NewBG: A surrogate corticosteroid-binding globulin with an unprecedentedly high ligand release efficacy

Bernd R. Gardill[a, b, c], Karin Schmidt[a, c] and Yves A. Muller[a,*]

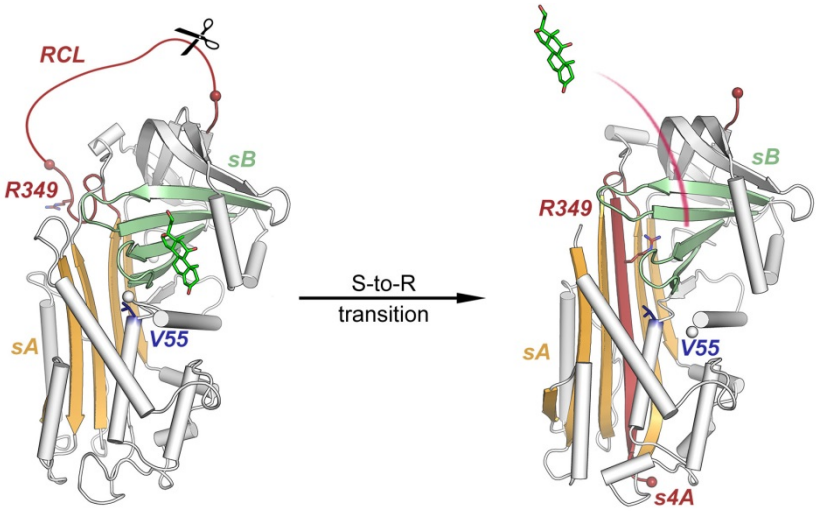
[a] Division of Biotechnology, Department of Biology, Friedrich-Alexander-University Erlangen-Nuremberg, Henkestr. 91, D-91052 Erlangen, Germany

[b] Present address: The University of British Columbia, Department of Biochemistry and Molecular Biology, 2350 Health Sciences Mall, Vancouver, BC Canada V6T 1Z3

[c] These authors contributed equally to this study

* Correspondence to Yves A. Muller: Division of Biotechnology, Department of Biology, Friedrich-Alexander University Erlangen-Nuremberg, Henkestr. 91, D-91052 Erlangen, Germany. yves.muller@fau.de. Tel: +49-9131-8523082. Fax: +49-9131-8523080

Graphical abstract



Highlights:

- Ligand-binding site design remains challenging
- So does the design of allosteric coupling pathways
- Incorporation of 16 mutations into human ACT created a surrogate CBG protein
- NewBG exhibits a 20-fold increased ligand release efficacy compared to CBG
- The simultaneous design of multiple protein properties is possible and desirable

ABSTRACT

The introduction of ligand-binding sites into proteins and the engineering of molecular allosteric coupling pathways are topical issues in protein design. Here, we show that these issues can be addressed concurrently, using the serpin human α1-antichymotrypsin (ACT) as a model. We have introduced up to 15 amino acid substitutions into ACT, converting it into a surrogate corticosteroid-binding globulin (CBG), thereby creating a new binding globulin (NewBG). Human CBG and ACT share 46 % sequence identity, and CBG served as the blue-print for our design, which was guided by side-chain-packing calculations, ITC measurements and crystal structure determinations. Upon transfer of ligand-interacting residues from CBG to ACT and mutation of specific second shell residues, a NewBG variant was obtained, which binds cortisol with 1.5 μM affinity. This novel serpin (NewBG-III) binds cortisol with a 33-fold lower affinity than CBG, but shares a similar ligand-binding profile and binding mode when probed with different steroid ligands and site-directed mutagenesis. An additional substitution, i.e. A349R, created NewBG-III-allo, which introduced an allosteric coupling between ligand-binding and the serpin-like S-to-R transition in ACT. In NewBG-III-allo, the proteinase-triggered S-to-R transition leads to a greater than 200-fold reduction in ligand affinity, and crystal structures suggest that this is mediated by the L55V and A349R substitutions. This reduction significantly exceeds the 10-fold reduction in binding affinity observed in human CBG.

Keywords

Serpin; computational protein design; allostery; ligand-binding site design; crystal structure

Abbreviations

ACT, α1-antichymotrypsin; CBG, corticosteroid-binding globulin; TBG, thyroxin-binding globulin; NewBG, new binding globulin; serpin, serine proteinase inhibitor; S-to-R, stressed-to-relaxed; RCL, reactive center loop; ITC, isothermal titration calorimetry; CD, circular dichroism; DBS, drug-binding serpin.

1. Introduction

Protein design and redesign is an innovative research field aimed at *de novo* engineering and reengineering the functional properties of proteins. Documented successes include the *de novo* design of protein folds, protein assemblies, protein-protein interactions, enzyme design and ligand-binding sites (Arai, 2018; Huang et al., 2016; Polizzi et al., 2017; Regan et al., 2015; Tinberg et al., 2013). In addition to designing proteins with novel properties, the incorporation and interlinking of multiple new properties within the same protein remains a challenge. This is of particular interest since the vast majority of natural proteins exert their functions in multiple ways. For example, key metabolic enzymes do not only function as enzymes. They often form protein assemblies that harbor allosteric mechanisms that influence their activities through the binding of small effector molecules to distant ligand-binding sites (Monod et al., 1963). Hence, the interplay of multiple properties, rather than individual properties, represents a key functional aspect of a protein.

In this report, we describe how *de novo* engineering has led to the introduction of a cortisol-binding site into human α1-antichymotrypsin (ACT, SERPINA3). Wild-type ACT is a member of the serine proteinase inhibitor (serpin) family that is devoid of any ligand-binding activity. A structural hallmark of the serpin family is the so-called stressed-to-relaxed (S-to-R) transition (Gettins, 2002; Whisstock et al., 2010). Cleavage of an exposed loop segment called reactive center loop (RCL) triggers the S-to-R transition. As a result, the entire RCL sequence is incorporated as an additional β-strand into the central β-sheet A of the serpin fold (Fig. S1). The present study also reports how further engineering of ACT enabled the coupling of the S-to-R transition to ligand binding, thereby allowing for an allosteric modulation of the ligand-binding affinity.

The proteinase-triggered S-to-R transition enables serpins to fulfill different roles (Gettins, 2002; Whisstock et al., 2010). Serpins, like ACT, act as proteinase inhibitors. As a result of the S-to-R transition, the target proteinase becomes irreversibly inhibited and remains permanently attached to the RCL. Other serpins, like corticosteroid-binding globulin (CBG, SERPINA6) and thyroxin-binding globulin (TBG, SERPINA7), do not act as inhibitors but bind and transport small molecule ligands in the blood. In these proteins, the S-to-R transition is linked *via* an allosteric coupling mechanism to the respective ligand-binding site. As a result of the S-to-R transition the ligand-binding affinity decreases in CBG and TBG. Human CBG binds cortisol with 45 nM affinity and TBG binds thyroxin (more precisely a fluorescent T4 conjugate) with 2.2 nM affinity (Gardill et al., 2012; Qi et al.,

2011). The S-to-R transition leads to a 10 and 5.5-fold reduction in binding affinity in CBG and TBG, respectively (Gardill et al., 2012; Qi et al., 2011). Although these decreases in affinity appear modest, they are considered important for the physiological function of CBG and TBG *in vivo* (Henley et al., 2016; Lin et al., 2010; Robbins, 2000).

Human ACT and human CBG are both members of the same clade A of the serpin superfamily and share 46 % overall sequence identity (Fig. S2) (Gettins, 2002). We have designed a way to transform ACT into a CBG-like-binding globulin. We also sought to incorporate an allosteric coupling mechanism into these novel ACT variants to achieve a comparable proteinase-triggered modulation of the ligand-binding affinity as observed in CBG (Gardill et al., 2012). We show that non-ligand-binding ACT can be reengineered to bind cortisol with μM affinities, and have named the resulting variants new binding globulins (NewBGs). These new variants initially lacked modulation of the ligand-binding affinity upon proteinase-triggered S-to-R transition, but additional engineering allowed for the incorporation of an allosteric coupling mechanism that triggers a greater than 200-fold reduction in ligand-binding affinity, which is more than 20 times more efficient than the coupling mechanism observed in human CBG or TBG.

2. Material and Methods

2.1. Computational design

Computational design calculations were performed with the in-house program MUMBO (Fig. S3) (Stiebritz and Muller, 2006). The program is freely available from the authors upon request (https://www.biotechnik.nat.fau.de/research/downloads/). MUMBO uses side-chain-packing algorithms and energy calculations to identify the best combination of amino acids, their orientations (side-chain rotamers) and ligand positions (ligand poses) to bind a certain ligand. Diversity in side-chain positioning was achieved using a backbone-dependent rotamer library in combination with the ‘backrub’ algorithm. The latter allows for the introduction of small shifts in the main-chain trace (Davis et al., 2006; Shapovalov and Dunbrack, 2011). Diversity in ligand positioning was generated by applying random translational and rotational shifts to a starting ligand pose. Interaction energies were calculated using the CHARM19 force field (Brooks et al., 1983). The global minimum energy conformation (GMEC) was identified using first dead-end elimination followed by a Metropolis Monte Carlo search (Stiebritz and Muller, 2006). Each MUMBO calculation

provides a unique solution. A single ligand pose is selected, and one amino acid type is retained at each position of the protein. The calculations were performed very similarly to those previously reported for the identification of doxycycline and doxorubicin-binding ACT variants (Schmidt et al., 2018).

2.2. Mutagenesis

A pQE-T7 expression vector, encoding for amino acids 3 to 400 of mature human ACT (UNIPROT entry P01011, offset = -23 residues), was purchased from the company Qiagen (Hilden, Germany). In this vector, the protein sequence is preceded by an N-terminal hexahistidine-tag with the sequence MKHHHHHHMKQ, and the coding region has been optimized for expression in *E. coli* (codon usage, mRNA stability and mRNA secondary structure). All mutations were introduced sequentially into the template using a two-stage mutagenesis protocol similar to the QuikChange™ Site-directed mutagenesis protocol (Stratagene, Santa Clara) (Wang and Malcolm, 1999). After each step the successful mutagenesis was confirmed *via* DNA sequencing. The primers are listed in Table S1.

2.3. Protein production and purification

All protein variants were produced and purified as previously described with the exception of a few minor changes in the purification protocol (Schmidt et al., 2018). In short, a single colony of *E. coli* BL21 Star™ (DE3) (Invitrogen), carrying the expression plasmid, was used to inoculate 150 ml of LB medium supplemented with 15 µg/ml kanamycin. After incubation overnight at 33 °C, 10x 850 ml of LB medium in 2 L shaker flasks were inoculated to an OD_{600} of 0.03 with the starter culture. The cells were allowed to grow at 37 °C until their density reached an OD_{600} of 0.2. The temperature was subsequently reduced to 20 °C. Expression was induced at an OD_{600} of 0.3 upon addition of 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After an additional 24-h-long incubation period, cell pellets were harvested by centrifugation and stored at -20 °C until further use.

The cell pellet was resuspended with 6 ml of lysis buffer per gram of pellet (buffer composition: 50 mM NaH_2PO_4 pH 8.0, 500 mM NaCl, 20 mM imidazole). 0.25 mM AEBSF was added, and the cells lysed by sonication. The raw extract was centrifuged at 95,000x g and 4 °C for 1 h in order to remove insoluble cell debris. The supernatant was filtered and applied to a HisTrap FF column (GE Healthcare, Freiburg, Germany). The target protein was

eluted from the column with a ten-column-volumes-lng linear gradient, increasing the imidazole concentration in the buffer from 20 to 500 mM. Fractions, containing the target protein, were pooled, and the buffer exchanged to 50 mM Tris/HCl pH 7.9, 50 mM NaCl using a HiPrep 26/10 desalting column. The protein was subsequently loaded onto a 12 ml XK16 Q-Sepharose FF column (GE Healthcare) and eluted *via* a twenty-column-volumes-long gradient, increasing the NaCl concentration from 50 to 1000 mM. Fractions, containing mainly pure target protein, were pooled and loaded onto a HiLoad 26/60 Superdex 75 column (GE Healthcare) equilibrated with 20 mM Tris/HCl pH 7.9, 150 mM NaCl. The purification protocol yielded highly pure protein samples of all ACT variants, as judged by SDS-PAGE.

2.4. Reactive center loop cleavage

Bovine pancreas α-chymotrypsin (Sigma-Aldrich) was dissolved in deionized water and added at a 1:20 molar ratio (proteinase to protein variant) to the purified proteins. The cleavage reaction was incubated at room temperature at a protein concentration of 1 mg/ml in 20 mM Tris/HCl pH 7.9, 150 mM NaCl for 2-3 days. The samples were checked by SDS-PAGE for completion of the cleavage reaction, which is accompanied by a characteristic shift to lower molecular weight and the release of a 4.7 kDa peptide. Once the reaction was complete, 1.5 mM AEBSF was added to inhibit the proteinase. The cleaved samples were purified on a HiLoad 16/60 Superdex 75 column equilibrated with 20 mM Tris/HCl pH 7.9, 150 mM NaCl before being used in further experiments. The RCL cleavage of DBS-II-allo-L55V, which was designed to be cleaved by the highly specific human rhinovirus 3C proteinase (HRV3CP), was performed as described earlier (Schmidt et al., 2018).

2.5. Circular dichroism spectroscopy

A Jasco J-815 CD spectrometer with a PTC-423S peltier element (Jasco, Tokyo, Japan) was used for the recording of the circular dichroism (CD) spectra. The far-UV CD spectra were collected at 20 °C with a protein concentration of 2.5 µM in a 10 mM potassium phosphate pH 7.4 buffer and using a 0.1 cm quartz cuvette. The data pitch was set to 0.1 nm, the digital integration time (D.I.T.) to 1 s and the band width to 1.0 nm. The spectra were measured with a scanning speed of 20 nm/min and accumulated 8 times.

Thermal denaturation measurements were recorded at 222 nm using a 1 cm cuvette and a protein concentration of 0.7 µM in 10 mM potassium phosphate pH 7.4. Data points

were continuously recorded every 0.2 °C while heating the sample with a rate of 1 °C/min. Band width was set to 1.0 nm and D.I.T. to 8 sec.

2.6. Native PAGE

The temperature induced polymerization of uncleaved protein samples was analyzed on a native PAGE. Protein solutions with concentrations of 5 µM in a 10 mM potassium phosphate pH 7.4 buffer were heated to the desired temperatures with a heating rate of 0.1° C/s, and the temperature kept constant for 5 min before cooling to 4 °C. The samples were mixed with sample buffer, containing 4.5 % sucrose, 0.25 % bromophenol blue, and separated on a 7.5 % native polyacrylamide gel at constant 5 mA for 4-5 h at 8 °C.

2.7. Isothermal titration calorimetry

The protein samples were dialyzed against a 20 mM potassium phosphate pH 7.4, 50 mM NaCl buffer. The steroid ligands cortisol, corticosterone and aldosterone (Sigma-Aldrich) were dissolved in the same dialysis buffer *via* sonication. All solutions were centrifuged (16,000x g, 15 min, 4 °C) and degassed under vacuum (900 rpm, 20 min, 25 °C) prior to the titration experiment. The protein samples NewBG-0 and NewBG-I were titrated with cortisol in a VP-ITC microcalorimeter (Microcal, Piscataway) at 25 °C. The titration experiments consisted of 25 consecutive injections of a 10 µl ligand solution with a concentration of 240 µM cortisol into the sample cell filled with either 20 µM NewBG-0 or 15 µM NewBG-I. The injection rate was set to 0.5 µl/s with 240 s spacing between the injections and constant stirring at 307 rpm. The titration experiments with NewBG-II, NewBG-III and NewBG-III-allo were performed in a NanoITC standard cell (TA Instruments, New Castle, USA) at 25 °C and consisted of 34x 5 µl injections of ligand solution (240 µM cortisol, 270 µM corticosterone and 270 µM aldosterone) into 15-20 µM protein solution with a 240 s spacing between injections.

The affinity of doxorubicin to uncleaved and cleaved DBS-II-allo-L55V was determined using a titration setup, consisting of 24x 4 µl injections of 650 µM doxorubicin in a 20 mM HEPES pH 7.4 buffer into the cell filled with 25-30 µM protein solution in 20 mM HEPES pH 7.4. A volume of either 5 µl (VP-ITC) or 2 µl (NanoITC) was used in the first titration step in order to remove mixed reactants from the needle tip. All titration experiments were corrected with an individual ligand-into-buffer titration. Integrated heat data were

analyzed using nonlinear regression by using a one-site-binding model in the Origin-7^{TM} software (VP-ITC) or by using an independent binding model in the NanoAnalyze software version 3.6.0 (NanoITC). The stoichiometry was fixed (n = 1) for Wiseman c-values below 5, in order to not overparameterize the data during curve fitting (Turnbull and Daranas, 2003). The changes in free energy and entropy of binding were calculated using the following equations: $\Delta G° = -RT\,lnK = \Delta H° - T\Delta S°$ (Wiseman et al., 1989).

2.8. Crystallization

The surface lysine residues of uncleaved NewBG-I were methylated in order improve protein crystallization (Chongyun et al., 2007). The reaction was conducted in a 20 mM potassium phosphate, pH 7.4, 100 mM NaCl buffer with a protein concentration of 3.4 mg/ml. The reaction was quenched upon addition of 125 µl of 1 M Tris/HCl pH 7.5, thereby deviating from the original protocol (Chongyun et al., 2007). The modified protein was purified on a HiLoad 16/60 Superdex 75, equilibrated with 20 mM Tris/HCl pH 7.9 and 150 mM NaCl. Crystals were obtained by using 0.1 µl of methylated NewBG-I at a concentration of 13.1 mg/ml in a 20 mM Tris/HCl pH 7.9, 150 mM NaCl buffer mixed with 0.2 µl of a reservoir solution, consisting of 0.2 M NaCl, 0.1 M Tris/HCl, pH 8.5, 25 % w/v PEG 3350 supplemented with 10 % of a 0.5 M NaF solution, and additional 0.1 µl of a 'silver bullets' bio reagent mixture, consisting of 0.16 % w/v thymidine, 0.16 % w/v adenosine 3',5'-cyclic monophosphate sodium salt monohydrate, 0.16 % w/v sarcosine, 0.16 % w/v 4-aminobenzoic acid, 0.16 % w/v acarbose, 0.16 % w/v inosine and 0.02 M HEPES pH 6.8 buffer. Crystals of methylated NewBG-I grew within 4 weeks at 19 °C and were cryo-protected by soaking the crystals in reservoir solution, supplemented with 20 % ethylene glycol, prior to flash-cooling the crystals in liquid nitrogen.

Uncleaved NewBG-0 was crystallized at a concentration of 12.5 mg/ml in 20 mM Tris/HCl pH 7.9, 150 mM NaCl. This protein solution was mixed with a reservoir solution, consisting of 0.2 M magnesium chloride hexahydrate, 0.1 M BIS-Tris pH 6.5, 25 % w/v PEG 3350, at a protein to reservoir volume ratio of 2:1. Crystals grew in a sitting drop setup at 19 °C. The protein NewBG-0 crystallized in the R-conformation, and, presumably, the RCL was cleaved by an unknown proteinase during the incubation time span.

Co-crystallization of NewBG variants with different steroid ligands was accomplished with ligand-saturated protein and reservoir solutions. Steroids were added as powders to both the protein solution (20 mM Tris/HCl pH 7.9, 150 mM NaCl) and the reservoir solution (see

below). The ligand-reservoir mixtures were sonicated for 15 min. The saturated protein and reservoir solutions were centrifuged at either 4 or 19 °C after incubation overnight in order to remove any residual solid ligand particles prior to crystallization.

Cleaved NewBG-II was co-crystallized with cortisol, using a ligand-saturated protein solution and a protein concentration of 6.5 mg/ml. The protein solution was mixed at a protein-to-reservoir volume ratio of 1:1 with ligand-saturated reservoir solution, consisting of 0.2 M ammonium sulfate and 20 % w/v PEG 3350. Crystals grew within 4 days at 19 °C with a hanging drop setup.

Cleaved NewBG-III was co-crystallized with either cortisol, corticosterone, aldosterone, progesterone or testosterone, using ligand-saturated protein solutions with protein concentrations of 6.2 mg/ml. Crystals of cleaved NewBG-III in complex with cortisol, corticosterone, aldosterone and testosterone grew with the same reservoir solution (2 % v/v Tacsimate pH 5.0, 0.1 M sodium citrate tribasic dihydrate pH 5.6, 16 % w/v PEG 3350). Saturation of the reservoir solution with the respective steroid ligands was achieved as described above. The protein-to-reservoir volume ratios were 1:2, 1:1, 1:2 and 2:1 for the co-crystallization with cortisol, corticosterone, aldosterone and testosterone, respectively. Crystals grew in hanging drop setups at 19 °C.

Co-crystallization of cleaved NewBG-III with progesterone was achieved using a ligand-saturated 2 % Tacsimate pH 6.0, 0.1 M BIS-Tris pH 6.5, 20 % w/v PEG 3350 solution as a reservoir solution. The latter was mixed 1:1 with a ligand-saturated protein solution (6.2 mg/ml) and setup for crystallization at 19 °C using the hanging drop method.

Cleaved NewBG-III was crystallized without any ligand in a sitting drop setup at 19 °C by using a reservoir solution, consisting of 4 % Tacsimate pH 4.0 and 12 % w/v PEG 3350, a protein-to-reservoir ratio of 2:1 and a protein solution with a concentration of 6.2 mg/ml.

Crystals of cleaved NewBG-III-allo were obtained by using a reservoir solution, consisting of 0.3 M calcium chloride dehydrate and 20 % w/v PEG 3350, a protein solution with a concentration of 6 mg/ml and a protein-to-reservoir volume ratio of 2:1. Crystals grew within 5 days at 19 °C in a hanging drop setup.

RCL-cleaved DBS-II-allo-L55V was crystallized in the presence of doxorubicin. The protein was treated with HRV3CP for RCL cleavage, purified via size exclusion chromatography, concentrated to 10 mg/ml and supplemented with a 5-molar excess of doxorubicin dissolved in 20 mM HEPES pH 7.4. Crystals were obtained by using a reservoir solution, consisting of 2 % Tacsimate pH 4.0, 0.1 M sodium acetate trihydrate pH 4.6 and

16 % w/v PEG 3,350, and a protein-to-reservoir volume ratio of 1:1. If not otherwise mentioned, all crystals were cryo-protected by using a mixture of 30 % (v/v) ethylene glycol and 70 % reservoir solution.

2.9. Data collection and structure determination

X-ray diffraction datasets were collected at beamline BL14.1 at BESSY II, Berlin, and at beamline P13 at PETRA III, Hamburg (Mueller et al., 2012). Diffraction data were processed with programs XDS and XSCALE (Kabsch, 1993). The phase problem was solved with the molecular replacement method with program PHASER (McCoy et al., 2007). For the dataset of uncleaved NewBG-I, the structure of murine SERPINA3N (PDB id 1YXA, (Rose et al., 2017)) was used as a search model. In case of cleaved NewBG-0, the structure of cleaved α1-antichymotrypsin (pdb id 1AS4) was used as a search model. The datasets of all subsequent NewBG variants were phased with the structure of cleaved NewBG-0 as a search model. Refinement of cleaved NewBG-0 and uncleaved NewBG-I was carried out with program REFMAC5 (Vagin et al., 2004). All other NewBG variants were refined with program PHENIX (Adams et al., 2010). Between cycles of restrained refinement, the models were manually rebuilt and corrected with program COOT (Emsley et al., 2010). The crystallographic data collection and refinement statistics are summarized in Table S2.

3. Results

3.1. Ligand-binding site transfer from CBG to ACT

The introduction of a cortisol-binding site on top of β-sheet B of ACT started with detailed structural comparisons and extensive side-chain-packing calculations with the in-house computer program MUMBO (Fig. S3) (Gettins, 2002; Stiebritz and Muller, 2006). A comparison between the crystal structure of human ACT (PDB id 1AS4, (Lukacs et al., 1998)) and those of human and rat CBG (PDB ids 2VDY and 2V95, respectively (Klieber et al., 2007; Zhou et al., 2008)), highlighted structurally homologous residues that directly contact the steroid in CBG (Fig. S2, Fig. S4). A superposition of the structures revealed an extensive difference in the conformation of the loop that interconnects β-strands s4B to s5B in ACT and CBG (secondary structure nomenclature as in (Gettins, 2002; Klieber et al.,

2007)) (Fig. 1a). This led to the construction of a modified ACT structure model in which the s4B to s5B loop conformation from CBG was pruned onto ACT. This modified model was subsequently used in all computational design calculations. During these calculations multiple ligand poses were considered and all side-chains located within a 20 Å distance from the anticipated steroid-binding site were allowed to repack in program MUMBO (Stiebritz and Muller, 2006). In total, 88 residue positions were considered for amino acid substitutions and/or a repacking of the side-chain rotamers (Fig. S2). At this stage, the types of amino acids considered for a given residue position were strictly restricted to amino acids present at this position in either ACT or CBG. The exception was position 277, where in addition to serine and arginine, a glycine residue was also considered (Fig. S2). When considering only the 88 residues, the sequence identity between CBG and ACT is 47 %, while the overall sequence identity between all CBG and ACT residues is 46 %. Thus, the evolutionary distance between residues forming the ligand-binding site in CBG and the equivalent residues in ACT is as large as between all other segments in CBG and ATC (Fig. S2). During the initial design calculations, 9 ACT residues were mutated with program MUMBO, and this led to the creation of NewBG-0 (Table 1). Disappointingly, ligand-binding experiments revealed that NewBG-0 lacked cortisol-binding activity (Table 2, Table S3, Fig. S5).

Table 1. Protein variants.

Variant	Total number of amino acid substitutions	Number of newly introduced changes[a]	Newly introduced substitutions[a]	Substitutions in comparison to UNIPROT entry P01011[b]
ACT[c]	0	0	0	0
NewBG-0	9	9	L24R, E242Q, K244N, K274N, R277G, P382D, T383H, D384F, Q386W	L24R, E242Q, K244N, K274N, R277G, P382D, T383H, D384F, Q386W
NewBG-I	12	3	L269S, P270R, N387S	L24R, E242Q, K244N, L269S, P270R, K274N, R277G, P382D, T383H, D384F, Q386W, N387S
NewBG-II[d]	12	1	N274A	L24R, E242Q, K244N, L269S, P270R, K274A, R277G, P382D, T383H, D384F, Q386W, N387S
NewBG-III	15	3	L55V, A251V, L252F	L24R, L55V, E242Q, K244N, A251V, L252F, L269S, P270R, K274A, R277G, P382D, T383H, D384F, Q386W, N387S
NewBG-III-allo	16	1	A349R	L24R, L55V, E242Q, K244N, A251V, L252F, L269S, P270R, K274A, R277G, A349R, P382D, T383H, D384F, Q386W, N387S
DBS-II-allo[e]	25	n. a.[f]	n. a.	L24R, W194F,W215Y, E242Q, K244N, L269S, P270Q, K274S, W276F, R277F, D278E, A349R, V355L, K356E, I357V, T358L, L359F, L360Q, S361G, A362P, P382D, T383N, D384F, Q386W, N387S
DBS-II-allo-L55V	26	1	L55V	L24R, L55V, W194F,W215Y, E242Q, K244N, L269S, P270Q, K274S, W276F, R277F, D278E, A349R, V355L, K356E, I357V, T358L, L359F, L360Q, S361G, A362P, P382D, T383N, D384F, Q386W, N387S

[a] In comparison to the protein variant described in the immediately preceding line.
[b] Residue numbering refers to the mature human α1-antichymotrypsin (ACT) protein without the signal peptide (offset in comparison to P01011 = -23).
[c] Human α1-antichymotrypsin (ACT).

[d] NewBG-II is synonymous to variant DBS-0 in (Schmidt et al., 2018), where it served as a starting point for generating drug-binding serpins (DBS) displaying doxorubicin and doxycycline-binding affinities.
[e] This variant corresponds to the engineered doxorubicin drug-binding serpin described in (Schmidt et al., 2018).
[f] Not applicable.

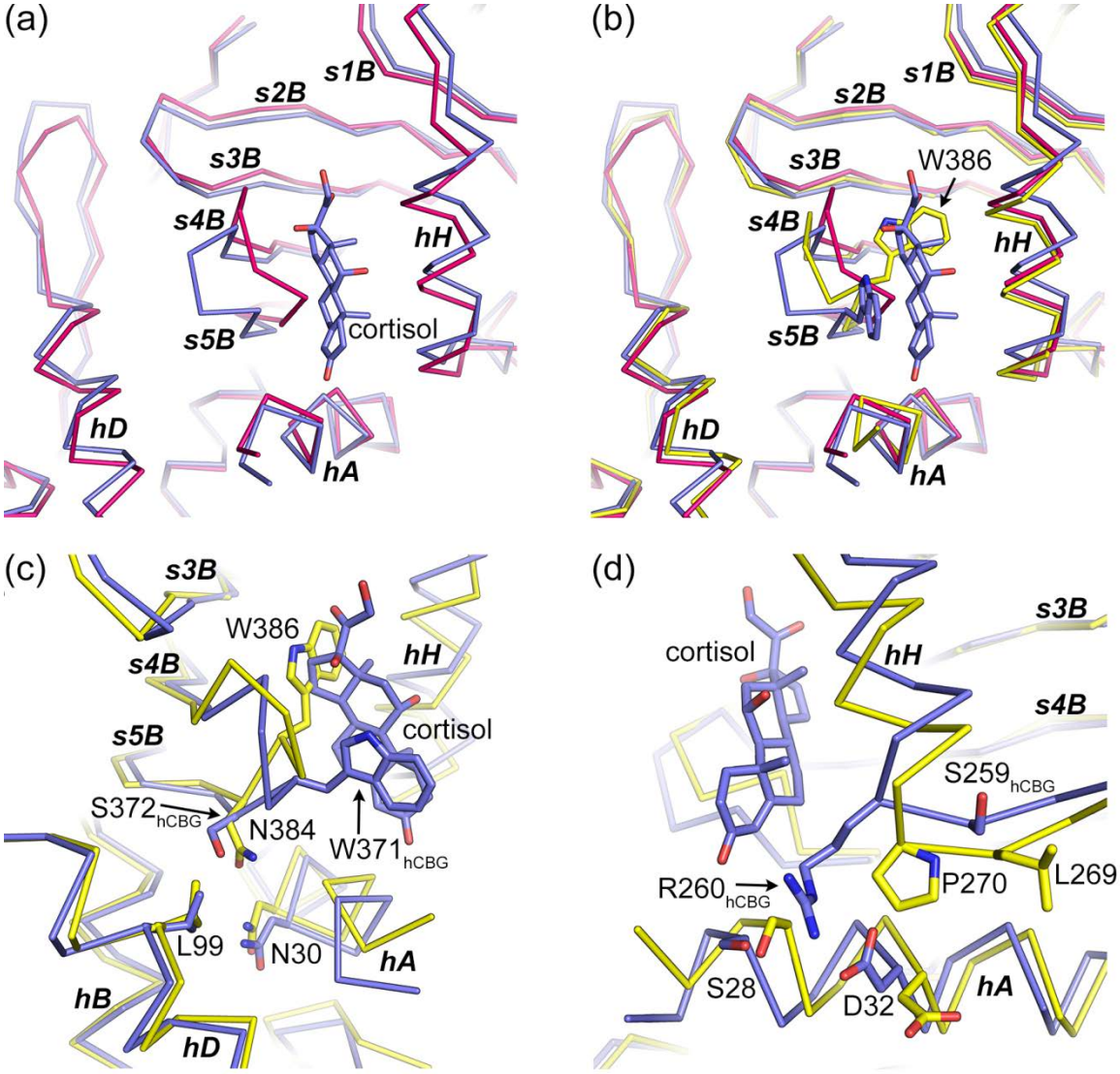


Fig. 1.

Lessons from a first futile attempt to transfer the steroid-binding site from human CBG onto human ACT. (a) Comparison of the main-chain conformation of human CBG (in prune) and human ACT (in magenta). (b) Comparison between the main-chain conformations of CBG, ACT and the crystal structure of NewBG-0 (in yellow). In NewBG, the side-chain of Trp371 packs across the top of β-sheet B and occludes the binding site. (c) Asn387, as a target for subsequent design steps. (d) Residues Pro270 and Leu269 as target positions. The superpositions were calculated with program COOT using entire chains. Secondary structure elements (s2B, s3B, hA, hD, etc) are labelled in bold and italic in all panels.

Table 2. Cortisol-binding affinities and thermodynamic parameters determined via ITC.

Protein variant[a]	Conformation	K_d[b] (μM)	Δ*H* (kJ/mol)	-TΔ*S* (kJ/mol)	Δ*G* (kJ/mol)	Affinity change: S *versus* R
ACT	S (uncleaved)	no measurable binding heats				-
CBG	S (uncleaved)	0.045[c]	-	-	- 42.1	-
CBG	R (cleaved)	0.442[c]	-	-	- 36.3	9.8
NewBG-0	S (uncleaved)	no measurable binding heats				-
NewBG-I	S (uncleaved)	13.7	-23.5	-4.3	-27.8	-
NewBG-I-S387N	S (uncleaved)	no measurable binding heats				-
NewBG-I-S269L-R270P	S (uncleaved)	40.0	-33.1	-8.0	-25.1	-
NewBG-II	S (uncleaved)	7.7	-71.5	42.3	-29.2	-
NewBG-II	R (cleaved)	9.6	-55.7	27.0	-28.7	1.2
NewBG-III	S (uncleaved)	1.5	-66.7	33.5	-33.3	-
NewBG-III	R (cleaved)	3.8	-41.8	10.8	-31.0	2.5
NewBG-III-allo	S (uncleaved)	1.9	-59.4	26.7	-32.7	-
NewBG-III-allo	R (cleaved)	no measurable binding heats				> 200[d]
DBS-II-allo-L55V [e]	S (uncleaved)	1.9	-18.1	-14.6	-32.7	
DBS-II-allo-L55V	R (cleaved)	no measurable binding heats				> 300[d]

[a] Please see also Table S3 for individual measurements.
[b] Measured at 25 °C.
[c] Values for human CBG were taken from (Gardill et al., 2012).
[d] Binding heats are expected to be observable for Wiseman c-values as low as 0.05 (Turnbull and Daranas, 2003). Therefore, it can be anticipated that affinities up to Kd values of 400 μM (cortisol) and Kd values of 600 μM (doxorubicin) could be measured when considering the experimental setups used in the present study. From this it can be concluded that the affinity loss of variants that display no measurable binding heats anymore after completion of the S-to-R transition, is higher than 200-fold in case of cortisol and 300-fold in case of doxorubicin.
[e] DBS-II-allo-L55V is a further evolved variant of a previously described doxorubicin-binding ACT variant (Schmidt et al., 2018). In this case, ligand-binding affinity values report doxorubicin (and not cortisol) binding to uncleaved and cleaved DBS-II-allo-L55V.

To understand the shortcomings of the NewBG-0 design, its crystal structure was determined at a resolution of 1.9 Å (Table S2). NewBG-0 adopts the R-conformational state in the crystal, and the RCL has become cleaved prior to crystal formation albeit that the protein was set up for crystallization in the uncleaved native state. Such a fortuitous cleavage of the RCL during crystallization is not unprecedented for serpins and can be caused by a bacterial contamination of the buffers and a concomitant preferential crystallization of serpins in the R-state (Gardill et al., 2012).

The NewBG-0 crystal structure reveals that the loop, interconnecting β-strands s4B to s5B, adopts a conformation that lies in-between the conformations observed for the corresponding loop in ACT and CBG (Fig. 1b). Hence, the transfer of the s4B to s5B loop conformation from CBG to ACT had only been partially successful in NewBG-0. At the same time, residue Trp386, which was introduced to provide hydrophobic interactions with the ligand, adopts an unexpected conformation in which it makes contacts with β-strand s4B while occluding the ligand-binding space (Fig. 1b). Inspection of the crystal structure hints that ACT residue Asn387 could be responsible for the incomplete repositioning of the loop. Asn387 is replaced by serine in human CBG, and surrounding residues, namely Asn30 and Leu99, are conserved in ACT and CBG (Fig. 1c). A structural comparison hints that the presence of a less space-filling side-chain at position 387 might allow for a slight inward shift of the main-chain at this position and thereby favor a more CBG-like conformation in the s4B to s5B loop.

In addition, helix hH, which also directly participates in ligand binding, is slightly displaced in NewBG-0 (and ACT) in comparison to CBG. Space-filling considerations similar to those above suggest that a substitution of ACT residues Leu269 and Pro270 by the corresponding human CBG residues, namely serine and arginine, might help adjusting the positioning of helix hH (Fig. 1d). Replacement of Pro270 with the corresponding CBG residue arginine might shift the entire helix and correctly anchor helix hH to helix hA *via* the formation of a salt bridge with residue Asp32 from helix hA (Fig. 1d). Such a salt bridge exits in CBG at this position, and Asp32 is conserved in ACT and CBG. The residue that precedes the arginine residue in CBG corresponds to a serine, and its side-chain participates in a classical helix N-cap interaction (Fig. 1d) (Presta and Rose, 1988). Thus, the corresponding ACT residue Leu269 was considered for mutation into a serine residue because of the expected stabilizing effect on helix hH.

These combined considerations led to the generation of a NewBG-0 triple-mutant with the substitutions L269S, P270R and N387S. The resulting variant, termed NewBG-I, displays a cortisol-binding affinity of 13.7 µM (Table 1, Table 2, Fig. S5). To better understand, which of these three substitutions was primarily responsible for the newly emerged binding affinity, two NewBG-I mutant variants were generated, namely NewBG-I-S387N and NewBG-I-S269L-R270P. In these variants, all three newly-introduced substitutions were mutated back to the corresponding wild-type ACT residues. Whereas NewBG-I-S387N does not display any measurable binding affinity, the mutant NewBG-I-S269L-R270P displays a binding affinity of 40 µM, which is about three times lower than the

affinity of NewBG-I (Table 2). These results show that both substitution pairs contribute to binding but that the introduction of Ser387 in NewBG-I can be regarded as primarily responsible for the emergence of binding in NewBG-I. Ser387 does not directly point towards the binding pocket and hence is not expected to directly interact with cortisol (see also below). Its contribution to the binding affinity of NewBG-I provides a nice example for how second-shell amino acids determine ligand-binding affinity.

3.2. Validation of the ligand-binding site by alanine mutagenesis

A substitution of a total of 12 residues in ACT by amino acids displayed from structurally equivalent positions in CBG gave rise to the cortisol-binding competent variant NewBG-I (Fig. S5, Table 1). The crystal structure of a NewBG-I-cortisol complex could not be determined. Hence, a number of alanine substitutions were produced in order to investigate to what extent ligand binding to NewBG-I resembles cortisol binding to CBG (Table 3). For rat CBG, it is known that mutating any of the rat residues to alanine that are equivalent to residues Arg24, Gln242, Asn274 and Trp386 in NewBG-I completely abolishes cortisol binding (Klieber et al., 2007).

Table 3. Cortisol-binding affinities of selected NewBG-I alanine substitutions.

Protein variant[a]	K_d (µM)
NewBG-I	13.7
NewBG-I-R24A	no measurable binding heats
NewBG-I-Q242A	22.4
NewBG-I-N274A (= NewBG-II)[b]	7.7
NewBG-I-W386A	no measurable binding heats

[a] All variants were investigated in the uncleaved S-state at 25 °C.
[b] This variant is identical to NewBG-II (see also Table 2).

Mutation of either Arg24 or Trp386 to alanine led to a complete loss of binding affinity in NewBG-I (Table 3) thereby mirroring the behavior of rat CBG. Substitution of Gln242 by alanine resulted in a slight decrease in affinity, only (K_d= 22.4 µM, 1.6-fold decrease, Table 3). Moreover, the Asn274Ala substitution unexpectedly led to an about 2-fold increase in affinity (K_d= 7.7 µM, Table 3). These data hint to what was subsequently confirmed by a crystal structure analysis (see below), namely that NewBG-I binds the ligand *via* an arginine-backed tryptophan-ligand stacking interaction similar to rat and human CBG

(see below). At the same time, however, other fine details of the interaction such as the formation of hydrogen bonds *via* Gln242 and Asn274 are not realized in NewBG-I. This is also reflected by the 300-fold lower cortisol-binding affinity of NewBG-I when compared to human CBG. Because of the remarkable binding affinity increase of NewBG-I mutant N274A, this variant was retained and termed NewBG-II in the following (Table 1). In parallel to the present study, NewBG-II also served as a starting point for creating ACT variants with altered ligand-binding specificity, namely variants that bind specifically doxycycline and doxorubicin (Schmidt et al., 2018).

3.3. Structural insight into ligand binding to NewBG-II

The affinity of NewBG-II for cortisol is about 2-fold higher than that of NewBG-I but still about 170-fold lower than in human CBG (Table 2) (Gardill et al., 2012). In order to visualize the cortisol-binding mode and to remedy possible design shortcomings, the crystal structure of NewBG-II in complex with cortisol was determined. Serpins crystallize more readily in the R-state and therefore NewBG-II was cleaved with chymotrypsin prior to crystallization (Gettins, 2002). The protease-triggered S-to-R transition induces a slight reduction in the cortisol-binding affinity of NewBG-II, only (9.5 versus 7.7 µM). Hence, it can be anticipated that highly similar cortisol-binding interactions occur in the S and R-state of NewBG-II (Table 2).

The 1.65 Å crystal structure of NewBG-II in complex with cortisol shows that cortisol binds to the same site and in the same orientation as in human CBG (Fig. 2a). As intended by the design effort, the steroid stacks against Trp386, and the latter aligns with Arg24. The s4B to s5B loop, displaying Trp386, adopts a conformation that is highly similar to the conformation seen in human CBG (Fig. 2b). The position of the ligand differs slightly between NewBG-II and human CBG, and this might be linked to a slight displacement of helix hH. Replacing Ser269 and Pro270 in NewBG-0 with the corresponding human CBG residues enables the formation of a highly similar interaction network between helices hA and hH in NewBG-II and human CBG. However, and in contrary to the intention, these substitutions did not induce a repositioning of helix H (Fig. 2a).

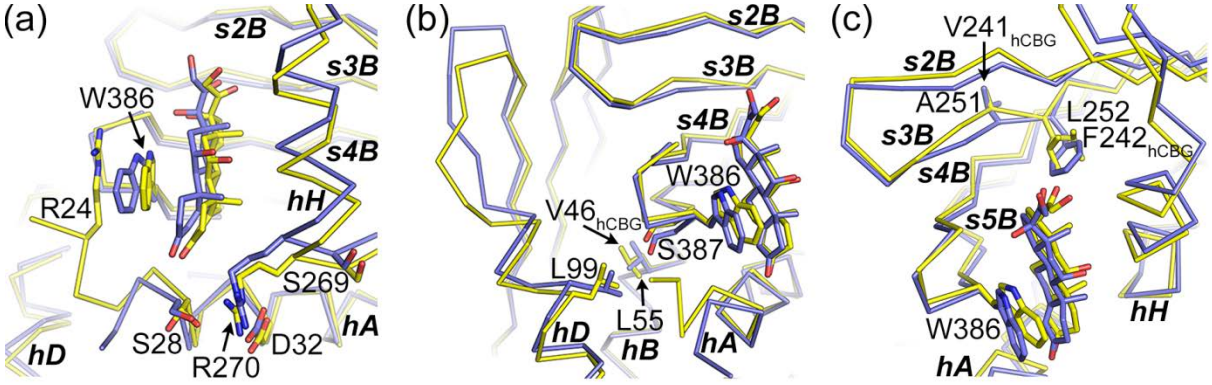


Fig. 2.

Binding of cortisol to NewBG-II mimics cortisol binding to human CBG. (a) Comparison of cortisol binding to NewBG-II (in yellow) and human CBG (in prune). (b) Packing differences between positions 55, 99 and 387 in comparison to the corresponding positions in human CBG. (c) Conformational differences in β-strand s3B of NewBG-II and human CBG.

The structure of NewBG-II was inspected in order to identify further mutations that could lead to an increased binding affinity. Thus, the surrounding of position 387 was explored anew since a substitution of Asn387 by a serine has proved essential for the emergence of the initial binding affinity (see above, Fig. 2b). Ser387 packs against Leu99 from helix hD and Leu55 from the beginning of helix hB. Whereas Leu99 is conserved in human CBG, Leu55 is replaced by a valine in human CBG (Fig. S2). The structure of NewBG-II hinted that a L55V mutation might provide for a better anchoring of β-sheet B to the core of the protein and more precisely lead to tighter side-chain-packing interactions between strand S5B and helices hB and hD. Possibly, this could stabilize the conformation of the ligand-free binding site and thus reduce the loss of conformational entropy anticipated for ligand binding.

Residue Leu252 displayed from strand s3B was also investigated for improving ligand binding (Fig. 2c). In NewBG-II, Leu252 displays two alternative side-chain conformations, whereas in rat and human CBG a phenylalanine is displayed at this position and forms tight contacts with the ligand (Gardill et al., 2012; Klieber et al., 2007; Zhou et al., 2008). Replacing the phenylalanine residue by alanine completely abolishes cortisol binding in rat CBG (Klieber et al., 2007). Side-chain-packing calculations with program MUMBO hinted that a L252F mutation would not lead to more favorable protein-ligand interactions unless the main-chain atoms of residue 252 are repositioned, too. Visual inspection of the NewBG-II structure in combination with further side-chain-packing calculations suggested that a concurrent A251V mutation would lead to both a slight repositioning of the main-chain conformation of residues 251 and 252 and to favorable interactions between Phe252 and the

ligand. Following these considerations, a NewBG-II triple-mutant was generated with the substitutions L55V, A251V and L252F (NewBG-III, Table 1).

3.4. NewBG-III displays a 1.5 µM ligand-binding affinity

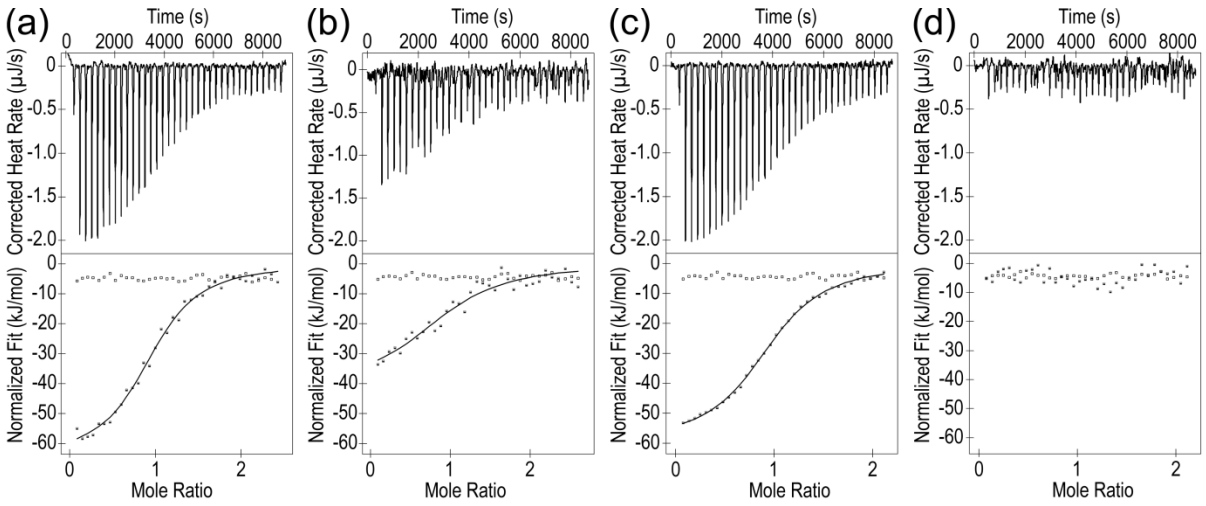


Fig. 3.

Binding of cortisol to NewBG variants investigated with ITC. Titration of cortisol to (a) uncleaved NewBG-III, (b) cleaved NewBG-III (c) uncleaved NewBG-III-allo and (d) cleaved NewBG-III-allo.

NewBG-III binds cortisol with a K_d of 1.5 µM and displays a 5-fold higher binding affinity than NewBG-II (Table 2). This affinity is only 16 and 35 times lower than the affinity of cortisol for rat and human CBG, respectively (Table 2) (Gardill et al., 2012). The affinity improvement also increased the Wiseman c-value from 2 to 10 in the ITC experiments (Turnbull and Daranas, 2003; Wiseman et al., 1989). The almost perfectly sigmoidal titration curves allow for the determination of the stoichiometry of the complex without any danger of data overfitting (Fig. 3a) (Turnbull and Daranas, 2003). Binding of cortisol to NewBG-III fits perfectly to a one ligand-binding site model. Upon cleavage of the RCL, cleaved NewBG-III still retains its affinity for cortisol, and, as was the case for NewBG-II before, only a very slight 2.5-fold reduction in binding affinity is observed (Table 2, Fig. 3b). At the same time, cleaved and uncleaved NewBG-III display circular dichroism spectra of well-folded proteins, and their secondary structure content and thermal stabilities are similar to those of cleaved and uncleaved ACT (Fig. S6) (Schmidt et al., 2018).

In serpins, the S-to-R transition typically goes in hand with an increase in the thermal stability and changes in the secondary structure content of the protein. Both of these changes

can be readily monitored *via* circular dichroism (CD) measurements (Dafforn et al., 2004). At the same time, native PAGE can be used to visualize the reduced tendency of the serpin R-state to aggregate upon heating (Dafforn et al., 2004). Combining CD measurements with PAGE experiments confirmed that NewBG-III adopts the R conformation upon RCL cleavage and thus retains the ability of ACT to undergo an S-to-R transition (Fig. S6).

NewBG-III displays the highest cortisol-binding affinity of all variants in the present study. The crystal structure of NewBG-III in complex with cortisol reveals that its ligand-binding pocket is very similar to those of human and rat CBG (Fig. 4, Fig. S7 and Table S2) (Klieber et al., 2007; Zhou et al., 2008). Four of the five oxygen atoms present in cortisol are involved in highly similar molecular interactions in all complexes (Table 4). Thus, a direct hydrogen bond, formed between atom O20 of cortisol and a glutamine residue displayed from β-strand s2B, is fully conserved in all proteins (Gln242, Gln232 and Gln224 in NewBG-III, human CBG and rat CBG, respectively) (Moss, 1989). Strictly conserved is also a water-mediated hydrogen-bonding interaction between atom O17 of cortisol and a tryptophan side-chain that stacks against the steroid ring system. A direct hydrogen bond between O17 and a histidine from strand s4B is realized only in human CBG and not in NewBG-III (His383 in NewBG-III, Fig. 4a,b). In rat CBG, the corresponding residue is substituted by a lysine residue and, as is the case for NewBG-III, no hydrogen bond is formed with atom O17 of cortisol (Fig. S7).

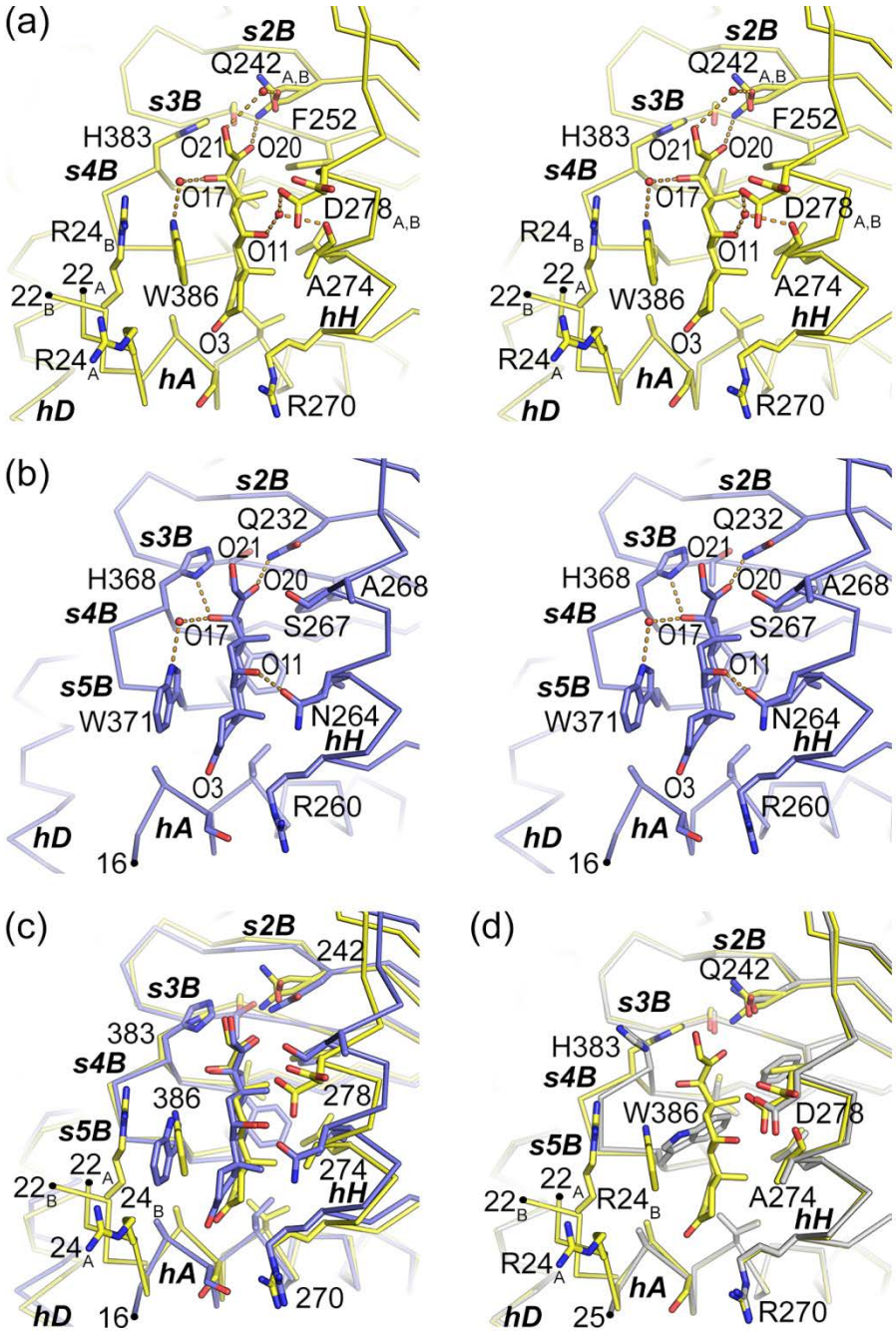


Fig. 4.

Comparison of the ligand interactions in NewBG-III and human CBG. (a) Stereo representation of NewBG-III in complex with cortisol. (b) Stereo representation of the human CBG-cortisol complex. (c) Superposition of the binding sites of NewBG-III (in yellow) and human CBG (in prune). Residue numbering according to NewBG-III. (d) Superposition of the cortisol-NewBG-III complex (in yellow) with ligand-free NewBG-III (in light grey).

.

Table 4. Cortisol hydrogen-bonding network.

Cortisol	**NewBG-III**			**Human CBG**			**Rat CBG**		
Ligand atom[a]	Protein atom	Distance [Å][b]	Remark	Protein atom	Distance [Å]	Remark	Protein atom	Distance [Å]	Remark
O3	-	-	Not involved in any polar interactions	-	-	Not involved in any polar interactions	-	-	Not involved in any polar interactions
O11	Asp278 -OD1	2.5/2.7	Indirect interaction *via* bridging water molecule	Asn264 -OD1	2.6	Direct interaction	Asp256 -OD1	2.6	Direct interaction
O11	Ala274 -O	2.5/2.8	Indirect interaction *via* bridging water molecule	-	-	-	Asp256 -O	2.8/2.8	Indirect interaction *via* bridging water molecule
O17	Trp386 -NE1	2.6/3.1	Indirect interaction *via* bridging water molecule	Trp371 -NE1	2.7/2.9	Indirect interaction *via* bridging water molecule	Trp362 -NE1	2.6/2.9	Indirect interaction *via* bridging water molecule
O17	-	-	-	His368- ND1	3.0	Direct interaction	-	-	-
O20	Gln242 -NE2	3.1	Direct interaction	Gln232 -NE2	2.8	Direct interaction	Gln224 -NE2	2.9	Direct interaction
O21	Gln242 -NE2	3.3/2.7	Indirect interaction *via* bridging water molecule	-	-	-	Gln224 -OE1	2.7/2.5	Indirect interaction *via* bridging water molecule

[a] Please see Fig. S8 for cortisol atom numbering.
[b] Distance pairs describe the distance between ligand atom and water molecule position plus distance between water molecule position and protein atom.

Additional differences between NewBG-III and human CBG occur in the polar interactions in which atoms O11 and O21 of cortisol participate. In NewBG-III and rat CBG, atom O21 makes an indirect hydrogen-bonding interaction *via* a bridging water molecule with the amide oxygen of the glutamine residue mentioned above (Fig. 4a, Fig. S7). In human CBG, this water-mediated interaction is missing. In case of atom O11, direct hydrogen bonds are formed with residues from helix hH in human and rat CBG, only, but not in NewBG-III. In the latter, only water-bridged interactions occur with helix hH (Fig. 4). From these considerations it appears that a slightly higher number of direct protein to ligand hydrogen bonds are formed in human CBG as compared to NewBG-III (and to some extent also to rat CBG, Table 4). Conversely, a slightly higher number of indirect interactions exist in NewBG-III.

A hallmark of the cortisol-binding mode is the tight stacking of the steroid against Trp386. This tryptophan residue appears to be hold in place *via* the above described hydrogen-bonding network but also *via* a cation-π-stacking interaction with an arginine side-chain displayed from helix A. This stacking interaction was first observed in rat CBG (Fig. S7) (Klieber et al., 2007). In the crystal structure of human CBG, the entire protein N-terminus (up to residue 16) is disordered and no arginine-tryptophan stacking interaction is observed (Fig. 4b). In the cortisol-NewBG-III complex, two alternative main-chain traces can be observed for the segment 22 to 27. Whereas in one conformation the side-chain of Arg24 participates in a stacking interaction with Trp386, the side-chain of Arg24 points away from the binding site in the second conformation (Fig. 4a). The conformational disorder in the segment 22 to 27 appears surprising since the R24A mutation shows that Arg24 is absolutely required for ligand binding (Table 3).

Conformational disorder or more precisely alternative side-chain orientations are also observed for residues Gln242 and Asp278 in NewBG-III, whereas the corresponding residues in human and rat CBG display single conformations (Fig. 4, Fig. S7). This could hint that, even in the presence of the ligand, the binding site in NewBG-III displays a heightened flexibility, the cause of which could be an imperfect binding site shape complementarity. This would also explain why the ligand-binding affinity of NewBG-III is lower than that of human and rat CBG (Table 2) (Gardill et al., 2012). At the same time, the binding sites of NewBG-III and human CBG differ in a number of additional details, making it difficult to explain exactly how structural and binding affinity differences interrelate (Fig. 4c).

Additionally, the crystal structure of ligand-free NewBG-III was determined and provides further insight into ligand binding (1.9 Å resolution, Table S2). Differences between

the ligand-bound and ligand-free state appear to be limited to an inward rotation of the side-chain of Trp386 and the absence of any electron density for the segment 22 to 24. In the ligand-bound structure, this segment contributes Arg24 that stacks against Trp386 (Fig. 4d). Other differences, such as a slight inward movement of the s4B to s5B loop, appear to be subtle. Overall, the binding site of ligand-free NewBG-III displays a conformation that appears to be largely prepared for ligand binding.

3.5. Steroid-binding profile of NewBG-III in comparison to human CBG

Table 5. Steroid-binding affinities (K_d in µM).

Ligand	NewBG-III	human CBG	
	this study[a]	Ref: (Gardill et al., 2012)[a]	Ref: (Mickelson et al., 1981)[b]
Cortisol[c]	1.5	0.045	0.00141
Corticosterone	0.6	n. d.[d]	0.00104
Progesterone	n. d.	n. d.	0.00169
Testosterone	n. d.	n. d.	0.02000
Aldosterone	41.6	n. d.	0.12500

[a] Measured at 25 °C using ITC, serpins in the S-state.
[b] Measured at 4°C using a competitive equilibrium dialysis assay.
[c] See also Table 2.
[d] Not determined.

A wealth of affinity data is available for the binding of steroids to human CBG. When plasma-isolated CBG is tested in a competitive equilibrium dialysis assay then the steroids corticosterone, cortisol, progesterone, testosterone and aldosterone bind with decreasing affinities ranging from 1.0 to 125 nM to human CBG (Table 5) (Mickelson et al., 1981). A substantially lower affinity is observed for cortisol binding to human CBG when investigated with samples produced in *E. coli* (45 nM *versus* 1.4 nM), and it has been shown that this discrepancy is caused by the absence of protein glycosylation in *E. coli*-produced samples (Table 5) (Chan et al., 2013; Simard et al., 2018). In case of NewBG-III, ITC-derived binding affinities could only be determined for cortisol, corticosterone and aldosterone and not for progesterone and testosterone because of ligand solubility issues (Table 5). Interestingly, the ligand affinity profile of NewBG-III is highly similar to that of human CBG despite the fact that all individual affinities are on the whole reduced in NewBG-III (Table 5). Thus, for

example, corticosterone binds with a 1.9-fold higher affinity to NewBG-III than cortisol (1.4-fold higher affinity in case of human CBG). In contrast, binding of aldosterone to NewBG-III displays a 27-fold reduced affinity (89-fold reduced affinity in human CBG).

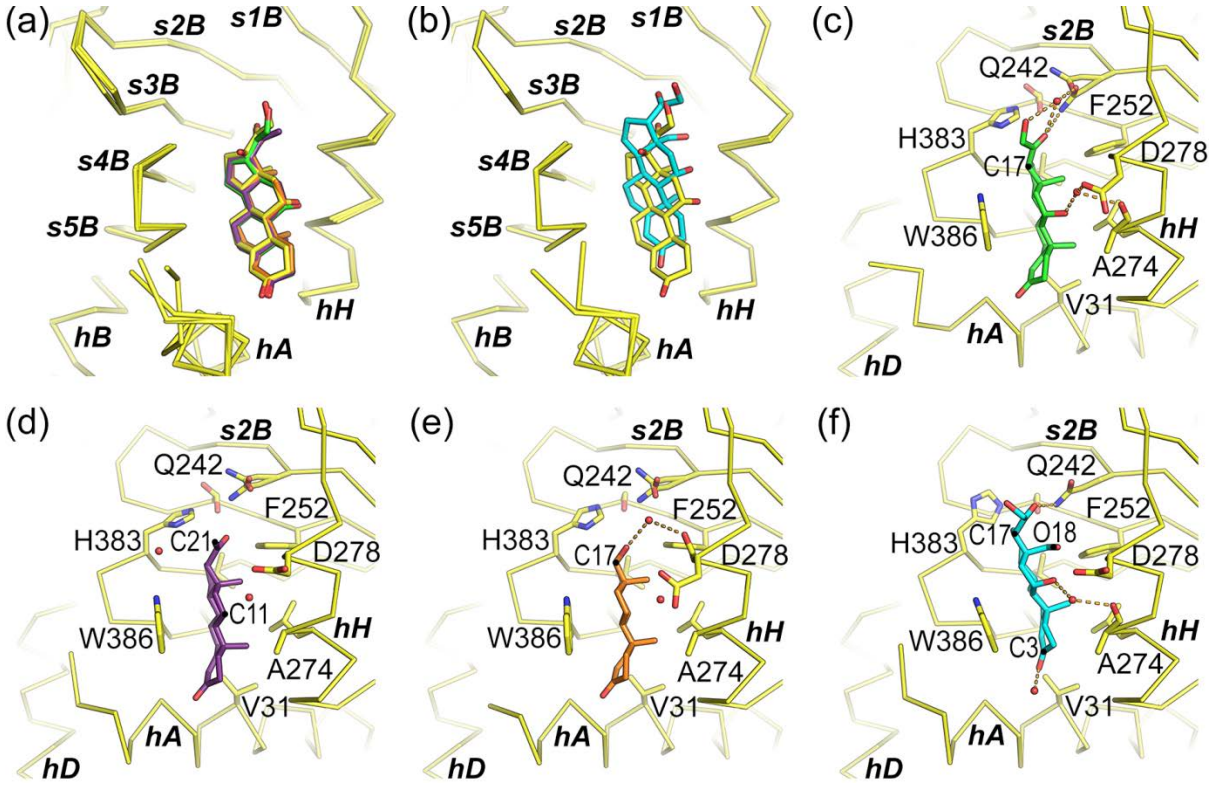


Fig. 5.

Ligand-binding pocket interactions of NewBG-III with different steroid ligands. (a) Superposition of the crystal structures of NewBG-III in complex with cortisol, corticosterone, progesterone and testosterone. (b) Superposition of the crystal structures of NewBG-III in complex with cortisol and aldosterone. (c) Details of the interaction of NewBG-III with corticosterone, (d) with progesterone, (e) with testosterone and (f) aldosterone. The same ligand color-coding is used in all panels. All polar interactions within 3.5 Å between any ligand atoms and protein atoms, whether water-bridged or not, are shown in panels (c) to (f). The naming of the ligand atoms adheres to the standard steroid atom numbering (Moss, 1989).

To further investigate the ligand-binding specificity of NewBG-III, crystal structures of NewBG-III in complex with cortisol, corticosterone, progesterone, testosterone and aldosterone were determined, (Table S2). All structures can be superimposed with RMSD values between 0.15 and 0.31 Å (Cα-atoms), showing that the main-chain folds of all NewBG-III complexes are highly similar (Fig. 5a). With the exception of aldosterone, the

steroid ligands also align well; aldosterone is shifted by about 2.5 Å in the binding site with respect to all other steroids (Fig. 5a,b).

Corticosterone displays the highest binding affinity of all NewBG-III ligands. Its chemical structure differs from that of cortisol by the absence of a hydroxyl group (O17-H) attached to atom C17 (Table 5, Fig. S8). With the exception of the interactions involving the missing hydroxyl group, all other interactions are highly similar in the cortisol and corticosterone complexes of NewBG-III (Fig. 4a, Fig. 5c). Since loss of the C17-attached hydroxyl group increases the binding affinity, it can be concluded that any putative affinity gains originating from interactions involving the C17-attached hydroxyl group of cortisol are outmatched by other effects such as a reduction of the hydrophobic effect or poorer shape complementarity when comparing cortisol and corticosterone binding to NewBG-III

With the stipulation that parallels in the ligand affinity profile of NewBG-III and human CBG also extend to the ligands progesterone and testosterone than it can be expected that progesterone binds with a similar affinity to NewBG-III than cortisol or corticosterone, whereas the affinity of NewBG-III for testosterone is expected to be 10 to 20-fold lower. Progesterone lacks two hydroxyl groups in comparison to corticosterone and three when compared to cortisol (Fig. S8). Consequently, the associated polar interactions are absent in NewBG-III in complex with progesterone with possibly again no detrimental effect on the binding affinity (Fig. 5d). In comparison to progesterone, testosterone lacks two carbon atoms, namely atoms C20 and C21, and contains, in addition, a hydroxyl group attached to carbon atom C17 (Fig. S8). The crystal structure of the testosterone-NewBG-III complex shows that the void that is introduced upon binding of this smaller steroid is filled by a water molecule (Fig. 5e).

The crystal structure of the aldosterone-NewBG-III complex shows that aldosterone is not accommodated at the same positon as the other steroids (Fig. 5f). Aldosterone differs from corticosterone by an additional hydroxyl group attached to carbon atom C18 (Fig. S8). The structure suggests that this additional hydroxyl group might cause steric clashes at the original binding position and therefore cause the ligand to become shifted by about 2.5 Å in direction of Gln242. At the same time a void is introduced near the carbon oxygen attached to position C3, and this void is filled by a water molecule (Fig. 5f). This alternative ligand positioning goes in hand with an experimentally determined 27 and 51-fold reduction in binding affinity when compared to cortisol and corticosterone, respectively (Table 5). Overall, these observations suggest that shape complementarity plays a major role in ligand recognition in NewBG-III.

3.6. NewBG-III-allo displays an unprecedentedly high ligand release efficacy

NewBG-III binds cortisol with similar affinities irrespective of whether the RCL segment of NewBG-III is cleaved or not (Table 2). This is in contrast to human CBG, where cleavage of the RCL and concomitant S-to-R transition leads to an about 10-fold reduction in ligand-binding affinity (Table 2) (Gardill et al., 2012). In human CBG, the proteinase-triggered S-to-R transition provides for an active mechanism for the release of anti-inflammatory corticosteroids at sites of inflammation, i.e. at sites where neutrophil elastase, the proteinase that cleaves the RCL of human CBG, occurs at elevated concentrations (Hammond et al., 1990; Pemberton et al., 1988). Even though it was possible to transfer the cortisol-binding site from CBG onto ACT, no coupling between ligand binding and S-to-R transition is observed in NewBG-III. Thus, the allosteric mechanism that couples the RCL insertion site to the ligand-binding site in CBG is missing in ACT albeit the fact that human CBG and ACT share a sequence identity that is as high as 46 %.

In our quest to create a surrogate CBG protein and to implement a CBG-like ligand-binding affinity modulation into NewBG, we first determined the crystal structure of a NewBG variant in the pre-cleavage S-state (Table S2). This was only successful for ligand-free NewBG-I, and the 2.8 Å resolution crystal structure of uncleaved NewBG-I shows that the RCL is partially inserted into β-sheet A in the S-state (Fig. S1a). This conformation is similar to the S-state conformation of human TBG and the murine orthologue of ACT (Horvath et al., 2005; Zhou et al., 2006). In this conformation, the NewBG-I residues, corresponding to positions P14 and P15 of the RCL (residues 346 and 347), participate in hydrogen-bonding interactions with strands s3A and s5A from β-sheet A, whereas preceding residues P1 to P13 do not (P-nomenclature according to Schechter & Berger, (Schechter and Berger, 1967)). By contrast, the RCL is fully expelled in the S-state conformation in many other serpins and among these also rat CBG (Klieber et al., 2007). In the crystal structure of uncleaved NewBG-I, the binding site is not entirely empty. It is occupied by the side-chain of a lysine residue from a neighboring molecule in the crystal. Because such crystal-packing interactions may influence the topology of the binding site, no conclusions can be drawn from this structure in view of the organization of the ligand-free binding site in the S-conformational state.

In a parallel study we previously showed that mutation of ACT residue P12 against arginine led to a pronounced modulation of the ligand-binding affinity in ACT variants

engineered for doxycycline and doxorubicin binding (Schmidt et al., 2018). The S-to-R transition causes the RCL positions P3 to P15 to become inserted into β-sheet A and gives rise to new β-strand s4A. As a result of the transition, the arginine side-chain at position P12 points towards the interior of the serpin fold, pushes against β-sheet B and distorts the ligand-binding site (Fig. S1) (Schmidt et al., 2018). Please note that the effect of such a A349R substitution onto the structure of ACT has been studied by others before, albeit in an entirely different context (Lukacs et al., 1998). The P12 substitution A349R was now also introduced into NewBG-III, and the ligand-binding affinity of this new variant termed NewBG-III-allo investigated before and after cleavage of the RCL (Table 1).

NewBG-III-allo displays an unexpectedly dynamic change in binding affinity upon protease cleavage of the RCL, and the change significantly exceeds those seen in human CBG and other serpins (Gardill et al., 2012; Qi et al., 2011; Schmidt et al., 2018). Whereas the binding affinity of uncleaved NewBG-III-allo (1.9 µM) is very similar to that of uncleaved NewBG-III (1.5 µM), cleaved NewBG-III-allo no longer exhibits ITC-recordable ligand-binding affinity (Table 2, Fig. 3c,d). When considering the signal detection limits of the ITC experiment, it can be estimated that the proteinase-triggered S-to-R transition leads to a higher than 200-fold reduction in binding affinity in NewBG-III-allo. At the same time, the CD spectra of uncleaved and cleaved NewBG-III-allo show that both proteins remain properly folded (Fig. S6).

The affinity modulation in NewBG-III-allo exceeds those documented for doxycycline and doxorubicin-binding ACT variants harboring the same A349R substitution (Schmidt et al., 2018). Similar to human CBG, these variants display only an about ten-fold modulation of the ligand-binding affinity (Table 2) (Gardill et al., 2012). An analysis of amino acid differences between NewBG-III-allo and these variants highlights a specific residue that does not directly participate in ligand binding. NewBG-III-allo displays a valine residue at position 55, whereas in the previously studied doxycycline and doxorubicin-binding variants of ACT the wild-type residue leucine was retained. Position 55 seems of particular interest since this residue is in contact with residues from s5B from the ligand-binding site and helix hD (Fig. 2b). In previous publications, helix hD has been discussed as playing a major role in the mechanism that allosterically couples the S-to-R transition to the ligand-binding affinity in CBG and TBG (Klieber et al., 2007; Qi et al., 2011; Zhou et al., 2008). In order to test the importance of this position, we introduced the L55V substitution into a previously characterized doxorubicin-binding variant (Schmidt et al., 2018). The resulting variant, termed DBS-II-allo-L55V, now also displays a greater than 300-fold

reduction in doxorubicin-binding affinity upon the S-to-R transition (Table 2, Fig. S5) whereas only a 9.4-fold reduction of the binding affinity was observed in DBS-II-allo in the absence of the L55V mutation (Schmidt et al., 2018).

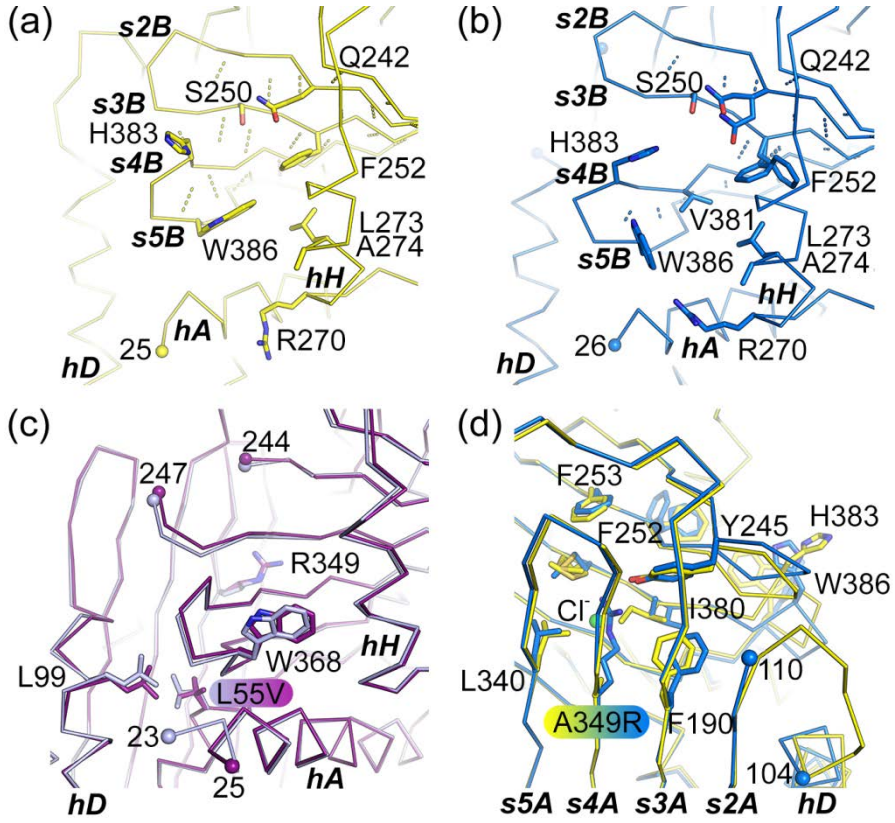


Fig. 6.

Allosteric regulation of ligand release in NewBG-III-allo and DBS-II-allo-L55V. (a) Display of the orientations of the ligand-binding site residues and of the main-chain hydrogen bond network in β-sheet B of ligand-free NewBG-III (yellow). (b) In NewBG-III-allo (blue) a breach is formed between strands s3B and s4B that is paralleled by the ablation of 3 main-chain hydrogen bonds between s3B and s4B and rearrangements in the ligand-binding residues. (c) In cleaved DBS-II-allo-L55V (purple) the ligand-binding site distortions or highly similar to those observed in cleaved DBS-II-allo (light blue, PDB ID 5OM8). (d) Structural rearrangements around position 349 in NewBG-III-allo. The A349R substitution and the concomitant binding of a chloride ion lead to distinct structural rearrangements such as a repositioning of Ile380 and Tyr245 and appearance of a breach.

The crystal structure of cleaved ligand-free NewBG-III-allo and its comparison with the structure of cleaved ligand-free NewBG-III provide atomic insight into the allosteric coupling mechanism (Table S2). The most obvious structural feature in cleaved NewBG-III-allo is the appearance of a breach in β-sheet B (Fig. 6a,b). This breach is formed between β-strands s3B and s4B and is characterized by the ablation of three main-chain hydrogen bonds

between residues 248 to 250 from strand s3B and residues 381 to 383 from s4B. The loss of binding affinity in cleaved NewBG-III-allo can be readily associated with the increase in the separation of strands s3B and s4B. Upon formation of the breach, the binding site becomes considerably distorted since residues, participating in steroid binding, are displayed from both edges of the breach (Fig. 6a,b).

Breach formation in NewBG-III-allo appears to be a direct consequence of the A349R substitution. Upon completion of the S-to-R transition spatial rearrangements are required to alleviate Van der Waals clashes introduced by the arginine substitution. These appear augmented by the fact that accommodation of the arginine side-chain goes hand in hand with the binding of a chloride anion that acts as a counter ion to the positively charged arginine side-chain (Fig. 6d) (Lukacs et al., 1998). In the R-conformational state, the newly introduced arginine side-chain makes direct contact with residues from β-sheet B. As a consequence, Ile380 from s4B is displaced by about 1.3 Å (Cα-position) and Tyr245 from s2B by about 0.8 Å in NewBG-III-allo (Fig. 6d). Thus, formation of the breach appears to be a direct consequence of the structural strain arising from the arginine substitution in NewBG-III-allo.

Interestingly, in the crystal structure of cleaved DBS-II-allo-L55V no such breach is formed (Table S2). Moreover, the overall topology of the ligand-binding site is highly similar to that previously observed in ligand-free DBS-II-allo despite the fact that the doxorubicin-binding affinity decreases by more than 300-fold in DBS-II-allo-L55V and by only a mere 9.4-fold in DBS-II-allo upon completion of the S-to-R transition (Fig. 6c). Nevertheless, we have previously shown that the ligand-binding site in ligand-free DBS-II-allo is also very distorted when compared to the ligand-binding site observed in doxorubicin-bound DBS-II-allo (Schmidt et al., 2018). Thus, it appears that in DBS-II-allo-L55V the binding affinity reduction stems primarily from an increased stabilization of the distorted binding site conformation, thereby significantly shifting the equilibrium from the ligand-binding competent to the binding incompetent conformation.

4. Discussion

The high sequence identity between CBG and ACT underpinned the rationale for our protein engineering experiments, designed to create a novel ligand-binding. However, a simple transfer of the ligand-binding residues from CBG onto the structurally homologous segments of ACT failed to produce a ligand-binding ACT variant. Only further substitutions informed by experimentally determined crystal structures and side-chain-packing calculations allowed creation of NewBG-III, a variant that binds cortisol with 1.5 µM affinity. Among the required substitutions were many so-called second-shell residues. The observation that residues, which do not directly participate in ligand-binding interactions, can be critical for generating ligand affinity has been highlighted before, but to the present day it remains difficult to explain and anticipate their effect on ligand-binding properties (Polizzi et al., 2017; Spindler et al., 2014; Tziridis et al., 2014). The present study emphasizes an important role of second-shell residues in improving the ligand shape complementarity of the binding site.

It could be argued that the present study is limited in scope because a mere incremental substitution of ACT residues against CBG residues would sooner or later have led to a cortisol-binding competent ACT variant without the need for any side-chain-packing calculations and validations *via* crystal structure determinations. However, such an exercise would not inform how second shell residues are important in this process, such as the role of Asn387 in improving the positioning of the s4B to s5B loop or the role of Val251 in positioning neighboring residue Phe252. Moreover, such an approach would not allow for the discovery of variants with properties that surpass those of the blueprint protein, such as the unprecedentedly high ligand release efficacy observed in NewBG-III-allo.

NewBG-III, which displays the highest cortisol-binding affinity (1.5 µM) of all the NewBG variants we created, still binds cortisol with a 33-fold reduced affinity in comparison to human CBG (45 nM). Nevertheless, the µM ligand-binding affinity compares well to affinities observed in computational binding site design studies, and higher affinities are most often only obtained when such a design effort is followed by experimental selection methods such as those provided by molecular evolution techniques (Tinberg et al., 2013). Recently, an exceptionally high 45 nM binding affinity was observed for the binding of an abiological porphyrin ligand to a computationally designed ligand-binding protein (Polizzi et al., 2017). Clearly, additional research is required to better understand the relationship between binding site geometry and thermodynamic binding properties (Martin and Clements, 2013).

Crystal structure analyses show that the ligand-binding mode in NewBG-III is comparable to that in human and rat CBG. Residues, directly participating in ligand binding, adopt similar functions in NewBG and CBG. Thus, in NewBG and CBG replacement of Trp386 or Arg24 by alanine abolished ligand binding. Differences, however, also exist. For example, residues, corresponding to Gln242 and Asn274 in NewBG-II, were identified as providing crucial contributions to ligand binding in rat CBG (Klieber et al., 2007). Yet, in NewBG-II, the Q242A substitution only leads to a 1.6-fold reduction in binding affinity, whereas the N247A substitution even leads to a 2-fold increase in binding affinity. This suggests that in the NewBG variants, hydrophobic stacking interactions and binding site shape complementary might dominate ligand binding and thereby surpass contributions from polar interactions. In contrast, in rat and human CBG different physicochemical interaction types contribute equally to ligand binding. In support of the assumption of weak polar ligand-protein interactions in NewBG variants is the observation that residues Gln242 and Asp278 display multiple side-chain orientations in NewBG-III, whereas the corresponding residues in human and rat CBG do not (Fig. 4 and Fig. S7).

The conclusions from above are further supported by the crystal structures of NewBG-III in complex with the additional steroids corticosterone, progesterone, testosterone and aldosterone. The steroid-NewBG-III complexes can be divided into three classes: (i) a group consisting of pregnane-derived steroids (cortisol, corticosterone and progesterone) with highly similar binding modes, (ii) the pregnane aldosterone with a shifted binding mode and (iii) the androstane testosterone. The ligands of the first class bind highly similarly and their binding affinities are anticipated to be highly similar. Yet they differ considerably in the number of possible polar interactions between the ligand and the protein. The binding position of aldosterone differs significantly from that of the other ligands, and this is paralleled by a reduced binding affinity. The general positioning of testosterone is again similar to that of the high-affinity-binding ligands, however, because of a reduced number of carbon atoms, testosterone does not fill the binding pocket adequately and a water molecule fills the unoccupied space. In our view, these observations suggest anew that hydrophobic interactions and binding site shape complementary rule ligand binding in NewBG-III. Conversely, further improvements in shape complementarity could lead to improved binding affinities. At the same time, a more precise engineering of hydrophilic interactions could provide a route to increase ligand-binding specificity with the possibility to even surpass that observed in human CBG (Table 5).

The high sequence identity between human ACT and human CBG raised the question whether the allosteric coupling mechanism that links completion of the S-to-R transition to a 10-fold reduction of the cortisol-binding affinity in CBG might have been preserved in ACT. Clearly this is not the case since we could not observe a significant difference in the binding affinity between native and cleaved NewBG-III. Only upon further engineering, the variant NewBG-III-allo was obtained. This observation fits the generally accepted view that allosteric regulatory mechanisms emerged late during protein evolution and are highly protein specific (Kuriyan and Eisenberg, 2007). As a result, highly homologous proteins can display considerably differences in their allosteric mechanisms (Royer et al., 2005).

In NewBG-III-allo, introduction of an arginine at position P12 of the RCL led to an extraordinary large modulation of the ligand-binding affinity (> 200-fold). In human CBG and TBG, this modulation is significantly weaker (10-fold at most). In the latter proteins, the allosteric coupling between the RCL insertion site and the ligand-binding site has been proposed to proceed *via* helix hD (Klieber et al., 2007; Qi et al., 2011; Zhou et al., 2008). In CBG and TBG, insertion of the RCL into the central β-sheet A leads to a partial unwinding of helix hD, and this change in secondary structure is perpetuated to the ligand-binding site. Helix hD is in direct contact with residues from the loops that interconnect β-strands s2B and s3B as well as s4B and s5B and that participate in ligand binding (Fig. 2b). Guided by this observation and with the aim of introducing a similar affinity modulation into the NewBG variants, several substitutions, i.e. N387S and L55V, were introduced to improve the packing of β-sheet B against helix hD (Fig. 2b). Unexpectedly, variant NewBG-III, harboring both substitutions, displayed similar affinities in the S and R-state. Only when incorporating results from an own previous study, a NewBG variant (NewBG-III-allo) was obtained that displays an unexpectedly high modulation of the ligand-binding affinity (Schmidt et al., 2018). In NewBG-III-allo, the alanine residue at RCL position P12 is substituted against arginine, and the allosteric modulation of its binding affinity exceeds all those seen before (Schmidt et al., 2018).

Our data strongly imply that the cooperative effect of two substitutions, namely, A349R and L55V, is responsible for this unprecedentedly strong allosteric coupling effect. Indeed, upon addition of the L55V mutation to a previously described doxorubicin-binding variant called DBS-II-allo (and resulting in DBS-II-allo-L55V) a similar marked change in binding affinity modulation is observed as upon addition of the A349R substitution to NewBG-III (resulting in NewBG-III-allo). At the same time, the starting variant DBS-II-allo, harboring the A349R mutation, only, and the starting variant NewBG-III, harboring only the

L55V mutation, displayed either a moderate or almost no modulation of the binding affinity upon proteolytic cleavage (Table 2) (Schmidt et al., 2018).

The observation of the formation of a breach in the ligand-binding site of the ligand-binding incompetent NewBG-III-allo structure provides an explanation for the marked reduction in binding affinity. However, this explanation appears to fall short since no such breach is formed in DBS-II-allo-L55V. However, in both proteins a significant distortion of the ligand-binding site becomes apparent when comparing the structures of these proteins to either our crystal structure of ligand-free NewBG-III or to the ligand-bound structure of DBS-II-allo (Schmidt et al., 2018). The extensive reduction of the binding affinity is synonymous to an extensive shift in the equilibrium between a ligand-binding competent and incompetent conformation in direction of the ligand-binding incompetent conformation. In both cases, Val55 appears to improve the packing of β-strands β4B and β5B against helix hD and apparently this interaction stabilizes the ligand-binding incompetent 'breached' conformation of NewBG-III-allo and the severely distorted ligand-binding site in DBS-II-allo-L55V. With the available high-resolution crystal structures of these impaired ligand-binding conformations, new opportunities might arise to enhance the affinity modulation through additional engineering and specific stabilization of these ligand-binding incompetent conformations.

In the present study, engineering of a ligand-binding site into a non-ligand-binding protein and the introduction of a highly effective allosteric coupling mechanism were addressed concurrently in a single protein. As a result, a CBG-like surrogate protein was obtained that differs at 16 amino acid positions from human wild-type ACT sequence and that displays an allosteric ligand affinity modulation that significantly exceeds that observed in CBG. Overall, human ACT appears to be an excellent model system to further explore and improve the development of artificial binding proteins (Schmidt et al., 2018). We propose that such ACT-derived proteins could find an application in the active delivery of pharmacological compounds to sites of interest, i.e. to sites where the S-to-R transition is triggered by a site-specific proteinase, thereby causing the release and delivery of the cargo compound.

5. Footnotes

Coordinates and structure factor amplitudes have been deposited with the protein data bank with accession numbers 6HGD (cleaved NewBG-0), 6HGE (uncleaved NewBG-I), 6HGF (cortisol-bound cleaved NewBG-II), 6HGG (cortisol-bound cleaved NewBG-III), 6HGH (ligand-free cleaved NewBG-III), 6HGI (corticosterone-bound cleaved NewBG-III), 6HGJ (aldosterone-bound cleaved NewBG-III), 6HGK (progesterone-bound cleaved NewBG-III), 6HGL (testosterone-bound cleaved NewBG-III), 6HGM (cleaved NewBG-III-allo) and 6HGN (cleaved DBS-II-allo-L55V).

Acknowledgements

We would like to thank Uwe Müller and Manfred Weiss for help with data collection at BESSY-II synchrotron in Berlin and Thomas Schneider from the EMBL beamlines for help with data collection at DESY synchrotron in Hamburg. We thank Norbert Sträter (University of Leipzig) Carolin Kisker and Hermann Schindelin (both at University of Würzburg) for access to ITC equipment and Franz X. Schmid and Stephanie Thoms (University of Bayreuth) for access to a CD spectrometer at the beginning of this study. We also would like to thank Geoffrey Hammond (University of British Columbia, Vancouver) for discussions and critical reading of the manuscript and Madhumati Sevvanna (University Erlangen-Nuremberg) for her advice in crystallographic data processing and refinement as well as Andrea Decker and Sigrid Weiler (both at University Erlangen-Nuremberg) for technical assistance.

Funding

This work was supported in part by Elitenetzwerk Bayern (BIGSS) and by Deutsche Forschungsgemeinschaft
via GRK1962 and MU 1477/5.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66, 213-221.

Arai, R., 2018. Hierarchical design of artificial proteins and complexes toward synthetic structural biology. Biophys Rev 10, 391-410.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M., 1983. CHARMM: a program for macromolecular energy minimization and dynamics calculations. J. Comput. Chem. 4, 187-217.

Chan, W.L., Carrell, R.W., Zhou, A., Read, R.J., 2013. How changes in affinity of corticosteroid-binding globulin modulate free cortisol concentration. J Clin Endocrinol Metab 98, 3315-3322.

Chongyun, C., Neil, S., Zhi-Jie, L., 2007. Procedure for reductive methylation of protein to improve crystallizability. doi:10.1038/nprot.2007.287.

Dafforn, T.R., Pike, R.N., Bottomley, S.P., 2004. Physical characterization of serpin conformations. Methods 32, 150-158.

Davis, I.W., Arendall, W.B., 3rd, Richardson, D.C., Richardson, J.S., 2006. The backrub motion: how protein backbone shrugs when a sidechain dances. Structure 14, 265-274.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66, 486-501.

Gardill, B.R., Vogl, M.R., Lin, H.Y., Hammond, G.L., Muller, Y.A., 2012. Corticosteroid-binding globulin: structure-function implications from species differences. PLoS One 7, e52759.

Gettins, P.G., 2002. Serpin structure, mechanism, and function. Chem Rev 102, 4751-4804.

Hammond, G.L., Smith, C.L., Paterson, N.A., Sibbald, W.J., 1990. A role for corticosteroid-binding globulin in delivery of cortisol to activated neutrophils. J Clin Endocrinol Metab 71, 34-39.

Henley, D., Lightman, S., Carrell, R., 2016. Cortisol and CBG - Getting cortisol to the right place at the right time. Pharmacol Ther 166, 128-135.

Horvath, A.J., Irving, J.A., Rossjohn, J., Law, R.H., Bottomley, S.P., Quinsey, N.S., Pike, R.N., Coughlin, P.B., Whisstock, J.C., 2005. The murine orthologue of human antichymotrypsin: a structural paradigm for clade A3 serpins. J Biol Chem 280, 43168-43178.

Huang, P.S., Boyken, S.E., Baker, D., 2016. The coming of age of de novo protein design. Nature 537, 320-327.

Kabsch, W., 1993. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J Appl Crystallogr 26, 795--800.

Klieber, M.A., Underhill, C., Hammond, G.L., Muller, Y.A., 2007. Corticosteroid-binding globulin, a structural basis for steroid transport and proteinase-triggered release. J Biol Chem 282, 29594-29603.

Kuriyan, J., Eisenberg, D., 2007. The origin of protein interactions and allostery in colocalization. Nature 450, 983-990.

Lin, H.Y., Muller, Y.A., Hammond, G.L., 2010. Molecular and structural basis of steroid hormone binding and release from corticosteroid-binding globulin. Mol Cell Endocrinol 316, 3-12.

Lukacs, C.M., Rubin, H., Christianson, D.W., 1998. Engineering an anion-binding cavity in antichymotrypsin modulates the "spring-loaded" serpin-protease interaction. Biochemistry 37, 3297-3304.
Martin, S.F., Clements, J.H., 2013. Correlating structure and energetics in protein-ligand interactions: paradigms and paradoxes. Annu Rev Biochem 82, 267-293.
McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. J Appl Crystallogr 40, 658-674.
Mickelson, K.E., Forsthoefel, J., Westphal, U., 1981. Steroid-protein interactions. Human corticosteroid binding globulin: some physicochemical properties and binding specificity. Biochemistry 20, 6211-6218.
Monod, J., Changeux, J.P., Jacob, F., 1963. Allosteric proteins and cellular control systems. J Mol Biol 6, 306-329.
Moss, G.P., 1989. The Nomenclature of Steroids - Recommendations 1989. Eur J Biochem 186, 429-458.
Mueller, U., Darowski, N., Fuchs, M.R., Forster, R., Hellmig, M., Paithankar, K.S., Puhringer, S., Steffien, M., Zocher, G., Weiss, M.S., 2012. Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. J Synchrotron Radiat 19, 442-449.
Pemberton, P.A., Stein, P.E., Pepys, M.B., Potter, J.M., Carrell, R.W., 1988. Hormone binding globulins undergo serpin conformational change in inflammation. Nature 336, 257-258.
Polizzi, N.F., Wu, Y., Lemmin, T., Maxwell, A.M., Zhang, S.Q., Rawson, J., Beratan, D.N., Therien, M.J., DeGrado, W.F., 2017. De novo design of a hyperstable non-natural protein-ligand complex with sub-A accuracy. Nat Chem 9, 1157-1164.
Presta, L.G., Rose, G.D., 1988. Helix signals in proteins. Science 240, 1632-1641.
Qi, X., Loiseau, F., Chan, W.L., Yan, Y., Wei, Z., Milroy, L.G., Myers, R.M., Ley, S.V., Read, R.J., Carrell, R.W., Zhou, A., 2011. Allosteric modulation of hormone release from thyroxine and corticosteroid-binding globulins. J Biol Chem 286, 16163-16173.
Regan, L., Caballero, D., Hinrichsen, M.R., Virrueta, A., Williams, D.M., O'Hern, C.S., 2015. Protein design: Past, present, and future. Biopolymers 104, 334-350.
Robbins, J., 2000. Editorial: new ideas in thyroxine-binding globulin biology. J Clin Endocrinol Metab 85, 3994-3995.
Rose, P.W., Prlic, A., Altunkaya, A., Bi, C., Bradley, A.R., Christie, C.H., Costanzo, L.D., Duarte, J.M., Dutta, S., Feng, Z., Green, R.K., Goodsell, D.S., Hudson, B., Kalro, T., Lowe, R., Peisach, E., Randle, C., Rose, A.S., Shao, C., Tao, Y.P., Valasatava, Y., Voigt, M., Westbrook, J.D., Woo, J., Yang, H., Young, J.Y., Zardecki, C., Berman, H.M., Burley, S.K., 2017. The RCSB protein data bank: integrative view of protein, gene and 3D structural information. Nucleic Acids Res 45, D271-D281.
Royer, W.E., Jr., Zhu, H., Gorr, T.A., Flores, J.F., Knapp, J.E., 2005. Allosteric hemoglobin assembly: diversity and similarity. J Biol Chem 280, 27477-27480.
Schechter, I., Berger, A., 1967. On the size of the active site in proteases. I. Papain. Biochem Biophys Res Commun 27, 157-162.
Schmidt, K., Gardill, B.R., Kern, A., Kirchweger, P., Borsch, M., Muller, Y.A., 2018. Design of an allosterically modulated doxycycline and doxorubicin drug-binding protein. Proc Natl Acad Sci U S A 115, 5744-5749.
Shapovalov, M.V., Dunbrack, R.L., Jr., 2011. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. Structure 19, 844-858.
Simard, M., Underhill, C., Hammond, G.L., 2018. Functional implications of corticosteroid-binding globulin N-glycosylation. J Mol Endocrinol 60, 71-84.

Spindler, N., Diestel, U., Stump, J.D., Wiegers, A.K., Winkler, T.H., Sticht, H., Mach, M., Muller, Y.A., 2014. Structural basis for the recognition of human cytomegalovirus glycoprotein B by a neutralizing human antibody. PLoS Pathog 10, e1004377.

Stiebritz, M.T., Muller, Y.A., 2006. MUMBO: a protein-design approach to crystallographic model building and refinement. Acta Crystallogr D Biol Crystallogr 62, 648-658.

Tinberg, C.E., Khare, S.D., Dou, J., Doyle, L., Nelson, J.W., Schena, A., Jankowski, W., Kalodimos, C.G., Johnsson, K., Stoddard, B.L., Baker, D., 2013. Computational design of ligand-binding proteins with high affinity and selectivity. Nature 501, 212-216.

Turnbull, W.B., Daranas, A.H., 2003. On the value of c: can low affinity systems be studied by isothermal titration calorimetry? J Am Chem Soc 125, 14859-14866.

Tziridis, A., Rauh, D., Neumann, P., Kolenko, P., Menzel, A., Brauer, U., Ursel, C., Steinmetzer, P., Sturzebecher, J., Schweinitz, A., Steinmetzer, T., Stubbs, M.T., 2014. Correlating structure and ligand affinity in drug discovery: a cautionary tale involving second shell residues. Biol Chem 395, 891-903.

Vagin, A.A., Steiner, R.A., Lebedev, A.A., Potterton, L., McNicholas, S., Long, F., Murshudov, G.N., 2004. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. Acta Crystallogr D Biol Crystallogr 60, 2184-2195.

Wang, W., Malcolm, B.A., 1999. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis. Biotechniques 26, 680-682.

Whisstock, J.C., Silverman, G.A., Bird, P.I., Bottomley, S.P., Kaiserman, D., Luke, C.J., Pak, S.C., Reichhart, J.M., Huntington, J.A., 2010. Serpins flex their muscle: II. Structural insights into target peptidase recognition, polymerization, and transport functions. J Biol Chem 285, 24307-24312.

Wiseman, T., Williston, S., Brandts, J.F., Lin, L.N., 1989. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Anal Biochem 179, 131-137.

Zhou, A., Wei, Z., Read, R.J., Carrell, R.W., 2006. Structural mechanism for the carriage and release of thyroxine in the blood. Proc Natl Acad Sci U S A 103, 13321-13326.

Zhou, A., Wei, Z., Stanley, P.L., Read, R.J., Stein, P.E., Carrell, R.W., 2008. The S-to-R transition of corticosteroid-binding globulin and the mechanism of hormone release. J Mol Biol 380, 244-251.

Table S1. Primer used for mutagenesis.

Mutation	Nucleotide sequence of forward primers[a]
L24R	5’-tttgcgcttgccagaccgcgatcaacatgtgtaccg-3’
R277G	5’-acgaaattccaggctatcgccccaacgtttcagggtttc-3’
Q386W	5’-tcactttgctcataaaaaaaatattccaggtatcggtcggaacaataatcatc-3’
P382D T383H D384F	5’-cgtccgtttctgatgattattgttgatcattttacctggaatatttttttttatgagc-3’
E242Q K244N	5’-actgagctgtaccgttgttcagctgaattataccggtaatgcaagcg-3’
K274N	5’-gccggaaaccctgaatcgttggggcg-3’
N387S	5’-gattattgttgatcattttacctggagtatttttttttatgagcaaagtgacca-3’
L269S P270R	5’-gaagaggttgaagcaatgctgagccgcgaaaccctgaatcgttggggc-3’
N274A	5’-gagccgcgaaaccttggcacgttggggcgatagcc-3’
L55V	5’-atgtgatttttagtccggtgagcattagcaccgca-3’
A251V L252F	5’-agctgaattataccggtaatgcaagcgtgtttttttattctgccggatcaggataaaatg-3’
A349R	5’-atgtttttgaagaaggcaccgaacgtagcgcagcaaccgc-3’
W194F	5’-gaattatatttttttttaaagccaaatttgaaatgccgtttgatccgcaggatac-3’
W215Y	5’-tcagagccgtttttatctgagcaaaaaaaaatatgtgatggttccgatgat-3’
W276F	5’-gagccgcgaaaccttgaatcgttttggcgatagcctg-3’
R270Q N274A S277F	5’-gaagaggttgaagcaatgctgagccaggaaaccttggcgcgttttttttgatagcctggaatttcgtgaaatt-3’
A274S	5’-gctgagccaggaaaccttgagccgtttttttgatagcctgg-3’
D278E	5’-aggaaaccttgagccgtttttttgaaagcctggaatttc-3’
H383N	5’-gttttaatcgtccgtttctgatgattattgttgataactttacctggagtatttttttttatga-3’
V355L K356E I357V T358L	5’-ccgaagcaagcgcagcaaccgcactggaagttctgctgctgagcgcactggttgaaac-3’
L359F L360Q S361G A362P	5’-cagcaaccgcactggaagttctgttccaggggcccctggttgaaacccgtaccattgt-3’

[a] The sequences of the reverse primers correspond to the complementary strand of the depicted forward primers.

Table S2. Crystallographic data collection and refinement statistics.

Protein sample	NewBG-0	NewBG-I[a]	NewBG-II	NewBG-III	NewBG-III-allo	DBS-II-allo-L55V
Features	cleaved RCL	uncleaved RCL	cleaved RCL	cleaved RCL	cleaved RCL	cleaved RCL
Bound ligand	–	–	cortisol	cortisol	–	–
PDB code	6HGD	6HGE	6HGF	6HGG	6HGM	6HGN
Data collection						
Space group	$P2_1$	$P2_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P3_121$
Unit cell parameters (Å)	41.6, 93.1, 49.3	58.0, 124.7, 127.5	72.3, 75.6, 79.9	70.0, 77.1, 79.2	71.3, 77.3, 79.3	84.8, 84.8, 97.1
Unit cell angles (°)	90, 110.4, 90	90, 90.9, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 120
Resolution (Å)[b]	41.4 – 1.90 (2.01 – 1.90)	34.5 – 2.80 (2.97 – 2.80)	34.2 – 1.65 (1.75 – 1.65)	43.4 – 1.79 (1.89 – 1.79)	38.7 – 1.37 (1.45 – 1.37)	42.4 – 1.48 (1.57 – 1.48)
R_{meas} (%)	5.2 (17.7)	9.1 (58.1)	6.3 (53.6)	9.9 (124.8)	6.7 (70.2)	6.9 (116.8)
$CC_{1/2}$ (%)			99.9 (82.8)	99.9 (80.0)	99.8 (75.7)	99.9 (86.4)
I / σ(I)	20.6 (7.9)	15.3 (2.6)	18.4 (3.0)	15.6 (1.8)	12.7 (2.2)	18.5 (1.9)
Completeness (%)	99.2 (98.2)	98.1 (95.9)	99.0 (97.5)	99.1 (97.5)	98.9 (97.7)	99.8 (99.2)
No. reflections	102682 (15894)	169460 (26372)	217281 (34344)	544951 (85882)	407814 (61742)	888891 (134885)
Redundancy	3.7 (3.6)	3.9 (3.8)	4.1 (4.2)	13.3 (13.3)	4.4 (4.3)	13.1 (12.5)
Refinement						
Resolution (Å)	38.9 – 1.90	34.5 – 2.80	34.2 – 1.65	35.2 – 1.79	38.7 – 1.37	42.4 – 1.48
R_{work} / R_{free}	14.7 / 19.8	18.4 / 26.2	15.7 / 18.7	17.6 / 21.6	16.8 / 18.1	17.9 / 19.9
Ramachandran (%)						
favored / outlier	98.6 / 0.0	95.1 / 0.2	99.5 / 0.0	98.9 / 0.0	98.0 / 0.0	98.9 / 0.0
No. atoms						
Protein	2993	11519	3092	3064	3056	3031
Ligand/ion	0	0	34	46	14	16
Water	385	58	482	259	413	260
B-factors						
Protein	19.4	45.9	18.3	35.9	21.3	32.0
Ligand/ion	–	–	31.2	50.8	31.1	51.2
Water	30.3	34.1	33.1	47.5	34.3	39.2
RMSD bond lengths (Å)	0.007	0.009	0.006	0.013	0.012	0.009
RMSD bond angles (°)	0.80	1.04	0.79	1.20	1.22	1.07

[a] Methylated protein sample.

[b] Values for the highest resolution shell are listed in parentheses.

Table S2. (Continued).

Protein sample Features Bound ligand	NewBG-III cleaved RCL –	NewBG-III cleaved RCL corticosterone	NewBG-III cleaved RCL aldosterone	NewBG-III cleaved RCL progesterone	NewBG-III cleaved RCL testosterone
PDB code	6HGH	6HGI	6HGJ	6HGK	6HGL
Data collection					
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters (Å)	48.2, 79.3, 102.2	72.6, 75.5, 79.1	69.4, 77.7, 79.5	69.4, 77.7, 79.2	71.0, 76.9, 80.1
Unit cell angles (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)[c]	38.2 – 1.9 (2.02 – 1.90)	37.7 – 1.52 (1.61 – 1.52)	43.4 – 1.83 (1.94 – 1.83)	43.3 – 1.86 (1.97 – 1.86)	38.5 – 1.92 (2.04 – 1.92)
R_{meas} (%)	4.1 (74.8)	8.1 (100.6)	9.8 (124.9)	10.5 (124.5)	7.0 (69.4)
$CC_{1/2}$ (%)	100 (87.5)	99.9 (80.9)	99.9 (79.0)	99.9 (80.8)	99.9 (76.7)
$I / \sigma(I)$	17.4 (2.0)	18.2 (2.2)	15.5 (1.5)	17.3 (1.9)	15.0 (2.1)
Completeness (%)	99.4 (98.9)	98.2 (93.6)	98.1 (96.3)	99.6 (98.2)	98.8 (96.5)
No. reflections	114011 (17830)	1727260 (253296)	508657 (81180)	474970 (72101)	124135 (19135)
Redundancy	3.6 (3.6)	13.5 (12.8)	13.3 (13.6)	12.9 (12.5)	3.7 (3.6)
Refinement					
Resolution (Å)	38.2 – 1.90	35.0 – 1.52	35.4 – 1.83	38.9 – 1.86	38.5 – 1.92
R_{work} / R_{free}	20.8 / 24.0	16.9 / 19.3	18.4 / 22.1	17.9 / 22.7	18.8 / 22.4
Ramachandran (%)					
favored / outlier	98.1 / 0.0	99.2 / 0.0	98.6 / 0.0	98.6 / 0.0	99.5 / 0.0
No. atoms					
Protein	2910	3104	2976	3037	3049
Ligand/ion	7	25	30	31	21
Water	109	353	185	202	239
B-factors					
Protein	49.6	26.7	37.7	34.3	34.9
Ligand/ion	59.2	41.7	54.6	49.2	54.8
Water	52.4	38.9	45.7	42.5	42.7
RMSD bond lengths (Å)	0.006	0.010	0.008	0.008	0.006
RMSD bond angles (°)	0.81	1.03	0.83	0.79	0.69

[c] Values of the highest resolution shell are listed within parentheses.

Table S3. Summary of all ITC measurements performed in the current study.

Protein variant	Conformation	titrated ligand	K_d (μM)	ΔH (kJ/mol)	$-T\Delta S$ (kJ/mol)	ΔG (kJ/mol)
ACT[a,b]	S (uncleaved)	cortisol	no measurable binding affinity			
NewBG-0[a]	S (uncleaved)	cortisol	no measurable binding affinity			
NewBG-I[a]	S (uncleaved)	cortisol	13.7	-23.5	-4.3	-27.8
NewBG-I-S387N[a]	S (uncleaved)	cortisol	no measurable binding affinity			
NewBG-I-S269L-R270P[a]	S (uncleaved)	cortisol	40.0	-33.1	8.0	-25.1
NewBG-I-R24A[a]	S (uncleaved)	cortisol	no measurable binding affinity			
NewBG-I-Q242A[a]	S (uncleaved)	cortisol	22.4	-22.9	-3.6	-26.5
NewBG-I-W386A[a]	S (uncleaved)	cortisol	no measurable binding affinity			
NewBG-II[b,c]	S (uncleaved)	cortisol	7.6	-74.1	44.9	-29.2
NewBG-II[b,c]	S (uncleaved)	cortisol	7.7	-68.9	39.7	-29.2
NewBG-II[b,c]	R (cleaved)	cortisol	9.7	-53.2	24.6	-28.6
NewBG-II[b,c]	R (cleaved)	cortisol	9.4	-58.1	29.4	-28.7
NewBG-III[b]	S (uncleaved)	cortisol	1.4	-64.6	31.2	-33.4
NewBG-III[b]	S (uncleaved)	cortisol	1.6	-68.8	35.7	-33.1
NewBG-III[b]	R (cleaved)	cortisol	3.6	-40.6	9.5	-31.1
NewBG-III[b]	R (cleaved)	cortisol	3.9	-42.9	12.0	-30.9
NewBG-III[b]	S (uncleaved)	corticosterone	0.60	-47.7	12.2	-35.5
NewBG-III[b]	S (uncleaved)	corticosterone	0.62	-51.1	15.7	-35.4
NewBG-III[b]	S (uncleaved)	aldosterone	44.3	-43.5	18.6	-24.9
NewBG-III[b]	S (uncleaved)	aldosterone	38.9	-39.8	14.6	-25.2
NewBG-III-allo[b]	S (uncleaved)	cortisol	1.9	-59.4	26.7	-32.7
NewBG-III-allo[b]	R (cleaved)	cortisol	no measurable binding affinity			
DBS-III-allo-L55V[b]	S (uncleaved)	doxorubicin	1.5	-17.7	-15.5	-33.2
DBS-III-allo-L55V[b]	S (uncleaved)	doxorubicin	2.3	-18.4	-13.7	-32.1
DBS-III-allo-L55V[b]	R (cleaved)	doxorubicin	no measurable binding affinity [d]			

[a] Measurements conducted with a VP-ITC microcalorimeter (Microcal, Piscataway).
[b] Measurements conducted with a NanoITC standard volume calorimeter (TA Instruments, Eschborn, Germany).
[c] Please see also Schmidt et al., 2018 (Schmidt et al., 2018). NewBG-II is synonymous to DBS-0.
[d] Two independent measurements.

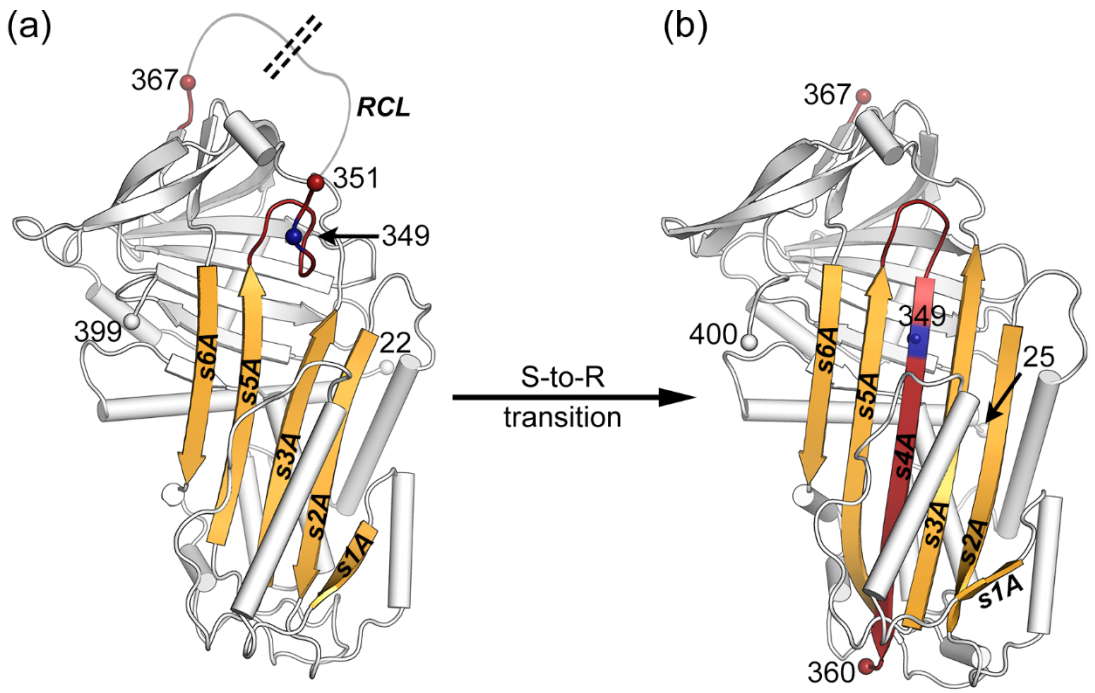


Fig. S1. Serpin typical S-to-R transition induced upon proteolytic cleavage of the reactive center loop (RCL) and illustrated with crystal structures from the present study. (a) Crystal structure of NewBG-I (see text) in the serpin-typical S conformational state. In this variant of human ACT the RCL sequence is partially inserted between β-strands s3A and s5A. (b) Crystal structure of ligand-free NewBG-III (see text) in the R conformational state with the RCL sequence fully inserted in the central β-sheet A and giving rise to new β-strand s4A. The amino acid residue located at RCL position 349 is highlighted in blue. After the S-to-R transition its side-chain points towards the interior of the protein. Introduction of an arginine at this position leads to an S-to-R transition-linked modulation of the ligand-binding affinity in NewBG-III-allo (see also main text).

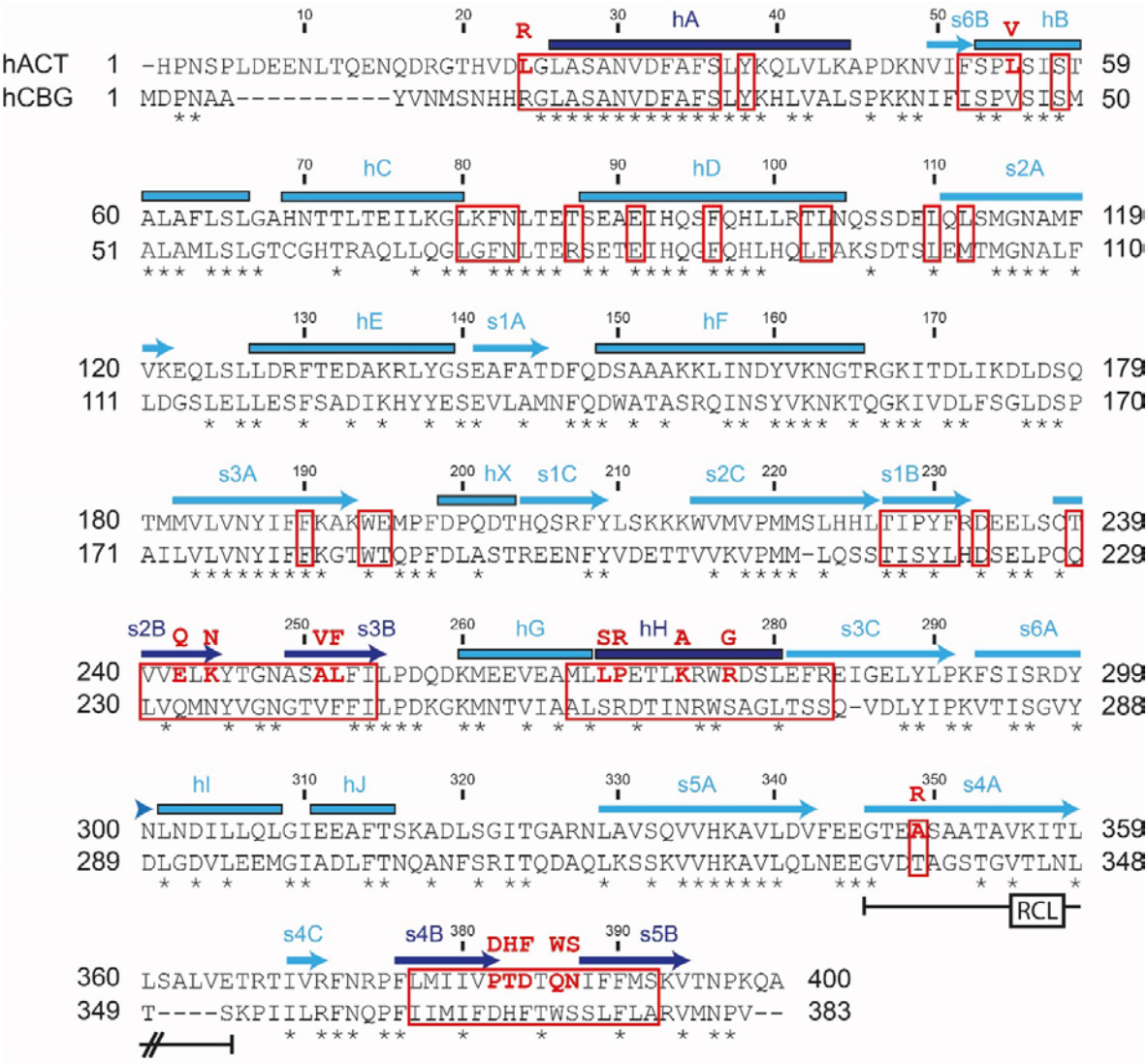


Fig. S2. Sequence alignment of human ACT and CBG. Human ACT (Uniprot entry P01011, (The UniProt Consortium, 2017)) and CBG (P08185) share 46 % sequence identity. Identical residues are marked with asterisks. Residues from ACT, which were included in the side-chain-packing calculations with program MUMBO, are boxed in red. Residues from ACT, which were replaced in any of the NewBG variants with residues from human CBG, are printed in red and in bold. Only three residues were substituted against residues not present in CBG, namely, K274A, R277G and A349R. Residues that are part of the reactive center loop (RCL, ACT residues 346-365) are underlined, and the RCL cleavage site (residue 360) is indicated. The secondary structure elements are depicted in blue and were taken from the human CBG structure (R-state, pdb entry 4BB2, (Gardill et al., 2012)). Helices are highlighted by blue boxes and β-strands with

blue arrows. The naming of the secondary structure elements adheres to the typical serpin naming with helices termed hA to hJ, strands s1A to s6A forming β-sheet A, s1B to s6B β-sheet B and s1C to s3C β-sheet C (Gettins, 2002; Klieber et al., 2007). Secondary structure elements that build up the immediate ligand-binding site are highlighted in dark blue. These include helices hA and hH as well as strands s2B, s3B, s4B and s5B from β-sheet B.

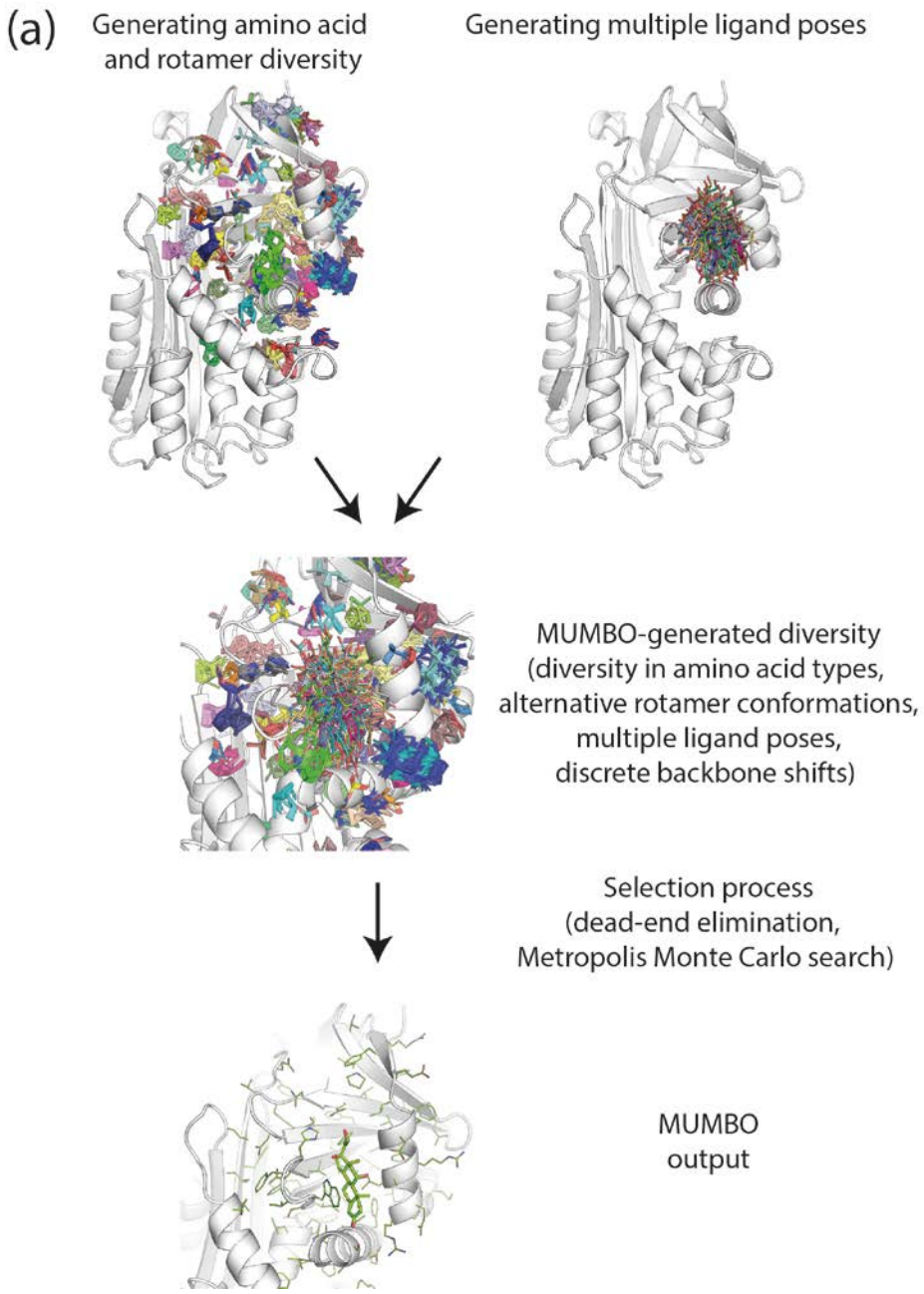


Fig. S3. Illustration of the side-chain-packing calculations. (a) Flow scheme of the side-chain-packing calculations performed with program MUMBO and (b) example input file for program MUMBO. A detailed description of all input key words is available at https://www.-biotechnik.nat.fau.de/research/downloads/.

(b)

Fig. S3. (continued)

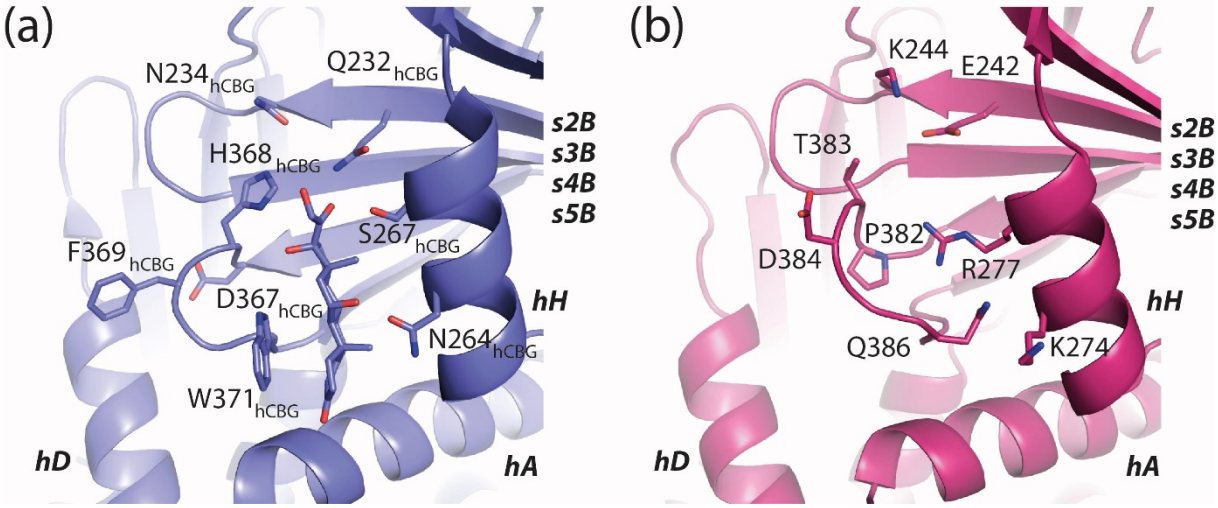


Fig. S4. Structure differences between (a) the corticoid-binding site in human CBG (in prune) and (b) the topologically equivalent site in human ACT (in magenta). In both panels eight out of nine positions are shown that were mutated to CBG residues in ACT in a first attempt to transfer the steroid-binding site from CBG to ACT. Additionally mutated residue Leu24 in ACT (Arg15 in CBG) is not shown. This residue is located at the beginning of helix hA and is not resolved in either the depicted ACT or CBG structure (PDB entries 1AS4 and 2VDY, respectively, (Lukacs et al., 1998; Zhou et al., 2008)). However, in the structure of rat CBG (PDB entry 2V95) this residue participates in a cation-π-interaction with Trp362 (Trp371 in human CBG) and is proposed to stabilize the orientation of the latter (Klieber et al., 2007). With the exception Arg277 that was mutated to a glycine residue present at the homologous position in rat CBG, all other mutations corresponded to substitutions with the homologous residues present in human CBG. Unfortunately, the nine-fold mutated ACT variant, termed NewBG-0, failed to bind any ligand.

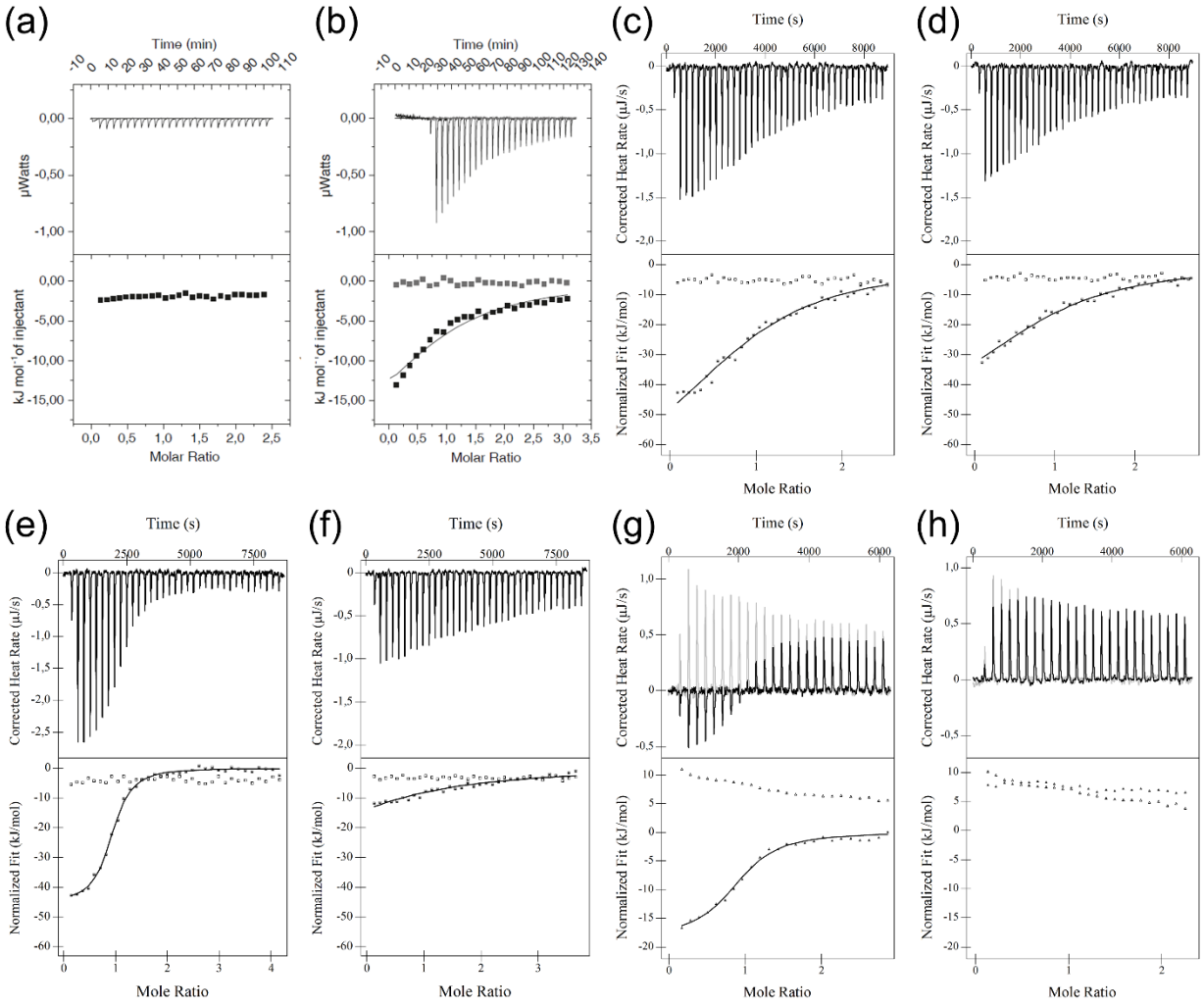


Fig. S5. Binding of steroids to NewBG variants and doxorubicin to DBS-II-allo-L55V variant investigated *via* ITC. Titration of (a) cortisol to uncleaved NewBG-0, (b) cortisol to uncleaved NewBG-I, (c) cortisol to uncleaved NewBG-II and (d) cleaved NewBG-II, (e) corticosterone to uncleaved NewBG-III, (f) aldosterone to uncleaved NewBG-III, (g) doxorubicin to uncleaved DBS-II-allo-L55V and (h) doxorubicin to cleaved DBS-II-allo-L55V. The titrations in (a) and (b) were performed with a VP-ITC microcalorimeter (Microcal, Piscataway), in (b) the reference cortisol into buffer titration is depicted as grey squares (bottom panel). All other titrations were conducted with a NanoITC standard volume calorimeter (TA Instruments, Eschborn, Germany). The corresponding reference titrations of the respective steroid ligand into buffer are shown as empty squares (bottom banels). In (g, h) the corresponding doxorubicin into buffer titration is depicted as gray line (upper panel) and empty triangles (bottom panel).

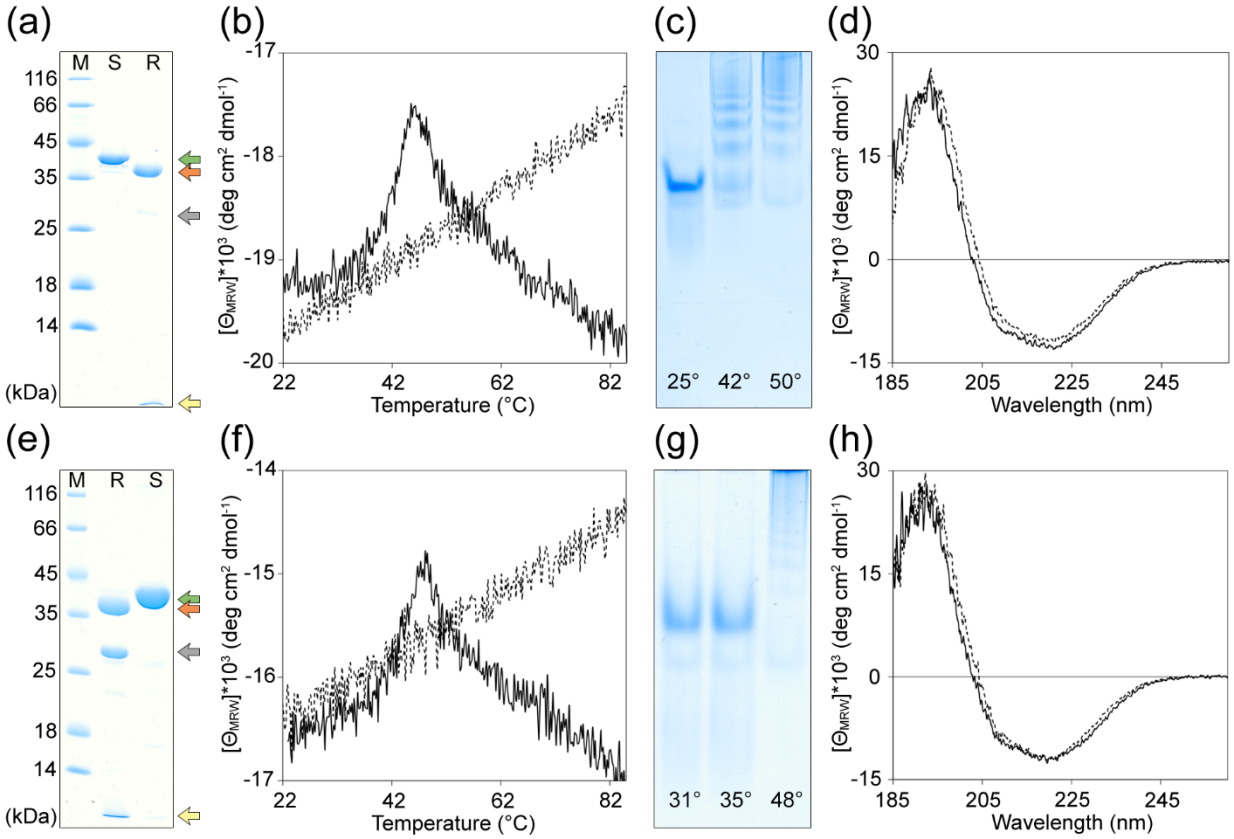


Fig. S6. Monitoring the S-to-R transition in NewBG variants. (a-d) Characterization of the S and R-conformations of NewBG-III and (e-h) NewBG-III-allo. (a, e) SDS-PAGE showing the shift to lower molecular weight upon RCL cleavage (uncleaved protein: green arrow, cleaved protein: orange arrow) and release of a 4.7 kDa peptide (yellow arrow). RCL cleavage with α-chymotrypsin produces an unspecific byproduct of 28 kDa (gray arrow). (b, f) Thermal stability of uncleaved (solid line) and cleaved NewBG variants (dotted line) monitored at 222 nm in a CD experiment. (c, g) Native PAGE of uncleaved NewBG variants after 5 min exposure at indicated temperatures. (d, h) Far UV CD spectra of uncleaved (solid line) and cleaved (dotted line) NewBG variants.

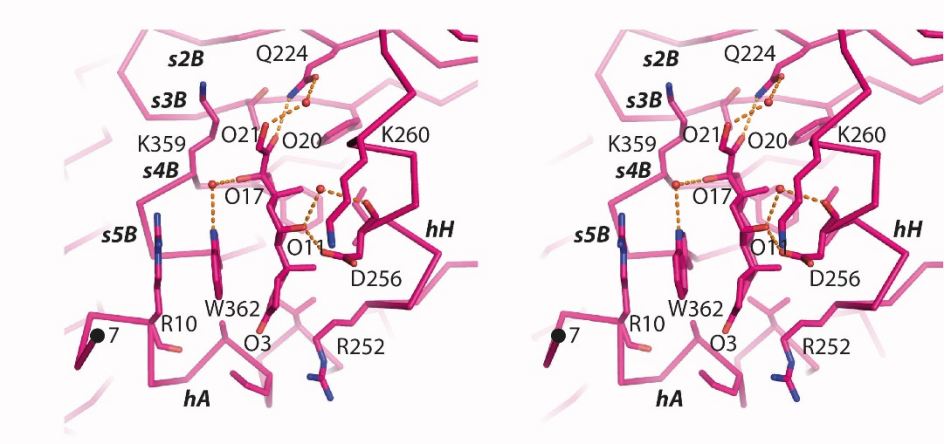


Fig. S7. Stereo representation of the cortisol-binding interactions in rat CBG as described in Klieber et al. (Klieber et al., 2007). Shown are all protein side-chains with atoms located within 5 Å of any ligand atoms plus the additional residues Arg10 and Lys260. Please see also Fig. S8 for cortisol atom numbering.

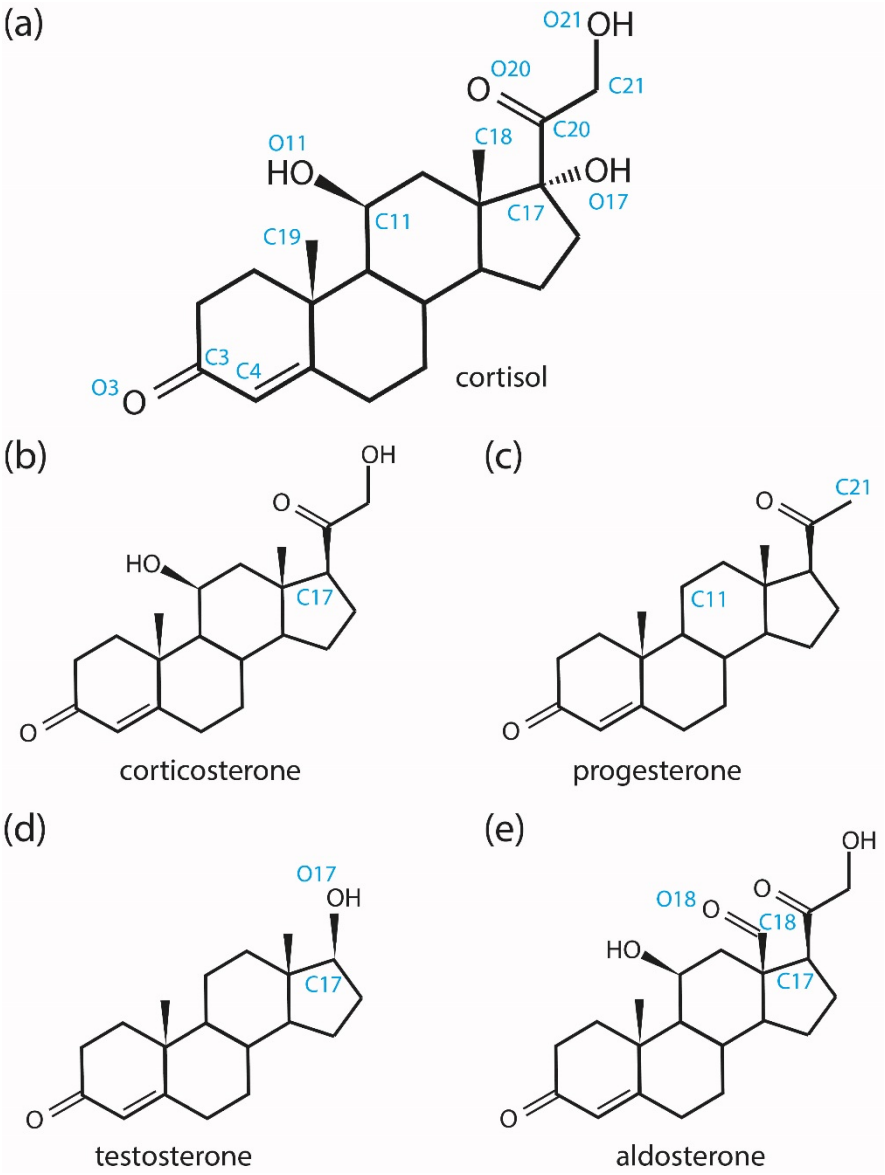


Fig. S8. Chemical structure of and atom numbering in the steroids used in this study. (a) Cortisol, (b) corticosterone , (c) progesterone, (d) testosterone and (e) aldosterone. In (b), (c), (d) and (e) only those atoms are highlighted that chemically differ from the corresponding atoms in cortisol. Atom numbering adheres to the common steroid nomenclature (Moss, 1989).

Supplementary references

Gardill, B.R., Vogl, M.R., Lin, H.Y., Hammond, G.L., Muller, Y.A., 2012. Corticosteroid-binding globulin: structure-function implications from species differences. PLoS One 7, e52759.

Gettins, P.G., 2002. Serpin structure, mechanism, and function. Chem Rev 102, 4751-4804.

Klieber, M.A., Underhill, C., Hammond, G.L., Muller, Y.A., 2007. Corticosteroid-binding globulin, a structural basis for steroid transport and proteinase-triggered release. J Biol Chem 282, 29594-29603.

Lukacs, C.M., Rubin, H., Christianson, D.W., 1998. Engineering an anion-binding cavity in antichymotrypsin modulates the "spring-loaded" serpin-protease interaction. Biochemistry 37, 3297-3304.

Moss, G.P., 1989. The Nomenclature of Steroids - Recommendations 1989. Eur J Biochem 186, 429-458.

Schmidt, K., Gardill, B.R., Kern, A., Kirchweger, P., Borsch, M., Muller, Y.A., 2018. Design of an allosterically modulated doxycycline and doxorubicin drug-binding protein. Proc Natl Acad Sci U S A 115, 5744-5749.

The UniProt Consortium, 2017. UniProt: the universal protein knowledgebase. Nucleic Acids Res 45, D158-D169.

Zhou, A., Wei, Z., Stanley, P.L., Read, R.J., Stein, P.E., Carrell, R.W., 2008. The S-to-R transition of corticosteroid-binding globulin and the mechanism of hormone release. J Mol Biol 380, 244-251.