Angewandte Chemie International Edition

DOI: 10.1002/anie.200906602

Protein Purification

Small-Molecule Affinity Ligands for Protein Purification: Combined Computational Enrichment and Automated In-line Screening of an Optically Encoded Library**

Jakob E. Rasmussen, Christine B. Schiødt, Søren F. Christensen, Leif Nørskov-Lauritsen, Morten Meldal, Phaedria M. St. Hilaire,* and Knud J. Jensen**

As a result of their complexity, the purification of protein biopharmaceuticals poses challenges not encountered with traditional small-molecule drugs and requires lengthy series of chromatographic steps. Affinity chromatography (AC) is perhaps the most powerful chromatographic technique available, but is limited in scope because of the lack of suitable ligands for proteins for which no natural small-molecule binding partner or inhibitor exists. Current AC ligands used for targeting protein surfaces are primarily other proteins, such as a monoclonal antibody (mAb) or another binding protein. These ligands are expensive and amenable to rapid degradation under common column-cleaning conditions. Selective, chemically robust small-molecule ligands would greatly expand the scope and use of AC for the purification of biopharmaceuticals.[1] However, small-molecule ligands targeting large protein surfaces are rare, despite significant recent achievements.[2] The reason remains the inherent difficulty in designing small molecules that bind tightly and specifically to a large, featureless, amphiphilic, and flexible protein surface.[3]

In this study we have focused on identifying small-molecule affinity ligands for human growth hormone (hGH) through the screening of a combinatorial library of compounds designed to mimic the interactions with its native receptor. Pioneering work has shown that the majority of the binding energy of the initial interaction between hGH and the extracellular part of its receptor (hGHbp) is conferred by a central functional epitope, or "hot spot", dominated by two tryptophan residues (Trp104 and Trp169) in close proximity on the receptor (Figure 1).[4,5] This ensemble of residues essential for binding suggests that small-molecule binders may be designed to mimic the interactions of the biological protein partner.

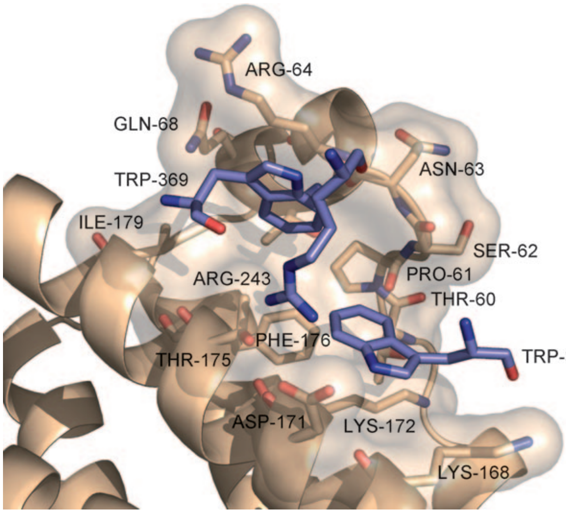


Figure 1. Interactions of the critical residues Trp104, Trp169, and Arg43 (blue, labeled TRP-304, TRP-369, and ARG-243) of hGHbp with site 1 of hGH (brown). Source: PDB ID 1A22.[5a]

Herein, we report a fully integrated methodology for the discovery and development of novel, small-molecule affinity resins, demonstrated by the development of the first non-protein affinity resin for the purification of hGH. The key feature of this research is the combination of virtual screening (VS) with "on-bead" automated, in-line library screening in which a novel optical encoding scheme is utilized to facilitate an exhaustive structure–affinity mapping of the entire one-bead-one-compound (OBOC) library.

First, computational screening of a large number of chemical building blocks was used to enrich the library and provide a more focused design. This automated approach also reduces the effect of human bias. The building blocks were extracted from a database of commercially available compounds (ACD) and incorporated into a virtual combinatorial library of more than 40000 compounds. The library was designed around a core diaminopropionic acid (DAPA)

---

[*] Dr. J. E. Rasmussen, Dr. C. B. Schiødt, Dr. S. F. Christensen,
Dr. L. Nørskov-Lauritsen, Dr. P. M. St. Hilaire
Protein Engineering, Novo Nordisk A/S
Novo Nordisk Park, 2760 Måløv (Denmark)
Fax: (+45) 4466-3939
E-mail: pshi@novonordisk.com
Dr. J. E. Rasmussen, Prof. Dr. K. J. Jensen
Faculty of Life Sciences, University of Copenhagen
Thorvaldsensvej 40, 1871 Frederiksberg (Denmark)
Fax: (+45) 3528-2398
E-mail: jera@life.ku.dk
kjj@life.ku.dk
Homepage: http://www.igm.life.ku.dk/kjj
Prof. Dr. M. Meldal
SPOCC Centre, Carlsberg Laboratory
Gamle Carlsberg Vej 10, 2500 Valby (Denmark)

[**] J.E.R. was supported by a PhD stipend from the Danish Ministry of Sciences and Technologies, Versamatrix A/S, and Novo Nordisk A/S (CORA)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200906602.

scaffold to which the 2715 carboxylic acid building blocks (R^3) and a short dipeptide (R^1 and R^2) were attached through amide bonds (Scheme 2). The peptide fragment was constructed from two identical sets of L-amino acids each consisting of Lys, Asp, Trp, and Asn. The carboxylic acid moiety of DAPA was substituted with a proton to avoid interference during the docking.

The intent of this design was to focus the majority of diversity at the R^3 position to allow screening of a more diverse set of building blocks within the technical constraints of the software. The virtual library was docked into a modified crystal structure of hGH (PDB ID 1A22)[5a] using the docking algorithm FlexX and evaluated by five different scoring functions and a consensus score. The docking was directed at the "hot spot" of site 1 to mimic the interactions of the critical tryptophan residues of the receptor.

Results from VS suggested a general preference for building blocks incorporating a bi- or tricyclic aromatic moiety in the R^3 position. Among the highest scoring building blocks were bicyclic aromatics, such as naphthalenes or heterocyclic *N*-derivatives, for example, quinolines and quinoxalines. The triheterocyclic compounds were characterized by a hydrophilic linear tricyclic structure with a hydrogen-bond acceptor heteroatom (N or O) in position 10, all of which shared a remarkably consistent binding motif. While one hydrophobic aromatic ring made an edge-to-face interaction with the phenyl moiety of Phe176, the other aromatic or aliphatic ring interacted with Ile179, the aliphatic chain of Arg64, and the methyl groups of Thr67 and Thr175 (Figure 2).

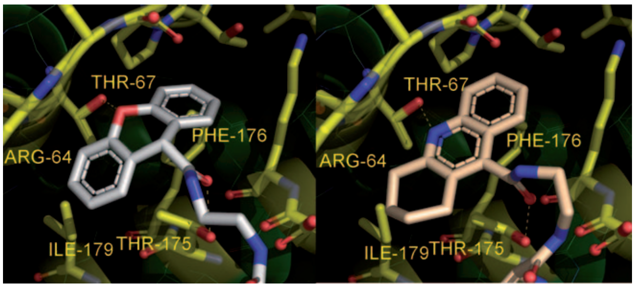
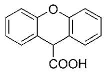
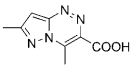
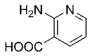
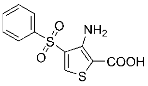
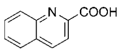
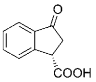
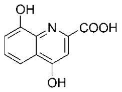
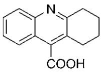
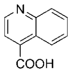
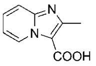
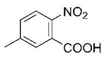
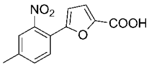
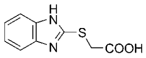
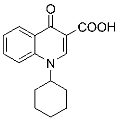
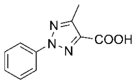
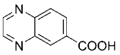
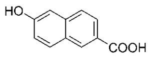
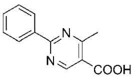


Figure 2. Interactions of ligands containing building blocks $\mathbf{R^3_1}$ (left, gray) and $\mathbf{R^3_{13}}$ (right, brown).

The heteroatom of the cyclic structure formed a hydrogen bond with the hydroxyl group of Thr67, while the carboxyl group of the amide connecting the building block to the scaffold formed a hydrogen bond to Thr175. Preference for amino acids varied within the different scoring functions, with Trp, not surprisingly, being the most favored followed by Lys, Asn, and Asp. A total of 18 building blocks for position R^3 were selected (Table 1) and combinatorialized with two identical sets of 11 L-amino acids (Trp, Lys, Asp, Asn, Tyr, His, Arg, Ile, Phe, Thr, and Pro) for positions R^1 and R^2 in an OBOC library,[6] thus providing a focused library of 2178 unique compounds.

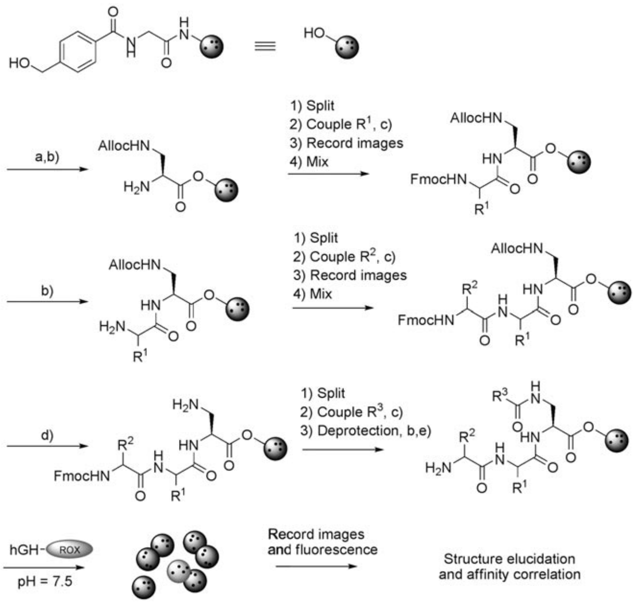
The OBOC library was encoded and deconvoluted using the optical bead-encoding protocol recently described by Meldal and Christensen.[7] The technology features facile and rapid structural elucidation of every library compound without the use of chemical tags or the destructive and laborious compound identification necessary with other methods.[8] The semiautomated protocol enables in-line fluorescence measurement of each library member, thus allowing an exhaustive mapping of the correlation between structure and affinity. In this way, structure–affinity information is presented not only for the most active compounds but also for all combinatorial products. The combination of 1) on-bead compound identification with 2) in-line screening/analysis of protein binding is a unique feature of this method.

Table 1: Building blocks (R^3) selected by virtual screening and incorporated into the OBOC library.

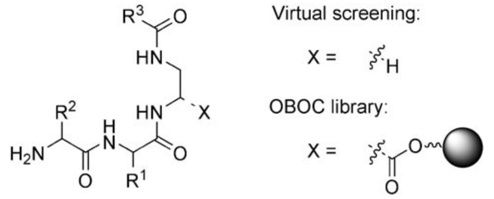
R^3_x	Structure	R^3_x	Structure
1		10	
2		11	
3		12	
4		13	
5		14	
6		15	
7		16	
8		17	
9		18	

The OBOC library was synthesized on approximately 50000 encoded poly(ethylene glycol)–poly(acrylamide) copolymer (PEGA) beads with a ligand density of 5.7 $\mu mol\,mL^{-1}$ (Scheme 1). The beads were incubated with rhodamine (ROX)-labeled hGH in a saline buffer at pH 7.5 and excess nonbinding protein was removed by washing. The ligand code and fluorescence intensity resulting from protein binding were measured in-line and simultaneously for each bead in the library at a rate of approximately three beads per second.

The effects of the individual building blocks on the overall affinity of the ligand were determined by analysis of variance (ANOVA). Interestingly, the two most active building blocks, $\mathbf{R^3_1}$ and $\mathbf{R^3_{13}}$, both belong to the group of tricyclic heterocyclic compounds and were the only such compounds in the set. This class of compounds was also found to show high affinity for hGH in the VS as detailed above. Of the 100 highest-scoring compounds in the library, 15 and 35 % contained $\mathbf{R^3_1}$ and $\mathbf{R^3_{13}}$,



Scheme 1. Synthesis, decoding, and affinity screening of the focused OBOC library. a) Fmoc-DAPA(Alloc)-OH, MSNT, MeIm, DCM; b) 20% piperidine/NMP; c) TBTU, NEM, DMF; d) $Pd(PPh_3)_4$, $Me_2NH \cdot BH_3$, DCM; e) $TFA/H_2O/TIS/EDT$. Fmoc = 9-fluorenylmethoxycarbonyl; Alloc = allyloxycarbonyl; MSNT = 1-(2-mesitylenesulfonyl)-3-nitro-1*H*-1,2,4-triazole; MeIm = 1-methylimidazole; DCM = dichloromethane; NMP = *N*-methylpyrrolidone; TBTU = *N*-[(1*H*-benzotriazol-1-yl)(dimethylamino)methylene]-*N*-methylmethanaminium tetrafluoroborate *N*-oxide; NEM = *N*-ethylmorpholine; TFA = trifluoroacetic acid; TIS = triisopropylsilane; EDT = 1,2-ethanedithiol.



Scheme 2. Ligand structures used for virtual screening and encoded OBOC library.

respectively. Most of the bicyclic compounds, however, exhibited low affinity for hGH. The amino acids Phe, Arg, and to a lesser extent Lys, Ile, and Tyr showed the highest affinity to hGH whereas Trp, surprisingly, did not affect affinity significantly.

Figure 3 shows the extensive nature of the data collected by using this protocol, which emphasizes the unique capabilities of this strategy. Figure 4 illustrates clearly that building blocks $\mathbf{R^3_1}$, $\mathbf{R^3_{11}}$, $\mathbf{R^3_{13}}$, and $\mathbf{R^3_{14}}$ contribute to ligands with significantly higher than average affinity, as indicated by the fluorescence intensity. Correlation of the screening results with docking predictions was good, with the scoring function Chemscore proving particularly accurate. Thus, $\mathbf{R^3_1}$, $\mathbf{R^3_{11}}$, and $\mathbf{R^3_{13}}$ were ranked 2, 6, and 4, respectively, among the total of 2715 docked building blocks, which indicates a positive enriching effect from the VS protocol.

Finally, a selection of the most promising ligands (Table 2) was resynthesized on amino-functionalized Fractogel. The resulting affinity resins were evaluated for their ability to selectively adsorb and elute hGH from a microfiltered *Escherichia coli* lysate. The hGH microfiltrate (2.6 $mg\,mL^{-1}$) was loaded onto the resin in 25 mM Tris buffer at pH 7.50 and eluted using a 1 M NaCl gradient in the same buffer. The purity of the eluent was qualitatively evaluated by SDS-PAGE. The

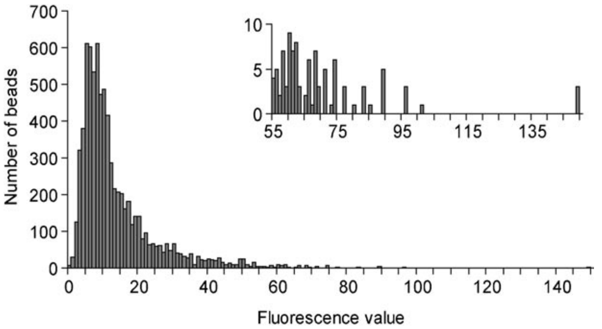


Figure 3. Distribution of measured fluorescence intensity, which shows a large number of beads with low fluorescence intensity close to background (8 units) and a few high-fluorescence hits (inset).

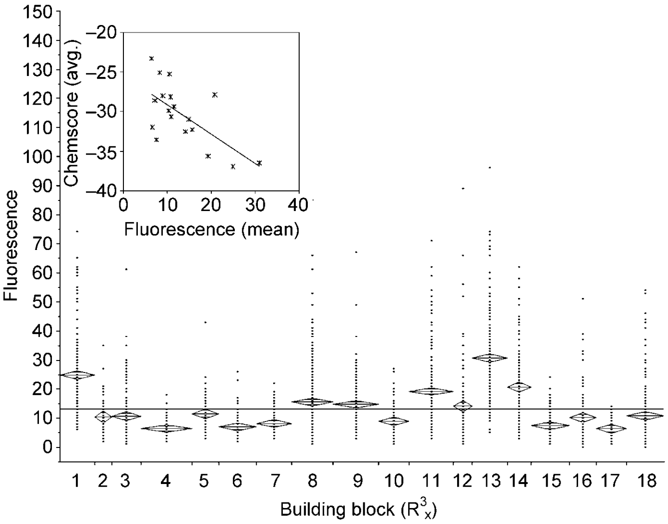


Figure 4. ANOVA results for the 18 different building blocks in position R^3. Each dot on the chart represents a bead and the diamonds the mean fluorescence for that building block. The horizontal line is the mean fluorescence intensity for the entire population, with a value of 13 units. The inset shows the correlation of the average docking score (Chemscore, lower is better) and the measured fluorescence for each of the 18 building blocks.

Table 2: Evaluation results and corresponding fluorescence intensity of selected affinity resins.

Resin	R^1	R^2	R^3	Flu[a]	DBC [$mg\,mL^{-1}$]	Selectivity[b]
F-1	Tyr	Arg	1	149	13.8	very poor
F-10	Phe	Ile	13	96	2.8	very good
F-15	Arg	Phe	14	62	< 5	good
F-18[c]	His	Asp	6	5	0	n/a

[a] Average fluorescence intensity of beads recorded during in-line screening. [b] SDS-PAGE evaluation. [c] Negative control. n/a = not applicable.

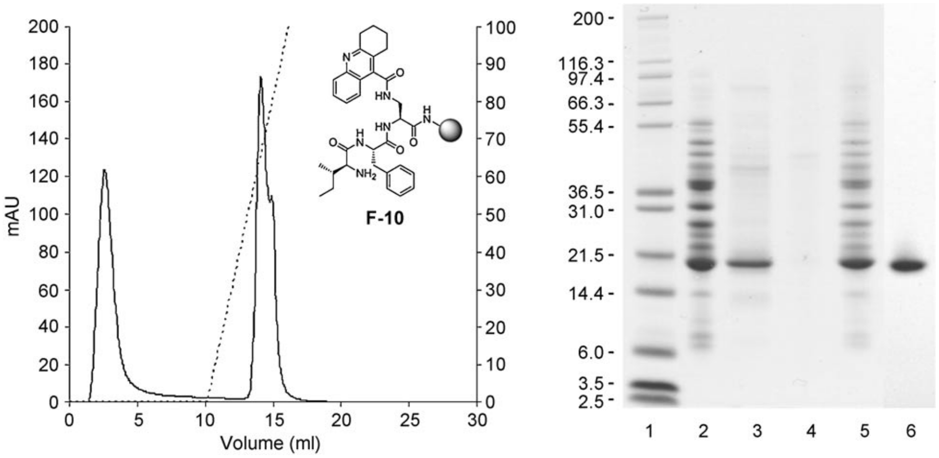


Figure 5. Left: Chromatogram showing the purification of hGH using affinity resin **F-10**; —— absorbance; ----- buffer gradient. Right: Coomassie-stained SDS-PAGE analysis of purification using a loading buffer of pH 6.25. Lane 1: protein marker; lane 2: flow-through; lane 3: eluate (pooled); lane 4: CIP (pooled); lane 5: hGH microfiltrate reference; lane 6: hGH reference (insert from separate gel).

resins were cleaned in place (CIP) with 0.2 M NaOH, thus underscoring the advantage of having a base-tolerant ligand attached to the resin. In particular, affinity resins containing $\mathbf{R}^3_{13}$ were found to selectively adsorb hGH and to provide good rates of recovery, whereas two negative controls selected from library members with the lowest fluorescence intensity after screening had no affinity for the hormone.

Optimization of the loading and elution conditions for the most promising affinity resin (**F-10**; Table 2) showed a distinct correlation between the pH of the loading buffer and the dynamic binding capacity (DBC) of the resin. The capacity for hGH thus declined sharply above pH 6.50 and was zero above pH 7.75. The optimal conditions were found to consist of a loading buffer of 50 mM 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol (Bis-Tris) at pH 6.25, at which a 10% DBC for pure hGH of 6.7 mg mL^{-1} resin was achieved (Figure 5). The 10% DBC of the hGH microfiltrate under these conditions was 3.2 mg mL^{-1}. Elution was best effected by using either a simple salt gradient or an increase in pH to values higher than 7.50. With the latter, we were able to isolate hGH from the untreated microfiltrate at a purity of 91% (HPLC) in a single purification step. Recovery was 61%, while higher rates of recovery up to 83% were obtained when eluting at pH 8.00 at the cost of lower purity (80%). The same ligand was synthesized, purified, and attached to a Fractogel resin (**F-21**) for validation of these results. Evaluation of this resin showed similar selectivity but slightly improved DBC, and the eluted hGH showed full biological activity in a BaF-3/hGHR cell proliferation assay.

In conclusion, we have demonstrated a versatile and expeditious methodology for the development of novel, small-molecule AC ligands for protein targets. By combining computational enrichment with automated, in-line screening and affinity mapping of an optically encoded combinatorial library, we discovered a novel set of AC ligands capable of efficiently isolating a difficult protein target from a complex fermentation broth. We believe that the generality and speed of this method holds great potential to broaden the scope and applicability of AC in protein purification.

Received: November 23, 2009
Revised: January 18, 2010

Keywords: affinity chromatography · combinatorial chemistry · hormones · protein–protein interactions · virtual screening

---

[1] a) A. C. A. Roque, C. R. Lowe, *Biotechnol. Adv.* **2006**, *24*, 17–26; b) N. E. Labrou, *J. Chromatogr. B* **2003**, *790*, 67–78.

[2] a) J. K. Murray, S. H. Gellman, *Pept. Sci.* **2007**, *88*, 657–686; b) H. Yin, A. D. Hamilton, *Angew. Chem.* **2005**, *117*, 4200–4235; *Angew. Chem. Int. Ed.* **2005**, *44*, 4130–4163; c) M. R. Arkin, J. A. Wells, *Nat. Rev. Drug Discovery* **2004**, *3*, 301–317.

[3] L. L. Conte, C. Chothia, J. Janin, *J. Mol. Biol.* **1999**, *285*, 2177–2198.

[4] a) W. L. DeLano, *Curr. Opin. Struct. Biol.* **2002**, *12*, 14–20; b) I. S. Moreira, P. A. Fernandes, M. J. Ramos, *Proteins Struct. Funct. Bioinf.* **2007**, *68*, 803–812.

[5] a) T. Clackson, M. H. Ultsch, J. A. Wells, A. M. de Vos, *J. Mol. Biol.* **1998**, *277*, 1111–1128; b) T. Clackson, J. A. Wells, *Science* **1995**, *267*, 383–386.

[6] a) K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, R. J. Knapp, *Nature* **1991**, *354*, 82–84; b) Á. Furka, F. Sebestyén, M. Asgedom, G. Dibó, *Int. J. Pept. Protein Res.* **1991**, *37*, 487–493.

[7] M. Meldal, S. F. Christensen, *Angew. Chem.* **2010**, DOI: 10.1002/ange.200906563; *Angew. Chem. Int. Ed.* **2010**, DOI: 10.1002/anie.200906563.

[8] a) K. Braeckmans, S. C. D. Smedt, M. Leblans, C. Roelant, J. Demeester, *Nat. Rev. Drug Discovery* **2002**, *1*, 1–10; b) C. Barnes, S. Balasubramanian, *Curr. Opin. Chem. Biol.* **2000**, *4*, 346–350.