## **Supporting Information**

Alvizo and Mayo 10.1073/pnas.0805858105

<br/>
<

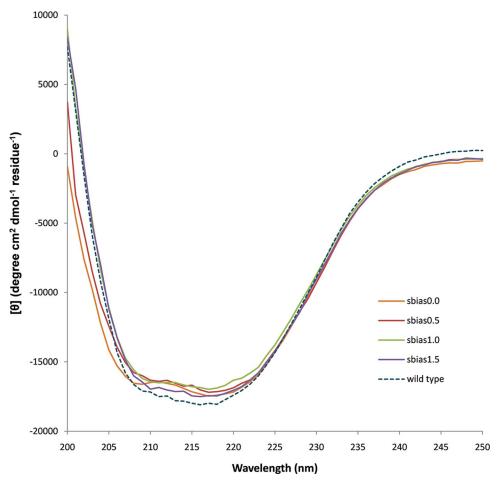


Fig. S1. Circular dichroism wavelength scans for G $\beta$ 1 and the various fixed composition proteins computed with the tuned force field.

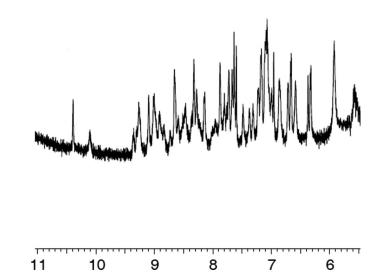


Fig. S2. The amide region of the 1D <sup>1</sup>H NMR spectrum of G $\beta$ 1's unbiased sequence design (sbias0.0) computed with the tuned force field.

A									lr	nit	tia	al	F	0	rc	e	F	i	el	d	-	E	Nł	H	Fi	X	ec	k (	C	or	n	20	S	it	io	n	D	e	si	g	ns	5							
					5					10					15					20				2	25				3	0				35	5			4	40				4	45				5	0
WILD	TYPE	T	A F	S	S	Е	Q	L	Α	R	L	κı	R	Е	F	Ν	Е	Ν	R	Y	L	Т	E	RI	RI	<b>२</b> (	2 (	ו ג	5	5 5	E	L	G	L	N	Е	А	Q	1	κ	1	W	F	Q	N	κ	R	AI	< 1
	0.0	S			Е	ĸ	Т	Е	R	F	1	. \	N	1	Q	Q	N	R	A	Е	1	S		ĸ	LI	L	ТΕ		A	١.	R					ĸ	Е	R	L	A	R	Q	Y		R	L		NO	2 F
ø	20	1	. 1		F	ĸ	т	F	R	1				F	0		Y	R		F	1				F /	4	. F	- 1	ι.		R					0		A	1		F	01	w			F		s	
Sequenci Bias	4.0		2.0				Δ	_																							W																		
quen Bias	6.0		• •					-														÷.	Ċ.,	•	Ľ /												-		÷.,	•		~		÷.		-	·		
еш	1000 0000	•	•••	•	•	•		1													•	•	•	•	•	•	• •				•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•••
S	8.0	•	•••	•	•	•	•	•	•	•	•	•	•	г	Ξ.	•	•	•	•	•	•	•	•	•	•	•	• •	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•••
	9.0	•	• •	•	•	٠	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		•		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
E	3							ŝ	Tı	ur	ne	d	F	0	ore	Ce	<b>e</b>	Fi	e	ld	-	Е	N	н	F	ix	e	d	С	0	m	p	09	sit	tic	on	C	)e	S	ig	n	S							
E	8				5					<b>ur</b> 10	ne	d	F	_	<b>)r</b> (	Ce	<b>e</b>	Fi		1 <b>d</b>	-	E	N		<b>F</b>	ix	e	d	<b>C</b>		m	p	09	sit		on	C		<b>S</b>	ig	n	S		45				5	0
E WILD		т,	A F	S	5 S	E	Q			10	1e		F		15	CE					- L		N E	2	25					0	m E E		<b>0</b> 9	35		on E	C			ig ĸ	<u>n</u>		F	45 Q	N	K	R	5 A 1	0
	TYPE	т /	A F	-	-	_	Q	L	A	10 R	L	κ	R	E	15 F	N	E	N	R	20 Y	L	т	E	R I	25 R I	२ <b>(</b>	ຊຸດ	וג	3	0	E	L	G	35 L	5 N	Е	C A E	Q	40 I	ĸ	L	W	F	Q	N	K	R	AI	< I
wild.	<b>TYPE</b>				Q	•		L •	A K	10 R	L •	к   •	R N	E Y	15 F	N •	E A	N •	R K	20 Y A	L •	T S	E •	R I E	25 R I		ג מ ג ו	Q L I F	3		EQQ	L ·	G •	35 L	N S	E I	A E	Q T	40 I	K R	I E	w	F •	Q E				A I Q I	K I R N
wild.	TYPE			:	QQ	:	•	L :	A K	10 R •	L •	K   •   	R N	E Y Y	15 F •	N	E A	N •	R K K	20 Y A N	L :	T S S	E •	R I E	25 R F • /		2 C K	2 L I F • F	3			L :	G •	35 L	N S S	E   	A E E	Q T T	40 I •	K R	I E E	w •	F •	Q E	R	•	•	A   Q   •	K I R N R N
wild.	0.0 0.5 1.0			:	QQ	:	:	L • •	A K •	10 R •	L • •	K   •   	R N •	E Y Y Y	15 F •	N •	E A •	N •	RKKK	20 Y A N E	L .	T S S S	E • •	2 R   E •	25 R I • /		2 C K	Q L I F F	3		E Q · ·	L •	G	3ť L •	N S S	E     .	A E E	Q T T T	40 I •	K R	I E E	w •	F • •	Q E ·	R •	:	:	A   Q   •	< 1 R N R N R •
equence Bias 	<b>TYPE</b> 0.0 0.5 1.0 1.5			:	QQ	:	:	L • •	A K •••	10 R •	L • •	K   •   •	R N •	E Y Y Y	15 F •	N •••••	E A	N • •	RKKKK	20 Y A N E	L	T S S S S	E • •	2 R   E •	25 R I • / • /			2 L I F I F	3		E Q	L • •	G • •	35 L •	N S S · ·	E     .	A E E • •	Q T T T T	40 I • •	K R •	 E •	w	F • •	Q E •	R •	:	:	A   Q   •	K I R N R N
wild.	0.0 0.5 1.0			:	QQ	:	:	L • •	A K •••	10 R •	L • •	K   •   •	R N •	E Y Y Y	15 F •	N • •	E A	N • •	RKKKK	20 Y A N E	L	T S S S S	E • •	2 R   E •	25 R I • / • /			2 L I F I F	3		E Q · ·	L • •	G • •	35 L •	N S S · ·	E     .	A E E • •	Q T T T T	40 I • •	K R •	 E •	w	F • •	Q E •	R •	:	:	A   Q   •	< 1 R N R N R •

Fig. S3. Computed sequences for ENH fixed composition designs. Designed core, boundary, and surface positions are shown in red, green, and blue, respectively. The wild-type sequence is shown followed by the sequences computed at increasing sequence bias values (kcal/mol per position). Dots represent wild-type amino acids. (A) Sequences obtained by using the initial parameters. (B) Sequences obtained by using the improved parameters.

PNAS PNAS