ChemMedChem

Full Papers
doi.org/10.1002/cmdc.202100458

**Chemistry Europe**
European Chemical
Societies Publishing

# A Step toward NRF2-DNA Interaction Inhibitors by Fragment-Based NMR Methods

Sven Brüschweiler,*[a] Julian E. Fuchs,[b] Gerd Bader,[b] Darryl B. McConnell,[b] Robert Konrat,[a] and Moriz Mayer[b]

The NRF2 transcription factor is a key regulator in cellular oxidative stress response, and acts as a tumor suppressor. Aberrant activation of NRF2 has been implicated in promoting chemo-resistance, tumor growth, and metastasis by activating its downstream target genes. Hence, inhibition of NRF2 promises to be an attractive therapeutic strategy to suppress cell proliferation and enhance cell apoptosis in cancer. Direct targeting of NRF2 with small-molecules to discover protein-DNA interaction inhibitors is challenging as it is a largely intrinsically disordered protein. To discover molecules that bind to NRF2 at the DNA binding interface, we performed an NMR-based fragment screen against its DNA-binding domain. We discovered several weakly binding fragment hits that bind to a region overlapping with the DNA binding site. Using SAR by catalogue we developed an initial structure-activity relationship for the most interesting initial hit series. By combining NMR chemical shift perturbations and data-driven docking, binding poses which agreed with NMR information and the observed SAR were elucidated. The herein discovered NRF2 hits and proposed binding modes form the basis for future structure-based optimization campaigns on this important but to date 'undrugged' cancer driver.

## Introduction

Nuclear factor erythroid 2-related factor 2 (NRF2) a cap'n' collar (CNC), basic leucine zipper (bZIP) transcription factor is one of the main regulators of cellular antioxidant response, thereby mitigating cell damage during chemical and oxidative stress. During basal conditions NRF2 is bound to its cytosolic repressor Kelch-like ECH-associated protein 1 (KEAP1), which is involved in marking NRF2 for proteasomal degradation, thereby maintaining low levels of NRF2. However, increased levels of oxidative or electrophilic stress lead to the dissociation of the KEAP1-NRF2 complex and NRF2 translocates to the nucleus where it forms a heterodimer with the small musculoaponeurotic fibrosarcoma oncogene homologue (Maf) proteins. Dimerization enables binding to the antioxidant-response elements (ARE) sequence within the promoter regulatory regions of cytoprotective genes. Binding to the promoter region induces transactivation of proteins involved in a diverse set of cellular functions such as xenobiotic metabolism and excretion, autophagy, protein homeostasis, DNA repair, cell survival, and mitochondrial function, giving activation of the NRF2 pathway a direct role in cancer prevention.[1]

In recent years it has become apparent that high levels of NRF2 can have a prooncogenic effect and lead to chemo-resistance; elevated NRF2 activity is mostly caused through genetic alterations in the KEAP1/NRF2 pathway.[2] Constitutive activation of NRF2 can be caused by promoter demethylation, copy number amplification of its gene, oncogene-induced transcription of NRF2 via c-Myc$^{ERT2}$, BRAF$^{V600E}$, and KRAS$^{G12D}$, or by gain-of-function somatic mutations in the regions responsible for KEAP1 interactions.[2b,3] Direct alterations of NRF2 have been documented for lung, colorectal, ovarian, pancreatic, and head and neck cancers.[1,4] Further, an increase of NRF2 through deletion, decreased transcription, and loss-of-function mutations for KEAP1 has been described in the literature.[4a,5] High levels of NRF2 in cancer cells enables increased proliferation and chemo- and radio resistance, making NRF2 inactivation a promising intervention strategy to enhance therapeutic effects for certain types of cancer.[6]

Human isoform 1 NRF2 consists of 605 amino acids which form seven conserved NRF2-ECH homology (Neh1-7) domains (Figure 1A).[7] The N-terminal Neh2 domain is responsible for KEAP1 binding; Neh6 functions as an additional negative regulatory domain through its interaction with beta-transducin repeat-containing domain (β-TrCP). The domains Neh3 to Neh5 are responsible for target gene transactivation, whereas the Neh7 domain has been shown to be involved in the inhibition of the NRF2-ARE-signaling pathway.[4b] Our study focuses on the Neh1 domain which forms a CNC basic leucine zipper DNA binding domain and is crucial for heterodimer formation with the small Maf (sMaf) proteins. NRF2 is predicted to be largely an intrinsically disordered protein in the absence of its binding

[a] Dr. S. Brüschweiler, Dr. R. Konrat
Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology,
Department of Structural and Computational Biology, Max Perutz Labs, University of Vienna
Campus Vienna Biocenter 5
1030 Vienna (Austria)
E-mail: sven.brueschweiler@univie.ac.at

[b] Dr. J. E. Fuchs, Dr. G. Bader, Dr. D. B. McConnell, Dr. M. Mayer
Boehringer Ingelheim RCV GmbH & Co. KG
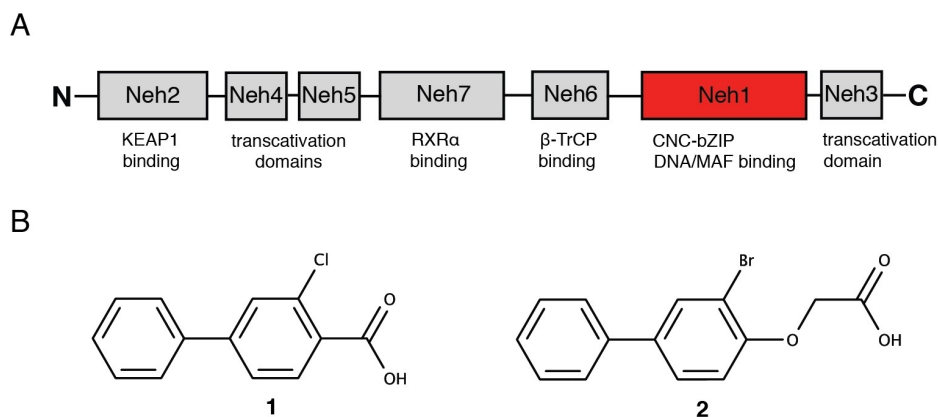Dr. Boehringer Gasse 5–11, 1121 Vienna (Austria)

**Figure 1.** NMR-based fragment screen identified Neh1 binders. (A) Schematic representation of the organization of the seven NRF2 domains, the Neh1 domain is highlighted in red. (B) Chemical structures of the biphenyl carboxylic acid compounds identified in the fragment screen.

partners with the exception of the CNC and parts of the basic leucine zipper region of Neh1 (Figure S1).

So far, only a few small molecule inhibitors of NRF2 have been reported, and only one of them, ML385, has been shown to directly interact with NRF2. ML385 is reported to interfere with small MafG heterodimer formation through binding to the leucine zipper region (LZIP),[8] although concern has been raised about ML385 acting as a broad inhibitor of various transcription factors.[1,4b]

The paucity of NRF2 inhibitors is mostly due to its lack of a well-defined three-dimensional structure. This, and its activation through protein-protein interactions (PPI) make NRF2 a challenging target for small-molecule drug discovery.[9] These circumstances apply to transcription factors (TFs) and intrinsically disordered proteins (IDPs) in general. Consequently, they were historically deemed to be 'undruggable'. Direct targeting of TFs, proposed by James Darnell almost 20 years ago,[10] with small-molecules that modulate deregulated TFs in cancer, requires the disruption of protein-protein or protein-DNA interactions. The tractability of modulating or disrupting PPIs by direct targeting of the frequently large and solvent exposed surface areas between TFs and their cofactors, as well as activators/repressors with drug-like small-molecules has been established in recent years.[11] The discovery of small-molecule binders to the typically convex and positively charged protein surface at the DNA binding site for direct modulation of their interactions, has however proven to be an even more challenging task.[12] Hence, only a very small number of molecules interfering competitively with DNA binding have been described in the literature.[13] Protein-protein and protein-DNA interfaces tend to have flat discontinuous cavities more amenable to fragment sized compounds which can be missed in HTS campaigns.[14]

Commercial so-called virtual make-on-demand chemical libraries can be an efficient way to widen and tailor the chemical space around a small-molecule hit, successfully demonstrated in several recent structure-based ligand docking screens using Enamine's virtual compound library (https://

enamine.net/).[15] Therefore, we set out to discover small-molecule hits to NRF2 with a protein NMR-based fragment screen - a method that has proven to be fruitful in the past for challenging drug targets[16] – to identify potential small-molecule binding pockets. In a second step we explore the structure-activity relationships (SARs) of the initial hits, for which we search the Enamine virtual make-on-demand compound library[17] to increase the number of fragment hits and to assess the feasibility of obtaining higher affinity binders without tying up in-house medicinal chemistry resources. Subsequently, to rationalize the SAR results and guide future structure-based analogous optimization we determined the NMR solution protein-ligand complex structure in a chemical shift perturbation driven docking simulation.

## Results and Discussion

**Fragment-based screen**. We set out to identify compounds that bind the Neh1 (residues 445–523, denoted by Neh1-ΔLZIP) domain of NRF2, a construct that does not include the LZIP part, by a fragment-based screening approach. Initially a library of 1800 commercial and internal fragments was screened by observing cross-peak shifts of uniformly $^{15}N$ labelled Neh1-ΔLZIP, indicative of ligand binding, in a 2 dimensional $^1H$-$^{15}N$ heteronuclear single-quantum correlation (HSQC) experiment. Two fragments (Figure 1B) found in the screen were characterized in more detail, where compound **2** showed significantly larger chemical shift perturbations (CSPs) (Figure 2A, B and Figure S2A) and the dissociation constant ($K_D$) was determined to be $1.7 \pm 0.1$ mM (average over R499, D500, R503) as obtained from the CSP values (Figure 2C). For several residues CSP values larger than 0.10 ppm were observed, namely A452, M484, R499, D500, and R503, with a standard deviation of 0.05 ppm for all assigned backbone amide chemical shift changes in the $^1H$-$^{15}N$ HSQC titration experiment for compound **2**. Furthermore, the side chain amide of N482 had similar dose-response to **2**; a chemical shift change of 0.3 ppm was recorded for the side
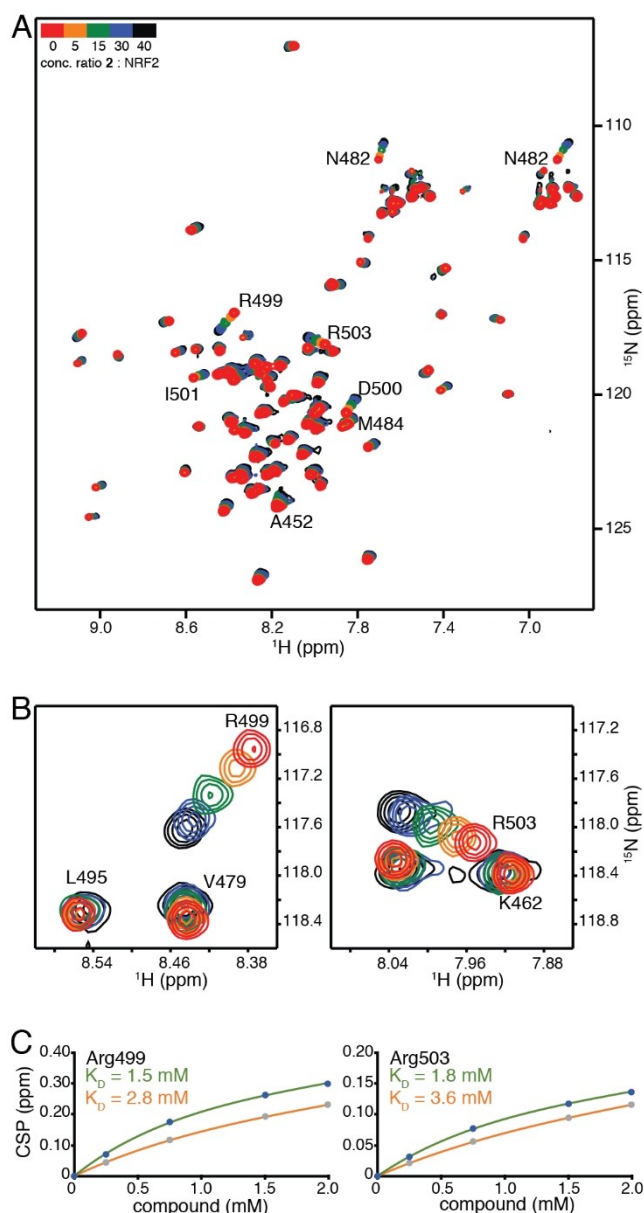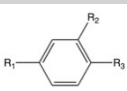
**ChemMedChem**

Full Papers
doi.org/10.1002/cmdc.202100458

Chemistry
Europe
European Chemical
Societies Publishing

**Figure 2.** NMR experiments of Neh1 binders. (A) Overlay of $^1$H-$^{15}$N HSQC spectra of uniformly labeled 50 μM $^{15}$N-Neh1-ΔLZIP (red) titrated with 250, 750, 1500, and 2000 μM compound **2** (from orange to black). Residues with normalized chemical shift perturbations larger than 0.1 ppm are indicated with residue numbers. (B) Selected regions of the $^1$H-$^{15}$N HSQC spectra overlay showing close up views of R499 and R503 and their concentration-dependent chemical shift changes upon increasing amounts of compound **2**. (C) Titration curves obtained from the $^1$H-$^{15}$N HSQC titration measurements for residues R499 and R503. $K_D$s were derived by fitting experimental CSP values to equation (1); experimental CSP values are shown as blue and grey points for compound **2** and **3**, respectively. The lines represent an individual residue fit to equation (1); fitted lines for compound **2** and **3** are colored green and orange, respectively.

chain and 0.04 ppm for the backbone amide, respectively. Due to the small magnitude of CSPs for compound **1** (Figure S2A) we focused our follow-up screen of commercial and internal compounds on molecules structurally similar to compound **2**.

Because of the low hit rate in the initial screen, often seen for proteins lacking an easily druggable binding pocket,[18] a close analogs search was carried out. Specifically, we searched for compounds with a core phenyl carboxylic acid scaffold with variable side chains branching of the phenyl ring(s) as well as derivatives of the linker connecting the aromatic ring and the carboxylic acid. The search resulted in 27 analogs from internal and commercial sources (e.g. compounds **3**, **4**, **10**, **11**, **12**, **13**, **14**, and **15** in Table 1). To further expand the SAR we searched the virtual make-on-demand library from Enamine, which contains 15.5 billion virtual compounds that can be synthesized out of 111 500 building blocks (https://enamine.net/library-synthesis/real-compounds/real-space-navigator). Compound **2** served as query structure in a REAL Space Navigator (bioSolveIT) search, a software tool (now infiniSee) that allows the exploration – based on pharmacophore similarity – of the virtual chemical space of the Enamine Real Space compounds.[17a] The chosen make-on-demand compounds were all composed of three Enamine building blocks; in total 8 compounds were delivered (e.g. compounds **5**, **6**, **7**, **8**, and **9** in Table 1). All 35 analogs of the follow-up search were tested by $^1$H-$^{15}$N HSQC against Neh1-ΔLZIP. 15 representative analogs are shown in Table 1, four of which (**3**, **5**, **6**, **9**) showed a similar CSP pattern to **2** (Figure S2B, C, D); compound **5** caused significantly larger CSP values than **2**, which could be due to the additional hydroxy group. The remaining compounds showed no or very small CSPs, furthermore compound **7** and **8** showed dose response as well. However, they lead to the precipitation of Neh1 at higher compound concentrations. Analysis of the CSP patterns indicated that the compounds most likely have similar binding sites and modes, however, they differ in affinities (Table 1). In a subsequent step we assessed the binding contribution of the negatively charged carboxyl group with four commercial and internal compounds lacking this functional group (e.g. compounds **16**, **17**, and **18** in Table 1). These molecules contained an amine or an amide instead of the carboxylic acid, none of which showed CSPs in $^1$H-$^{15}$N HSQC titration experiments against Neh1-ΔLZIP in the structured part of NRF2. Small chemical shift changes upon ligand addition were observed for an unassigned residue in the unstructured C-terminal part of the protein which are most likely caused by small changes in pH and/or ionic strength of the protein solution (Figure S2H, I).

The SAR analysis showed that the potency of the Neh1 binder **2**, 2-((3-bromo-[1,1′-biphenyl]-4-yl)oxy)acetic acid, is very sensitive to structural changes at several positions; changes of the bromo substituent at the three-position on the phenyl ring to either a bulky cyclohexane (**13**) or to a small hydrogen (**4**) led to no detectable chemical shift changes upon compound titration. Furthermore, substituting the bromine with a chlorine led to an approximately 2-fold decrease in ligand potency determined by $^1$H-$^{15}$N HSQC titration, confirming the importance of the halogen substituent in the binding mode (Figure 2C, Table 1). The contribution of the halogen bond could not be further investigated by an iodine containing compound due to the lack of availability. Special care must be taken as this is a $K_D$ difference close to the limit of distinguishability,

**ChemMedChem**

Full Papers
doi.org/10.1002/cmdc.202100458

**Chemistry Europe**
European Chemical
Societies Publishing

**Table 1.** Structure-activity relationships of analogs tested by $^{1}$H-$^{15}$N HSQC titration.

| Compound | R$_1$ | R$_2$ | R$_3$ | HSQC[a] | KD [mM][b] |
|---|---|---|---|---|---|
| 1 | phenyl | Cl | acetic acid | + | − |
| 2 | phenyl | Br | oxy-acetic acid | + | 1.7 ± 0.1 |
| 3 | phenyl | Cl | oxy-acetic acid | + | 3.0 ± 0.5 |
| 4 | phenyl | H | oxy-acetic acid | − | − |
| 5 | 2-hydroxyphenyl | Cl | oxy-acetic acid | + | 1.5 ± 0.1 |
| 6 | 3-hydroxyphenyl | Cl | oxy-acetic acid | + | 2.8 ± 0.4 |
| 7 | 2-chlorophenyl | Cl | oxy-acetic acid | + | precipitation |
| 8 | 3-chlorophenyl | Cl | oxy-acetic acid | + | precipitation |
| 9 | 3-aminophenyl | Cl | oxy-acetic acid | + | 3.2 ± 0.7 |
| 10 | CH$_3$ | Br | oxy-acetic acid | − | − |
| 11 | morpholine | Br | oxy-acetic acid | − | − |
| 12 | pyrimidine | Br | oxy-acetic acid | − | − |
| 13 | phenyl | cyclohexyl | oxy-acetic acid | − | − |
| 14 | phenyl | H | oxy-propanoic acid | − | − |
| 15 | phenyl | H | oxy-propanoic acid | − | − |
| 16 | phenyl | F | ethylamine | − | − |
| 17 | phenyl | H | amide | − | − |
| 18 | phenyl | Br | N-methyl amide | − | − |

[a] + indicates chemical shift perturbations in $^{1}$H-$^{15}$N HSQC spectra in presence of the compound. [b] K$_D$ values were determined from CSP values using equation (1).

therefore the titration experiments were carried out using the same protein stock solution and the ligand concentration was determined by 1D peak integration. On the one hand the 2-fold potency decrease could be reversed by changing the R$_3$ phenyl group to an ortho phenol group (2'-hydroxy-phenyl) **5** but not by a meta phenol (3'-hydroxy-phenyl) group **6**. On the other hand substituting the R$_3$ phenyl group to a methyl (**10**), morpholine (**11**), and a pyrimidine (**12**) group caused the compounds to lose their potency (Figure S2F, G, Table 1).

Based on our structure-activity relationship we derived a pharmacophore for small-molecule Neh1 binders; all of the better binders in this study contained three features, a biphenyl motif, a halogen at the three-position on the first phenyl ring and additionally an acetic acid para to the second aryl ring.

As none of the fragments in the follow up screen resulted in higher affinity binders we set out to determine the three-dimensional structure of the protein-small-molecule complex. This is key to advance the initial fragments hits in a structure-based design approach.

**ChemMedChem**

Full Papers
doi.org/10.1002/cmdc.202100458

**Chemistry
Europe**

European Chemical
Societies Publishing

**Structure determination of Neh1**. We determined the binding pose of fragment hit **2** in a two-step approach. In the first step we solved the ligand free Neh1 domain structure by

| Table 2. NMR and refinement statistics. | |
|---|---|
| NMR distance and dihedral restraints | Neh1-ΔLZIP (BMRB 34617, PDB 7O7B) |
| Distance restraints | |
| Total NOE | 337 |
| Intraresidual (i = j) | 182 |
| Sequential (i–j = 1) | 77 |
| Medium range (1 < i–j < 5) | 34 |
| Long range (i–j ≥ 5) | 44 |
| Total Dihedral restraints | 92 |
| φ | 46 |
| ψ | 46 |
| Structure Statistics | |
| Violations (mean ± SD) | |
| Distance restraints (Å) | 0.09 ± 0.01 |
| Dihedral restraints (°) | 0.97 ± 0.12 |
| Deviation from idealized geometry | |
| Bond length (Å) | 0.0034 ± 0.0001 |
| Bond angles (°) | 0.50 ± 0.02 |
| Ramachandran statistics | |
| Residues in allowed regions | 96.1% |
| Residues in disallowed regions | 3.9% |
| Average pairwise rmsd (Å) | |
| Heavy atoms (2° structure) | 1.07 ± 0.08 |
| Backbone atoms (2° structure) | 0.42 ± 0.07 |

solution state NMR spectroscopy and in a second step determined the binding pose of **2** by data-driven docking using the program HADDOCK.[19] The NMR solution structure of the Neh1-ΔLZIP domain of NRF2 (residues 445–523, PDB ID 7O7B) was determined using triple-resonance-resolved NMR spectroscopic techniques; backbone chemical shifts and NOEs were used as input restraints for the structure determination (Table 2).

The central scaffold of the Neh1-ΔLZIP domain is formed by a bundle of four α-helices (residues R456-L464, α1; residues V470-N475, α2; residues V478-K487, α3; residues E492-G505, α4) (Figure 3A, B). The structural orientation of the central scaffold is well defined by the 44 long range NOEs, which were determined in this study, as there are several NOEs between the structural elements that determine their relative arrangement to one another (Figure S3A, B). The C-terminal part of helix 1, the loop connecting helix 1 and helix 2 as well as terminal parts of helices 3–4 and their connecting loop pack together to form a hydrophobic core, which is capped by helix 2 on the C-terminal side of the protein. Helix 4 contains the basic region, starting from D500 and ending at K518, and is unstructured beyond residue G505 with the remaining residues likely only folded when DNA is bound, as seen for the DNA-binding domain of the CNC transcription factor Skn-1.[20]

The solution structure of a similar construct of the Neh1 domain was previously determined by the Northeast Structural Genomics Consortium (NESG). The NESG structure calculation and ours were conducted with a similar number of long range NOEs between unique residue pairs, 31 and 25, respectively. The here reported Neh1-ΔLZIP (PDB ID 7O7B) backbone
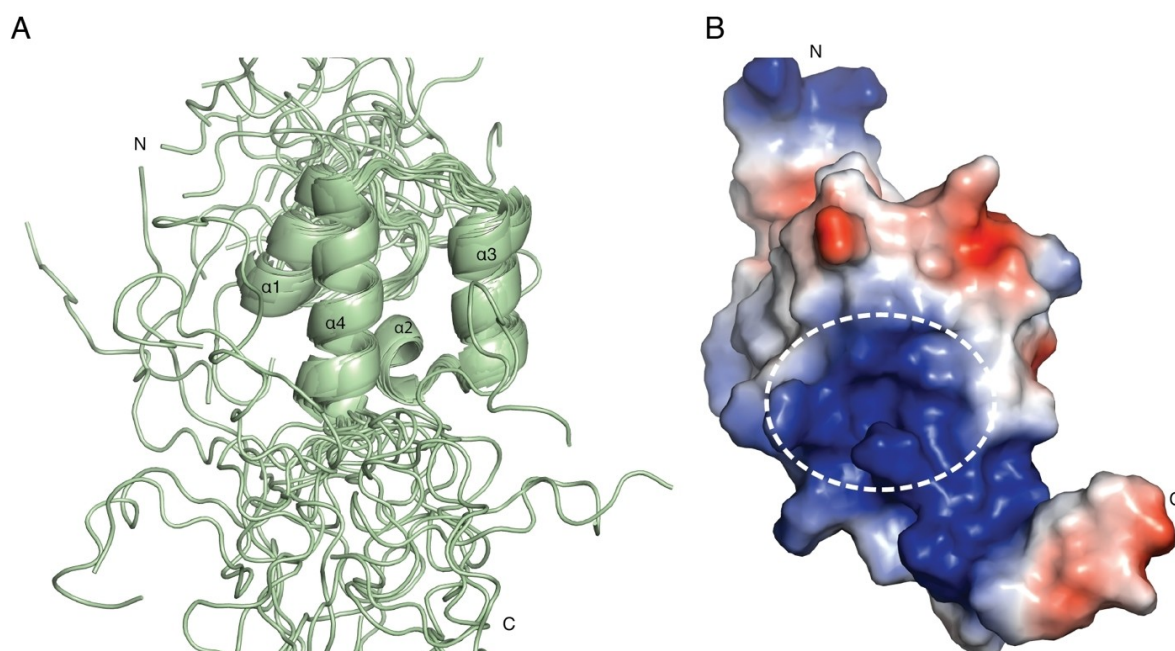


**Figure 3.** NMR solution structure of the Neh1 domain of NRF2. (A) NMR ensemble of 20 representative structures of the Neh1-ΔLZIP (PDB ID 7O7B) domain of NRF2; helices are displayed as ribbons. Secondary structure elements are labeled with increasing numbers from N- to C-terminus. (B) Lowest energy structure of Neh1-ΔLZIP (PDB ID 7O7B) shown in a surface representation colored by its electrostatic potential with red being negative and blue positive. The white dotted line outlines the binding area of the fragment hits.

# ChemMedChem

**Full Papers**
doi.org/10.1002/cmdc.202100458

**Chemistry
Europe**
European Chemical
Societies Publishing

structure can be superimposed with the Neh1 solution NMR structure (PDB ID 2LZ1) determined by NESG with a root mean square deviation for Cα atoms (rmsd$_{C\alpha}$) of 1.8 Å (Figure S4A). Helix 3 differs significantly between the two structures, in our structure helix 3 is in a straight conformation and lacks the bend that is observed in 2LZ1. This bend is also absent in the sMaf:DNA (PDB ID 3 A5T) and Skn-1:DNA (PDB ID 1SKN) complex structure, with the three proteins showing high sequence identity and presumably a conserved fold of their DNA-binding domains (Figure S4B, C).[20b,21] Structural alignment of the solution Neh1-ΔLZIP (PDB ID 7O7B) structure and the X-ray structures of the sMaf and Skn-1 DNA binding domains (PDB ID 3 A5T, 1SKN) yield a rmsd$_{C\alpha}$ of 1.6 and 1.3 for the helical regions, respectively.

**Data-driven docking of compound 2.** Mapping the chemical shift changes of the fragment hits onto the Neh1-ΔLZIP structure revealed that the largest changes occur in a well-defined surface patch, formed by the side chain amide of N482 on helix 3, and the C-terminal part of helix 4 for which a very shallow surface pocket formed by the arginine and lysine side chains can be delineated on several of the lowest energy Neh1 NMR solution structures (Figure 3B, indicated by a white dotted circle). The side chain amide of N482 showed CSP values of approximately 0.3 ppm, for R499, D500, and R503 values of 0.30, 0.23, and 0.14 ppm were observed upon addition of compound 2 (Figure 4A, B).

We sought to obtain the complex structure of compound 2 bound to Neh1-ΔLZIP, to determine key interactions that are responsible for ligand binding. Due to the weak affinity and surface location of the binding site no definite unambiguous NOE could be assigned, therefore, CSP values obtained from the $^1$H-$^{15}$N HSQC titration experiments, were converted into ambiguous restraints and used in a data-driven docking approach using HADDOCK.[19] Residues with CSP values equal or

larger than two standard deviations of the mean value were defined as residues actively involved in ligand binding, hence so-called ambiguous interaction restraints (AIRs) were generated for residues N482, R499, D500, and R503 to all polar atoms of the ligand.[19] HADDOCK docking calculations generated three clusters of binding poses with a similar HADDOCK score, in the three clusters compound 2 is oriented with its carboxylic acid toward helix 3 and the aryl rings toward helix 4 (Figure 5 and Figure S5). The orientation of compound 2 in the first and second cluster is rather perpendicular to helix 3/4 (Figure 5A, B and Figure S5A, B), whereas in cluster three the ligand is in an approximately 45 degree angle to the two helices (Figure 5C, D and Figure S5C, D). Compound 2 in cluster two is translated 1.5–2 Å away from helix 3 in comparison to the ligand binding pose in cluster one (Figure 5A, B, D). To determine if one of the three conformations is in better agreement with the experimental data, we applied post-docking filtering by comparing back-calculated ligand induced chemical shift changes of the backbone amide protons from our HADDOCK generated models to the experimental CSPs. This reintroduces the CSPs sign information which is otherwise lost in the HADDOCK calculations, as in HADDOCK CSPs are transformed into upper limit distance constraints,[19,22] and it was first shown by McCoy and Wyss[23] and later by others[22a,24] that this information can be crucial to identify the correct binding mode. In our case the sign of the CSPs constrain the relative orientation of the affected protein backbone amide protons and the two aromatic rings in the ligand, as it can be safely assumed due to the weak binding affinity that no or little protein backbone rearrangements occur upon ligand binding, thus the major contribution to the CSPs is caused by the direct ligand binding event. Hence, for the back-calculation in a first approximation the sign of a proton chemical shift perturbation is a function of the orientation of the ligand aromatic rings, and it has been shown
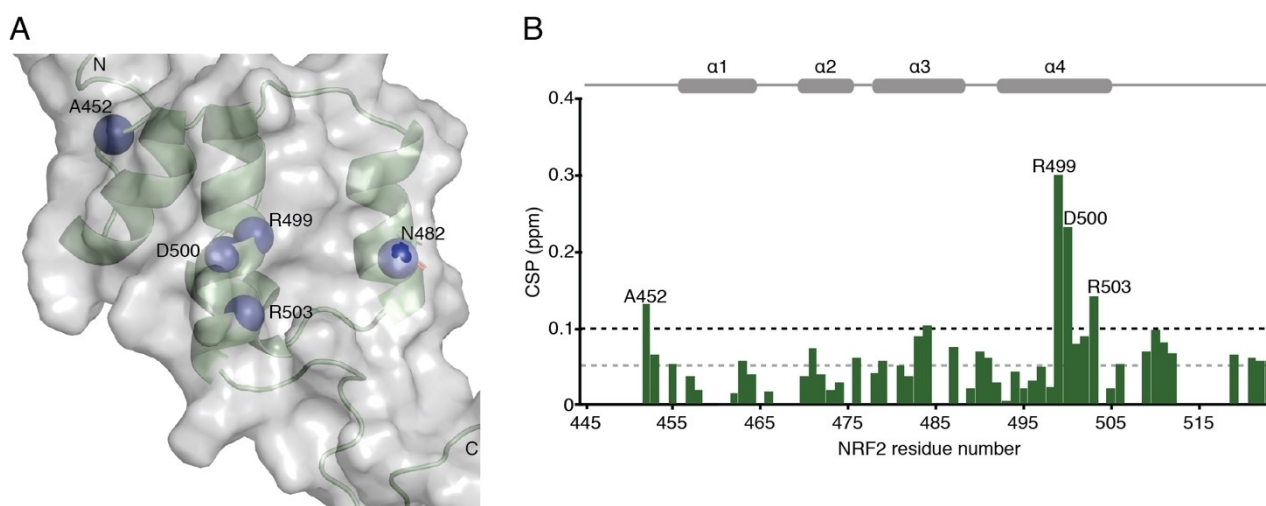


**Figure 4.** NMR determination of the fragment hits binding site. (A) Ribbon and surface representation of Neh1-ΔLZIP (PDB ID 7O7B) with nitrogen atoms shown as blue spheres for residues that showed CSP values larger than 2 standard deviations of the mean value of all CSPs in the NMR titration experiment. (B) Histogram showing CSP values for backbone amide groups of $^{15}$N labelled Neh1-ΔLZIP in the presence of 2 mM compound 2. Horizontal dotted grey and black lines indicate one and two standard deviations of the CSP values, respectively. Residues with backbone amide CSP values greater than two standard deviations are indicated with labels. Residues with without a bar were not assigned.

ChemMedChem

Full Papers
doi.org/10.1002/cmdc.202100458

Chemistry
Europe
European Chemical
Societies Publishing



**Figure 5.** Data-driven docking structures of fragment **2**; ribbon representation of NRF2 (residues R456-A510 are shown with the remaining residues omitted for clarity, PDB ID 7O7B) complexed to compound **2**. (A) Shown in sticks representation the binding pose of **2** with lowest Q-score in cluster 1. (B) Shown in sticks representation the binding pose of **2** with lowest Q-score in cluster 2. (C) Shown in sticks representation the binding pose of **2** with lowest Q-score in cluster 3. Black dotted lines indicate distances in Angstrom. (D) Superimposition of these docking poses in the binding site.

that they are to a lesser extent a function of the relative orientation between the protein proton and charged ligand groups.[22b,25]

In this study we simulated the sign and magnitude of backbone amide $^1H^N$ chemical shift changes, which were caused by the ring current effect induced by the two aromatic rings, with the empirical point-dipole model derived by Pople.[26] The electric field effect generated by electric charges on the carboxylic acid of the ligand was simulated with the empirical equation introduced by Buckingham.[27] The three clusters were assessed with the Q-score used by McCoy and Wyss[23] to quantify the difference between the experimental and calculated chemical shifts – with smaller Q-scores indicative of better agreement. This analysis showed that several ligand conformations can equally well reproduce the observed experimental chemical shift changes. Binding poses in cluster 1 and 2 had an average (five structures) Q-score for the ring current effect of $4.6 \pm 1.1$ and $4.9 \pm 0.8$, respectively, whereas the Q-score was $6.1 \pm 1.5$ for cluster 3. These values slightly increased when the electric field effect was additionally taken into account to $5.2 \pm 0.6$, $5.4 \pm 0.8$, and $8.1 \pm 0.7$. Furthermore, the predicted average chemical shift change for the side chain amide nitrogen N$\delta$2 of N482, assumed to be solely caused by the electric field effect, was $2.54 \pm 0.18$, $1.04 \pm 0.13$, and $1.12 \pm 0.26$ ppm respectively, with a measured value of 0.64 ppm at approximately 50–60% protein fraction bound. Thus, the pattern of calculated chemical

shift perturbations agree reasonably well with the experimentally determined values (Figure S6), with the discrepancy potentially coming from not including the bromine and the ether o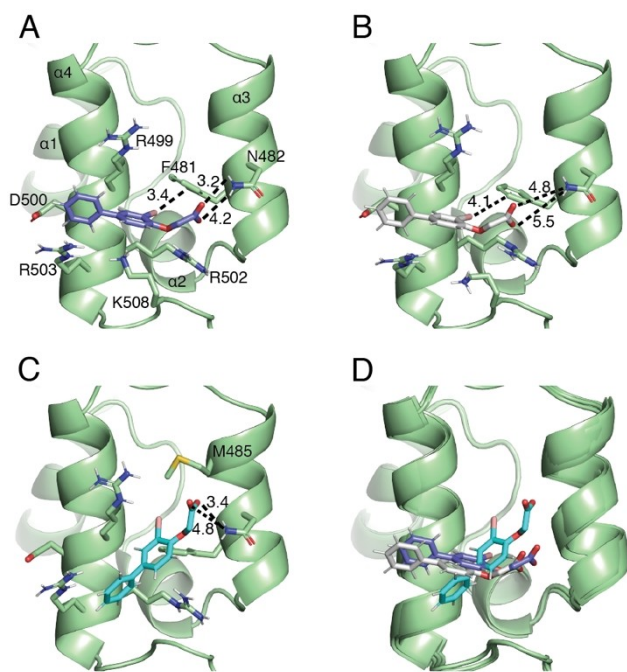xygen in the calculations, which were omitted from the calculation as the partial charges for these atoms are not readily described. Overall, the binding pose of compound **2** in cluster two showed the best agreement with the experimental CSP data when taking the chemical shift changes of the N$\delta$2 atom of N482 into consideration. However, due to the similar Q-values of the three binding poses a dynamic binding mode of the fragments cannot be ruled out, a phenomenon frequently seen for fragments even in less shallow binding sites.[28]

Based on the docking structures the structure activity relationship observed for the majority of screening hits can be rationalized. In cluster 1 and 2 the second aromatic ring is positioned above R503 allowing for a cation-π interaction between the protein and the ligand (Figure 5A, B and Figure S5A, B). This interaction explains how the absence of the second aromatic ring or an electron poor (π-deficient) heterocycle such as pyrimidine decreases affinity to the Neh1 domain of NRF2 (Table 1 and Figure S2F). The halogen in ortho position on the first phenyl ring, was shown by the SARs series to be crucial for ligand binding and is positioned in a way that it can favorably interact with F481, where presumably the positively charged sigma hole interacts with the partial negative charge on the aromatic ring plane forming a putative halogen bond.[29] This rationalizes the approximately 2-fold increase in affinity when chlorine is substituted with bromine at ring position three as the strength of interaction increases with the size of the halogen (Table 1 and Figure 2C).[29b,30] The carboxylic acid is in an orientation that allows for the acceptance of hydrogen bonds from the amide side chain of N482 and for a charge-charge interaction with the guanidinium group of R502 (Figure 5A, B). The ether oxygen could potentially function as a hydrogen bond acceptor donated from R502 and/or K508 and can additionally decrease the distance between the carboxyl group and N482, explaining the significant increase in chemical shift changes and binding affinity of compound **3** compared to its matched pair **1**, which lacks this oxygen atom (Table 1 and Figure S2A, B). Furthermore, the docking structures rationalize how the introduction of a phenolic hydroxy group, compound **5**, can lead to a 2-fold increase in affinity by functioning as a hydrogen donor to the backbone carbonyl of R499 compared to its nonphenolic matched molecular pair **3**.

In cluster 3 the carboxylic acid of compound **2** engages similarly in a hydrogen bonding network with N482. However, in this binding pose the bromine is positioned between F481 and M485 where it can participate in a hydrophobic core. The two aryl rings are potentially in an orientation which allows them to interact in a cation-π type interaction with the arginine residues of the basic region and a π-π interaction with F481 (Figure 5C).

From the binding models it becomes clear that the fragments bind at the putative Neh1 ARE DNA binding site (Figure S7). To further corroborate this finding we performed ARE DNA binding titration experiments.

ChemMedChem

Full Papers
doi.org/10.1002/cmdc.202100458

Chemistry
Europe
European Chemical
Societies Publishing

**DNA binding to Neh1**. In order to determine if Neh1 as a monomer possesses affinity for the ARE consensus sequence we performed a NMR titration experiment of unlabelled ARE (19bp) to $^{15}N$ labelled Neh1-ΔLZIP (Figure 6A). Indeed, Neh1-ΔLZIP showed specific chemical shift changes in the α-helical

and unstructured part of the basic region, several unstructured residues starting from G505 showed strong line broadening at higher DNA concentrations. In detail, the peaks of V479, D500, I501, R502, and R503 show CSPs larger than one standard deviation than the mean of all CSPs at a 4-fold excess of DNA in agreement with specific DNA binding in the basic region (Figure 6B). Residues G505, K506, and V509 as well as basic residues R512 and K518, which based on sequence alignment to Skn-1 are interacting with the phosphate groups of the DNA,[31] are broadened beyond detection at a 1:2 protein DNA ratio (Figure 6A). These findings are indicative of conformational rearrangement of the basic region upon DNA binding.[32] From the CSP data the dissociation constant of monomeric Neh1-ΔLZIP to ARE was determined to be in the higher micro-molar range, approximately $200 \pm 70 \ \mu M$ (average over R499, D500, R503), assuming a 1:1 stoichiometry (Figure 6C). In an earlier surface plasmon resonance (SPR) experiment by Yamamoto T *et al.* no binding could be observed to monomeric Neh1, whereas for the Neh1-sMafG heterodimer a $K_D$ of 20 nM was reported for the ARE consensus sequence (TGCTGACTCAGCA).[31] However, our measurements indicate that NRF2 can already preform a protein-DNA complex which could facilitate formation of the NRF2-sMaf heterodimer on the DNA.

Furthermore, comparison of the chemical shift perturbations caused by the ARE element and the fragment hits revealed an overlap in binding regions in the structured part of the basic region, which agrees with the superimposition of the Neh1-ΔLZIP:**2** docking structure and the Skn-1:ARE complex X-ray structure (Figure S7). To confirm this finding, we performed a competition binding experiment. In order to show specific replacement of **2** by ARE DNA, $^{15}N$ labelled Neh1-ΔLZIP in the presence of **2** was titrated with unlabeled ARE DNA. This $^1H$-$^{15}N$ HSQC titration experiment resulted in a final spectrum which is very close to the spectrum of Neh1-ΔLZIP bound to ARE (Figure S8), indicating that tighter binding ligands are expected to modulate DNA binding.
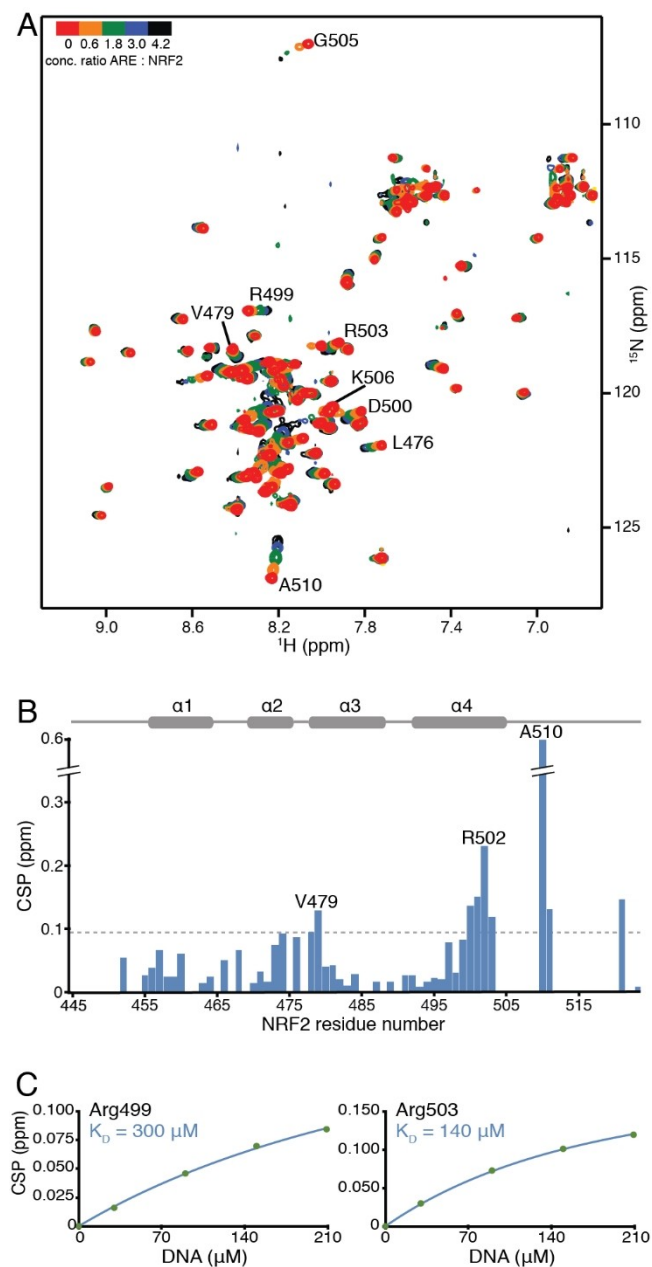


**Figure 6.** ARE DNA binding to NRF2 (residues 445–523). (A) Overlay of $^1H$-$^{15}N$ HSQC spectra of uniformly labeled 50 μM $^{15}N$-Neh1-ΔLZIP (red) titrated with 30, 90, 150, and 210 μM consensus ARE DNA (from orange to black). (B) Histogram showing chemical shift perturbation values for backbone amide groups of $^{15}N$-Neh1-ΔLZIP in the presence of 210 μM ARE DNA; the largest CSP value is 0.61 ppm (A510) and is truncated in the plot. Horizontal dotted grey line indicates the standard deviation (0.097 ppm) of the mean value of the CSP values. (C) Titration curves obtained from the $^1H$-$^{15}N$ HSQC titration measurements for residues R499, and R503. $K_D$s were derived by fitting experimental CSP values to equation (1). The lines represent an individual residue fit to equation (1).

## Conclusion

The role of NRF2 as a master regulator of cellular response to oxidative and electrophilic stresses bestows upon it a cancer suppressive function. This protective role is reversed in several types of cancer for which constitutive NRF2 activation has been found to cause chemo-resistance and tumor progression.[33] Therefore, several previous studies sought to find small molecule inhibitors of NRF2 to improve treatment of NRF2 maintained tumors. However, only a handful of small molecule NRF2 inhibitors, such as Brusatol,[34] Luteolin,[35] AEM1,[36] and ML385,[8] have been described in the literature to date. Furthermore, besides ML385, none of these compounds' putative mechanisms of action involve direct binding to NRF2.[1,4b,33] ML385 has been proposed to bind to the leucine zipper region of NRF2, however, no structural characterization of the interaction has been undertaken.[8] The leucine zipper motif is a common feature for dimer formation in many

**ChemMedChem**

Full Papers
doi.org/10.1002/cmdc.202100458

**Chemistry Europe**
European Chemical
Societies Publishing

transcription factors, hence, binding to this region could potentially cause inhibition of a large number of transcription factors.[1,4b] In this study we explored alternative binding sites on NRF2 by a fragment-based NMR screening approach, which led to the discovery of several small-molecule hits with a biphenyl phenoxy-acetic acid scaffold, which have binding affinities in the milli-molar range. In a 'SAR by catalogue' approach based on a pharmacophore similar to **2**, we searched for existing compounds (internal and commercial), and novel make-on-demand compounds from Enamine. In this way we exhaustively searched the commercially available chemical space for NRF2 binders in order to discover hits to the Neh1 domain. This resulted in compounds which allowed us to determine the initial SAR of our first fragment hits. The SAR study deciphered subtle differences in binding affinities due to ring substituents on the two phenyl rings; for example we determined that a 2′-hydroxy substituent in compound **5** decreased the dissociation constant 2-fold to its respective matched pair compound **3**. Additionally, we could show that substituting the second phenyl ring with an electron poor (π-deficient) pyrimidine abrogated ligand binding. The SAR data further established that the 3-halo group is crucial for anchoring the compounds in their binding site. Furthermore, the NMR data allowed us to elucidate the small-molecule binding site, which was composed of the CNC and basic region of NRF2; a region that is specific to transcription factors such as p45 NF—E2, and BACH proteins, as well as Maf transcription factors where the extended homology region has a similar fold. This likely decreases the probability of off-target binding to other transcription factor families. In a data-driven docking run we established three possible binding modes of compound **2**. These binding poses agree well with our experimentally measured chemical shift perturbations and our SAR data, with simulated CSPs for cluster 2 having the best agreement with the experimental values. In more detail, cluster 1 and 2 can equally well describe chemical shift changes which occur due to the ligand ring current effect on helix 4 with simulated values from cluster 3 being somewhat worse. However, the simulated CSP for the Nδ2 atom of N482 caused by the electric field effect of the ligand carboxyl group correlates best with the orientation of the acid in cluster 2 and 3, with predicted values in cluster 1 being larger than experimentally observed. Although the possibility of dynamic sampling of all three binding poses cannot be ruled out either, a feature rather common for low affinity binders. The binding conformations provided the basis for rationalizing the contribution of different ligand features to compound binding. Based on sequence alignment to Skn-1 and Maf proteins, it is apparent that this small and very shallow positively charged binding site is potentially engaged in DNA binding as well. We experimentally corroborated, by NMR ARE DNA titration measurements, that indeed the hit binding region coincides with the ARE promoter binding region on NRF2, therefore in theory high affinity binders could modulate NRF2 promoter binding.

This first high-resolution characterization of a NRF2 binder could serve as a starting point in a hit-to-lead campaign for optimizing binding affinity and selectivity for NRF2 protein-DNA inhibitors. The protein–ligand structure obtained by data-driven docking has potential for linker length optimization between the biphenyl scaffold and the carboxylic acid as well as its shape complementary to the binding pocket. Additionally, it could be envisioned that an optimal combination of phenyl ring side chains could lead to higher ligand potency. Optimized ligands could directly modulate NRF2 ARE promoter binding, and thereby modulate the transcription of NRF2 target genes. Mitigating the effects of constitutive NRF2 activation in several cancer types and help increase effectiveness of cancer treatment.

## Experimental Section

**Protein expression and purification**. Uniformly $^{15}$N and $^{13}$C, $^{15}$N labeled samples of Neh1 (residues 445–523, denoted by Neh1-ΔLZIP) of human NRF2 (isoform 1) were overexpressed in *Escherichia coli* BL21 (DE3) as a TEV cleavable N-terminal fusion to MBP, and additionally contained a C-terminal TEV cleavable hexahistidine tag. Cells were grown at 37 °C in M9 minimal media containing $^{15}$NH$_4$Cl and $^{13}$C$_6$-glucose as sole nitrogen and carbon sources in presence of ampicillin until OD$_{600}$≈0.6, then the temperature was lowered to 18 °C, and after 45 min protein synthesis was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM. Expression was carried out overnight.

Cells were lysed by sonication in lysis buffer containing 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 1 mM dithiothreitol (DTT) and afterward clarified by centrifugation. Subsequently, the supernatant was loaded onto a MBPTrap HP (cytiva) column. In the next step, MBP-Neh1 fusion protein was loaded onto a HisTrap FF crude (cytiva) column and after elution the MBP – and His – tag was cleaved by incubation with TEV. MBP and the His-tag were removed by running a Resource S column (cytiva). In a final purification step, Neh1 was purified to homogeneity by size exclusion chromatography using a Superdex 75 (cytiva).

**NMR spectroscopy**. NMR spectra were recorded at 25 °C on a Bruker AVANCE NEO 600 MHz (14.1 T) spectrometer and an Avance III HD 800 MHz (18.8 T) spectrometer. Data were processed using NMRPipe[37] and analyzed using CcpNmr.[38] Protein assignment and NOE experiments were carried out on a 0.5 mM $^{15}$N, $^{13}$C, uniformly labeled Neh1-ΔLZIP solution in 50 mM sodium phosphate buffer, pH 6.5, 150 mM NaCl, 1 mM DTT in 5 % D$_2$O/95 % H$_2$O. Backbone and side chain $^{1}$H, $^{13}$C, and $^{15}$N chemical shift assignment of Neh1-ΔLZIP was obtained with standard triple-resonance experiments: HNCA/HN(CO)CA, HNCO/HN(CA)CO, CBCA(CO)NH/HNCACB, (H)CCONH-TOCSY, and HCCH-TOCSY. Intramolecular distance restraints were obtained from 3D $^{1}$H−$^{1}$H-NOESY−$^{15}$N/$^{13}$C-HSQC (mixing time 120 ms) experiments.

Small-molecule titration experiments were carried out on a 50 μM $^{15}$N uniformly labeled Neh1-ΔLZIP solution in 50 mM sodium phosphate buffer, pH 6.5, 150 mM NaCl, 1 mM DTT in 10 % D$_2$O/95 % H$_2$O with 50 mM ligand stock solutions in dimethyl sulfoxid-d6 (DMSO-d6). 1D NMR spectroscopy was used to determine concentrations. For all titration experiments the same protein stock solution was used, the ligands were titrated to final concentrations of 0.250, 0.750, 1.5, and 2.0 mM. At each titration step chemical shift changes upon ligand addition were monitored with $^{1}$H-$^{15}$N HSQC experiments. The combined chemical shift perturbation of $^{1}$H$^{N}$ and $^{15}$N$^{H}$ was calculated as $\Delta\sigma = \sqrt{(\Delta 1H)^2 + 0.20(\Delta 15N)^2}$.

**ChemMedChem**

Full Papers
doi.org/10.1002/cmdc.202100458

**Chemistry
Europe**
European Chemical
Societies Publishing

Chemical shift perturbations at each titration point were used to calculate the dissociation constants ($K_D$) by nonlinear least square fit to equation 1.

$$\Delta\sigma = \Delta\sigma_{max} \frac{(K_D + [P] + [L]) - \sqrt{(K_D + [P] + [L])^2 - 4[P][L]}}{2[P]} \quad (1)$$

where $\Delta\sigma$ is the observed chemical shift perturbation, $\Delta\sigma_{max}$ the maximum chemical shift perturbation and $[P]$ and $[L]$ the protein and ligand concentrations, respectively.

For DNA titration measurements double stranded ARE DNA was obtained by mixing complementary oligonucleotides (5'-CCGGTGCTGAGTCAGCAGG and 5'-CCCTGCTGACTCAGCACCG) which were denatured at 95 °C and slowly annealed at 45 °C. The titration experiments were conducted with a 5 mM ARE DNA stock solution in 50 mM sodium phosphate buffer, pH 6.5, 150 mM NaCl, 1 mM DTT in 5% $D_2O$/95% $H_2O$. DNA was added to final concentrations of 0.03, 0.09, 0.150, and 0.210 mM to a 50 μM $^{15}N$ uniformly labeled Neh1-ΔLZIP solution. The $K_D$ was obtained by fitting the CSP data to equation 1.

The competition experiments were conducted with a 5 mM ARE DNA stock solution in 50 mM sodium phosphate buffer, pH 6.5, 150 mM NaCl, 1 mM DTT in 5% $D_2O$/95% $H_2O$. DNA was added to final concentrations of 0.10, 0.20, and 0.30 mM to a 50 μM $^{15}N$ uniformly labeled Neh1-ΔLZIP solution in the presence of 0.750 mM compound **2**.

**Structure calculation and refinement**. Backbone dihedral angle restraints were set to $\phi = -60$ ($\pm20$)° and $\psi = -45$ ($\pm20$)° for residues that were predicted to be α-helical based on $^{13}C\alpha$, $^{13}C\beta$, $^{13}C'$, $^{1}H\alpha$, $^{15}N$, and $^{1}H^N$ chemical shifts using the software TALOS–N.[39] For residues with a TALOS–N classification of good, outside of the α-helices, the respective predicted dihedral angles were used in the structure calculation. Intramolecular distance restraints were obtained from 3D $^{1}H$-$^{1}H$-NOESY-$^{15}N$/$^{13}C$-HSQC experiments. Ambiguity of NOE assignments were partially resolved, and distance restraints were calibrated, with the program Aria2.3.[40] The experimentally determined distance, and TALOS–N predicted dihedral restraints were used in a torsion angle simulated annealing protocol using CNS1.2/Aria2.3[40–41] to solve the solution structure of the Neh1 (residues 445–523) domain of NRF2. The final NMR ensemble was refined in an explicit water shell.[42] The 20 lowest-energy solution structures (out of 100 calculated) were selected as a final representative ensemble of Neh1-ΔLZIP (PDB ID 7O7B).

**Data-driven docking protocol**. The docking conformation of compound **2** was generated by minimizing the MMFF94[43] energy of the free molecule in LigandScout (version 4.4)[44] using default parameters. CNS parameter and topology files for compound **2** were generated using PRODRG[45] and the dihedral angles in the parameter file were modified manually to reflect the energy minimized ligand conformation. The lowest energy structure of Neh1-ΔLZIP was used as protein starting structure for the docking run. Ambiguous Interaction Restraints (AIRs) were generated for residues N482, R499, D500 and R503, as they showed CSP values equal or larger than two standard deviations of the mean value, upon compound **2** addition. Compound **2** was docked using the software HADDOCK (version 2.2)/CNS (version 1.3),[19,41,46] with recommended settings for small-molecule docking. Briefly, in a rigid body docking step 1000 structures were calculated, subsequently the 200 structures with the lowest HADDOCK scores were used in a final docking step with a fully flexible binding interface and flexible ligand. For each of the three HADDOCK clusters, with similar HADDOCK scores, five representative conformers with the

lowest energies and the least experimental restraint violations were picked for chemical shift perturbations back-calculation.

**CSP calculation**. There are two major through space contributions to protein $^{1}H^N$ chemical shift changes upon ligand binding, namely, the ring current (RC) and the electric filed (EF) effect, which are caused by aromatic rings and electric charges on the ligand.[47] The RC effect of the two aromatic rings of **2** was calculated by the empirical point-dipole model derived by Pople:[26,48]

$$\Delta\sigma_{RC} = 10^6 \frac{ne^2 a^2}{4\pi m c^2} \frac{3\cos^2\theta - 1}{r^3}$$

where $\Delta\sigma_{RC}$ is the change in the isotropic nuclear shielding constant, $n$ is the number of circulating aromatic ring electrons, $e$ is the elementary charge, $a$ is the radius of the aromatic ring, $m$ is the electron mass, $c$ is the speed of light, $\theta$ is the angle between the ring normal and the proton to ring center vector, and $r$ is the distance from the proton to the ring center.

The EF effect of the carboxylic acid point charges on **2**, which can introduce CSP at spatially close atoms, was calculated with an empirical equation derived by Buckingham:[27,49]

$$\Delta\sigma_{EF} = 10^6 \frac{1}{4\pi\varepsilon_0\varepsilon_r} A_{\parallel} \frac{q\cos\theta}{r^2}$$

Here, $r$ is the distance between the ligand atom with the partial charge $q$ and the protein $^{1}H^N$ or Nδ2 atom, the angle $\theta$ is between the distance vector $r$ and the vector of the $H^N$-$N^H$ amide bond or the side chain Cγ-Nδ2 bond. The partial charges $q$ for the carboxyl group were taken from the Amber force field.[50] The magnitude of the CSP due to the EF effect is proportional to the polarizability of the $H^N$-$N^H$ or Cγ-Nδ2 bond and is given by the nuclear shielding polarizability $A_{\parallel}$, in this study we used the values determined by Jens J. Led and co-workers[49] of 188 ppm au for the $^{1}H^N$ atom and 868 ppm au for the $^{15}N^H$ atom. For the dielectric constant $\varepsilon_r$ the value 6 was used in our calculations.[51] The CSP values were back-calculated for helices 3 and 4 as these are, based on the $^{1}H$-$^{15}N$ HSQC titration experiments, directly involved in ligand binding. The quality of agreement between experimentally measured and simulated CSPs was quantified with the Q-score suggested by McCoy and Wyss:[23]

$$Q_{score} = \sum_i \left( \frac{\Delta\sigma_{exp}^j}{\left|\Delta\sigma_{exp}^{max}\right|} - \frac{\Delta\sigma_{calc}^j}{\left|\Delta\sigma_{calc}^{max}\right|} \right)^2$$

where $\Delta\sigma_{exp}^j$ and $\Delta\sigma_{calc}^j$ are the experimental and calculated CSP value for the amide backbone proton i, respectively. $\left|\Delta\sigma_{exp}^{max}\right|$ and $\left|\Delta\sigma_{calc}^{max}\right|$ are the largest experimentally observed and simulated CSP value.

## Accession codes

The NMR structure ensemble determined in this article has been deposited in the Protein Data Bank under accession ID 7O7B. The chemical shift assignments have been deposited in the Biological Magnetic Resonance Bank (BMRB) under acces-

sion code 34617. The authors will release the atomic coordinates and experimental data upon article publication.

## Acknowledgements

## Conflict of Interest

The authors declare the following competing financial interest (s): Julian E. Fuchs, Gerd Bader, Darryl B. McConnell, and Moriz Mayer were fulltime employees for Boehringer Ingelheim at the time the data presented herein were collected.

**Keywords:** NRF2 · protein-DNA interactions · fragment screening · structure-activity relationships · ligand docking · NMR solution structures

[1] M. Dodson, M. R. de la Vega, A. B. Cholanians, C. J. Schmidlin, E. Chapman, D. D. Zhang, *Annu. Rev. Pharmacol. Toxicol.* **2019**, *59*, 555–575.

[2] a) N. Cancer Genome Atlas Research, *Nature* **2012**, *489*, 519–525; b) E. Panieri, L. Saso, *Oxid Med Cell Longev* **2019**, *2019*, 8592348.

[3] J. D. Campbell, A. Alexandrov, J. Kim, J. Wala, A. H. Berger, C. S. Pedamallu, S. A. Shukla, G. Guo, A. N. Brooks, B. A. Murray, M. Imielinski, X. Hu, S. Ling, R. Akbani, M. Rosenberg, C. Cibulskis, A. Ramachandran, E. A. Collisson, D. J. Kwiatkowski, M. S. Lawrence, J. N. Weinstein, R. G. Verhaak, C. J. Wu, P. S. Hammerman, A. D. Cherniack, G. Getz, N. Cancer Genome Atlas Research, M. N. Artyomov, R. Schreiber, R. Govindan, M. Meyerson, *Nat. Genet.* **2016**, *48*, 607–616.

[4] a) E. W. Cloer, D. Goldfarb, T. P. Schrank, B. E. Weissman, M. B. Major, *Cancer Res.* **2019**, *79*, 889–898; b) J. J. Qin, X. D. Cheng, J. Zhang, W. D. Zhang, *Cell Commun. Signaling* **2019**, *17*, 121.

[5] a) B. Padmanabhan, K. I. Tong, T. Ohta, Y. Nakamura, M. Scharlock, M. Ohtsuji, M. I. Kang, A. Kobayashi, S. Yokoyama, M. Yamamoto, *Mol. Cell* **2006**, *21*, 689–700; b) A. Singh, V. Misra, R. K. Thimmulappa, H. Lee, S. Ames, M. O. Hoque, J. G. Herman, S. B. Baylin, D. Sidransky, E. Gabrielson, M. V. Brock, S. Biswal, *PLoS Med.* **2006**, *3*, e420; c) V. D. Martinez, E. A. Vucic, L. A. Pikor, K. L. Thu, R. Hubaux, W. L. Lam, *Mol. Cancer* **2013**, *12*, 124.

[6] X. J. Wang, Z. Sun, N. F. Villeneuve, S. Zhang, F. Zhao, Y. Li, W. Chen, X. Yi, W. Zheng, G. T. Wondrak, P. K. Wong, D. D. Zhang, *Carcinogenesis* **2008**, *29*, 1235–1243.

[7] P. Moi, K. Chan, I. Asunis, A. Cao, Y. W. Kan, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 9926–9930.

[8] A. Singh, S. Venkannagari, K. H. Oh, Y. Q. Zhang, J. M. Rohde, L. Liu, S. Nimmagadda, K. Sudini, K. R. Brimacombe, S. Gajghate, J. Ma, A. Wang, X. Xu, S. A. Shahane, M. Xia, J. Woo, G. A. Mensah, Z. Wang, M. Ferrer, E. Gabrielson, Z. Li, F. Rastinejad, M. Shen, M. B. Boxer, S. Biswal, *ACS Chem. Biol.* **2016**, *11*, 3214–3225.

[9] M. R. Arkin, J. A. Wells, *Nat. Rev. Drug Discovery* **2004**, *3*, 301–317.

[10] J. E. Darnell, *Nature Reviews Cancer* **2002**, *2*, 740–749.

[11] a) M. R. Arkin, Y. Tang, J. A. Wells, *Chem. Biol.* **2014**, *21*, 1102–1114; b) A. S. Bhagwat, C. R. Vakoc, *Trends Cancer* **2015**, *1*, 53–65; c) A. Chen, A. N. Koehler, *Trends Mol. Med.* **2020**, *26*, 508–518.

[12] J. H. Bushweller, *Nat. Rev. Cancer* **2019**, *19*, 611–624.

[13] a) W. Huang, Z. Dong, F. Wang, H. Peng, J. Y. Liu, J. T. Zhang, *ACS Chem. Biol.* **2014**, *9*, 1188–1196; b) H. Li, F. Ban, K. Dalal, E. Leblanc, K. Frewin,

[14] a) M. A. Kuenemann, L. M. Bourbon, C. M. Labbe, B. O. Villoutreix, O. Sperandio, *J. Chem. Inf. Model.* **2014**, *54*, 3067–3079; b) L. Laraia, G. McKenzie, D. R. Spring, A. R. Venkitaraman, D. J. Huggins, *Chem. Biol.* **2015**, *22*, 689–703.

[15] a) J. Lyu, S. Wang, T. E. Balius, I. Singh, A. Levit, Y. S. Moroz, M. J. O'Meara, T. Che, E. Algaa, K. Tolmachova, A. A. Tolmachev, B. K. Shoichet, B. L. Roth, J. J. Irwin, *Nature* **2019**, *566*, 224–229; b) R. M. Stein, H. J. Kang, J. D. McCorvy, G. C. Glatfelter, A. J. Jones, T. Che, S. Slocum, X. P. Huang, O. Savych, Y. S. Moroz, B. Stauch, L. C. Johansson, V. Cherezov, T. Kenakin, J. J. Irwin, B. K. Shoichet, B. L. Roth, M. L. Dubocovich, *Nature* **2020**, *579*, 609–614.

[16] a) T. Maurer, L. S. Garrenton, A. Oh, K. Pitts, D. J. Anderson, N. J. Skelton, B. P. Fauber, B. Pan, S. Malek, D. Stokoe, M. J. Ludlam, K. K. Bowman, J. Wu, A. M. Giannetti, M. A. Starovasnik, I. Mellman, P. K. Jackson, J. Rudolph, W. Wang, G. Fang, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5299–5304; b) Q. Sun, J. P. Burke, J. Phan, M. C. Burns, E. T. Olejniczak, A. G. Waterson, T. Lee, O. W. Rossanese, S. W. Fesik, *Angew. Chem. Int. Ed.* **2012**, *51*, 6140–6143; *Angew. Chem.* **2012**, *124*, 6244–6247; c) D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, H. Jhoti, *Nat. Rev. Drug Discovery* **2016**, *15*, 605–619; d) D. Kessler, M. Gmachl, A. Mantoulidis, L. J. Martin, A. Zoephel, M. Mayer, A. Gollner, D. Covini, S. Fischer, T. Gerstberger, T. Gmaschitz, C. Goodwin, P. Greb, D. Haring, W. Hela, J. Hoffmann, J. Karolyi-Oezguer, P. Knesi, S. Kornigg, M. Koegl, R. Kousek, L. Lamarre, F. Moser, S. Munico-Martinez, C. Peinsipp, J. Phan, J. Rinnenthal, J. Q. Sai, C. Salamon, Y. Scherbantin, K. Schipany, R. Schnitzer, A. Schrenk, B. Sharps, G. Siszler, Q. Sun, A. Waterson, B. Wolkerstorfer, M. Zeeb, M. Pearson, S. W. Fesik, D. B. McConnell, *P Natl Acad Sci USA* **2019**, *116*, 15823–15829.

[17] a) T. Hoffmann, M. Gastreich, *Drug Discovery Today* **2019**, *24*, 1148–1156; b) W. P. Walters, *Journal of Medicinal Chemistry* **2019**, *62*, 1116–1124.

[18] a) P. J. Hajduk, J. R. Huth, S. W. Fesik, *J. Med. Chem.* **2005**, *48*, 2518–2525; b) F. N. Edfeldt, R. H. Folmer, A. L. Breeze, *Drug Discovery Today* **2011**, *16*, 284–287.

[19] C. Dominguez, R. Boelens, A. M. Bonvin, *J. Am. Chem. Soc.* **2003**, *125*, 1731–1737.

[20] a) M. C. Lo, S. Ha, I. Pelczer, S. Pal, S. Walker, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 8455–8460; b) P. B. Rupert, G. W. Daughdrill, B. Bowerman, B. W. Matthews, *Nat. Struct. Biol.* **1998**, *5*, 484–491.

[21] H. Kurokawa, H. Motohashi, S. Sueno, M. Kimura, H. Takagawa, Y. Kanno, M. Yamamoto, T. Tanaka, *Mol. Cell. Biol.* **2009**, *29*, 6232–6244.

[22] a) J. Stark, R. Powers, *J. Am. Chem. Soc.* **2008**, *130*, 535–545; b) T. Ten Brink, C. Aguirre, T. E. Exner, I. Krimm, *J. Chem. Inf. Model.* **2015**, *55*, 275–283.

[23] M. A. McCoy, D. F. Wyss, *J. Biomol. NMR* **2000**, *18*, 189–198.

[24] D. Gonzalez-Ruiz, H. Gohlke, *J. Chem. Inf. Model.* **2009**, *49*, 2260–2271.

[25] G. Cioffi, C. A. Hunter, M. J. Packer, A. Spitaleri, *J. Med. Chem.* **2008**, *51*, 2512–2517.

[26] J. A. Pople, *The Journal of Chemical Physics* **1956**, *24*, 1111–1112.

[27] A. D. Buckingham, *Canadian Journal of Chemistry* **1960**, *38*, 300–307.

[28] a) G. C. P. van Zundert, B. M. Hudson, S. H. P. de Oliveira, D. A. Keedy, R. Fonseca, A. Heliou, P. Suresh, K. Borrelli, T. Day, J. S. Fraser, H. van den Bedem, *J. Med. Chem.* **2018**, *61*, 11183–11198; b) M. L. Verteramo, O. Stenstrom, M. M. Ignjatovic, O. Caldararu, M. A. Olsson, F. Manzoni, H. Leffler, E. Oksanen, D. T. Logan, U. J. Nilsson, U. Ryde, M. Akke, *J. Am. Chem. Soc.* **2019**, *141*, 2012–2026.

[29] a) T. Clark, M. Hennemann, J. S. Murray, P. Politzer, *J. Mol. Model.* **2007**, *13*, 291–296; b) S. Sirimulla, J. B. Bailey, R. Vegesna, M. Narayan, *J. Chem. Inf. Model.* **2013**, *53*, 2781–2791.

[30] M. R. Scholfield, C. M. Zanden, M. Carter, P. S. Ho, *Protein Sci.* **2013**, *22*, 139–152.

[31] T. Yamamoto, M. Kyo, T. Kamiya, T. Tanaka, J. D. Engel, H. Motohashi, M. Yamamoto, *Genes Cells* **2006**, *11*, 575–591.

[32] H. J. Dyson, P. E. Wright, *Curr. Opin. Struct. Biol.* **2002**, *12*, 54–60.

[33] E. J. Moon, A. Giaccia, *Free Radical Biol. Med.* **2015**, *79*, 292–299.

[34] D. Ren, N. F. Villeneuve, T. Jiang, T. Wu, A. Lau, H. A. Toppin, D. D. Zhang, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1433–1438.

[35] X. Tang, H. Wang, L. Fan, X. Wu, A. Xin, H. Ren, X. J. Wang, *Free Radical Biol. Med.* **2011**, *50*, 1599–1609.

[36] M. J. Bollong, H. Yun, L. Sherwood, A. K. Woods, L. L. Lairson, P. G. Schultz, *ACS Chem. Biol.* **2015**, *10*, 2193–2198.

[37] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, *6*, 277–293.

**ChemMedChem**

Full Papers
doi.org/10.1002/cmdc.202100458

**Chemistry Europe**
European Chemical
Societies Publishing

[38] S. P. Skinner, R. H. Fogh, W. Boucher, T. J. Ragan, L. G. Mureddu, G. W. Vuister, *J. Biomol. NMR* **2016**, *66*, 111–124.

[39] Y. Shen, A. Bax, *J. Biomol. NMR* **2013**, *56*, 227–241.

[40] W. Rieping, M. Habeck, B. Bardiaux, A. Bernard, T. E. Malliavin, M. Nilges, *Bioinformatics* **2007**, *23*, 381–382.

[41] A. T. Brunger, *Nat. Protoc.* **2007**, *2*, 2728–2733.

[42] J. P. Linge, M. A. Williams, C. A. Spronk, A. M. Bonvin, M. Nilges, *Proteins* **2003**, *50*, 496–506.

[43] T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 490–519.

[44] G. Wolber, T. Langer, *Journal of Chemical Information and Modeling* **2005**, *45*, 160–169.

[45] A. W. Schuttelkopf, D. M. F. van Aalten, *Acta Crystallogr. Sect. D* **2004**, *60*, 1355–1363.

[46] A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. Warren, *Acta Crystallogr. Sect. D* **1998**, *54*, 905–921.

[47] M. P. Williamson, *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.

[48] D. A. Case, *J. Biomol. NMR* **1995**, *6*, 341–346.

[49] M. A. S. Hass, M. R. Jensen, J. J. Led, *Proteins* **2008**, *72*, 333–343.

[50] D. A. Case, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W. D. Cruzeiro, T. A. Darden, D. G. Duke, M. K. Gilson, H. Gohlke, A. W. Goetz, D. Greene, R. Harris, N. Homeyer, Y. Huang, S. Izadi, A. K. T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, D. J. Luo, K. M. M. Mermelstein, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, R. Pan, D. R. R. Qi, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, J. Smith, R. Salomon Ferrer, J. Swails, R. C. Walker, J. Wang, H. Wei, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman, *University of California, San Francisco*, **2018**.

[51] R. F. Alford, A. Leaver-Fay, J. R. Jeliazkov, M. J. O'Meara, F. P. DiMaio, H. Park, M. V. Shapovalov, P. D. Renfrew, V. K. Mulligan, K. Kappel, J. W. Labonte, M. S. Pacella, R. Bonneau, P. Bradley, R. L. Dunbrack, R. Das, D. Baker, B. Kuhlman, T. Kortemme, J. J. Gray, *J. Chem. Theory Comput.* **2017**, *13*, 3031–3048.