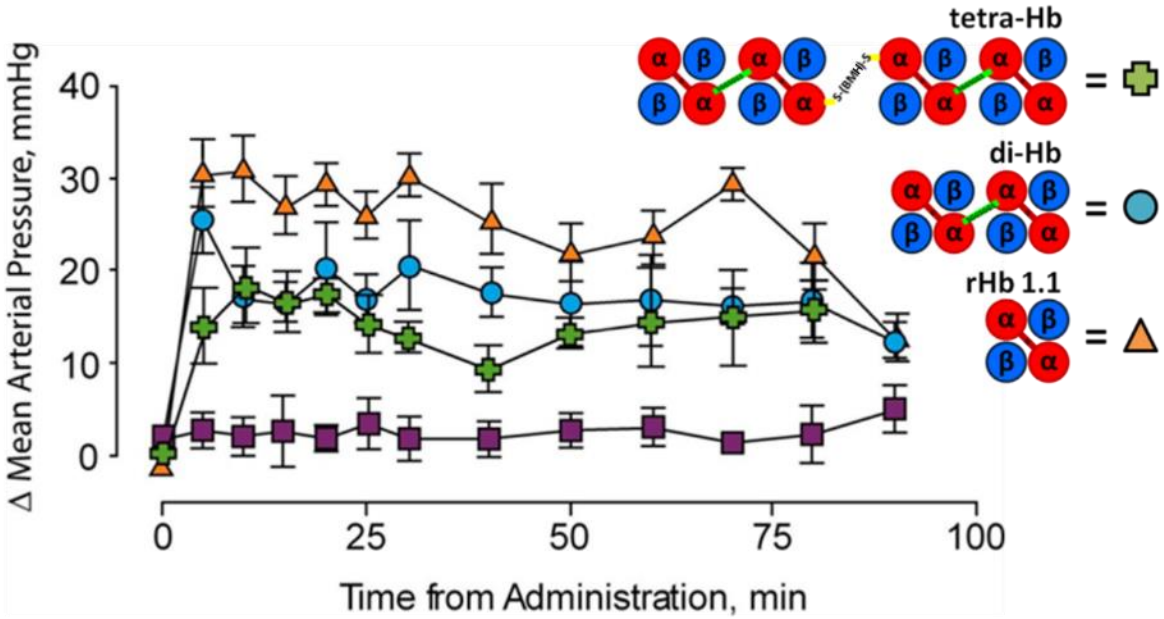

Figure 8. Change in mean arterial pressure (MAP) in conscious rats following intravenous injection of human serum albumin (purple squares), a 64 kDa hemoglobin (orange triangles), a 130 kDa polymeric hemoglobin (blue circles), or a 260 kDa polymeric hemoglobin (green cross) [8].

Development of a scHb will require covalent linkage of the α- & β-globin subunits within a single heterodimer and covalent linkage between the αβ-heterodimers across the dimer interface as illustrated in Figure 9. During the shift from the deoxy ("T") to the oxy ("R") state, these heterodimers rotate by 15° with respect to one another across the $\alpha_1\beta_2$ interface of the tetramer, while the $\alpha_1\beta_1$ interface between globins within a heterodimer remains relatively static. Our scHb design creates covalent connections across the $\alpha_1\beta_1$ and $\alpha_1\beta_2$ interfaces to minimize interference with allosteric switching of the tetramer. Linkage between heterodimers will be carried out as in rHb1.1, where the α_1-globin and α_2-globin are covalently linked with a single glycine residue (Figure 10).

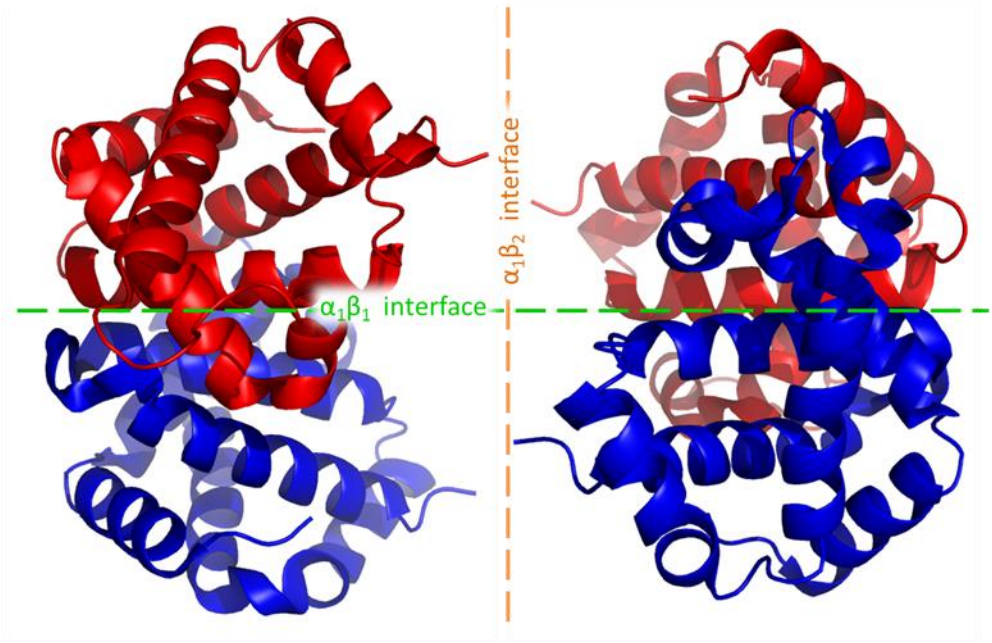

Figure 9. Illustration of the $\alpha_1\beta_1$ interface within a heterodimer (green) and the $\alpha_1\beta_2$ interface between heterodimers (orange). α-globins are in red and β-globins are in blue.

Covalent linkage within a heterodimer is considerably more challenging due to the relatively large distance between the α- & β-globin termini, approximately 50 Å through the center of

the wtHb tetramer, as illustrated in Figure 10. Linking the α_1- & β_2-globin termini in this manner would require a very long linker that would have to go around the outside of the molecule; this large modification at the interface of the heterodimers could negatively impact the folding and stability of the protein.

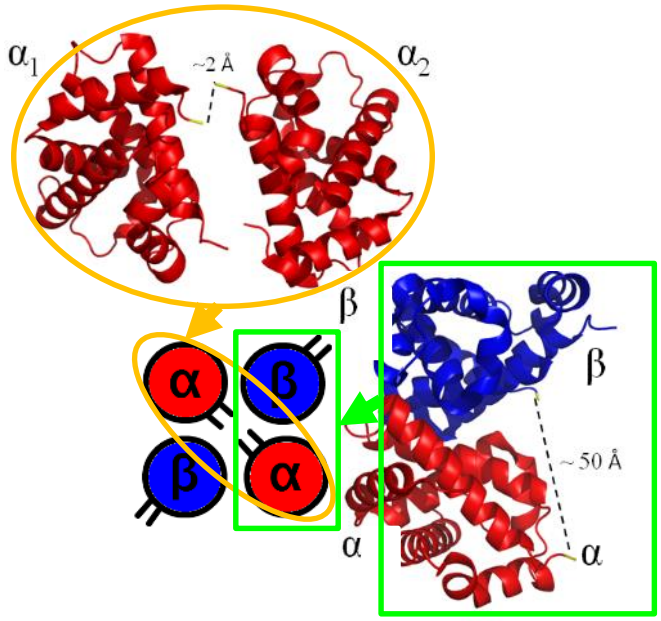

Figure 10. Illustration of the distances between the termini of the α_1- & α_2-globins (orange oval) and the termini of the α_1- & β_1-globins (green rectangle). Red circles are α-globins, the blue circles are β-globins, and short lines represent polypeptide termini for each globin.

The T-state conformation is dependent upon the formation of stabilizing interactions between the two αβ-heterodimers. Thus, the $\alpha_1\beta_2$ interface must be maintained in the design of the scHb. Our protein engineering efforts are targeted at designing a single-chain $\alpha_1\beta_1$ heterodimer, which can self-associate to form a sc-Hb. Covalent linkage via the insertion of a long linker between the α_1- and β_1-termini would likely enhance immunogenicity, and be a substrate for serum proteases due to its solvent-accessibility and flexible structure (dictated by the need to connect the native termini while minimally perturbing the R ↔ T switching of

the tetrameric structure). Relocation of the β-termini (the α-termini already tolerate covalent linkage as shown by rHb1.1) was identified as an alternative approach for covalent fusion of globins across the $\alpha_1\beta_1$ interface [9].

Circular permutation is the process by which existing protein termini can be relocated to an alternative site in a protein, as illustrated in Figure 11. We used circular permutation to relocate the β-globin termini. The site for insertion of new β-globin termini was chosen for its relative proximity to an external loop on the adjacent α-globin. As illustrated in Figure 12, this site in the loop between the β-globin G- and H-helices was selected for covalent linkage of the α- and β-globins, because it is close to the $\alpha_1\beta_1$ interface, and a similar permutation was well tolerated in myoglobin, which is a structural homologue of the Hb β-globin subunit [10]. Native β-globin termini were successfully fused by insertion of an eight amino acid linker (GSGGQGGG) and new termini were inserted via primer extension of a tandem β–globin gene to form a circularly permuted (cp) β-globin (cpβ) [11]

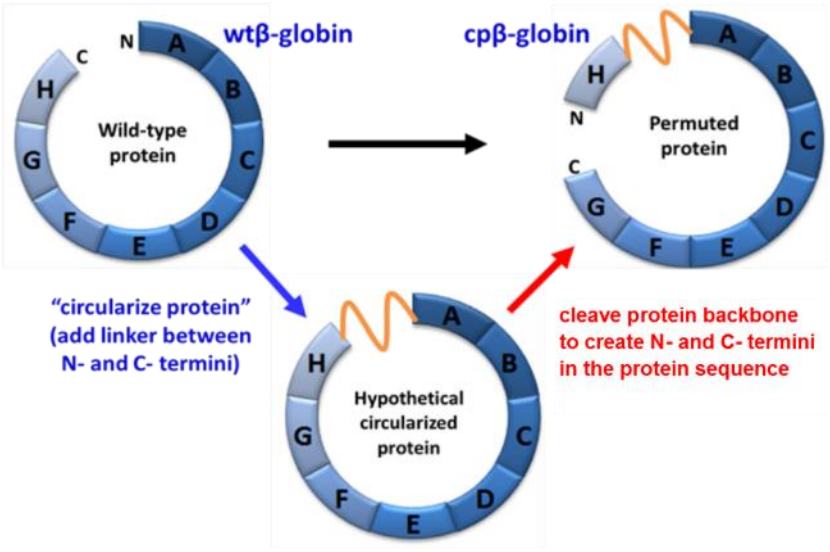

Figure 11. Illustration of the process of circular permutation of the β-globin. N & C identify the amino (N) and carboxyl (C) termini, respectively. The letters A-H represent helices A-H, and the yellow curve represents the inserted linker.

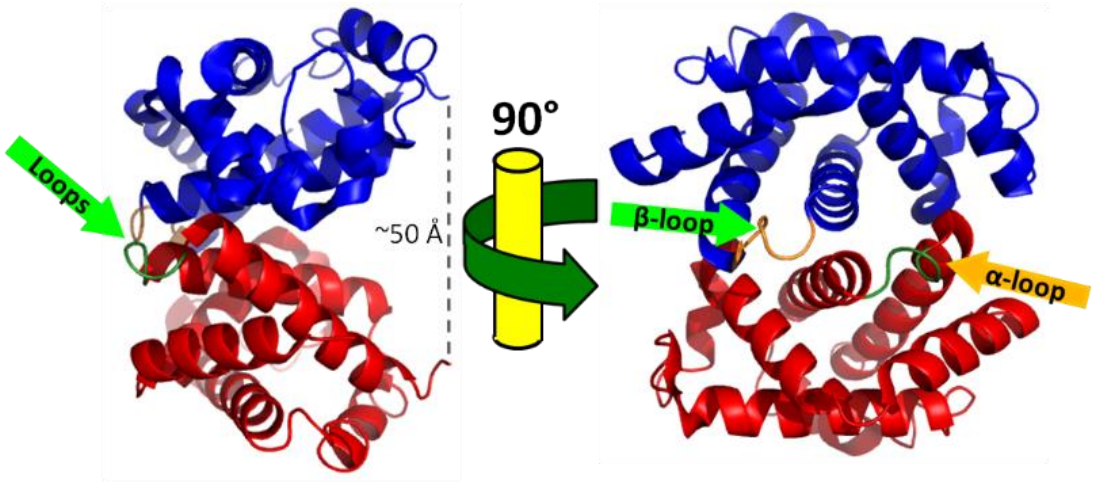

Figure 12. Illustration of the external loops of the α_1- (orange) & β_1-globins (green) and their relative proximity. α-Globins are in red and β-globins are in blue.

Recombinant α-globin and cpβ-globin (Figure 13) were co-expressed in *E. coli*, purified, and characterized structurally and functionally [11]. These subunits spontaneously associated to form α-cpβ heterodimers at low concentration, ~5 μM, which in-turn dimerized into an $\alpha_2(cp\beta)_2$ heterotetramer hemoglobin at higher concentrations, ~40 μM.

Alpha-cp beta:
Black: rHb 0.0(PDB 2dn1)
Red: wt alpha domain of α-cpβ
Blue: cp beta domain of α*cpβ*

Alpha globin amino acid sequence:

```
|M-LSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHFDLSHGSAQVKGHG
|MVLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHFDLSHGSAQVKGHG
 KKVADALTNAVAHVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTP
 KKVADALTNAVAHVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLVTLAAHLPAEFTP
 AVHASLDKFLASVSTVLTSKYR|
 AVHASLDKFLASVSTVLTSKYR|
```

CP beta globin amino acid sequence:

```
|MGEFTPPVQAAYQKVVAGVANALAHKYHGSGGQGGGVHLTPEEKSAVTALWGKVNVDEV
---EFTPPVQAAYQKVVAGVANALAHKYH|  |MVHLTPEEKSAVTALWGKVNVDEV
GGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKG
GGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGNPKVKAHGKKVLGAFSDGLAHLDNLKG
TFATLSELHCDKLHVDPENFRLLGNVLVCVLAHHFGK|
TFATLSELHCDKLHVDPENFRLLGNVLVCVLAHHFGK
```

Figure 13. Comparison of the sequences of wtHb and α-cpβ. In black: rHb 0.0 (PDB 2dn1), in red: wt alpha domain of α-cpβ, and in blue: cp beta domain of α-cpβ Hb.

Beyond relocation of the β-globin termini through circular permutation, two additional steps are required to produce a scHb: intradimer covalent linkage of the α_1- & β_1-globins and interdimer covalent linkage of the α_1- & α_2-globins as shown in Figure 14. Figure 15 shows the plan for adding covalent linkages between specific helices in the target sc-Hb, which will be single-chain α-cpβ (sc-α-cpβ). All the proteins shown in Figure 15 have been successfully expressed and characterized.

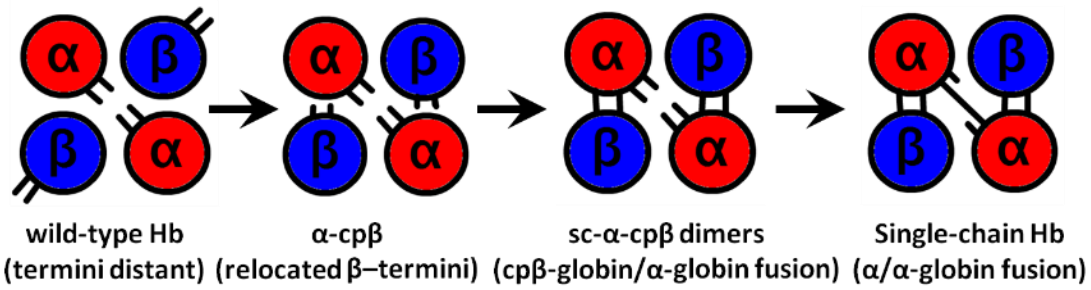

Figure 14. Scheme for the development of a single-chain hemoglobin. Red circles are α-globins, the blue circles are β-globins, the dashes represent polypeptide termini for each globin, and connected lines represent covalently bonded termini.

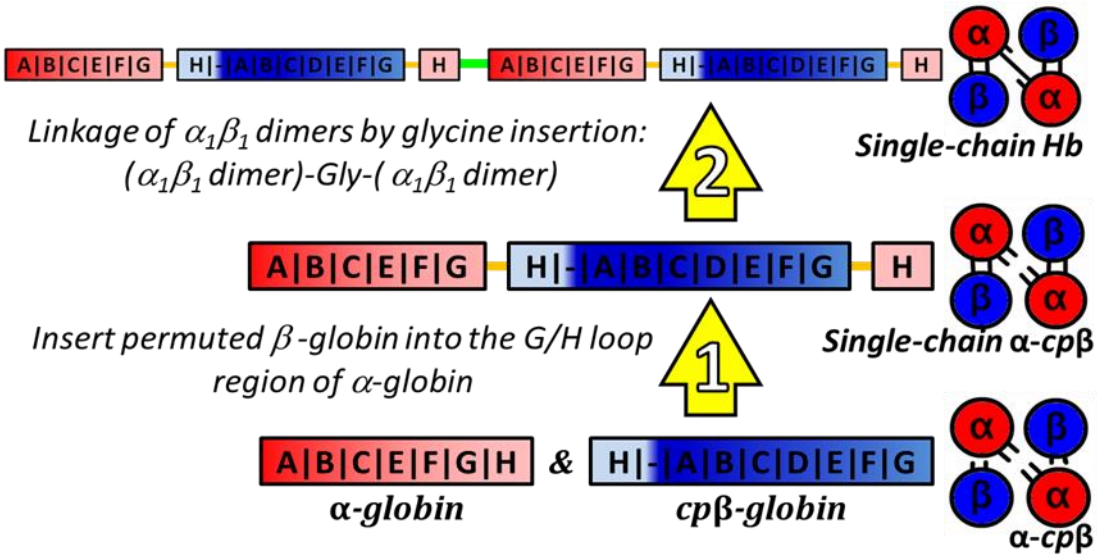

Figure 15. Proposed steps to introduce the covalent linkages necessary for fusion of the four globin subunits into a single-chain hemoglobin. Red circles are α-globins, the blue circles are β-globins, the dashes represent polypeptide termini for each globin, and connected lines represent covalently bonded termini.

Successful expression of α-cpβ enabled analysis of the effects of circular permutation on ligand binding affinity and protein structure. The mutant was characterized with high performance size exclusion chromatography (HPSEC), electrospray ionization mass spectrometry (ESI-MS), laser flash photolysis, and nuclear magnetic resonance (NMR) [12]. Characterization of α-cpβ by laser flash photolysis ligand binding assays revealed a significant increase in oxygen binding affinity when compared to that of wtHb [11]. This shift, associated with circular permutation, limits the oxygen transport capabilities of α-cpβ Hb. Efficient O_2 delivery to peripheral tissues is dependent upon the release of O_2 in low oxygen environments, e.g. peripheral tissues, which would be hindered by a shift to higher oxygen binding affinity. The shift in oxygen binding affinity is illustrated in Figure 16, a plot of log of partial pressure of O_2 (log P_{O_2}) versus fractional saturation of either wild-type Hb (HbA) or α-cpβ Hb. For large values of log P_{O_2} on the right side of the plot, the hemoglobins are predominantly in the oxygen bound (oxy- or R-) state. As O_2 partial pressure decreases, Hb releases bound oxygen to the surrounding solution; thus, the Hb molecules switch to the deoxygenated (deoxy- or T-) state. The midpoint of the transition between oxy- and deoxy-states indicates the oxygen binding affinity of the protein. In the case of α-cpβ Hb, the transition is shifted to the left such that it behaves as if stuck in a state similar to hemoglobin's high affinity R-state [11]. α-cpβ, due to its reduced oxygen binding cooperativity, has a similar oxygen binding affinity profile to that of myoglobin (a very high affinity oxygen transport protein) as illustrated in Figure 17.

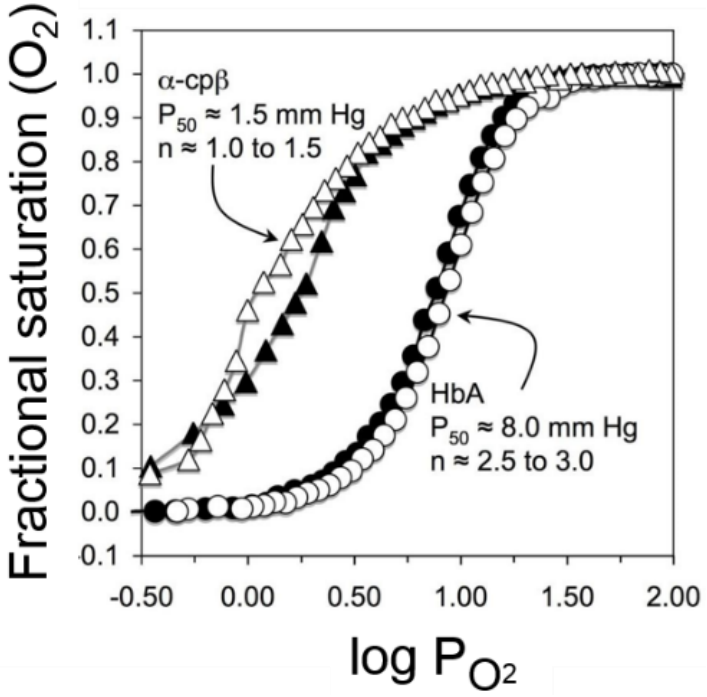

Figure 16. Plot of the log of the partial pressure of oxygen of HbA and α-cpβ versus fractional saturation with oxygen [11].

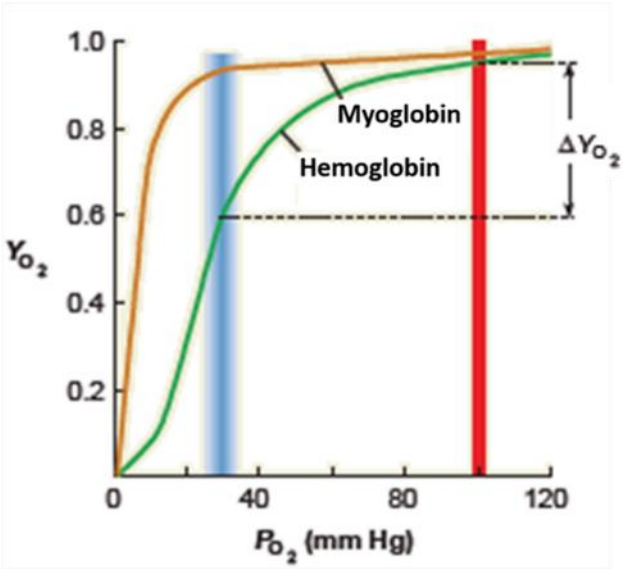

Figure 17. Oxygen binding affinity profiles for myoglobin and hemoglobin. ΔY_{O2} represents the fraction of bound O_2 delivered to peripheral tissues as Hb transits from the lungs to the capillaries.

High performance size exclusion chromatography (HPSEC) was employed to identify the extent to which the stability of the T-state had been perturbed. As shown in Figure 18, in the oxygenated form (in 50 mM PBS pH 7) the protein exists primarily as a ~30 kDa complex and in the deoxygenated form (in 50 mM PBS pH 7 + 5 mM dithionite [binds molecular oxygen]) it exists primarily as a ~60 kDa complex by comparing relative elution times to a molecular weight standard curve [11]. These data support the hypothesis that circular permutation of the β-globin has decreased the stability of the T-state tetrameric hemoglobin structure, but does not prevent association of heterodimers when conditions favor the T-state.

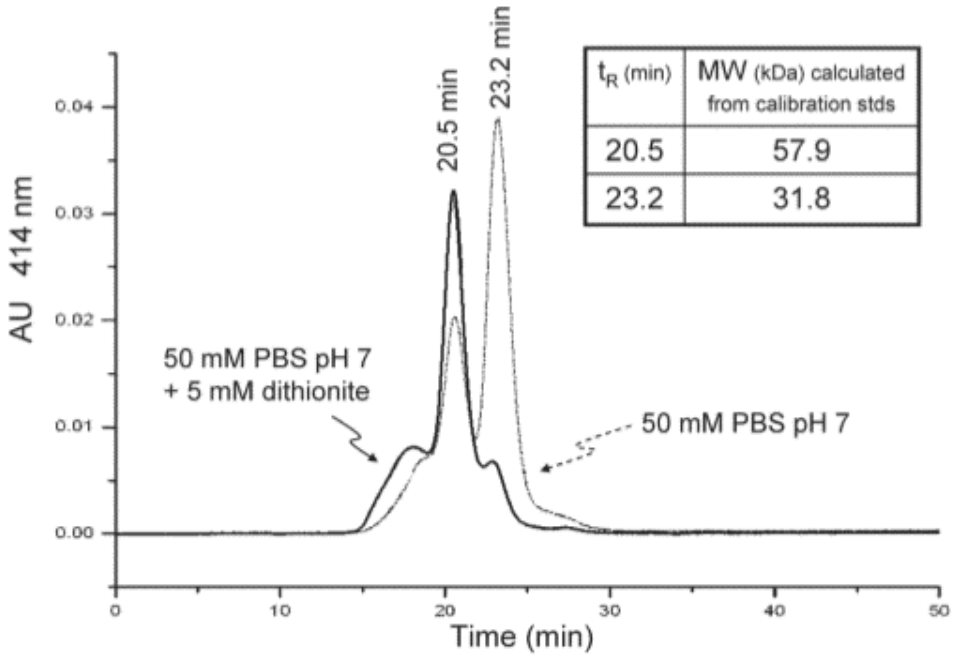

Figure 18. HPSEC data showing the relative proportion of hemoglobin species for α-cpβ , with tetramer ($\alpha_2\beta_2$) @ 20.5 min, and dimer ($\alpha_1\beta_1$) @ 23.2 min [11].

The function of α-cpβ Hb is adversely affected by circular permutation of the β-globin. Inspection of the structure of wtHb suggests an explanation for the apparent preference of α-cpβ for the high affinity conformation. The insertion of the new linker in the cpβ sequence,

results in the loss of a salt bridge interaction, which is critical for stabilization of the T state [6]. Figure 19 is a simplified schematic of hemoglobin's subunit domains superimposed with hydrogen bonding information obtained from X-ray crystallographic studies. The salt bridge between the ε-amino group of α_1 K40 and the carboxyl group of β_2 H146, the C-terminal residue of β-globin, must be absent in α-cpβ because the carboxylate of β_2 H146 is an amide in the cpβ linker. Without this stabilizing interaction across the $\alpha_1\beta_2$ interface, these α-cpβ heterodimers dissociate more readily. This was confirmed by sedimentation equilibrium (SE) studies which show a 5-6 kcal/mol (20-24 kJ/mol) loss or ~ 4 orders of magnitude decrease in tetramer stability across the α_1-β_2 interface (Table 1) [13].

	SE
Sample	**Kd**
rHb0.0	0.4-1.3 nM
alpha-cp-beta	4.3-9.8 μM

Table 1. Sedimentation equilibrium (SE) study results for wtHb (rHb0.0) and α-cpβ where K_d is the sedimentation equilibrium [13].

Dissociated heterodimers are known to adopt the R-conformation (and lose cooperativity) [11]. Thus the stability of the T-state, which is dependent on a tetramer structure, is reduced for α-cpβ Hb. This perturbation of the R to T equilibrium results in an alteration of the oxygen binding affinity such that α-cpβ Hb is a functional mimic of myoglobin (see Figure 17). Thus, increasing favorable interactions between heterodimers across the $\alpha_1\beta_2$ interface is key to recovering the desired oxygen binding cooperativity in a permuted Hb.

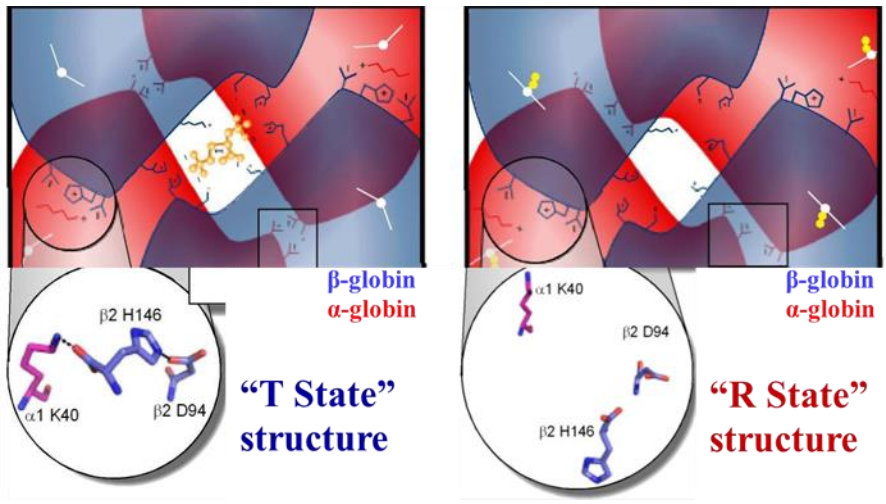

Figure 19. Illustration of important interactions that stabilize the T-state of hemoglobin [5 (used without permission)]. α-globins are in red, β-globins are in blue , and the circles show X-ray crystallographic detail at the $\alpha_1\beta_2$-interface.

It is not clear from published preclinical data (reviewed in [13] what the P_{50} and Hill coefficient (a measure of cooperativity) for an HBOC must be to enable effective oxygen transport *in vivo*. In the absence of definitive clinical data, our initial targets for O_2 binding affinity and cooperativity are those of whole blood. In contrast to the "directed evolution" of protein function, a more targeted approach to protein engineering is guided by structural information. Thus, we require an atomic-resolution structure of α-cpβ to identify mutations that might improve permuted Hb function. In the absence of such structural information for α-cpβ, published X-ray crystal structures were used to suggest stabilizing mutations. Two potential avenues of tetramer stabilization emerged, covalent linkage of the α-cpβ dimers through addition of a single glycine residue between the α_1- and α_2-globins (i.e., the Somatogen strategy shown as Step 1 of Figure 6), and introduction of the well-characterized stabilizing mutation βK82D. The first approach produced a permuted Hb composed of a covalently liked diα-globin that associated with two cpβ-globins (diα-cpβ); the second

approach resulted in a tetrameric hemoglobin containing two α-globins and two cpβ-globins with the K82D mutation (α-cpβ K82D), both are illustrated in Figure 20.

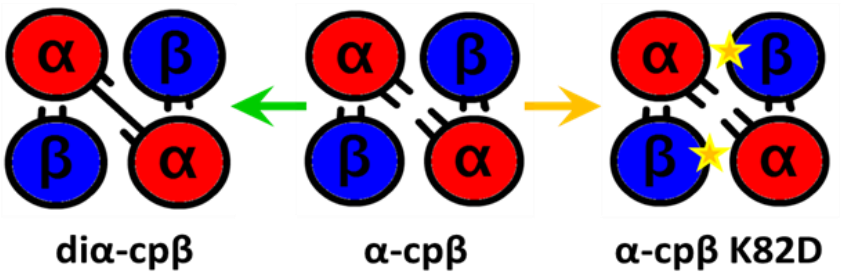

Figure 20. Protein engineering schemes to increase the stability of T-state hemoglobin. Red circles are α-globins, the blue circles are β-globins, the dashes represent polypeptide termini for each globin, golden stars represent the K82D mutation, and connecting lines represent covalently bonded termini.

NMR was employed to determine the effects of these mutations on the Hb structure. NMR spectra collected on various Hb isoforms (O_2-bound, CO-bound, and deoxy-Hb) provided evidence for $\alpha_1\beta_1$ and $\alpha_1\beta_2$ interactions similar to those in wild-type Hb [12]. It must be noted that the protein concentration in the NMR samples is ~ 1 mM, which is significantly higher than would be expected in circulation following transfusion of an HBOC. These high concentration conditions favor association of the relevant subunits to give the 64 kDa complex.

The proton NMR spectra shown in Figure 21 illustrate the peaks that correspond to interactions within (indicated by "R") and between ("T") heterodimers in the deoxygenated state of Hb [12]. The peaks labeled with a blue 'T' are indicative of interactions present only in the T-state tetramer. These 1-D ^{1}H NMR data suggest that association of the subunits in α-cpβ is altered and/or reduced compared to that in wt Hb, and that in diα-cpβ the T-state interactions are recovered (compare left and right panels of Figure 21).

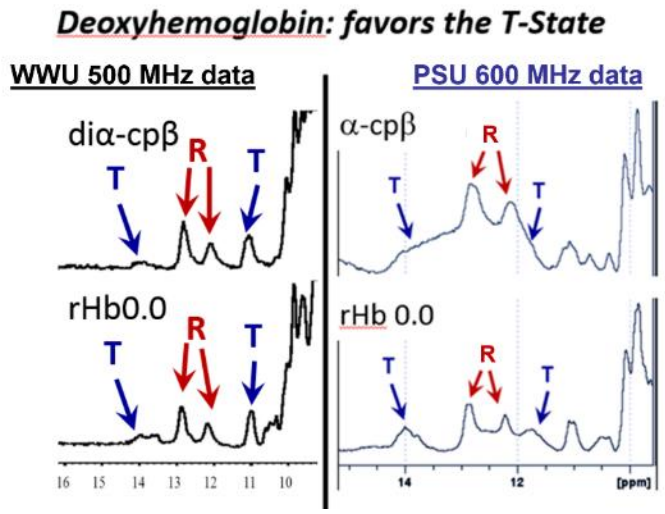

Figure 21. Proton NMR spectra of deoxy (T-state) form hemoglobins. 'R' in red indicates interactions within a heterodimer and 'T' in blue indicates interactions between heterodimers [15].

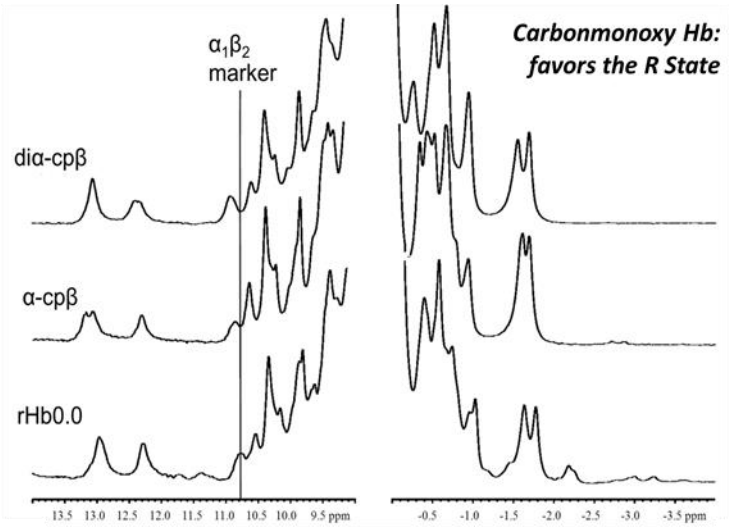

Figure 22. Proton NMR spectra of carbonmonoxy (R-state) form hemoglobins. The vertical line labeled with "$\alpha_1\beta_2$ marker" indicates a peak in rHb0.0 that is indicative of an interaction between groups on opposite heterodimers across the tetramer interface [15].

In addition to these structural data, functional data were collected for these mutants. Equilibrium oxygen binding curves (Figure 23) show the effects that our mutations had on recovering native-like oxygen binding affinity and cooperativity. It appears that, at best, modest gains were made with α-cpβ K82D Hb to return to a wtHb oxygen binding affinity

profile, but these data are preliminary. The binding curves for α-cpβ and diα-cpβ are nearly superimposable; thus, the functional data do not support the conclusions of the NMR studies shown above. This apparent discrepancy is likely due to the different protein concentrations used for the binding studies (10 μM) vs. the NMR experiments (300 μM).

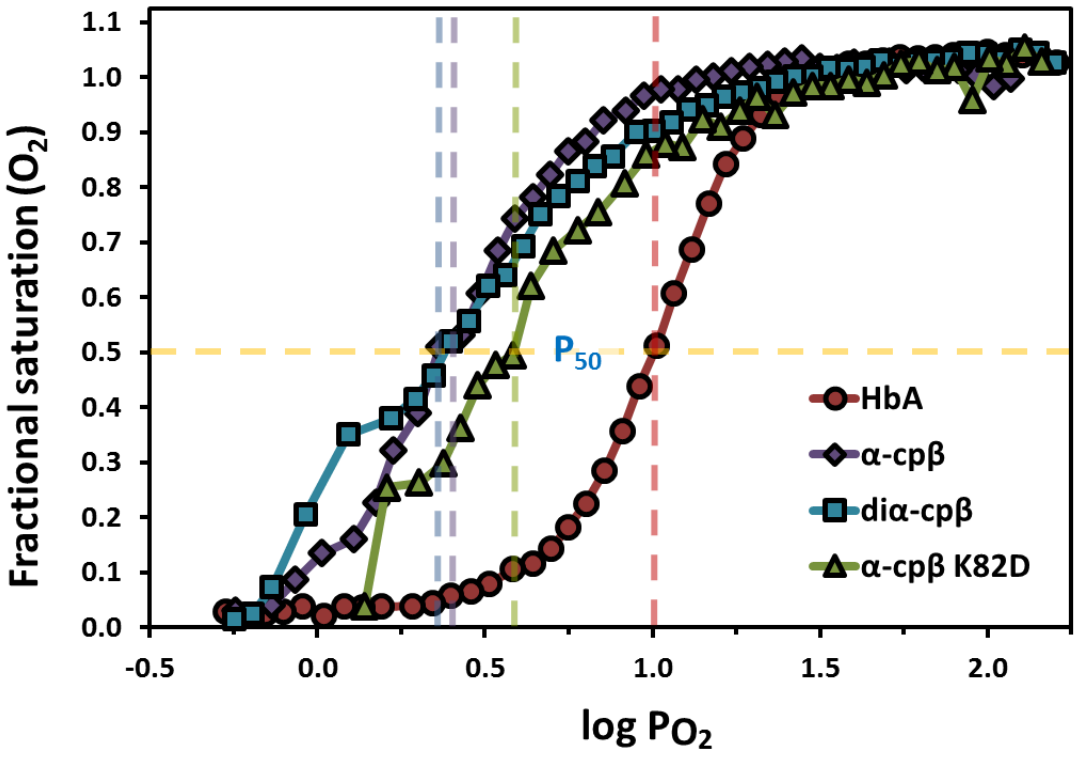

Figure 23. Plot of the log of the partial pressure of oxygen of HbA and mutant hemoglobins versus partial saturation with oxygen. [16].

Greater structural detail of the interactions between α-cpβ dimers is needed to better understand how circular permutation affects permuted Hb structure, stability, and function. The present work was undertaken to determine an atomic-level structure of α-cpβ, and thereby inform a rational protein engineering strategy aimed at recovering the stability of the T-state.

The permuted Hb proteins are of a size that would render structure determination by solution methods, e.g. NMR, highly challenging. Thus, we chose to approach elucidation of the 3D structure through the method of X-ray crystallography. X-ray crystallography is the process by which crystals of a given molecule may be analyzed with monochromatic, high flux X-rays to determine the electron density within the repeating unit of that crystal, as illustrated in Figure 24. The electron density maps produced from this technique can then be populated with a theoretical model of the molecule in the crystal. The resulting model is then modified to best fit the experimental data.

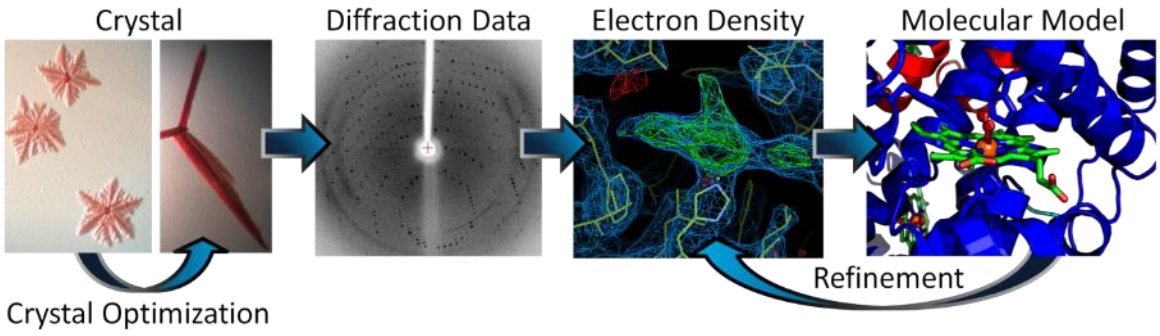

Figure 24. Structure determination by the X-ray crystallography process. Arrows indicate progression of steps in the process.

The first and arguably most arduous step in this process was the production of protein crystals with properties enabling the collection of high resolution diffraction data. The hanging drop method of protein crystallization was employed to screen crystallization conditions and to produce diffracting crystals. Driven by vapor diffusion, the hanging drop method facilitates the growth of crystals in a sealed system. The concentration of precipitants in the protein solution gradually increases via transfer of water vapor from hanging drop to mother liquor to the point that the protein precipitates, hopefully at a controlled rate such that crystals are formed [17]. A high degree of structural order (crystallinity) within the crystal,

monocrystalline composition, and sufficient size are characteristics that indicate a crystal should produce higher resolution diffraction patterns. The variability in the quality of the crystals combined with the difficulty of identifying conditions that promote growth of monocrystals, make X-ray crystallography a challenging and artful task.

Ambiguity in the interpretation of the diffraction data is engendered by the loss of phasing information. Phase information can be recovered partially through three common methods, multiple isomorphous replacement (employing integration of heavy atoms within the crystal to give phase guides), molecular replacement (utilizing a known homologous structure to assign phase computationally), and multi-wavelength anomalous dispersion (coherent X-ray light of different wavelengths are diffracted off the crystal at a synchrotron facility. As many high resolution hemoglobin structures exist and are readily accessible, molecular replacement was our chosen method to solve this phasing problem.

Once a diffracting crystal has been obtained, diffraction data collected, and the phase problem resolved, refinement of the molecular model for optimal fitting of the electron density begins. Refinement is an iterative process. As the model more closely fits the density, the resolution of the density maps improves, which in turn enhances the quality of the data available to fit the map. This iterative process is essential and is a function of the mathematical operations employed to ascertain the electron density maps from the pattern of X-rays diffracted by the crystal.

Materials and Methods:

Expression of Hemoglobins

Proteins were expressed in *E. coli* using plasmids developed previously [11]. One liter flasks of terrific broth (TB: 12 g tryptone, 24 g yeast extract, 4 ml glycerol, 2.31 g KH_2PO_4, and 12.54g K_2HPO_4, fill to 1000ml with dH_20) were inoculated with 10 mL of *E. coil* overnight culture (10 mL of LB in a 50mL flask shaken overnight at 37 °C), and incubated at 37 °C with shaking until an optical density at 600 nm of 0.7-0.9 was attained. The flasks were then cooled to 30°C and shaken until they reach an OD_{600} of 1.5 to 1.7, at which point IPTG was added to a final concentration of 1mM to induce protein expression. Shortly after, 1.5 g of hemin in 20mM sodium hydroxide was added to each flask, an identical aliquot of hemin was added 2 hours and 4 hours following induction. 6 hours after induction, the flasks were removed from the orbital shaker and the cells were harvested by centrifugation at $5{,}000 \times g$ for 10 minutes. The supernant was discarded and the pelleted cells were collected and flash frozen in liquid nitrogen. Frozen cells were stored at -80 °C until they were thawed for protein purification.

In the optimized process for production of the α-cpβ protein the *E. coli* were grown in TB to an OD_{600} of 0.7-0.9 at 37 °C, at which point the temperature was reduced to 30°C for α-cpβ or 32°C for diαcpβ and allowed to reach an OD600 of 1.5-1.7. At this point, protein expression was induced with 1.00 mM isopropyl β-thiogalactopyranoside (IPTG). Cells expressed α-cpβ for 5 hours for α-cpβ or 7 hours for diαcpβ at 30 °C before being harvested by centrifugation and flash freezing in liquid nitrogen.

Purification

Following the protein expression period, bacterial growth media was centrifuged at 5000 x g for 10 minutes to pellet the cells. Pelleted cells were harvested, flash frozen in liquid nitrogen, and resuspended in lysis buffer (50mM Tris-HCl pH 8.5, 17 mM NaCl). Cell suspensions were then lysed via sonication with a Branson Sonifer 450, fitted with a 1.8 cm diameter horn, for 90 seconds with a duty cycle of 50% and power output set to 6. The crude lysate (Figure 25) was then separated into soluble and insoluble fractions via centrifugation at 21,000 x g for 20 minutes, soluble fraction was collected and the pH was adjusted to between 8.0 and 8.5 to improve protein binding to the IMAC resin. Permuted and wild-type hemoglobins bind to zinc ions present on the charged IMAC column, 4 mM zinc acetate was added to the supernant to facilitate precipitation of contaminating cellular components that also possess Zn^{2+} binding affinity. Precipitated proteins were removed from the supernant via centrifugation at 40,000 x g for 20 minutes and filtration through a 0.45 µm pore size cellulose acetate membrane (Millipore).Purification procedures were adapted from published procedures [11] with the exclusion of the 250mM NaCl step in the IMAC phase of purification. In addition, the proteins were not frozen between each purification step, rather, they were kept at 4°C until purification was finished and then used immediately or flash frozen in liquid nitrogen for long-term storage.

Immobilized Metal Ion Affinity Chromatography (IMAC)

The IMAC column was hand-packed with ~ 10mL of Amersham-Pharmacia Biotech Chelation Sepharose Fast Flow Resin. Flow was maintained at 6 mL/min by a Bio-Rad EP-1 Econo Pump. Preparation and operation of the column is described in Table 2.

Table 2. IMAC wash steps. 1 CV = 10 mL.

Buffer	**Column volume (CV)**	**Flow rate**
Column preparation		
0.2 M NaCl	2	4 mL/min
20 mM $Zn(OAc)_2$	2	4 mL/min
0.2 M NaCl	2	4 mL/min
Protein purification		
Lysate	-	4 mL/min
20 mM Tris-HCl, 0.5 M NaCl, pH 8.5	8	4 mL/min
20 mM Tris-HCl, pH 8.5	6	4 mL/min
Protein elution		
20 mM Tris-HCl, 15 mM EDTA, pH 8.5	-	2 mL/min
Column cleaning		
50 mM EDTA	2	4 mL/min
0.2 M NaCl	2	4 mL/min
0.2 M NaOH	2	4 mL/min

Eluent fractions containing the mutant Hbs were collected and concentrated with Millipore Amicon Ultra Centrifugal Filter Devices (Amicon Ultra-15 or Amicon Ultra-4, 10 kDa filter) and filtered through Corning Spin-X Centrifuge Tube Filters (0.22 μm cellulose acetate). The Beer-Lambert law was employed to calculate preliminary protein expression yields from the heme Soret absorbance peak (410-418 nm) with ε_{heme} of 125,000 M^{-1} cm^{-1}.

Mono Q ion exchange (IEC)

Fractions pooled following IMAC were buffer exchanged into 20 mM Tris-HCl pH 8.0 (buffer A) with a Millipore Amicon Ultra Centrifugal Filter. Samples were then purified on an

Amersham Biosciences MonoQ 5/50 GL column connected to either a Waters 650E Advanced Protein Purification System and a Waters 486 Tunable Absorbance Detector or a GE Health Sciences Äktapurifer 900 system. The proteins were immobilized on the resin and subsequent washes with increasing salt concentration promoted elution of contaminants as described in Table 3. Fractions containing Hb were collected and concentrated as mentioned above.

Table 3. IEC (Mono Q) purification protocol

Buffer	**Column Volume (CV)**	**Flow rate**
Column preparation		
Buffer A	5	2 mL/min
Buffer B (20 mM Tris-HCl, 1 M NaCl, pH 8.0)	5	2 mL/min
Buffer A	5	2 mL/min
Protein loading		
Load IMAC pool	-	-
Buffer A	2	2 mL/min
Protein elution (linear gradient program) **Time course**		
100% Buffer A	0:00	2 mL/min
80% Buffer A, 20% buffer B	3:00	2 mL/min
30% Buffer A, 70% Buffer B	28:00	2 mL/min
100% Buffer B	29:00	2 mL/min
Column cleaning **Column Volume**		
Buffer B	2	2 mL/min
Buffer C (2 M NaCl)	2	2 mL/min
1 M NaOH	4	2 mL/min
ddH_2O	2	2 mL/min

Size Exclusion Chromatography (SEC)

Following purification by MonoQ, samples were further purified on an Amersham Biosciences Superdex 75 HR 10/300 GL size exclusion column connected to a Waters 650E Advanced Protein Purification System and a Waters 486 Tunable Absorbance Detector or the Äktapurifer 900 system. Purification by SEC was accomplished by the process described in Table 4. SEC-

purified samples were concentrated and final post purification yields were calculated from Soret band absorbance measurements as described above.

Table 4. SEC purification protocol

Buffer	**Column Volume**	**Flow rate**
Column preparation		
ddH_2O	2 CV	1 mL/min
20 mM ammonium acetate (NH_4OAc), pH 7.0	1 CV	1 mL/min
Protein elution		
20 mM NH_4OAc, pH 7.0	-	1 mL/min
Column cleaning		
20 mM NH_4OAc, pH 7.0	1-2 CV	1 mL/min
ddH_2O	1-2 CV	1 mL/min
0.5 M NaOH	1-2 CV	1 mL/min
ddH_2O	1-2 CV	1 mL/min
20% v/v EtOH	1-2 CV	1 mL/min

Protein Characterization

Purified proteins were analyzed for purity and molecular weight by SDS-PAGE. Electrospray Ionization Mass Spectroscopy was performed at the WWU Scientific Instrument Center using an Applied Biosystems API 2000 Triple Quadrupole. Samples were prepared in 50% (v/v) acetonitrile, with 25mM ammonium acetate buffer pH 7.0, and 0.1% formic acid.

Preparation of cyanomet hemoglobin isoform

Conversion of oxy-Hb to the cyanomet isoform, prior to analysis by ultracentrifugation, was achieved using Drabkin's reagent as described previously [26] with the single modification that a 5-fold excess of potassium ferricyanide to heme concentration of protein was used.

Preparation of protein for crystallization

Protein stocks of α-cpβ hemoglobin were purified to greater than 95% purity at a concentration of 10 mg/mL in 100 mM ammonium acetate, pH = 7, for crystallography. The hanging drop method of protein crystallization was applied in 24-well crystallography trays as follows: the rims of the wells in the crystallography plate were coated in white petrolatum to prevent dehydration of the experimental solution in the well, a 400 uL aliquot of precipitant solution (the mother liquor) was added to each well, 1 uL of concentrated protein solution and 1 uL of mother liquor were pipetted onto a siliconized glass cover slip, and the glass slip was inverted and sealed to the greased crystallization well. The hanging drops were examined for crystal formation at regular intervals with a compound microscope and crystal formation was recorded with a digital camera.

Shotgun approach

The shotgun approach was used to screen for crystal promoting conditions for mutant hemoglobins. Standard screening solutions were purchased from Emerald BioStructures and Hampton Research. The following Hampton Research solutions were evaluated with α-cpβ: Index, Crystal Screen, PEG/Ion, PEGRx, and Crystal Screen Cryo. Mixtures of polyethylene glycol 2250 (PEG2250) at concentrations of 20-30% w/v, polyethylene glycol monomethyl ether 1250 (PEGMME1250) at concentrations of 20-30% w/v, 0.2 M ammonium acetate, 0.2 M sodium chloride, 0.1 M Bis-Tris buffer at pH 5.8-7.0, and 0.1 M Tris-HCl at pH 8.5 to 9.0 produced crystals that were suitable for X-ray diffraction studies.

Identification of 'hits'

Any solutions from the initial shotgun screening that produced crystals were considered to be leads, or 'hits', that were used as starting conditions to determine optimal condition grids that would promote better crystal morphologies.

'Gridding' around hits

Precipitant solution composition from the initial hits obtained from the shotgun approach was varied to optimize crystal morphology and size. Optimum crystal morphology was identified by crystals that appeared clear and regularly geometric. PEG 3350 concentrations and pH were varied until a narrow range of conditions were identified as optimum for these two variables. Variation from standard conditions of 0.2 M salt and 2 uL drop volume failed to improve crystal morphology and quality.

Catalogue of crystals

Many distinct crystal morphologies were observed during the optimization process. Sheet, spike, cluster, and many other geometries were observed; each geometry exhibited a continuum of crystalline order from opaque multi-crystals to transparent single crystals, see Appendix A. Well conditions resulting in crystals of desirable morphologies (regular geometric in shape, single crystal, large, and optically transparent) were identified. Gradient trays were formulated in a fine grid around these optimal well conditions in an attempt to further improve resultant crystal morphologies. After fine optimization, whole trays would be plated with identical optimized well conditions to produce many optimal crystals.

Selection & harvesting of crystals for X-ray crystallography

Optically transparent, isolated crystals larger than 0.2 mm, were captured with a crystallography loop (Hampton Research catalogue #: HR4-957) . Preserved crystals were quickly transferred to a cryoprotectant solution identical to the mother liquor with the addition of 10-30% dimethylsulfoxide (DMSO), glycerol or Polyethylene glycol 400 (PEG400) and examined for stability in that solution. A trial and error approach was needed to identify which solution would protect the crystal without damaging its integrity. 20% glycerol was determined to be optimal for most crystals and so was used almost exclusively. Once successfully cryoprotected, the crystal was looped again and flash frozen by submerging in liquid nitrogen. Great care was taken to avoid contact with air from this point on as ice crystals will form rapidly on the frozen crystal; these ice crystals will interfere with collection of diffraction data, but can be removed by quickly thawing and refreezing the crystal. The crystals were stored in magnetically sealed tubes under liquid nitrogen until they were to be evaluated by X-ray diffraction.

X-ray crystallography

Eleven diffraction data sets were evaluated on the α-cpβ hemoglobin crystals with an AFC11 SATURN 944 Right system and Cu Kappa radiation at 1.54Å from a Rigaku rotating anode generator. Data collection for these crystals was performed at the Fred Hutchison Cancer Research Institute, Seattle, WA. Diffraction data was collected and processed by Dr. P. Clint Spiegel with HKL2000 and D*trek for the α-cpβ eight crystals evaluated. Crystals were removed from storage in LN_2, transferred to the mounting arm via a magnetic

crystallography loop and kept at LN_2 vapor temperature by a cryostream. If ice rings were detected in the diffraction pattern, crystals were removed from the cryostream

Evaluation of data set 11 (hb8) and molecular replacement

This portion of the work was performed by Dr. P. Clint Spiegel of Western Washington University, Department of Chemistry. The search model used was a high resolution oxy-hemoglobin structure from the PDB, PDB code 2dn1.

Calculation of electron density maps

Electron density maps were calculated with the Phenix.maps subroutine from the PHENIX macromolecular structure refinement suite [23]. Each refinement of the electron density produced two calculated maps, which guided model refinement. These calculated maps represented, respectively, the density of the model plus those features that had not been modeled yet (composite map) and density features that were not directly attributed to the model (difference map). The first of these maps was calculated by subtracting the absolute value of the $F_{calculated}$, or F_c, map from two times the absolute value of the $F_{observed}$, or F_o, map (2fo-fc) and the second is calculated by subtracting the absolute value of the F_c, map from the F_o map (f_o-f_c).

Automated structure refinement

The Phenix.refine subroutine of the Phenix software package was employed for automated refinement of the model. Alternating rounds of optimization weighted for biophysical parameters (bond angles and lengths) and for electron density fit were applied to the structure

to achieve a concurrent lowering of the R_{work} and R_{free} values. This alternation of fitting parameters enabled automated optimization of biophysical constraints and density fit while avoiding undesirable perturbations to the structure by the computer aided structure refinement package. These perturbations often remodeled difficult to define portions of the structure or placed residues into impossible orientations.

Manual structure refinement

WinCoot [24] was employed to manually fit many of the side-chains, terminal residues, ligands and waters. The full suite of tools and analysis packages available on the program was applied to develop the final model.

Validation of structure refinement

The Phenix.validate subroutine of the Phenix software package was employed to validate the structure by comparison with other models of similar resolution from the PDB.

Results and discussion:

To produce crystals of high quality, relatively large amounts of protein (>10 mg) at ~95% purity must be obtained. This requirement was initially challenging due to the low expression levels of the mutant proteins (approximately 1.2 mg of α-cpβ per 1 liter of growth media following the first purification step). Greater than 50% of this 1.2 mg was lost in the final purification steps, requiring upwards of 18-34 liters of bacterial culture to produce the final 10 mg of purified α-cpβ protein. Optimization of permutein expression reduced the media requirement to approximately 4 liters, this enabled production of material in quantities sufficient for crystallography.

Growth optimization

As shown in Table 5, yields of 1.2 mg α-cpβ per liter of culture, following immobilized metal ion affinity chromatography (IMAC) purification, were typical at the outset of this work. Other permutants expressed at even lower levels (0.1 mg diα-cpβ/L and 0.3 mg sc-α-cpβ/L), likely stemming from greatly reduced apoprotein stability [18]. These levels of protein expression are insufficient to support X-ray crystallography and NMR studies- particularly if isotopic labeling is required. Thus, protein yield enhancement was pursued through two avenues, optimization of the protein expression protocol conditions and insertion of a heme-uptake gene (*hug*) cassette [19].

Table 5. Growth optimization data for hemoglobin mutants. Data collected from many past and present research group members.

hug	Medium	OD 600nm	Temp. (°C)	Growth (hr)	Yield (mg/L)
Wild Type Hb					
N	LB	0.5 - 0.7	25	4	16
N	LB	0.9-1.3	25	4	30
N	LB	1.2-1.5	25	4	20
N	LB	1.2-1.5	25	4	37
αcpβ Hb					
N	LB	0.5 - 0.7	25	4	1.26
N	LB	0.5 - 0.7	25	4	1.07
N	LB	0.5 - 0.7	25	4	0.90
N	LB	0.5 - 0.7	25	4	1.30
Y	TB	1.0-1.5	25	4	3.73
Y	TB	1.0-1.5	25	4	3.35
Y	TB	1.5-1.7	30	4	4.21
Y	TB	1.5-1.7	30	5	5.28
Y	TB	1.5-1.7	37	4	5.71
N	TB	1.5-1.7	30	5	5.89
N	TB	1.5-1.7	37	4	2.98
diαcpβ Hb					
N	LB	0.5-0.7	25	4	0.10
N	LB	0.5-0.7	25	4	0.08
N	LB	0.9-1.3	25	4	0.30
N	LB	0.9-1.3	25	4	0.50
N	LB	0.9-1.3	25	4	0.62
N	TB	0.9-1.3	25	4	0.41
N	TB	0.9-1.3	25	4	0.42
Y	TB	0.9-1.3	25	4	0.14
Y	TB	0.9-1.3	25	4	0.46
Y	TB	0.9-1.3	25	4	0.37
Y	TB	0.9-1.3	25	4	0.27
N	TB/2.5g G	1.5-1.7	30	5	4.78
N	TB/2.5g G	1.5-1.7	30	5	4.67
N	TB/5g G	1.5-1.7	30	6	3.78
Y	TB	1.5-1.7	32	5	3.66
Y	TB	1.5-1.7	32	6	4.83
Y	TB	1.5-1.7	32	7	5.36
scα-cpβ Hb					
Y	TB	1.0-1.5	25	4	0.26
Y	TB	1.0-1.5	25	4	0.26
Y	TB	1.0-1.5	25	4	0.28
Y	TB	1.0-1.5	25	4	0.31
Y	TB/2.5g G	1.0-1.5	30	4	0.29
Y	TB/2.5g G	1.0-1.5	30	4	0.30
Y	TB/2.5g G	1.0-1.5	37	4	0.25
Y	TB	1.5-1.7	30	5	1.80
Y	TB	1.5-1.7	30	5	1.81
Y	TB	1.5-1.7	30	6	2.24
Y	TB	1.5-1.7	30	6	2.02
Y	TB	1.5-1.7	30	7	1.87
Y	TB	1.5-1.7	37	5	1.15
Y	TB	1.5-1.7	37	6	1.47
Y	TB	1.5-1.7	37	6	1.42
Y	TB	1.5-1.7	37	7	0.71

Protein expression protocol optimization was accomplished through modulation of: medium formulation, target OD_{600} for induction, concentration of isopropyl β-D-1-thiogalactopyranoside (IPTG) added at induction, concentration and number of hemin additions post-induction, free glucose concentration, the temperature during protein expression, and length of induction. An expression yield enhancement of 4-50 fold was observed for the suite of proteins listed in Table 5 (1.2 to 5.8 mg/L for α-cpβ, 0.1 to 5.4 mg/L for diα-cpβ, and 0.3 to 2.1 mg/L for sc-α-cpβ). Use of a richer medium with more available nutrients (Terrific broth [TB] vs. LB), a higher concentration of IPTG at induction, and an increased concentration of added hemin, collectively resulted in a 3-4 fold yield enhancement of α-cpβ post-IMAC purification step. Optimization of target OD_{600} for induction, the temperature during protein expression, and length of protein expression period increased protein yields again by approximately 60% for αcpβ.

Multiple variables in the expression protocol for α-cpβ were simultaneously modified to improve protein yields. Thus, the effect of the *hug* cassette on protein yield enhancement was difficult to isolate. However, improved expression yields of α-cpβ and diαcpβ were generally observed when the *hug* cassette was present in the Hb expression vector. Yield data obtained from later experiments with other variants of α-cpβ (which contain stabilizing point mutations [20]) failed to correlate the observed enhancement of permuted Hb expression levels with the coexpression of the *hug* cassette (S. Jones, J. Sigurjonsson, unpublished data). In the case of diαcpβ, significant yield enhancement was not observed until other variables were also altered, namely temperature of expression, target OD_{600} for induction, and length of induction. Further testing to determine the need for the *hug* cassette was not undertaken as

our major goals were structure determination, and we had improved expression yields sufficiently to pursue those goals.

Yield enhancements observed for diα-cpβ suggest that target OD_{600} for induction and length of induction are among the most critical parameters to optimize. Protein yield data obtained from optimization of sc-α-cpβ expression (Figures 14 and 15; Table 5) also showed that significant enhancement in protein yields was not observed until the OD at induction and length of induction were increased.

Purification Results

Proteins were purified from clarified lysate by Immobilized Metal Ion Affinity Chromatography (IMAC), Ion Exchange Chromatography (IEC), and Size Exclusion Chromatography (SEC) [see Materials and Methods].

Protein yields

The rigorous purification procedure required to attain greater than 95% purity (Figure 25) resulted in low protein recoveries. Post IMAC yields, detailed in Table 5, were reduced by the MonoQ and SEC steps by ~70%. Typical protein recoveries from a 6L α-cpβ expression are as follows: 20.1 mg of protein following IMAC, 5.8 mg from the combination of the MonoQ and SEC purification steps (29%). Such losses between the first and final purification step further emphasized the need for optimization of protein expression conditions.

Protein purity was determined through sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and electrospray mass ionization mass spectrometry (ESI-MS). Densitometry of protein gels was performed using the ImageJ software [21]. An example of high purity α-cpβ (98%) and α-cpβ-*hug* (96%) characterized by this technique is shown in Figure 25. ESI-MS characterization of α-cpβ and α-cpβ-*hug* is shown in Figures 26 and 27. Molecular weights (MW) were calculated from ESI-MS m/z ratios with Equations 1 and 2 (m_H = 1.00794) and are compared to their expected MWs in Table 6.

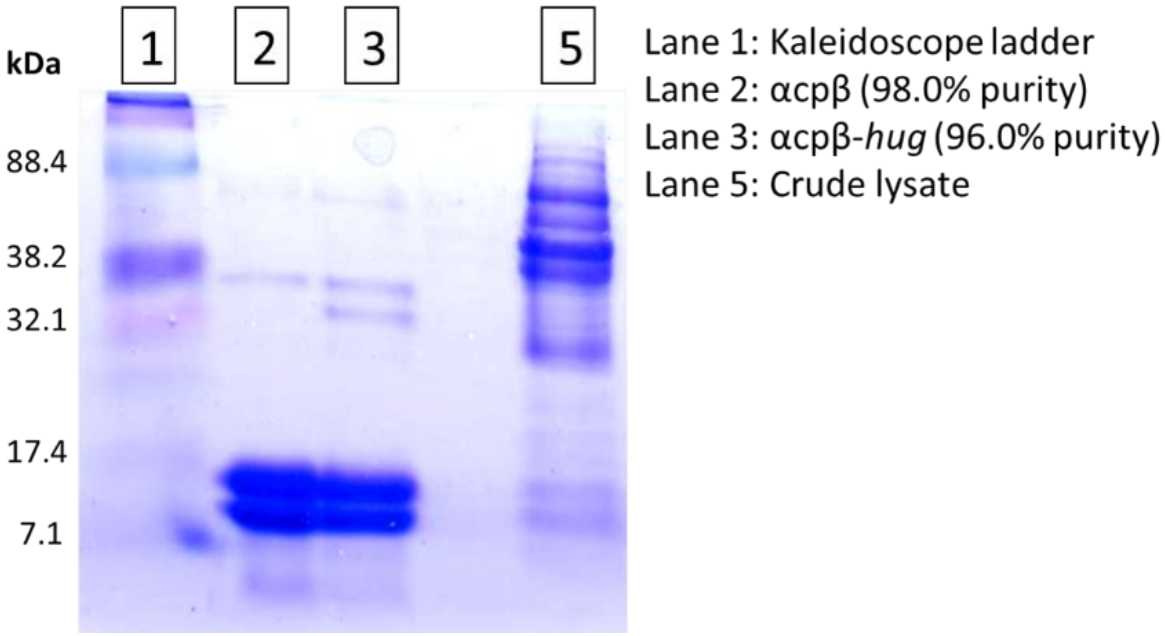

Figure 25. SDS-PAGE of α-cpβ proteins. Lane 1 is a protein ladder, standard molecular weights of select bands are indicated to the left. Lane 2 and 3 are α-cpβ and α-cpβ-*hug* samples at high purity (98.0 and 96.0 %). Lane 5 is a crude lysate of α-cpβ prior to purification efforts.

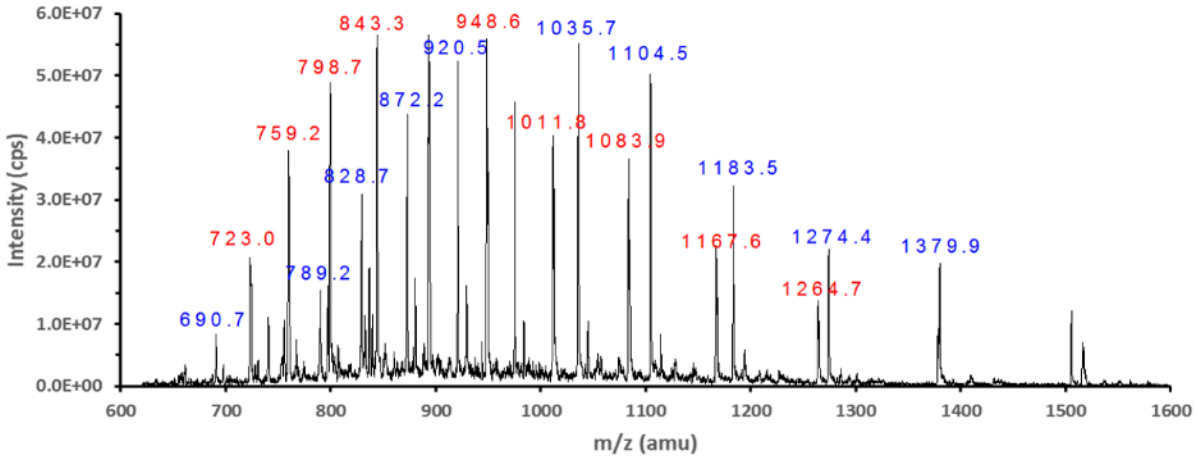

Figure 26. ESI-MS of α-cpβ at high purity (98.0%).

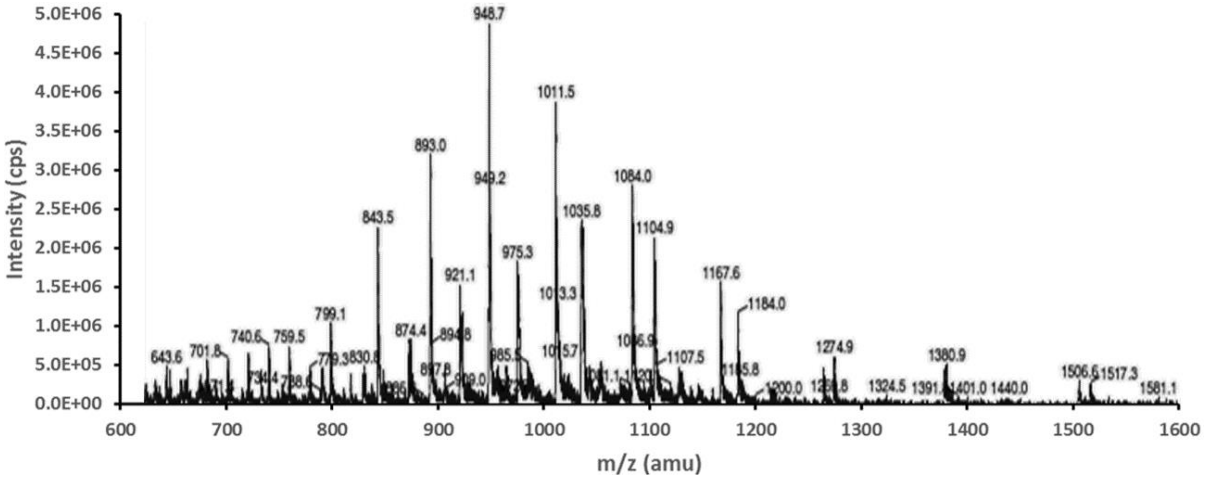

Figure 27. ESI-MS of α-cpβ-*hug* at high purity (96.0%).

$$z = x_2 - m_{H/}x_1 - x_2. \quad \text{(Equation 1)}$$

$$M = z \cdot x_1 - z \cdot m_H = (z+1) \cdot x_2 - (z+1) \cdot m_H \quad \text{(Equation 2)}$$

Table 6. Expected and observed MWs for α-cpβ calculated from ESI-MS m/z ratios.

Expected MW α-cpβ	Observed MW α-cpβ
15,159.42 (α-globin)	15,158.75 (α-globin)
16,556.94 (cpβ-globin)	16,571.39 (cpβ-globin)

Unexpected higher MW contaminants

As shown in Figure 28, SDS-PAGE gels of α-cpβ that had been purified with freeze/thaw cycles between each purification step exhibited four bands of varying intensity at expected MWs for α-cpβ (15 and 16.5 kDa) and at unexpected higher MWs (30 and 33 kDa. Bands at 30 and 33 kDa suggest the presence of cross-linked proteins that are twice the mass of the α- (15 kDa) and cpβ- (16.5 kDa) globins, respectively. The intensity of these higher MW bands was shown to increase as a result of protein conversion to the cyano-met isoform and after repeated flash freeze/thaw cycles. Thus, the purification process was altered to exclude the flash freezing between purification steps. Samples were instead refrigerated at 4°C between chromatography steps. This modification reduced the presence of these bands to below detection levels (Figure 29).

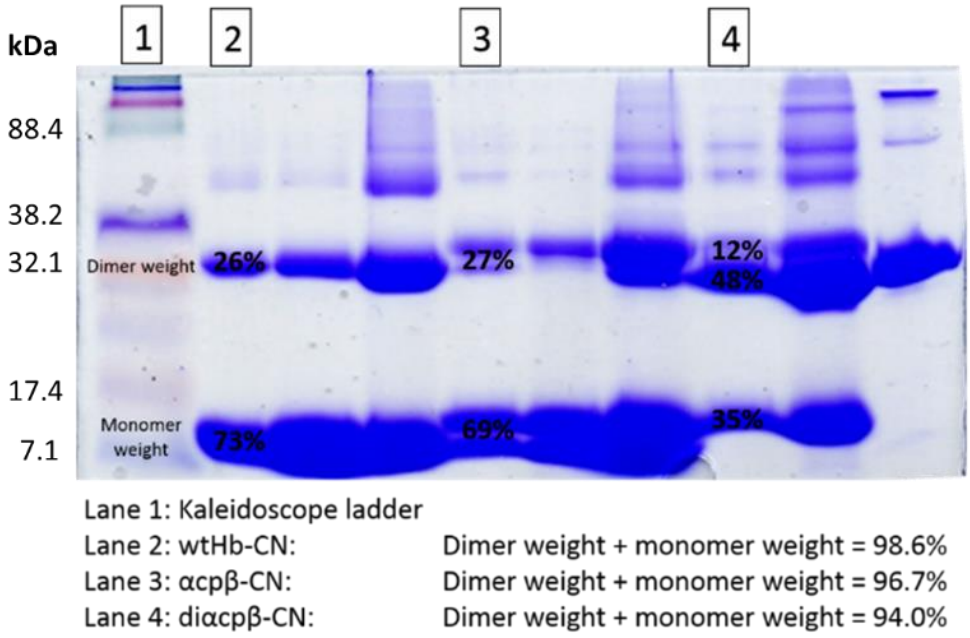

Figure 28. SDS-PAGE of cyano-met isoform hemoglobins. Lanes were overloaded to better assay impurity levels. 'Monomer weight' indicates approximate MW of monomeric globins and 'Dimer weight' indicates approximate MW of dimeric globins. wtHb should have 100% monomeric globins, α-cpβ should have 100% monomeric globins, and diα-cpβ should have 50% monomeric and 50% dimeric globins.

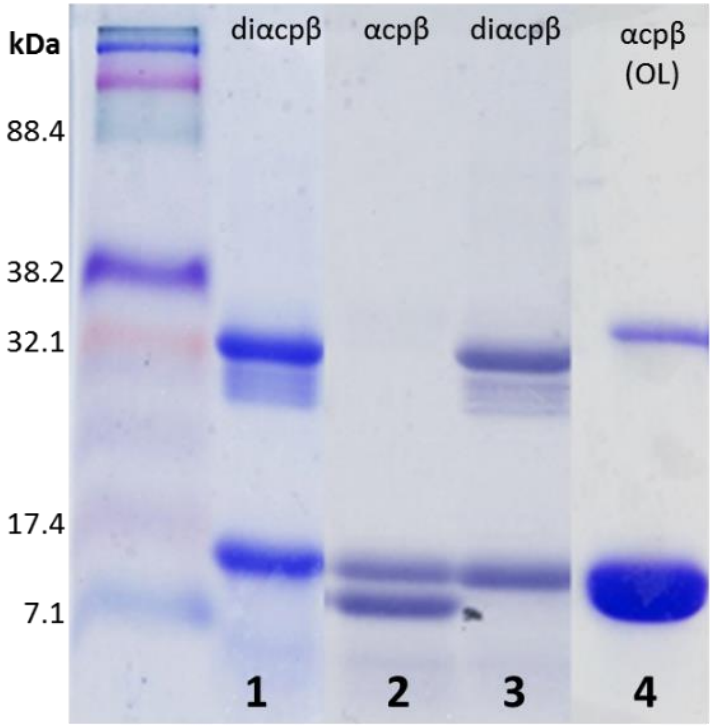

Figure 29. SDS-PAGE gel of α-cpβ and diα-cpβ at normal loading (lanes 2 and 3 respectively) and at over loaded concentrations to illustrate purity (lane 4) following elimination of freeze/thaw cycles.

Sparse matrices

Crystallization trials were performed, as describe in Materials and Methods, to produce diffraction quality crystals of α-cpβ and di α-cpβ, an example image can be seen in Figure 30. Roughly 1500 individual crystallization experiments were performed in ~70 individual 24-well crystallization plates. Before attempting to crystalize α-cpβ, data published for hemoglobin structures deposited in the Protein Data Bank (PDB) were examined for conditions that promote production of diffraction quality crystals (Table 7). Neutral pH and 20° C appeared to be common variables, which proved to promote production of diffraction quality crystals in α-cpβ and diα-cpβ as well. Sparse matrix gradients resulted in many wells containing microcrystals that appeared too small for diffraction studies directly, but indicate good conditions to screen around. The α-cpβ was tested against the following sparse matrices

(the number of promising hits from each matrix are given in parentheses): Cryo II (1), Hamilton Crystal Screen (3), Index (9), and PEG/ION II (1). Many hits resulting from the shotgun approach were very small micro crystals of less than 0.01 mm in diameter. Fourteen hits were identified for further optimization by forming a gradient of conditions around the mother liquor composition for each hit. The Index screen from Hampton Research produced the best results, closest to the ideal single crystals of greater than 0.1 mm in diameter free from visible defects, for all of the permutants.

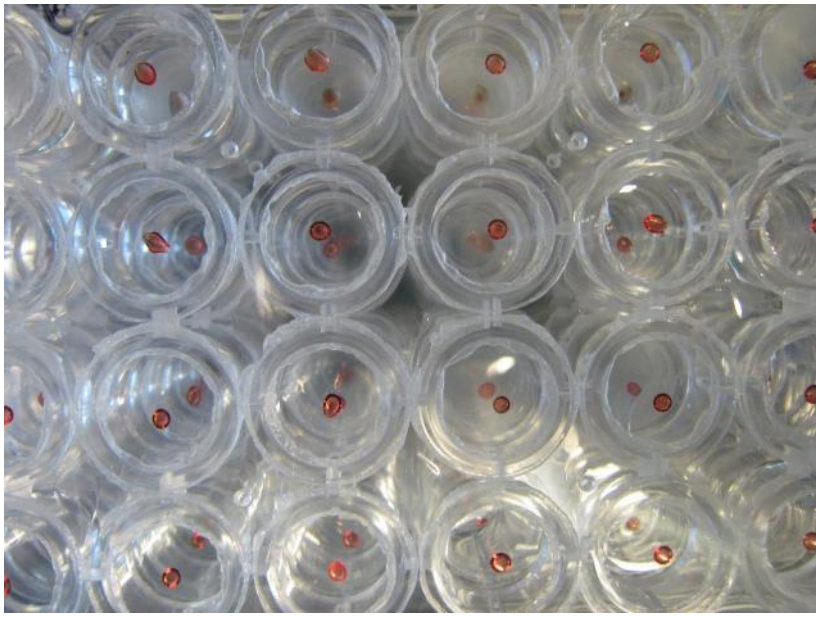

Figure 30. Image of stacked 24-well crystallization trays with 2µL α-cpβ crystallization trials (red dots).

Table 7. Crystallization conditions reported for hemoglobin structures in the PDB.

PDB ID:	Crystallant [M]	Buffer [M]	pH	Additive [M]	Temp. (K)
1bz0	$(NH_4)_2SO_4$ [2.3]	$(NH_4)_3PO_4$ [0.3]	6.5	Ferrous citrate [0.01]	
3qjb		$(Na/K)_3PO_4$ [2.3]	6.7		293
3nl7		$(Na/K)_3PO_4$ [2.3]	6.8		293
3nmm		$(NH_4)_{3/2}(PO_4/SO_4)$ [2.5]	7.0		293
2w6v	$(NH_4)_2SO_4$ [3.2]	$(NH_4)_{2/1}H_{1/2}PO_4$ [0.4]	6.5	Dithionite (0.2 %)	
2dn1	Glycerol	$(Na/K)_3PO_4$	6.7		277
2dn2	$(NH_4)_2SO_4$	$(NH_4)_3PO_4$	6.5		293

Gradient trays

Gradient trays (Figure 31) were formulated (see Table 8) around these hits to explore the effects of small variations in crystallization conditions on crystal morphology. Trays labeled with a number only were stored at ambient room conditions (~25°C), while those with a number and the letter 'x' were stored at a constant 20°C. These gradient conditions produced numerous crystals with a variety of morphologies illustrated in Figures 32 – 39, crystals identified by blue arrows had desirable morphologies for X-ray diffraction.

Table 8. Tabulation of gradient crystallization conditions extrapolated from sparse matrix hits for α-cpβ.

Tray #	Sparse Matrix 'Hit'	Crystallant	Buffer [0.1M]	pH	Additive	[Additive]
1	HCS* 22	25-35% w/v PEG 4,000	Tris-HCl	7.5-9.0	NaOAc	0.2 M
2	Index 43	20-30% w/v PEG 3,350	Bis-Tris	5.8-7.0		
3	Index 44	20-30% w/v PEG 3,350	HEPES & Tris-HCl	7.5-9.0		
4	Index 76 & 77	20-30% w/v PEG 3,350	HEPES & Tris-HCl	7.0-8.5	$Li_2SO_4 \bullet H_2O$	0.2 M
5	Index 80 & 81	20-30% w/v PEG 3,350	HEPES & Tris-HCl	7.0-8.5	NH_4OAc	0.2 M
6	Index 78 & 79	22-32% w/v PEG 3,350	Bis-Tris	5.7-7.0	NH_4OAc	0.2 M
7	Index 70 &71	22-32% w/v PEG 3,350	Bis-Tris	5.7-7.0	NaCl	0.2 M
8	Index 82 & 83	22-32% w/v PEG 3,350	Bis-Tris	5.7-7.0	$MgCl_2$	0.2 M
9	Index 43	20-30% w/v PEG 3,350	Bis-Tris & Tris	6.5-8.0		
10	Index 43 & 44	14-24% w/v PEG 3,350	Bis-Tris & HEPES	6.5-8.0		
11	Index 78 & 79	22-32% w/v PEG 3,350	Bis-Tris	6.0	NH_4OAc	0.1-0.35 M
12	Index 70 &71	22-32% w/v PEG 3,350	Bis-Tris	6.0	NaCl	0.1-0.35 M
13	Index 43	21-26% w/v PEG 3,350	Bis-Tris	5.9-6.5		
14	Index 43	21-26% w/v PEG 3,350	Bis-Tris	6.7-7.3		

* HCS = Hamilton Crystal Screen

	PEG 3.35K (w/w) %					
pH	20%	22%	24%	26%	28%	30%
6.5	*1*	*2*	*3*	*4*	*5*	*6*
7.0	*7*	*8*	*9*	*10*	*11*	*12*
7.5	*13*	*14*	*15*	*16*	*17*	*18*
8.0	*19*	*20*	*21*	*22*	*23*	*24*

Figure 31. General layout of a gradient tray. Concentration of PEG 3,350 (w/w)% varies along the horizontal axis and pH varies along the vertical axis to create a grid of finely tuned conditions.

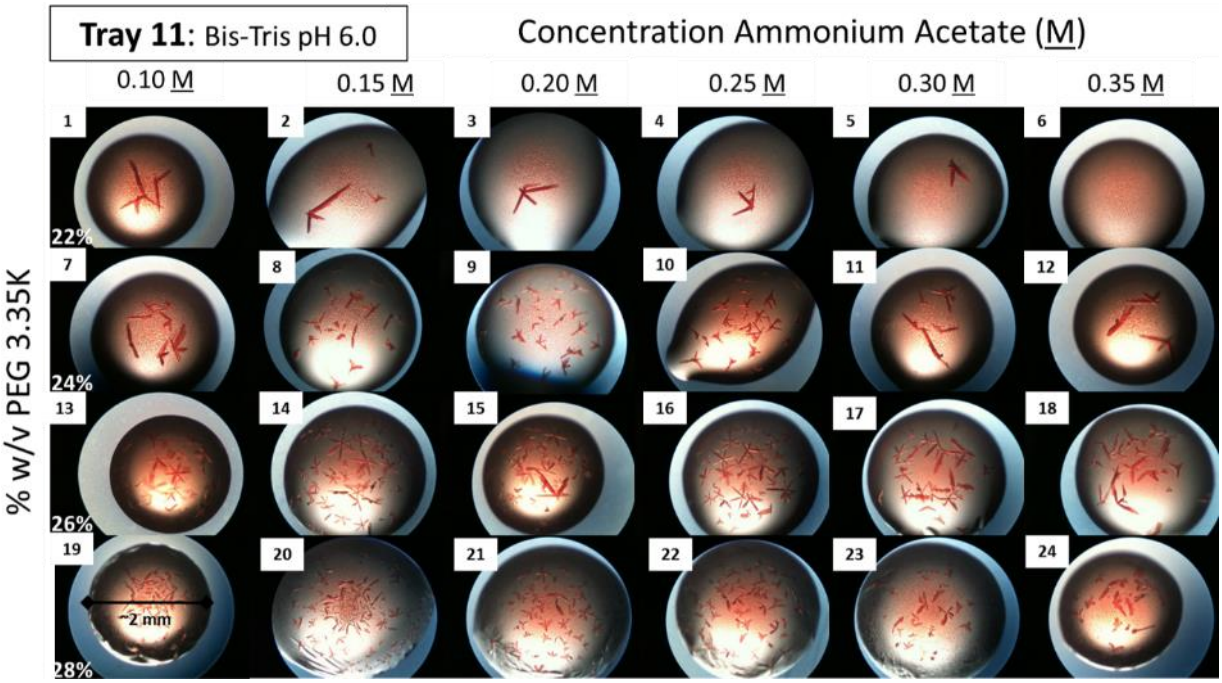

Figure 32. Tray 11 crystal results (100x light microscope) for α-cpβ.

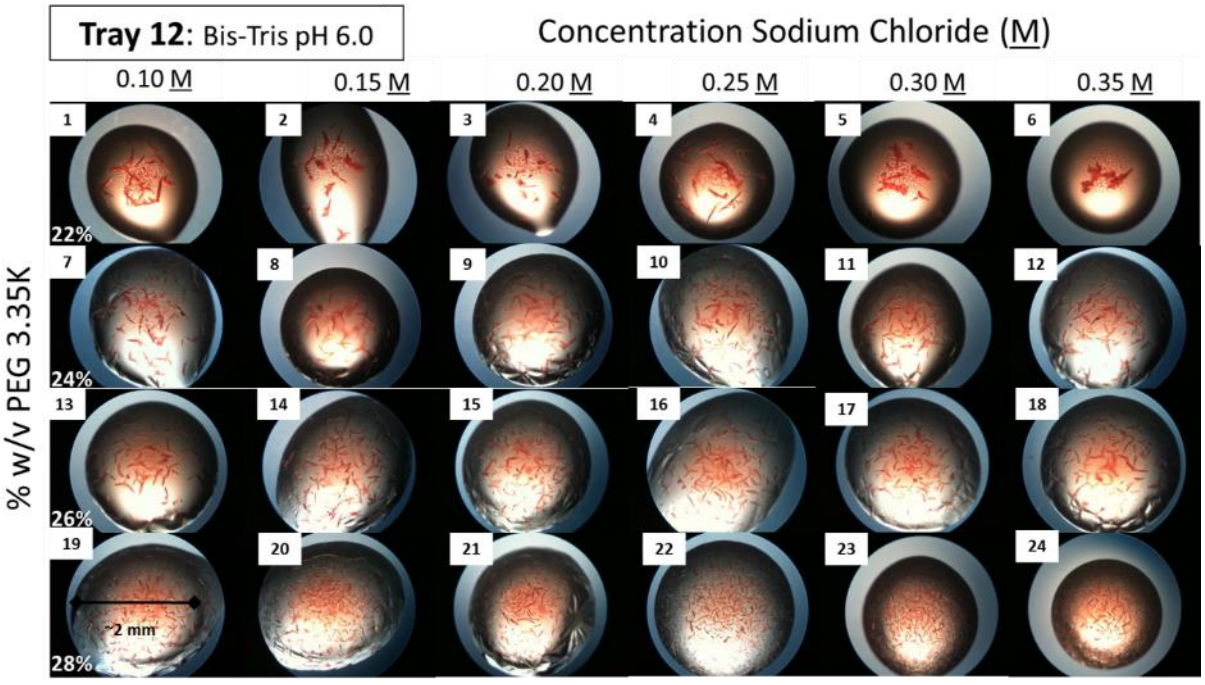

Figure 33. Tray 12 crystal results (100x light microscope) for α-cpβ.

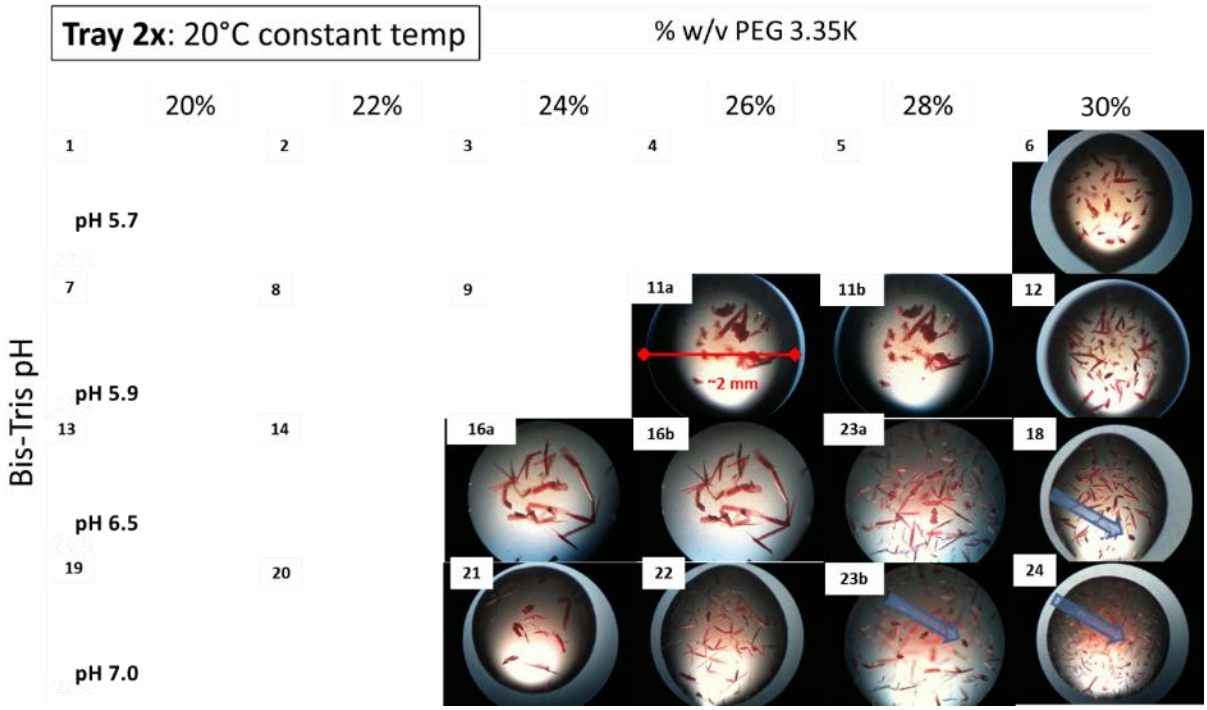

Figure 34. Tray 2x crystal results (100x light microscope) for α-cpβ.

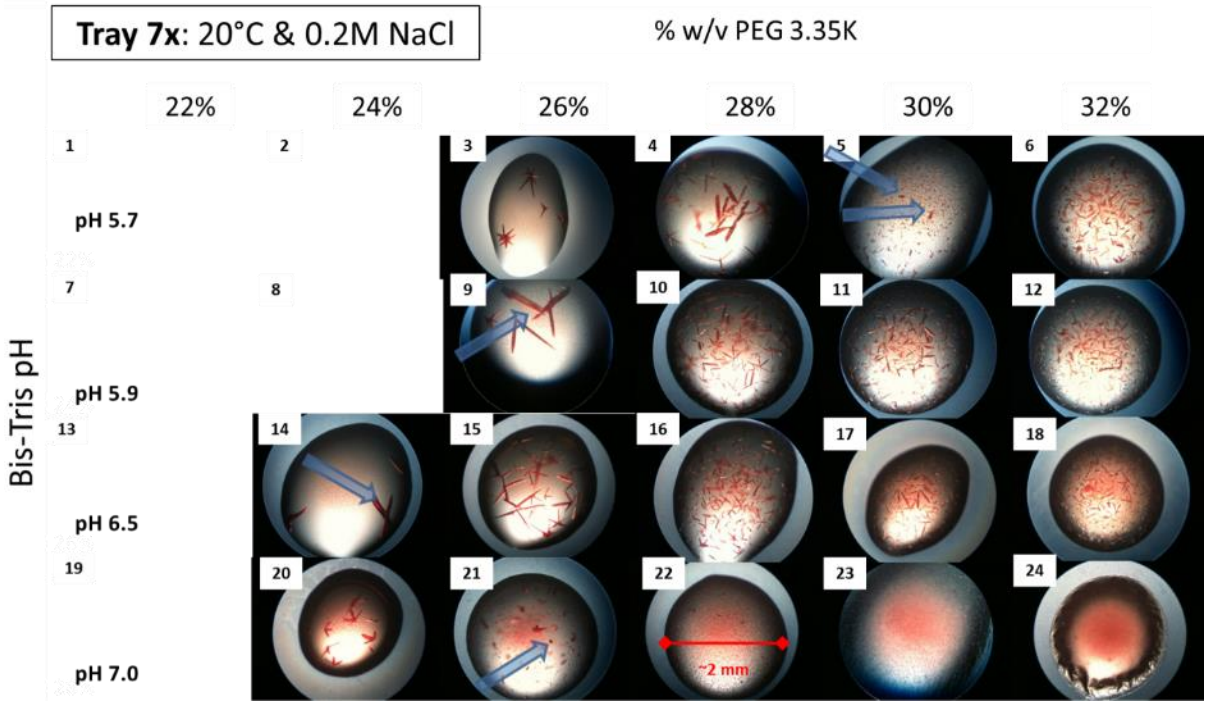

Figure 35. Tray 7x crystal results (100x light microscope) for α-cpβ.

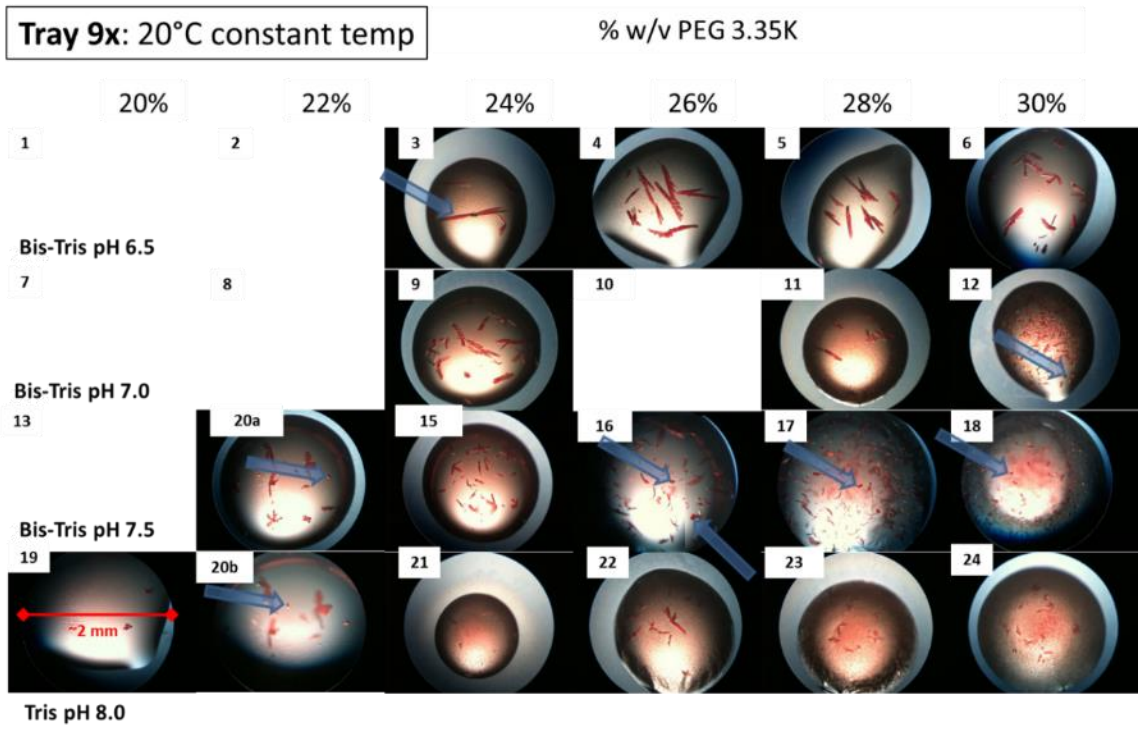

Figure 36. Tray 9x crystal results (100x light microscope) for α-cpβ.

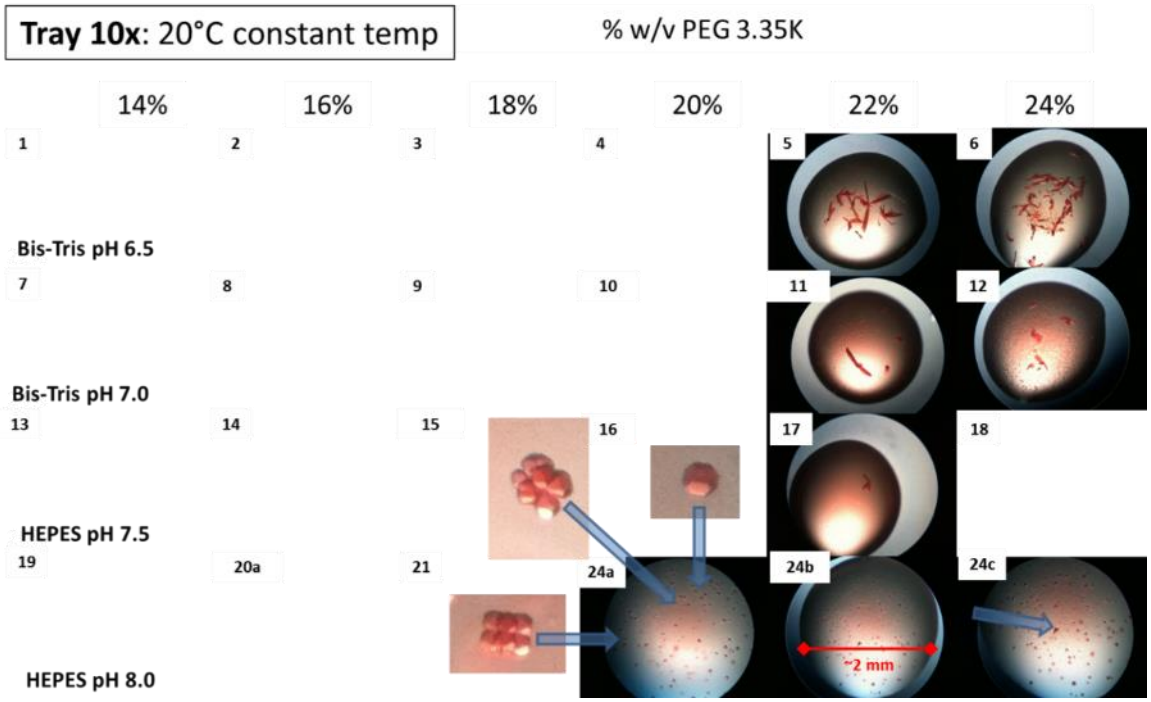

Figure 37. Tray 10x crystal results (100x light microscope) for α-cpβ. Three enlarged examples of microcrystals are shown for well 24a.

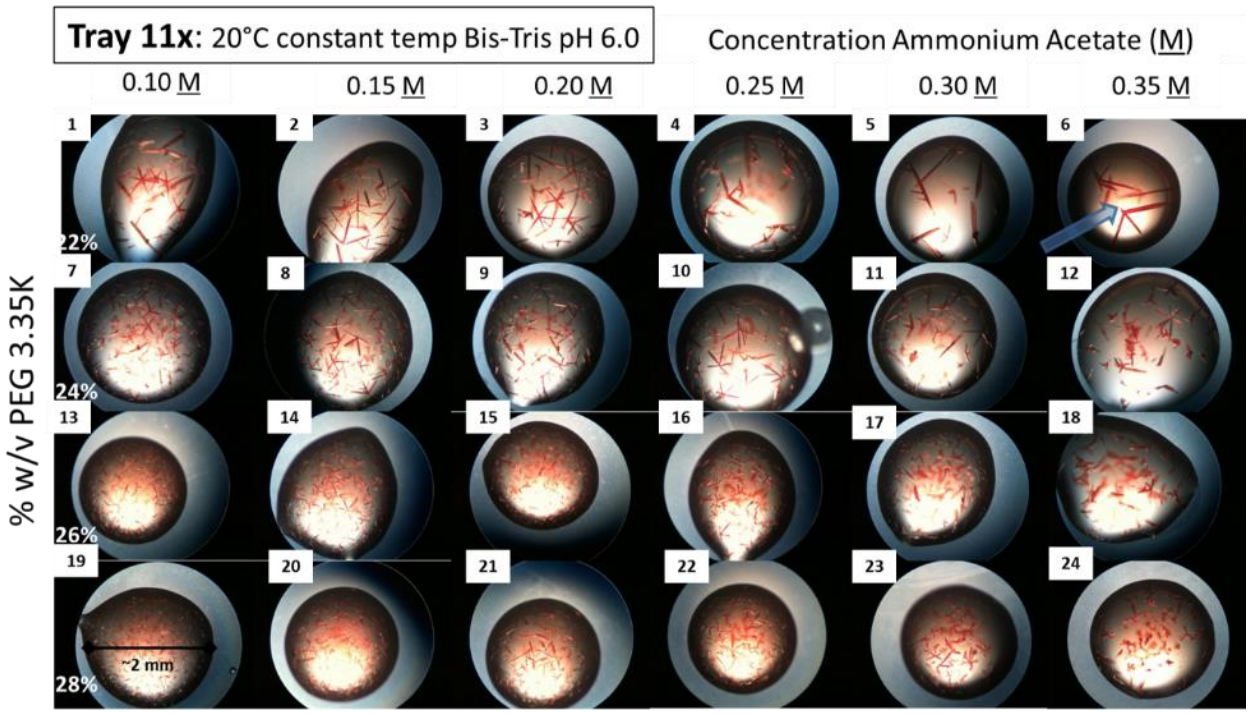

Figure 38. Tray 11x crystal results (100x light microscope) for α-cpβ.

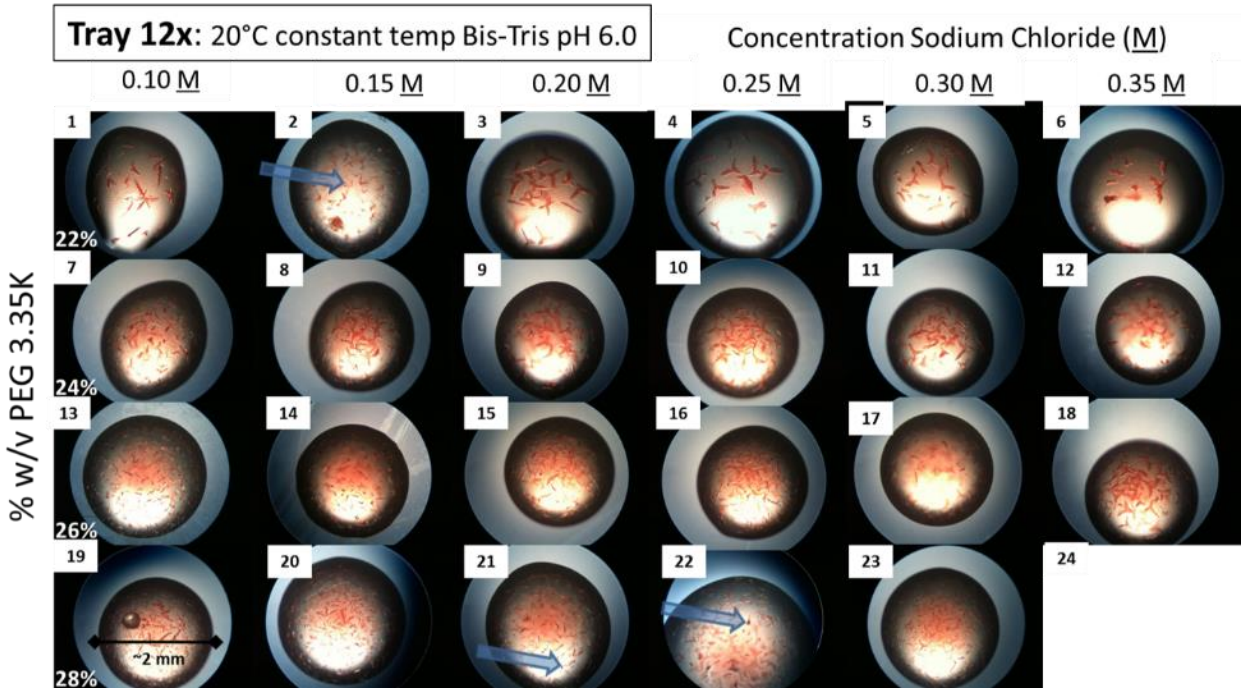

Figure 39. Tray 12x crystal results (100x light microscope) for α-cpβ

Attempted analysis of birefringence by polarized light microscopy

Polarized light microscopy can be used to evaluate the degree of crystallinity of target crystals [22]. The birefringence of crystalline materials can be observed with polarized light microscopy as a shift in the color of the crystalline material. Evaluation of crystals via this technique and selection of those most strongly birefringent, improves the chances of selecting a crystal for X-ray diffraction that will diffract at high resolution [22]. Application of polarized light microscopy met with varying degrees of success. The birefringence of the polystyrene crystallography trays severely impacted the sensitivity of the technique as the birefringence of protein crystals had to be detected over the background birefringence of the plastic, as illustrated in Figure 40. Even with this interference, detection of protein crystal birefringence was possible and strongly observed in some cases (Figures 41 and 42). Many promising protein crystals were identified with this technique; unfortunately the heat produced by the microscope resulted in destruction of all crystals analyzed. Though application of this technique has many hurdles to overcome, if non-birefringent trays could be obtained and if the crystals could be evaluated in a refrigerated area this technique could increase efficiency of crystal screening and enable collection of crystals more likely to diffract at high resolution [22].

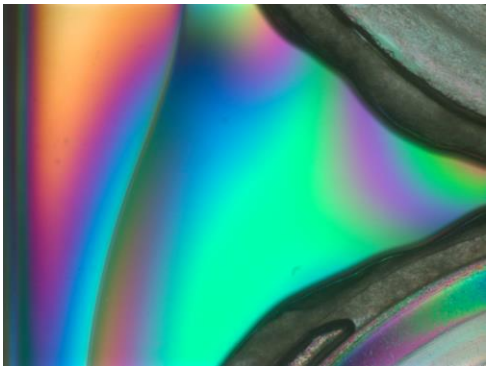

Figure 40. Example of birefringence of the polystyrene trays.

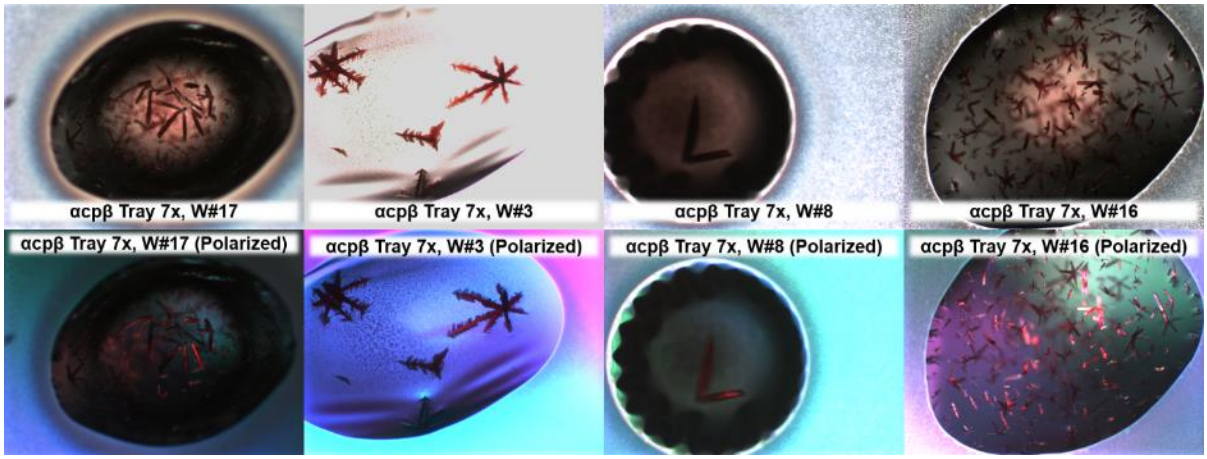

Figure 41. Examples of α-cpβ crystals under normal light (top) and polarized light with a cross polarized filter (bottom).Brighter sections in the polarized images are caused by higher birefringence which indicated greater order of the crystal.

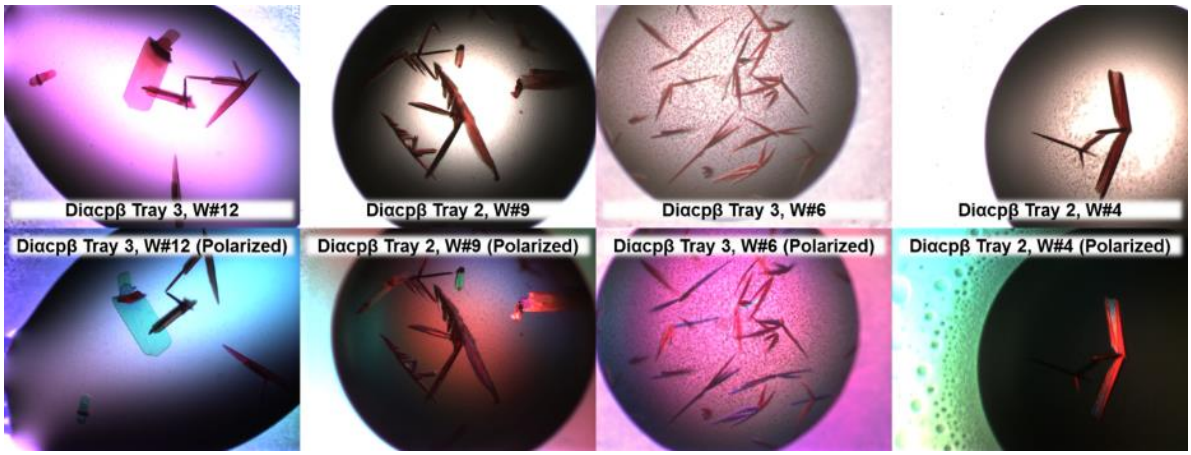

Figure 42. Examples of diα-cpβ crystals under normal light (top) and polarized light with a cross polarized filter (bottom). Brighter sections in the polarized images are caused by higher birefringence which indicated greater order of the crystal.

Catalogue of crystal morphologies

Many of the observed crystal morphologies were transient in nature, forming and disappearing at distinct time points. In an effort to track these morphologies and the conditions that promoted their formation, the crystals were imaged and organized into categories with qualifiers. A color coded system was developed to enable quick visual

assessment of a tray's potential to develop diffracting crystals. In addition to facilitating review of past results, this system enabled observation of trends in crystal forming ability as a function of mother liquor composition and time.

Optimized mother liquor

The conditions that promoted single crystal formation were similar to those that had previously crystallized wild-type hemoglobin and mutants developed by other labs. The ideal condition range, identified by the property of producing diffracting crystals of α cpβ, was a of pH 6.5 – 8.0 and 22-24% (w/w) PEG 3350. The condition that produced the crystal that was ultimately used to collect diffraction data to 2.7 Å was 0.1 M Bis-Tris at pH 6.5 with 24% (w/w) PEG 3350.

Selection of crystals to send for X-ray diffraction

Crystals were selected for qualities that are commonly associated with higher levels of diffraction, namely clarity, size, geometric regularity, and stability in cryobuffer solutions. Selection for crystals with these properties generally increases the chances of obtaining higher resolution diffraction data, but counterintuitive examples abound and a seemingly 'perfect' crystal is not guaranteed to diffract well. The crystals described in Table 9 were selected for X-ray analysis based on their likelihood to diffract. Of the 21 crystals sent for X-ray diffraction at the Fred Hutchison Cancer Research Center, 8 individual crystals were analyzed over 11 X-ray diffraction experiments (Table 10). Two crystals diffracted above 4 Å, five above 3 Å, two at 2.8 Å and one at 2.7 Å. The 2.7 Å data set was selected to calculate the electron density maps, and ultimately the structure of α-cpβ.

Table 9. Tray origin of α-cpβ crystals and cane position for X-ray diffraction. "Crystal" is abbreviated 'xtal'.

Crystal #	Origin Tray/Well	Cane ID	Position on Cane	Cryobuffer	Notes
1	T7/#7	MM1	1	20% Glycerol	
2	T7/#7	MM1	2	20% Glycerol	
3	T7/#7	MM1	3	20% Glycerol	
4	T7/#7	MM1	4	20% Glycerol	
5	T7/#7	MM1	5	20% Glycerol	
6	T7/#7	MM2	1	20% PEG 400	
7	T7/#7	MM2	2	20% PEG 400	
8	T7/#7	MM2	3	20% PEG 400	
9	T7/#7	MM2	4	20% PEG 400	
10	T10/#24	MM3	1	20% Glycerol	large xtal
11	T10/#24	MM3	2	20% Glycerol	small xtal
12	T10/#24	MM3	3	20% Glycerol	small xtal
13	T10/#24	MM3	4	20% Glycerol	small tetramer
14	T10/#24	MM3	5	20% Glycerol	small xtal
15	T9x/#3	MM4	1	20% PEG 400	big xtal w/small
16	T9x/#3	MM4	2	20% PEG 400	big xtal
17	T9x/#3	MM4	3	20% PEG 400	big xtal
18	T9x/#3	MM4	4	20% PEG 400	medium xtal
19	T9x/#3	MM4	5	20% PEG 400	medium xtal
20	T9x/#3	MM5	1	30% Glycerol	
21	T9x/#3	MM5	2	30% Glycerol	

Table 10. Crystal diffraction information for α-cpβ crystals.

File name	Buffer (0.1 M)	Crystallant (% PEG 3350)	Additive (0.2 M)	Crystal morphology	Cryobuffer	Detector distance	Scan time	Angles	Diffraction resolution	Space group	α,β,γ	a,b,c
hb1	Bis-Tris, pH 6.5	24%	N/A	small spike	30% glycerol	75 mm	5 sec	0°, 30°	~4 Å	C_2	192,128,152	90,105,90
hb2	Bis-Tris, pH 6.5	24%	N/A	large spike	20% PEG400	80 mm	5 sec	0°, 30°	~4 Å	C_2		
hb2_2						80 mm	5 sec	0°, 30°	~3 Å	C_1	119,137,137	69.5,69.7,80.7
hb2_3						70 mm	7 sec	0°, 30°	3.3 Å			
hb3	HEPES, pH8.0	24%	N/A	single xtal	30% glycerol	70 mm	5 sec	0°, 30°	10 Å			
hb4	HEPES, pH8.0	24%	N/A	single xtal	30% glycerol	70 mm	5 sec	0°, 30°	low	C_2		
hb5	Bis-Tris, pH 6.5	24%	N/A	large spike	20% PEG400	70 mm	5 sec	0°, 30°	~2.8 Å	C_2		
hb5_2						70 mm	5 sec	0°, 30°	~2.8 Å		222,52,128	90,125,90
hb6	Bis-Tris, pH 6.5	24%	N/A	large spike	20% PEG400	70 mm	5 sec	0°, 30°	~4 Å			
hb7	Bis-Tris, pH 5.9	22%	NaCl	large spike	30% glycerol	70 mm	5 sec	0°, 30°	~4 Å	C_2	235,55,136	90,124,90
hb8	Bis-Tris, pH 6.5	24%	N/A	tetramer	30% glycerol	75 mm	5 sec	0°, 30°, 60°	2.7 Å	D_2	62,77,97	90,90,90

Diffraction Data

The crystal hb8 diffracted to the highest resolution and yielded space group dimensions expected for a molecule the size of a hemoglobin tetramer (64kDa). A high resolution (1.25 Å) human oxy-hemoglobin structure (PDB ID: 2dn1) was employed to calculate the phasing information of the α-cpβ model and to solve the phase problem with the Phenix software suite [23]. Diffraction data were collected at lower resolution (Figure 43) and ultimately to 2.7 Å resolution (Figure 44), which enabled identification of the main chain atoms, the shape of side chains, and tightly bound water molecules. This level of resolution provides clear evidence of the expected locations of the new linker and termini in the cpβ globin, as well as native-like association of the α-cpβ heterodimers within a 64 kDa complex.

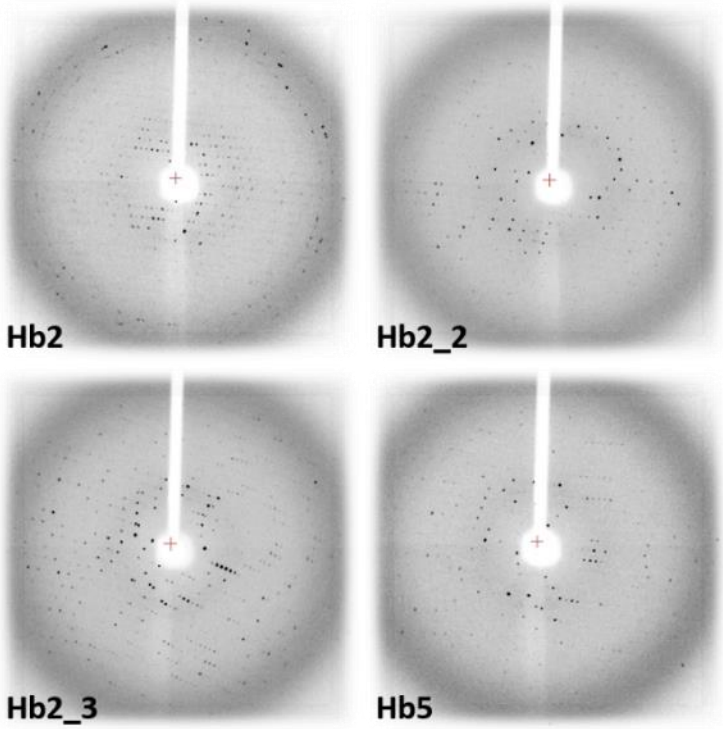

Figure 43. Diffraction patterns of α-cpβ crystals at lower resolution. Patterns originated from crystals indicated in lower left of each diffraction pattern.

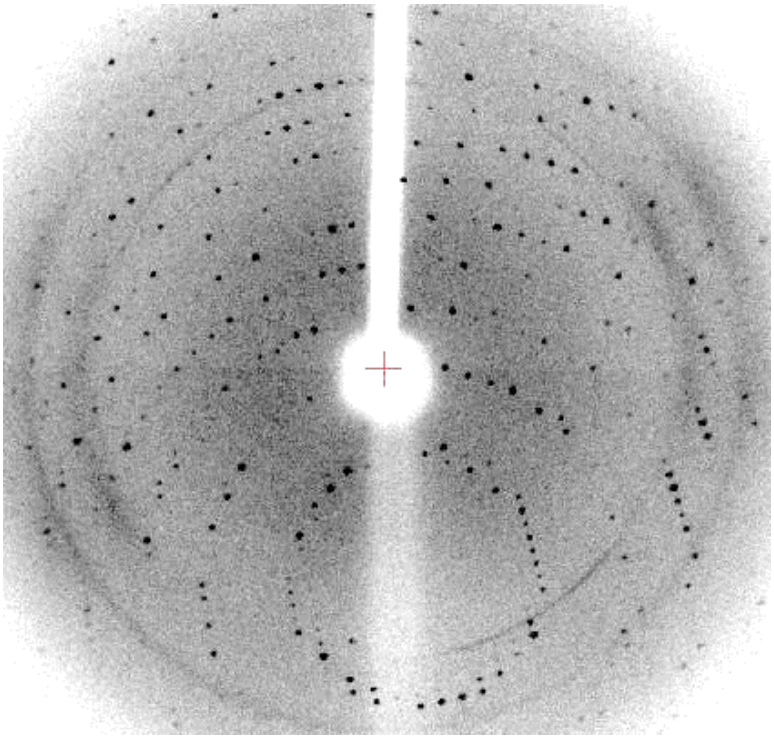

Figure 44. One diffraction pattern collected for the data set used to calculate electron density maps used to refine the α-cpβ model from the highest resolution diffraction data set.

Refinement

The first step in the process of refining the initial model was to look for areas of mismatch between the electron density and of the density predicted by the model. If the density was absent for a portion of the model (side chain or residue), it was reoriented to a position supported by electron density, or it was deleted for subsequent modeling when the structure was more highly refined. If there was density present that was not occupied by the model, it was populated with a misplaced residue or water molecule. The molecular modeling program, WinCoot, [24] has different levels of user control that enable either direct manual manipulation or computer aided optimization of the model best fit to electron density using biophysical constraints to minimize intramolecular strain within the model.

The goal of refinement is to optimize bond angles, bond lengths, and occupation of electron density. This requires compromise between strain minimization (achieved by alteration of the covalent bonding parameters) and maximization of electron density occupation. In the initial stages of model refinement, significant bias was placed on fitting the model to the electron density. After the initial optimization for electron density this bias was removed, and the model was adjusted to conform to standard parameters for bond angles and lengths (within an acceptable range of variation).

The original model was modified through a process combining manual manipulation of the structure and computer aided optimization. The general process was to examine the model produced by the Phenix software, with the molecular modeling software WinCoot. WinCoot enabled the visualization of the electron density plots produced by Phenix and fitting of the preliminary models to this electron density.

Further refinement of the initial model was accomplished by comparing density predicted by the model to density supported by the data. The putative quality of intermediate models was assessed by determining how accurately the model predicted the observed data. There is a bias created by this process because it is possible to over fit the model to the calculated electron density maps. To prevent this over fitting of the model to a structure that is not correct, a percentage of the original data points are reserved and omitted from computationally aided refinement steps. These omitted data points are employed to validate the model refinements through a process that discerns if they could be produced by the electron density of the refined model. The degree to which these parameters validate the

current model are indicated by a score, R_{free}, which has a value between 1 and zero, with zero being a perfect fit. An R_{free} value of less than 0.27 was our target for the refinement.

Refinement of the α-cpβ model was initiated by fitting a model of wild type tetrameric hemoglobin (PDB ID: 2dn1), omitting all ligands, to the electron density maps produced by the 2.7 Å resolution diffraction data. This was followed by deletion of any residues that were not supported by the electron density maps. The residues linking the G- and H- helices of the cpβ-globin had very weak density, indicating that they were not well resolved, or absent in the crystal structure, and so were deleted. This observation is not surprising given the expectation that α-cpβ should have new termini at this location. Following this initial truncation of the model, the electron density maps were recalculated and used to adjust the secondary structural elements of the model to better fit the electron density.

The initial search model lacked the functionally essential heme group. The correct placement of this iron-containing ligand in the α-cpβ structure greatly improved the quality of the electron density maps calculated with the model, due to the relatively dense iron atom at the core of the heme group. Placement of the heme ligand was strongly supported by the observation of electron density, which was isolated in the putative α-cpβ heme binding pocket, and had a volume that closely matched the molecular dimensions of heme. Once in place, the quality of the electron density maps markedly improved, and density supporting a ligand bound to the heme group became apparent. This density was fit with both a single oxygen atom to simulate the aquomet, or water bound, form of Fe^{3+} hemoglobin and dioxygen to simulate the Fe^{2+}- O_2 bound form. The dioxygen better occupied the density

provided by the diffraction data and yielded a greater reduction in R_{free}. Thus, we concluded that our structure was of an oxy (Fe^{2+}) hemoglobin, which was expected as the α-cpβ protein was crystallized under ambient atmospheric conditions.

The process of fitting additional resides to this intermediate model began by constructing a generic peptide backbone in the areas of unfilled density using alanine or glycine residues. After the backbone had been correctly modeled with alanine and glycine residues, the density maps were recalculated and the correct amino acid was exchanged for the substitute residue. This process was much more effective, on average, than attempting to insert the intended amino acid directly. The electron density of the side chain atoms was relatively low compared to those of the peptide backbone, especially for residues with flexible side chains (lysine, aspartic acid, etc.). The majority of the amino acids predicted by the α-cpβ gene were supported in their correct order through the application of this process, with the exception of C-terminal residues on the α-globin and the cpβ-linker residues on the cpβ-globin. It is likely that the density of these residues was difficult to extract from the data due to the relatively high conformational flexibility expected for residues in these positions.

After 48 rounds of manual and 172 rounds of automated structure and map refinement, the peptide backbone and side chains for these residues were placed with good to poor fits to the electron density data. Residues with poor fits in the final model include those residues in the middle of the cpβ-linker (B31Gly and C28Gly) and a terminal residue (A140Arg) where there was very little electron density to constrain the fit. The linkers were not designed to adopt a particular structure, conformational flexibility is a common characteristic of such

synthetic linkers. The lack of electron density for this region of the protein structure suggests it is dynamic and lacks significant interactions with other parts of the hemoglobin structure. The addition of water molecules to the model in later stages of refinement led to better agreement between the linker and the density map. The 157 water molecules that were added to the structure were validated by the density map and application of standard biophysical constraints (e.g., H-bonding distances, angles, etc.). Many more waters could have been placed, but they violated the maximum distance possible for a hydrogen bonded water molecule.

We speculate that it may be possible to design stabilizing interactions between the linker residues and other elements within the protein. The location of the linker at the interface between cpβ-globins presents both potential problems as well as potential opportunities. Linker–linker interactions may be difficult or impossible to design with current computational design methods. The symmetry of the Hb tetramer may promote stabilizing complementary linker-linker interactions that favor the T state by filling the cavity where 2,3-BPG normally binds.

Validation parameters were utilized to evaluate the model as an accurate structure, supported by the diffraction data. These validation parameters, shown in Figure 45 were R_{work}, R_{free}, RMSD(bonds), RMSD(angles), Clashscore, and Average B. The R_{work} value quantifies the level of agreement between the X-ray diffraction data and the crystallographic model (Equation 3) by relating calculated (F_{calc}) and observed (F_{obs}) structure factors (Equation 4), where '*I*' is intensity and '*hkl'* are Miller indices. The R_{free} value behaves in the same fashion

as R_{work}, but is employed to assay accuracy of the model [25]. This is accomplished by comparing computed structure factors with a random batch of reflections that were excluded from the modeling and refinement process [25]. 'RMSD(bonds)' indicates root mean squared deviation from standard bond lengths. 'RMSD(angles)' indicates deviation from standard bond angles. 'Clashscore' quantifies the degree of clashing between the calculated electron clouds of adjacent atoms. 'Average B' indicates the average temperature of atoms within the model by averaging their individual temperatures, which are quantified by a B-factor value (Equation 5, where '*u*' is the displacement from the scattering center and '*B'* is the calculated B-factor value). The final model agreed very well with other published models with similar resolution. Comparison between values of the validation parameters for our structural model and the published data is presented in Figure 45.

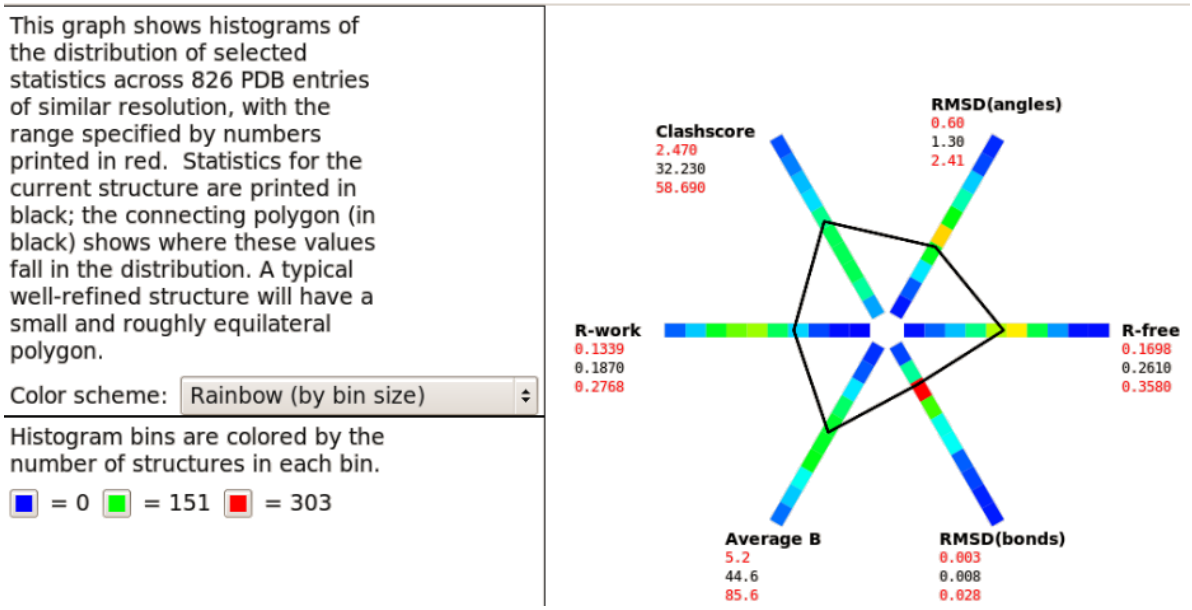

Figure 45. Histograms of relevant validation parameters for molecular models with similar resolution to the α-cpβ model.

$$R = \frac{\sum ||F_{\mathrm{obs}}| - |F_{\mathrm{calc}}||}{\sum |F_{\mathrm{obs}}|}$$

(Equation 3)

$$I_{hkl} \propto |F(hkl)|^2$$

(Equation 4)

$$B = 8\pi^2 \langle u^2 \rangle$$

(Equation 5)

The initial R_{work} and R_{free} values were 0.3663 and 0.375, respectively, indicating that the search model was a reasonable starting structure, but significant modifications would be needed to make it acceptable. A lower number for both parameters is desirable as it indicates the degree to which the model fits the data and the level of bias that has been imposed on the interpretation of the data by the model. The refinement goal was an R_{work} of 0.22 and an R_{free} of 0.27, significantly lower than the starting values. After 18 rounds of manual refinement combined with >170 rounds of computer aided refinement, an R_{work} of 0.1833 and an R_{free} of 0.2460 was achieved (Tables 8 and 9). An artificially low R_{work} can be achieved by over fitting the model to the data, we do not believe that that is the case with our model as the R_{free} is within acceptable limits, and the validation parameters of our structure compared very well to those of structures with similar resolution and number of residues. These values indicate that we have produced a model that very closely represents the true structure of the α-cpβ hemoglobin protein in complex with heme and oxygen.

Table 11. Progress of model refinement as a function of model fitting/electron density map refinement iterations.

	Initial	Round 1	Round 4	Round 8	Round 90	Round 188
Rwork:	0.3663	0.2727	0.2750	0.2103	0.1870	**0.1833**
Rfree:	0.3753	0.3050	0.3022	0.2955	0.2610	**0.2460**
Description:	PDB ID: 2dn1 w/out ligands	Trimmed termini	Backbone fit to maps	Sidechains fit to maps	Iterative revisions	**Final model**

Table 12. X-ray data collection and refinement statistics for the final α-cpβ model

X-ray Data Collection and Refinement Statistics	
data set name	α-cpβ-Hb
beamline	
space group	P 21 21 21
resolution (Å)	28.61 - 2.72
no. of unique reflections	14234 (1376)
unit cell (Å)	**a** = 64.454
	b = 79.444
	c = 99.751
angle (deg)	**α** = 90
	β = 90
	γ = 90
completeness (%)	99.87 (98.85)
⟨I⟩/⟨σ(I)⟩	23.52 (5.87)
Refinement Statistics	
resolution reflection used (Å)	30-2.25
no. of protein residues in the model	587
no. of water molecules	286
R-factor(%)	0.18330 (0.2095)
R_{free} (%)	0.2460 (0.2989)
rmsd for bond legths (Å)	0.0189
rmsd for bond angles (deg)	1.53
φ, ψ most favored (%)	95.63
φ, ψ additionally favored (%)	4.2
φ, ψ generously favored (%)	0
φ, ψ disallowed (%)	0.1
Wilson *B*-factor	47.04
average *B*-factor (Å^2)	45.3
protein	45.2
ligand	38.7
solvent	51.7

*Statistics for the highest-resolution shell are shown in parentheses

The final structural model is illustrated in Figure 46. Recall that $\alpha_1\beta_2$ interfacial interactions are essential for T state stability and optimal Hb function. A primary goal of this work was to establish the degree to which this interface had been perturbed by circular permutation of the

β-globin. Figure 47 shows the superposition of αcpβ and wtHb backbone structures with an overall RMSD of 2.67 Å. Though the inter-heterodimer binding affinity was greatly reduced by circular permutation of the β-globin (see Table 1), the crystal structure shows that the heterodimers still associate in a native-like tetramer. This was not unexpected given previous observations that α-cpβ exists predominantly in the tetramer state in solution at concentrations > 40 μM (see Figure 18; [11]). Sixteen heme-ligand states were identified for Hb structures deposited in the PDB (Table 13). Each of these was compared to the modeled heme-ligand state, and it was determined that the oxygen bound R-state present in PDB ID 2dn1 most closely matched that of the model. The refined structure also shows new density consistent with the expected location of the linker (Figure 48). Figure 49 shows electron density consistent with O_2 bound to the heme Fe^{2+} ion. Minor changes in the binding pocket environment can be observed (unfilled green density), which may have an effect on ligand binding affinity. Figure 50 shows that the cpβ-globin termini were relocated as expected, and that the N-terminal Met is removed by methionine aminopeptidase (consistent with ESI-MS data shown in Figure). The α-globin termini remain in their wtHb orientation as shown in Figure 51.

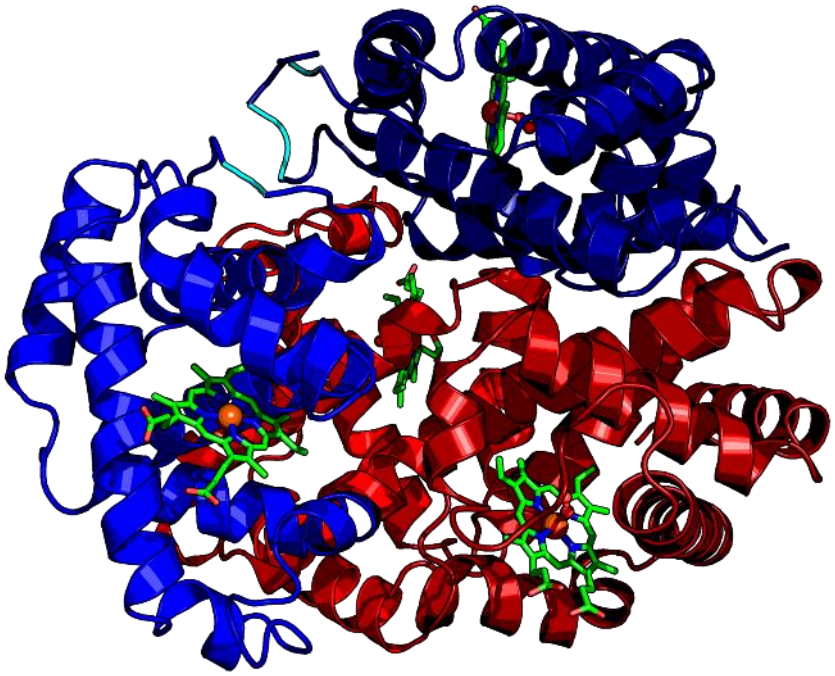

Figure 46. Cartoon model of the final α-cpβ structural model with heme groups and ligands modeled in stick format. α-Globins are in red and cpβ-globins are in blue.

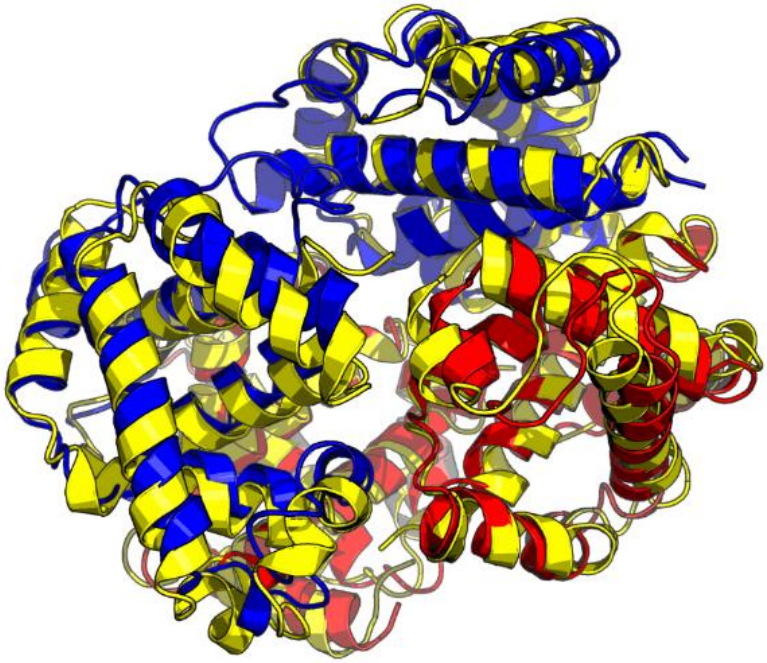

Figure 47. Overlaid **c**artoon models of the final α-cpβ structure and wtHb. α-Globins are in red, cpβ-globins are in blue, and the wtHb (PDB ID: 1gzx) is in yellow, RMSD (Å): 2.67.

Table 13. Example structures collected from the PDB of different R and T state configurations.

PDB ID	Bound Ligand	Resolution	State (R or T)	Space Group	a	b	c	α	β	γ
3p5q	aquo met	2.00	R	P41212	53.55	53.55	192.76	90	90	90
1hgb	aquomet	2.10	T	P21212	95.78	97.78	65.49	90	90	90
2dn3	carbonmonoxy	1.25	R	P41212	53.27	53.27	189.78	90	90	90
1mko	carbonmonoxy	2.18	RR2 (novel)	P21212	65.53	154.57	55.27	90	90	90
1yzi	carbonmonoxy	2.07	R3	P4122	61.46	61.46	174.75	90	90	90
1bbb	carbonmonoxy	1.70	R2 (novel)	P212121	97.5	101.5	61.1	90	90	90
1ird	carbonmonoxy	1.25		P41212	53.27	53.27	190.78	90	90	90
1aby	cyano met	2.60		P212121	102.54	115.17	56.7	90	90	90
2dn2	deoxy	1.25	T?	P21	63.1	82.91	53.65	90	99.29	90
1hga	deoxy	2.10	T	P21212	95.78	97.78	65.49	90	90	90
2dn1	oxy	1.25	R	P41212	53.44	53.44	191.71	90	90	90
1hho	oxy	2.10	R	P41212	53.7	53.7	193.8	90	90	90
1gzx	oxy	2.10	T	P21212A	97.05	99.5	66.11	90	90	90
1thb	oxy/deoxy	1.50	T	P21212	95.8	97.8	65.5	90	90	90
1hgc	oxy/deoxy	2.10	T	P21212	95.78	97.78	65.49	90	90	90
1bz1	deoxy	1.59		P21	63.3	83.6	53.8	90	99.4	90

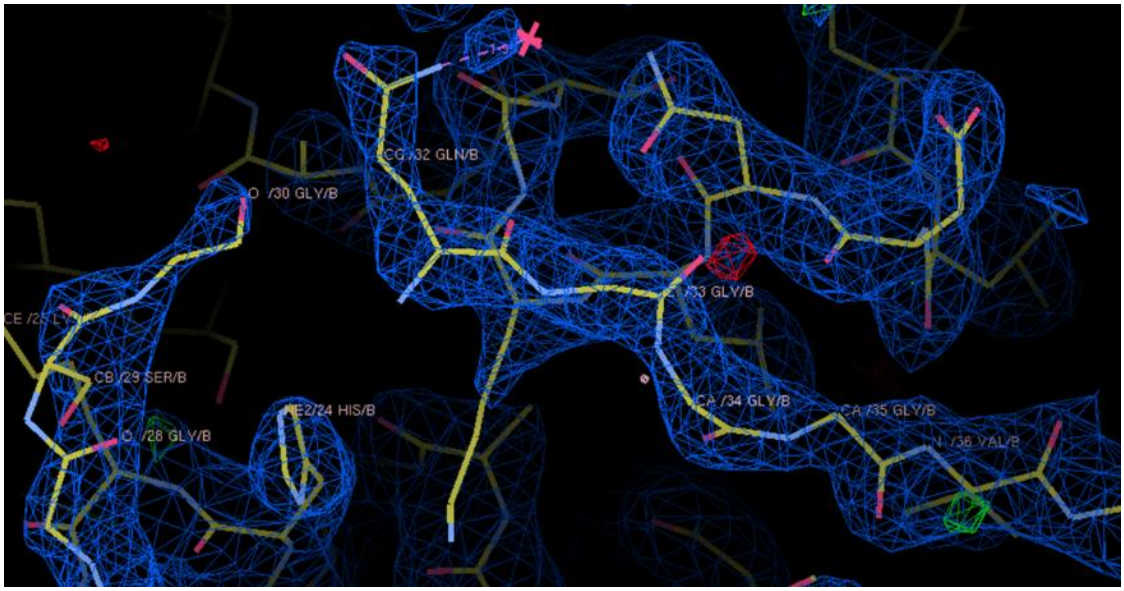

Figure 48. Electron density and stick model of the final α-cpβ structure of the cpβ linker. The composite map is in blue (calculated density) and the difference map is in green (missing density) and red (excess density). cpβ-Globin residues, heme ligand, and oxygen ligand are modeled in stick form. Green polygons indicate minor unresolved model deficiencies.

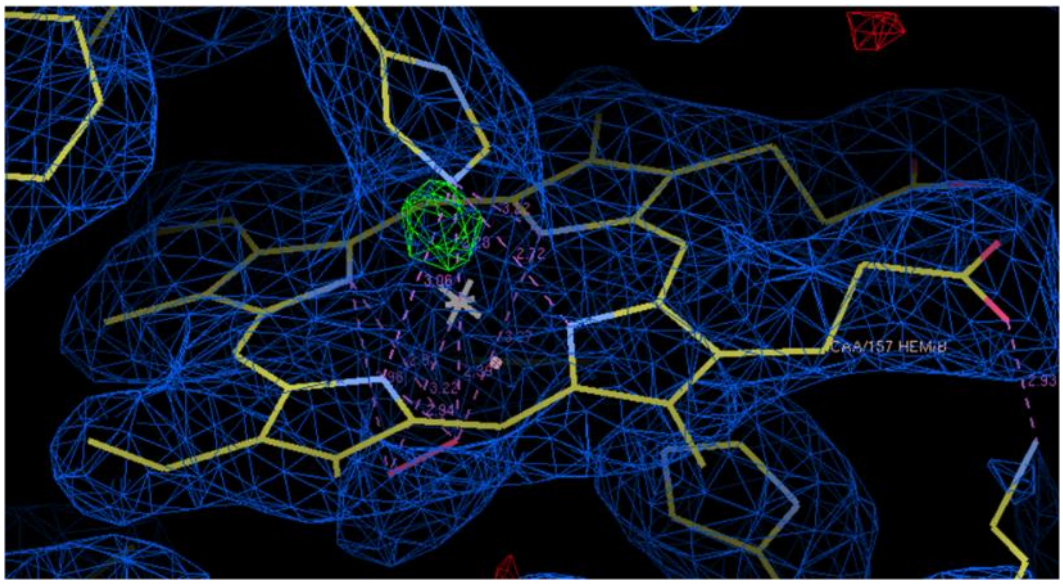

Figure 49. Structure refinement view of one cpβ-globin heme-binding pocket. The composite map is in blue (calculated density) and the difference map is in green (missing density) and red (excess density). cpβ-Globin residues, heme ligand, and oxygen ligand are modeled in stick form. Green polygons indicate minor unresolved model deficiencies.

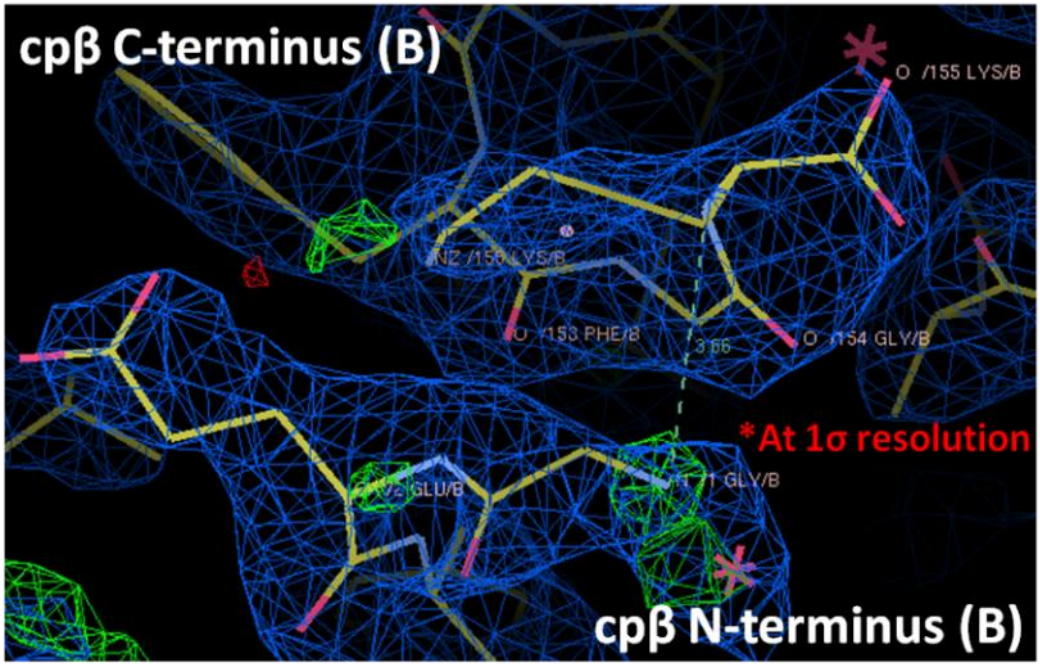

Figure 50. Structure refinement view of the cpβ-globin termini. The composite map is in blue (calculated density) and the difference map is in green (missing density) and red (excess density). cpβ-Globin residues and waters are modeled in stick form.

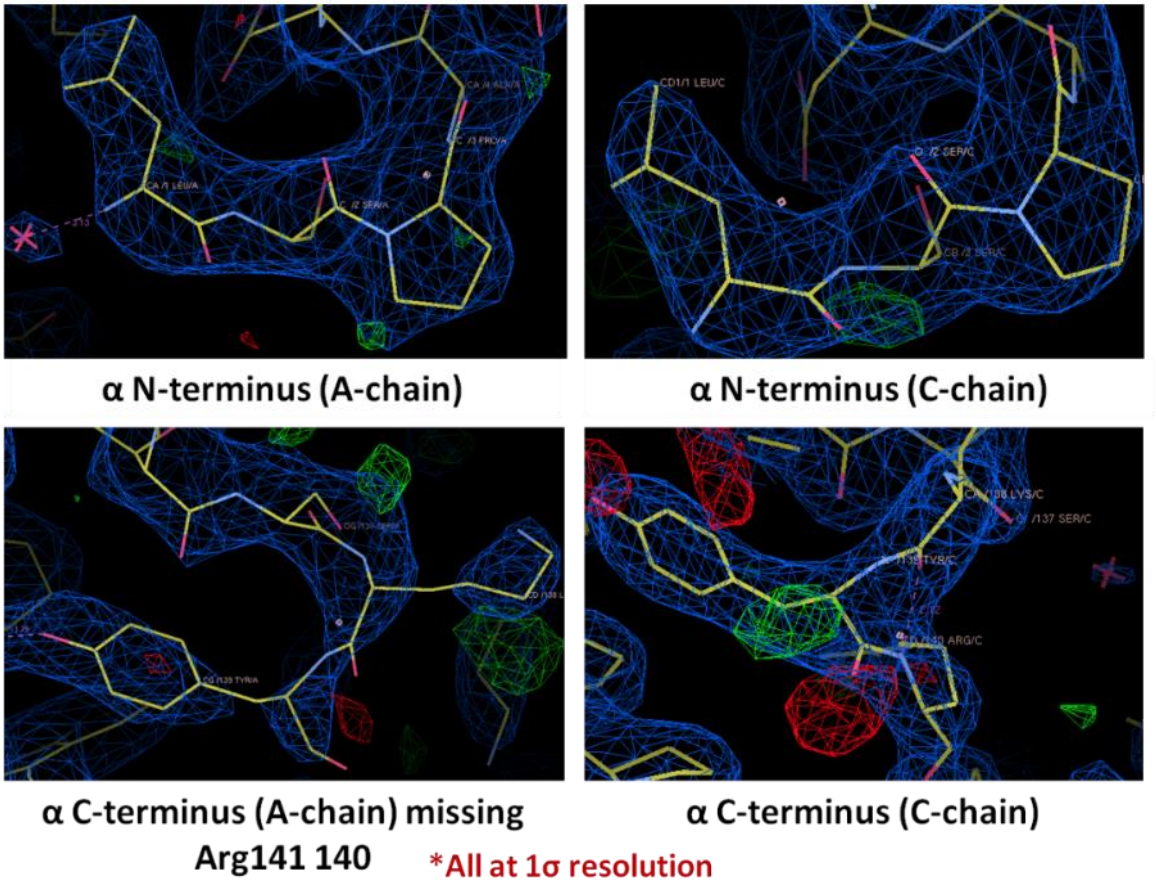

Figure 51. Structure refinement view of the α–globin termini. The composite map is in blue (calculated density) and the difference map is in green (missing density) and red (excess density). α–Globin residues and waters are modeled in stick form.

Ligand binding data show that α-cpβ has predominantly R state character, thus, we expected that the protein would crystallize in an R-like conformation. Figure 52 shows the superposition of α-cpβ subunits with high resolution structures of the corresponding wild type hemoglobin subunits in R- and T-states (PDB ID: 2dn1 and 2dn2). The cpβ subunits show greater similarity to wt Hb R state β globins; however, the α globins in α-cpβ adopt a conformation that is equally similar to wt Hb R- and T- state conformations (Figure 53). Figure 54 indicated that the new termini in the cpβ subunit are appropriately placed to allow insertion of the cpβ sequence into the G-H loop of the α globin using a pair of shorter (4-5 residue) linkers.

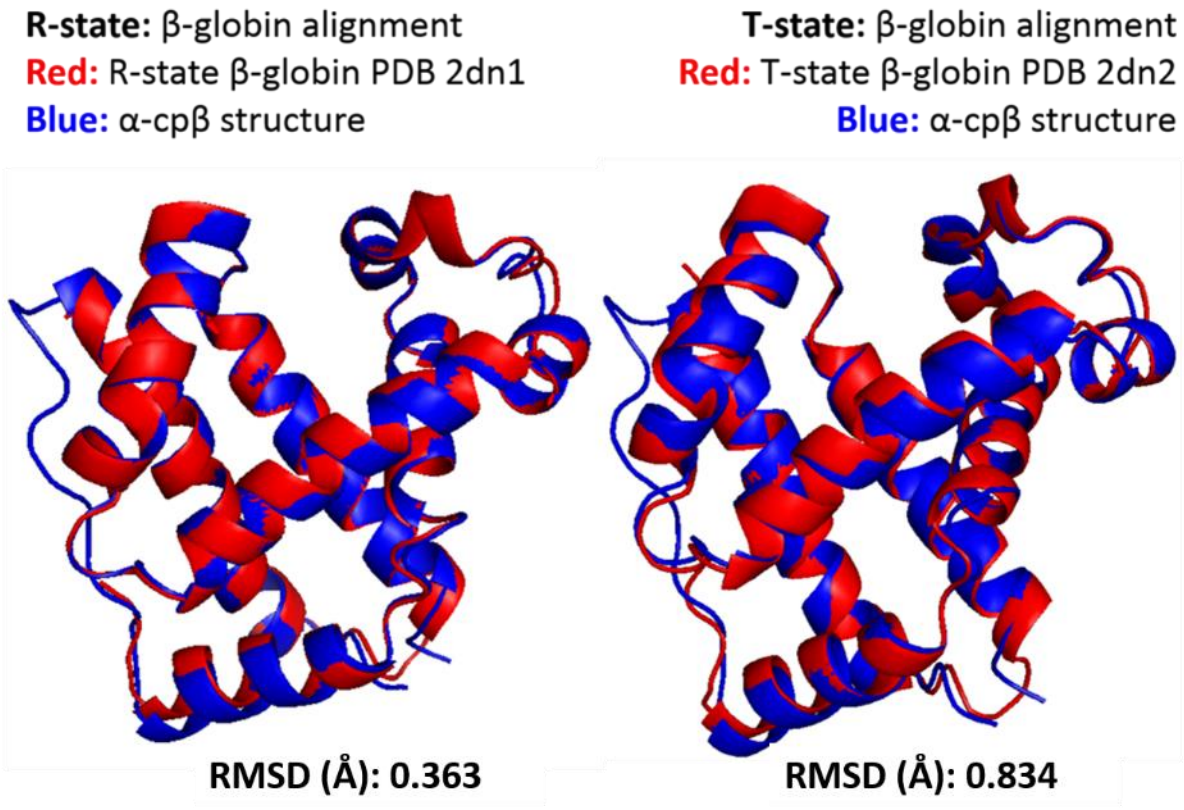

Figure 52. Structural homology of cpβ-globin to the oxy (R-state) and deoxy (T-state) wtHb structures. RMSD = root mean squared deviation, calculated with PyMol.

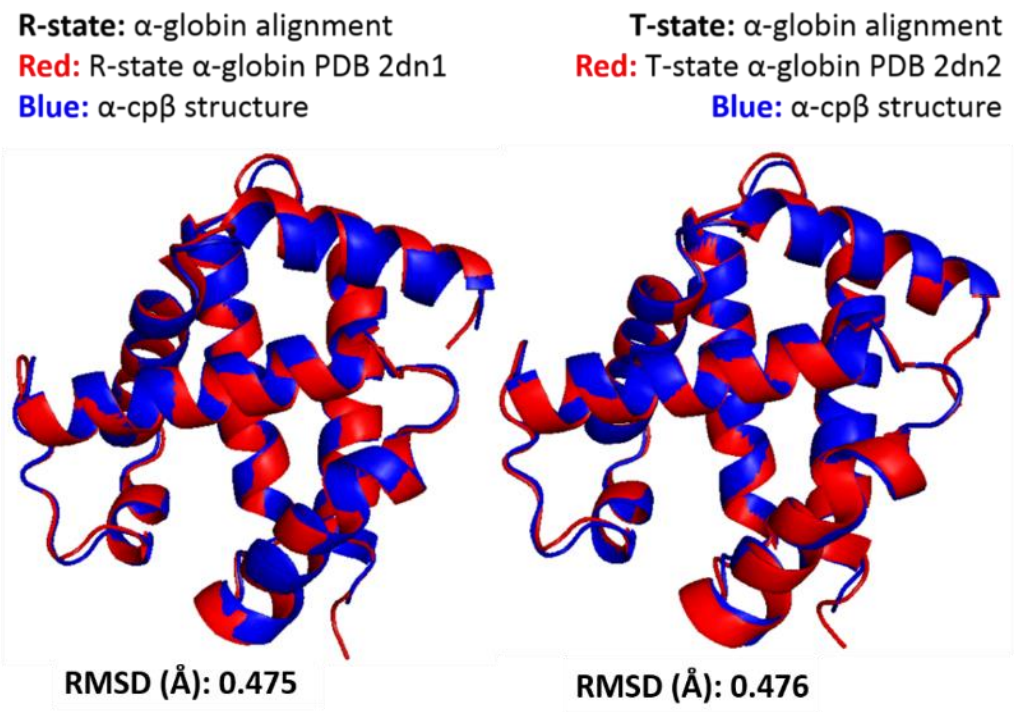

Figure 53. Structural homology of α-cpβ α-globins to the Oxy (R-state) and Deoxy (T-state) wtHb globin structures. RMSD = root mean squared deviation, calculated with PyMol.

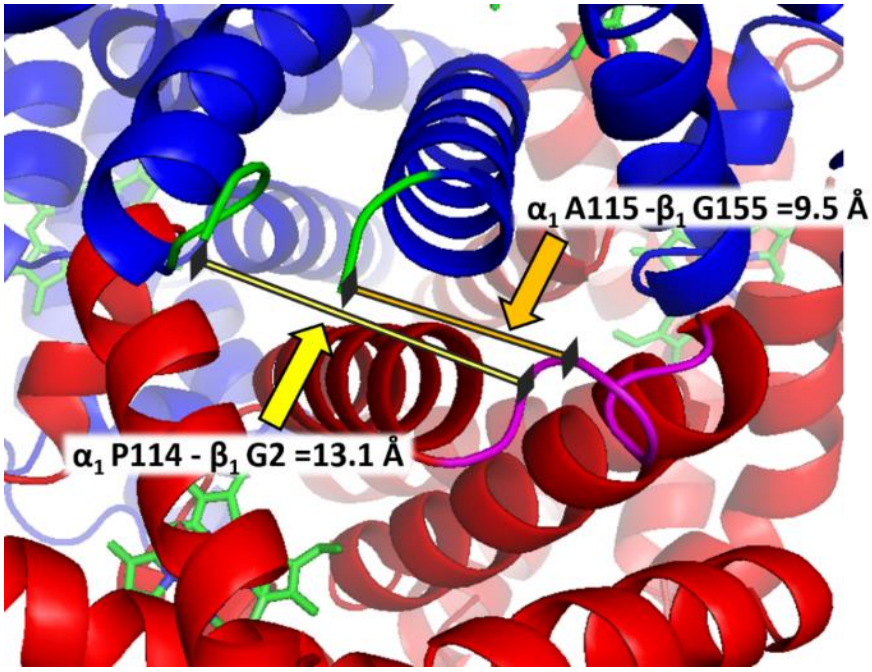

Figure 54. Cartoon model of the final α-cpβ structure, showing the potential for forming new covalent linkages between the α_1 and cpβ_1 globins, to yield a single-chain heterodimer (See Figure 14).

Conclusions:

Circular permutation was achieved without significantly perturbing the tertiary structure of α-cpβ hemoglobin. Circular permutation of the β-globin has been verified by this model, as the original termini clearly reside at altered positions in the structure and are connected by the intended linker (GSGGQGGG). The lack of the T state stabilizing interaction between α Lys40 and β His146 (the carboxyl terminal residue of wt β-globin) was also apparent in the α-cpβ model. The carboxyl group of β His146 has been incorporated into a new peptide bond in the engineered cpβ linker. This residue has also been pulled away from its native position by a distance of approximately 10 Å, rendering interaction with the amine group on α lysine 40 impossible. The engineered cpβ-termini are readily visible in the model and are in close proximity (~10 Å) to the α G-H helix loop, the linkage site for the fusion of the α- and β-globins in sc-α-cpβ. In related work from our lab, a scHb with a GAAG linker between α P114 and β G2 and a GGTGG between α A115 and β G155 has been successfully expressed.

Modeling of the cpβ linker revealed that the middle of the linker is highly flexible and that it appears to lack significant stabilizing secondary structural elements (β-sheets, α-helices, salt bridges, or hydrogen bonds). It is interesting to note that the two cpβ linkers appear to occupy different three dimensional states. Optimization of the interactions between these two linkers could improve T state stability of the permuted Hb. The electron density map for the cpβ-linkers is not complete in the current model. In addition, a higher resolution model will be needed to give better definition to side chain structures involved in interactions across the α_1-$cp\beta_2$ interface, which are critical to T state stabilization. Cyanomet and carbonmonoxy

isoforms of α-cpβ are more stable than the oxy-isoform crystallized in this study. Crystallization of these two isoforms might enable isolation of crystals that diffract to higher resolution than the current data set.

References:

1. Whitaker, B.I., and Sullivan, M., *Historical Perspectives in the 2005 Nationwide Blood Collection and Utilization Survey Report*, 2005, American Association of Blood Banks, p. 51-54.

2. US Department of Health and Human Services. (2011). *The 2009 national blood collection and utilization survey report*. Washington, DC: US Department of Health and Human Services, Office of the Assistant Secretary for Health.

3. Kjellstrom, B.T. Blood substitutes: where do we stand today? *Journal of Internal Medicine* 2003; **253**: 495–497.

4. Stowell, C.P., J. Levin, B.D. Spiess, and R.M. Winslow. Progress in the development of RBC substitutes. *Transfusion* 2001; 41: 287-299.

5. Hemoglobin Molecule of the Month May 2003, URL http://www.rcsb.org/pdb/101/motm.do?momID=41.

6. N. Mathews, C.K., van Holde, K.E., Appling, D.R., Anthony-Cahill, S.J., *Biochemistry*, 4th Ed. p 249.

7. Doyle, M.P.; Apostol, I.; Kerwin, B.A. Glutaraldehyde modification of recombinant hemoglobin alters its hemodynamic properties. *J. Biol. Chem.* 1999; 274: 2582–2591.

8. Marquardt, D. A., Doyle, M. P., Davidson, J. S., Epp, J. K., Aitken, J., Lemon, D. D., Anthony-Cahill. S. J. "Monodisperse 130 kDa and 260 kDa Recombinant Human Hemoglobin Polymers as Scaffolds for Protein Engineering of Hemoglobin-Based Oxygen Carriers." *J. Functional Biomater*. 2012; 3: 61-78.

9. Looker, D.; Abbott-Brown, D.; Cozart, P.; Durfee, S.; Hoffman, S.; Mathews, A.J.; Miller-Roehrich, J.; Shoemaker, S.; Trimble, S.; Fermi, G.; et al. A human recombinant haemoglobin designed for use as a blood substitute. *Nature* 1992, 356, 258–260.

10. Fishburn, A.L., J.R. Keeffe, A.V. Lissounov, D.H. Peyton, and S.J. Anthony-Cahill, A circularly permuted myoglobin possesses a folded structure and ligand binding similar to those of the wild-type protein but with a reduced thermodynamic stability, *Biochemistry* 2002; 41: 13318-13327.

11. Asmundson. A.; Taber, A.M.; van der Walde, A.; Lin, D.; Olson, J.S.; Anthony-Cahill, S. Co-expression of α- and circularly permuted β-globins yields a hemoglobin with normal R state but modified T state properties. *Biochemistry* 2009; 48: 5456–5465.

12. N.-A. Huynh, D. Peyton unpublished data

13. B. Kokona, R. Fairman, unpublished data

14. P.W. Buehler and F. D'Agnillo "Toxological consequences of extracellular hemoglobin: Biochemical and physiological perspectives" *Antioxidants & Redox Signaling* 2010; 12: 275-291.

15. Huynh, N.-A., Anthony-Cahill, S.J. unpublished data.

16. Olson, J., Anthony-Cahill, S.J. unpublished data.

17. DeTitta, G. T. & Luft, J. R.. Rate of water equilibration in vapor-diffusion crystallization: dependence on the residual pressure of air in the vapor space. *Acta Cryst.* 1995; D51: 786-791.

18. Mark S. Hargrove, Szymon Krzywda, Anthony J. Wilkinson, Yi Dou, Masao Ikeda-Saito, John S. Olson. "Stability of Myoglobin: A Model for the Folding of Heme Proteins" Biochemistry 1994; 33 (39):11767–11775

19. Villarreal, DM, CL Phillips, AM Kelley, S Villarreal, P Hernandez, JL Olson and DP Henderson. "Enhancing recombinant hemoglobin production in Escherichia coli BL21(DE3) containing the Plesiomonas shigelloides heme transport system" *Appl and Environ Microbiol.* 2008; 74: 2854-5856.

20. Graves, Philip E., Douglas P. Henderson, Molly J. Horstman, Brian J. Solomon, and John S. Olson. Enhancing stability and expression of recombinant human hemoglobin in *E. coli*: Progress in the development of a recombinant HBOC source. *Biochimica et Biophysica Acta (BBA)-Proteins & Proteomics* 2008; 1784: 1471-1479.

21. Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997-2012.

22. Echalier, A., R. L. Glazer, V. Fulop, and M. A. Geday. "Assessing crystallization droplets using birefringence." Acta Crystallographica Section D: Biological Crystallography 60, no. 4 (2004): 696-702.

23. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.* 2010; D66: 213-221.

24. Emsley, P. and Cowtan, K. *Acta Cryst* 2004; D60: 2126-2132

25. Briinger, Axel T. "Free R value: a novel statistical quantity for assessing the accuracy of crystal structures." *Nature* 355 (1992): 472-475.

26. Drabkin, D. L., and Austin, J. H.. Spectrophotometric studies II. Preparations from washed blood cells; nitric oxide hemoglobin, and sulfhemoglobin. *J. biol. Chem.* 1935; 112: 51-65.

Appendix A: Catalog of crystal morphologies

1	**Clear & no crystals**
2	**Precipitated**
3	**Starfish**
4	**Clusters**
5	**Small multidomain**
6	**Large multidomain**
7	**Small monocrystalline**
8	**Large spike monocrystalline**
9	**Plate monocrystalline**