Published in final edited form as:
Biochemistry. 2012 December 18; 51(50): 10035–10043. doi:10.1021/bi3011785.

Identification of a Hydrophobic Cleft in the LytTR Domain of AgrA as a Locus for Small Molecule Interactions that Inhibit DNA Binding

Paul G. Leonard[†,‡], **Ian F. Bezar**[‡,†,§], **David J. Sidote**[†,‡], and **Ann M. Stock**[†,‖,‡,*]

[†]Center for Advanced Biotechnology and Medicine, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, Piscataway, New Jersey 08854-5635

[‖]Department of Biochemistry and Molecular Biology, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, Piscataway, New Jersey 08854-5635

[§]Graduate School of Biomedical Sciences, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, Piscataway, New Jersey 08854-5635

[‡]University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, Piscataway, New Jersey 08854-5635

Abstract

The AgrA transcription factor regulates the quorum-sensing response in *Staphylococcus aureus*, controlling the production of hemolysins and other virulence factors. AgrA binds to DNA via its C-terminal LytTR domain, a domain not found in humans but common in many pathogenic bacteria, making it a potential target for antimicrobial development. We have determined the crystal structure of the apo AgrA LytTR domain and screened a library of 500 fragment compounds to find inhibitors of AgrA DNA-binding activity. Using NMR, the binding site for five compounds has been mapped to a common locus at the C-terminal end of the LytTR domain, a site known to be important for DNA-binding activity. Three of these compounds inhibit AgrA DNA binding. These results provide the first evidence that LytTR domains can be targeted by small organic compounds.

S. aureus typically causes skin or soft tissue infections at a localized lesion.[1] Patients are at risk of more serious life threatening diseases such as pneumonia, osteomyelitis, bacteremia, endocarditis and toxic shock syndrome[2–5] if *S. aureus* is able to cross into the bloodstream. The treatment of *S. aureus* infections has become increasingly problematic due to the emergence of antibiotic resistant strains. Methicillin resistant *S. aureus* (MRSA) is now the most common antibiotic resistant pathogen identified in developed world hospitals[6] and poses a significant threat to human health.

The pathogenicity of *S. aureus* requires the coordinated expression of a large number of virulence factors. The *agr* quorum-sensing system, which allows the pathogen to sense both

[*]Corresponding Author: Center for Advanced Biotechnology and Medicine, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, 679 Hoes Lane, Piscataway, NJ 08854-5627. Tel.: 732-235-4844; stock@cabm.rutgers.edu.
[§]I.F.B. is supported in part by a National Institutes of Health Graduate Training grant GM008360

The authors declare no competing financial interest.

Structural Data

The atomic coordinates and structure factors for apo $AgrA_C$ have been deposited in the Protein Data Bank. (PDB accession code: 4G4K) and NMR resonance assignments have been deposited into the Biological Magnetic Resonance Data Bank (BMRB accession number 18598).

the density of the local *S. aureus* population and its degree of confinement, plays a central role in the regulation of the *S. aureus* virulon. Activation of the *agr* system represses the production of cell surface adhesins and promotes the secretion of extracellular toxins.

The *agr* operon consists of four genes, designated *agrBDCA*. The propeptide, AgrD, is processed and exported by AgrB. The secreted cyclic peptide is the autoinducing peptide (AIP) that functions as the quorum-sensing signal. When the extracellular AIP concentration is high, the histidine kinase, AgrC, is activated, resulting in increased phosphorylation of the cytoplasmic response regulator protein, AgrA. Phosphorylated AgrA activates transcription of the *agrBDCA* operon and genes encoding phenol soluble modulins[7] and the effector RNA molecule, RNAIII.[8] In addition to its function as the messenger RNA for hemolysin δ,[9] RNAIII uses antisense RNA mechanisms to down-regulate adhesins and activate the transcription of genes encoding hemolysins, Panton-Valentine leukocidin (PVL) and enterotoxins.[10–13]

Deletion of the *agr* operon attenuates the *S. aureus* infection in mouse and rabbit animal models of infection demonstrating the importance of *agr* quorum sensing for pathogenesis.[14–19] Indeed, AIP analogues that inhibit the AgrC histidine kinase are effective in reducing the severity of *S. aureus* infection.[17,19] However, attempts to identify inhibitors of response regulator AgrA have not been reported, despite the attractiveness of AgrA as a target because of the absence of LytTR DNA-binding domain proteins in mammalian proteomes.[20]

We have determined the high-resolution crystal structure of the AgrA C-terminal LytTR domain ($AgrA_C$) and used a fragment screening approach to search for small molecule binding sites on the DNA-binding surface of this domain. Fragment-screening approaches have been widely used to efficiently screen a broad area of potential chemical space, using a relatively small library of compounds. Although the affinity of small fragment compounds for a target protein is expected to be relatively low, due to the low molecular weight of the compounds used (<300 $g \cdot mol^{-1}$), fragment screening can identify energetic focal points on a protein surface that can be targeted by small molecule compounds.[21]

EXPERIMENTAL PROCEDURES

Protein expression and purification

Unlabeled *S. aureus* $AgrA_C$ protein (AgrA residues Asp137 to Ile238) samples were produced in *E. coli* grown in terrific broth (TB) media as previously described.[22] For NMR studies, ^{15}N and $^{15}N/^{13}C$ isotopically-enriched protein samples of $AgrA_C$ were prepared by expressing the $AgrA_C$ protein in *E. coli* BL21 (DE3) pLysS grown in M9 minimal media at 18 °C using 2.5 g L^{-1} $(^{15}NH_4)_2SO_4$ and 2 g L^{-1} $[^{13}C]$-glucose (Cambridge Isotope Laboratories) as appropriate. The M9 media was supplemented with 50 μg mL^{-1} kanamycin, 50 μM $FeCl_2$, 2 μM $CuCl_2$, 2 μM Na_2MoO_4, 2 μM $NiCl_2$, 2 μM $CoCl_2$, 2 μM H_3BO_3, 10 μM $MnCl_2$, 10 μM $ZnSO_4$, 20 μM $CaCl_2$, and 1 μM of each of the following micronutrients: nicotininc acid, pyridoxine, thiamine, biotin, riboflavin, folic acid, D-pantothenic acid and myo-inositol. Recombinant protein expression was induced using 0.3 mM isopropyl-β-D-thiogalactopyranoside.

All $AgrA_C$ protein samples were purified from *E. coli* lysates following the previously published procedure[22] using HiTrap SP HP cation exchange, HiLoad Phenyl Sepharose HP hydrophobic interaction and HiLoad Superdex 75 gel filtration chromatograpy (GE Healthcare). Purified protein was transferred into appropriate buffers by dialysis.

Crystallization of $AgrA_C$

Initial $AgrA_C$ crystals were prepared at 4 °C by hanging drop vapor diffusion by mixing 1 μL of 1 mM $AgrA_C$ (dissolved in 20 mM Bis Tris, 100 mM NaCl and 10 mM DTT at pH 6.0) with 1 μL of reservoir solution (100 mM Tris, 150 mM $LiSO_4$ and 11% (w/v) PEG 4000 at pH 8.0). The hanging drop was suspended above a 1-mL reservoir. The crystals were improved by streak seeding after 24 h incubation at the same condition except the PEG 4000 concentration was lowered to 8% (w/v).

Data collection and structure refinement

$AgrA_C$ crystals were soaked for 30 s in 50 mM Tris, 75 mM $LiSO_4$, 8% PEG 4000 and 20% glycerol at pH 8 before flash freezing in liquid nitrogen. A native data set was collected at 100 K using a Rigaku MicroMax-007 HF generator equipped with RAXIS-IV++ detector. Data were processed and scaled with DENZO and SCALEPACK.[23] The structure of apo $AgrA_C$ was solved by molecular replacement using Phaser 2.0[24] with the DNA-bound state of $AgrA_C$ as the search model (PDB 3BS1). The *R* factor after rigid body refinement of the best molecular replacement solution was 0.440. From the initial molecular replacement solution, the structure was rebuilt from scratch using RESOLVE[25] before iterative refinement using COOT 6.02[26] and Phenix.[27] The refined model contains two molecules of $AgrA_C$ within the asymmetric unit; chains A and B contain residues Glu141 to Ile238 and Ser139 to Ile238, respectively. The model was refined to 1.52 Å with *R* factor and R_{free} values of 0.180 and 0.209 respectively. All residues lie within the allowed regions of the Ramachandran plot and exhibit favorable bond angles and bond lengths. An overview of the data collection and refinement statistics is provided in Table 1.

NMR spectroscopy

All NMR data were acquired on Varian INOVA 600 MHz or Varian 500 MHz spectrometers equipped with triple resonance probes. The NMR spectra were referenced to an internal 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) standard as described previously.[28] All spectra were processed using NMRpipe[29] and analyzed using CCPNMR analysis 2.1.[30]

For the backbone NMR assignments, ^1H-^{15}N HSQC[31], HNCACB, HN(CO)CACB, HNCA, HN(CO)CA and HNCO[32] spectra were acquired at 298 K from a sample containing 1 mM $AgrA_C$ in 20 mM sodium phosphate, 100 mM NaCl and 5% D_2O at pH 5.8. The backbone resonance assignments for $AgrA_C$ were established for all residues between Asn138 and Ile238 except for Tyr229.

^{15}N heteronuclear NOE[33] spectra were recorded with 3.0 s ^1H saturation in the latter part of a 3.5-s preparation period delay, which was also used without radio frequency pulses for the reference 2D spectrum without NOE.

A library of 500 fragment compounds (Maybridge Ro3 500 fragment library) was screened for binding to $AgrA_C$ using a WATERGATE W5 LOGSY NMR experiment.[34] Each sample contained a mixture of 10 compounds (400 μM of each compound) and 15 μM $AgrA_C$, dissolved in 20 mM sodium phosphate, 100 mM NaCl, 10 mM DTT and 5% D_2O at pH 7.0. As a negative control, spectra were acquired for a second sample without $AgrA_C$. Hit compounds were identified by the inversion of the compound spectra when $AgrA_C$ was present and validated using single compound repeats of the WATERGATE W5 LOGSY experiment. The WATERGATE W5 LOGSY spectra were acquired at 20 °C using a mixing time of 1.2 s, 8192 complex data points, spectral width of 10000 Hz and number of scans set to 128. The acquisition time was 0.819 s with a relaxation delay of 3.0 s. The total experiment time was 11 min.

To identify the compound-binding sites, ^{1}H- ^{15}N HSQC spectra were acquired for samples containing 300 μM $AgrA_C$ dissolved in 20 mM sodium phosphate, 100 mM NaCl, 10 mM DTT and 5% D_2O at pH 5.8 after each addition of compound to the protein sample. 200 mM stock solutions of each compound, dissolved in DMSO-d_6 were used for the titrations. A control experiment where 3% DMSO-d_6 was added to the $AgrA_C$ NMR sample in the absence of compound was used to establish which chemical shift perturbations were compound dependent.

The combined chemical shift was calculated for each amide resonance after the addition of compound, using the established method,[35] and mapped onto the apo $AgrA_C$ crystal structure using Pymol v1.2r2. For each compound the highest spectra obtained with the highest concentration of compound tested in the NMR titrations was used to determine the combined chemical shift perturbation.

Compound docking

In-silico docking of the compounds binding to $AgrA_C$ was performed using Autodock Vina.[36]

Inhibition Assays

The activities of five compounds as inhibitors of $AgrA_C$ DNA binding were tested using an electrophoretic mobility shift assay performed in the presence of increasing amounts of each compound (from 0 mM to 5 mM) to test whether the presence of compound inhibits the DNA-binding activity of $AgrA_C$. A 19-bp DNA duplex was prepared by annealing 5′-ATTTAACAGTTAAGTATTT-3′ with its complementary oligonucleotide in 21 mM sodium phosphate, 105 mM NaCl, 10.5 mM DTT at pH 5.8 and purified by size exclusion chromatography using a Superdex 75 HR column. Solutions were prepared containing 2 μM DNA duplex, 20 μM $AgrA_C$ and various compound concentrations in 20 mM sodium phosphate, 100 mM NaCl, 10 mM DTT and 2.5% DMSO-d_6 at pH 5.8. After incubation at room temperature for 30 min, 12% (w/v) Ficoll 400 was added to each sample to give a 2% (w/v) final concentration before running the samples on an 8% acrylamide/bis-acrylamide 40 mM Tris·HCl, 20 mM acetic acid, 1 mM EDTA at pH 8.0 (TAE) gel at 70 V and 4 °C. The gels were stained in TAE supplemented with 0.1 mg mL^{-1} ethidium bromide for 10 min prior to destaining in TAE for 10 min. DNA was visualized with UV light and images were captured using a CCD camera imager (Alpha Innotech).

RESULTS

Backbone Resonance Assignments of $AgrA_C$

Backbone resonance assignments were determined for $AgrA_C$ using a ^{15}N,^{13}C-labeled sample of AgrA LytTR domain (Asp137 to Ile238). The protein sample gave excellent quality NMR spectra with well-dispersed chemical shifts (Fig. 1.) allowing for assignment of the backbone resonances. Assignments could be made for all of the residues between Asn138 and Ile238 except for Tyr229 whose resonances could not be confidently assigned, probably due to overlapping crosspeaks with another amide crosspeak. The unusual downfield shift of the serine 202 backbone amide proton resonance is caused by the hydrogen bonding of this proton with the imidazole ring of His200, as observed in the crystal structure of $AgrA_C$. A similar downfield shift has been described for the hydrogen bonding of Arg7 backbone amide to His106 side chain in *B. subtilis* chorismate mutase.[37]

$AgrA_C$ in the Absence of DNA

The structure of $AgrA_C$ bound to its target DNA duplex has been reported previously,[22] but any structural changes that occur within this domain upon DNA binding were unknown. To

establish if there are structural differences between the apo and DNA-bound states of $AgrA_C$ both NMR and X-ray crystallographic analyses were used to investigate the structure of $AgrA_C$ protein in the absence of DNA.

$AgrA_C$ was crystallized in its apo state, forming $P2_12_12_1$ space group crystals that diffracted to 1.5-Å resolution. The structure of apo $AgrA_C$ was solved by molecular replacement. To remove model bias from the molecular replacement solution, the $AgrA_C$ protein was rebuilt from scratch using RESOLVE and then iteratively refined. The final structure contains two molecules of $AgrA_C$ in the asymmetric unit of the crystals: chain A contains residues Glu141 to Ile238 and chain B contains Ser139 to Ile238).

For both $AgrA_C$ protomers, the N-terminal residues are not observed in the electron density map, suggesting that this region is disordered, consistent with the negative $\{^1H\}^{15}N$ heteronuclear NOE values observed for these amino acids (Fig. 2A) under solution conditions. The apo $AgrA_C$ protein adopts a β-β-β-sandwich fold as observed previously for the DNA-bound state. The fold is stabilized by a set of conserved salt bridges: Glu141-Arg195, Asp157-His208, Asp193-Arg195, Asp157-Arg195 and His174-Glu226. The domain is very stable with heteronuclear NOE values for the backbone amide bond vectors between 0.7 and 1.0 for the majority of the protein. However, the β1-β2 loop, the β3-β4 loop and the C terminus have $^1H^{15}N$-heteronuclear NOE values between 0.6 and 0.7, suggesting greater conformational flexibility for these regions relative to the rest of the LytTR domain (Fig. 2A). In addition, alternate rotamer conformations are observed in the electron density map for several amino acid sidechains within the $AgrA_C$ crystal structure (Ser164, Ser215 and Glu217 from protomer chain A and Ile143, Ser149 and Asp176 from chain B), indicative of localized conformational flexibility within the domain. When the two protomer chains in the $AgrA_C$ crystal are superposed, the only differences in the protein backbone are observed for the N and C termini as well as the β1-β2, and β3-β4 loops, in excellent agreement with the heteronuclear NOE data. The root mean square (RMS) deviation for the two protomer chains in the apo $AgrA_C$ crystal is 0.19 Å for all backbone atoms, but a much greater difference is seen when the DNA-bound state is compared with the apo $AgrA_C$ protein chains A or B (RMS deviations of 0.41 or 0.46 respectively), suggesting that a conformational change takes place upon DNA binding.

When the apo and DNA-bound structures of $AgrA_C$ are compared, changes in the conformation of the DNA-binding surface are observed for residues that are known to interact with the DNA backbone (Ser 168, Arg170, Tyr183, Lys187, Ser202, Arg218, Asn234) as well as for three residues that make direct, base-specific contacts (His169, Asn201 and Arg233)[22] (Fig. 2B). The largest change between the DNA-bound and apo states of the $AgrA_C$ protein is observed for the β3-β4 loop (Ser165 to Arg170). There are no crystal contacts for this region for the chain B protomer in the apo protein structure, so crystal lattice packing does not explain the different configuration of the β3-β4 loop relative to that in the DNA-bound state of $AgrA_C$. The conformational differences between the apo and DNA-bound structures suggest that DNA binding traps the protein into a conformation that represents only one conformation from the population of states adopted by the apo $AgrA_C$ protein. The conformational freedom of the DNA-binding surface would allow some flexibility for AgrA binding to its DNA target sequences, albeit with an entropic penalty for binding.

Fragment Screening for Ligands of $AgrA_C$

LytTR domains are considered attractive drug targets for the development of new antibiotics because these domains are not present in humans and are often found to activate virulence pathways in bacterial pathogens.[20] NMR fragment screening methods were chosen because they allow the entire molecular surface of the target protein to be probed by a fragment

compound library, in contrast to X-ray diffraction screening where some surfaces are occluded by lattice contacts. Fragment screening by NMR has proven to be a powerful tool for identifying regions of a protein that can be targeted by low complexity, small molecule ligands (fragment compounds), which subsequently can be used as starting points for drug development.[38,39] A library of 500 compounds was screened for binding to $AgrA_C$ using a WATERGATE W5 LOGSY NMR experiment. Five compounds, 4-phenoxyphenol, 9H-xanthene-9-carboxylic acid, 2-(4-methylphenyl)-1,3-thiazole-4-carboxylic acid, [5-(2-thienyl)-3-isoxazolyl]methanol and 4-hydroxy-2,6-dimethylbenzonitrile, were identified as binding to $AgrA_C$ as indicated by inversion of the compound spectrum when $AgrA_C$ was present in the solution.

Mapping the Compound-Binding Site

Having identified five compounds that interacted with $AgrA_C$ in the initial screen, increasing amounts of the each compound were titrated into a ^{15}N-labeled sample of $AgrA_C$ and chemical shift perturbations in the ^{15}N-^{1}H-HSQC spectrum of the protein were monitored (Fig. 3.). Addition of each of the five compounds resulted in the perturbation of the same chemical shifts corresponding to the backbone amides for residues Ser231, Val232, Arg233, Asn234, Lys236, Lys237 and Ile238, at the C-terminal end of the protein (Fig. 4). A control experiment confirmed that the chemical shift perturbations were dependent on the presence of the compounds and not caused by the addition of DMSO-d_6 to the NMR sample (Fig. 4F). Mapping the chemical shift perturbations onto the surface of $AgrA_C$ revealed a surface corresponding to a shallow groove formed by the β10-α2 loop and the α2 helix. Interestingly, the β10-α2 loop is known to be important for DNA binding by AgrA[22] so ligands that occlude this site would be expected to reduce the DNA-binding activity of AgrA.

Compound Binding Predictions

In silico docking was used to predict the binding mode of the compounds using AutoDock Vina.[36] The search box was restricted to the region at the C-terminal end of $AgrA_C$, identified by NMR as the binding site for the compounds (Fig. 5). All five compounds docked into a shallow groove located between Val232 and Lys236. 4-Phenoxyphenol is predicted to form hydrogen bond interactions with Lys236 and the backbone carbonyl of Arg233 and bends around Val232 to make tight Van der Waals interactions with this amino acid sidechain. 9H-xanthene-9-carboxylic acid binds with the planar xanthene ring structure between Val232 and Lys236. The predicted interaction is further stabilized by a salt bridge between the carboxylic acid group of the compound and the sidechain amine of Lys236. 2-(4-Methylphenyl)-1,3-thiazole-4-carboxylic acid binding is predicted to be stabilized by a hydrogen bond between the thiazole nitrogen and the carbonyl group of Val232 and a salt bridge between Lys237 and the carboxylic acid group in the compound. [5-(2-thienyl)-3-isoxazolyl]-methanol is predicted to bind with the isoxazolyl ring oxygen forming a hydrogen bond with Lys236 amine. The putative compound binding mode is further stabilized by a hydrogen bond between Val232 backbone carbonyl group and the hydroxyl group of the compound. The 4-hydroxy-2,6-dimethylbenzonitrile compound is predicted to bind in the shallow groove formed between Val232, Arg233, Lys236 and Lys237, with the hydroxyl group of the compound making hydrogen bond interactions with the backbone of Val232 and Lys236. The calculated binding energies for the best docking poses for 4-phenoxyphenol, 9H-xanthene-9-carboxylic acid, 2-(4-methylphenyl)-1,3-thiazole-4-carboxylic acid, [5-(2-thienyl)-3-isoxazolyl]methanol and 4-hydroxy-2,6-dimethylbenzonitrile were −4.1 kcal mol^{-1}, −4.1 kcal mol^{-1}, −4.0 kcal mol^{-1}, −3.5 kcal mol^{-1} and −3.6 kcal mol^{-1}, respectively, in agreement with the millimolar affinity observed by NMR.

Three Fragment Compounds Inhibit the DNA-binding Activity of $AgrA_C$

The five compounds identified by fragment screening all target the C-terminal end of $AgrA_C$, a region known to be involved in DNA binding.[22] It was therefore important to establish whether the compounds alter the DNA-binding activity of AgrA. The DNA-binding activity of $AgrA_C$ was tested using electrophoretic mobility shift assays in the presence of various compound concentrations from 75 μM to 5000 μM (Fig. 6). For three compounds (4-phenoxyphenol, 9H-xanthene-9-carboxylic acid and 2-(4-methylphenyl)-1,3-thiazole-4-carboxylic acid), the intensity of the band corresponding to the $AgrA_C$:DNA complex decreases as the compound concentration is increased; no $AgrA_C$:DNA complex is observed when 5 mM compound is present. Two other compounds ([5-(2-thienyl)-3-isoxazolyl]methanol and 4-hydroxy-2,6-dimethylbenzonitrile) had no apparent effect on the DNA-binding activity of $AgrA_C$ at concentrations up to 5 mM. The DMSO concentration in all binding reactions was constant so the disruption to the DNA-binding activity of $AgrA_C$ is dependent only on the presence of the compound. In the control experiments where no $AgrA_C$ was present, there is no effect of the compounds on the migration of DNA through the gel. Some of the compounds that were identified to bind at the C terminus of $AgrA_C$ can disrupt AgrA DNA-binding activity.

DISCUSSION

The *agr* quorum sensing system has been suggested as a possible drug target because deletion of the *agr* system attenuates *S. aureus* virulence in animal models of infection.[14–16,18] Furthermore, a functional *agr* system has been found to be important for bacterial survival in the early stages of abscess formation in animal models.[19] On the other hand, isolation of *agr*-deficient mutants from patients with chronic *S. aureus* infections[40] suggests that *agr* is not important for maintaining the infection once it has been established. Thus substantial controversy surrounds the validity of the *agr* system as a target for antimicrobial drug development. *agr* inhibitors would be especially useful tools in addressing the role of the *agr* system during the course of infection in animal models of infection. Furthermore, *agr* inhibitors might be effective for prophylactic treatment. Indeed, co-inoculation of *S. aureus* with an *agr* inhibitor was sufficient to reduce virulence in a mouse skin infection model.[19]

To date, research on inhibitors of *S. aureus agr* quorum sensing has focused on analogues of the autoinducing peptide as inhibitors of AgrC activation.[17,19,41–43] Targeting *agr* quorum sensing using AIP analogues is complicated by the existence of several Staphylococcal AgrC types that are inhibited or activated by different AIP molecules.[17,44,45]

The results presented here represent the first attempt to inhibit the *agr* quorum sensing pathway by targeting the DNA-binding surface of the AgrA LytTR domain. Fragment screening can be used to identify an energetic focal point on the surface of a protein that is favorable to compound binding.[21] While small fragments are efficient in covering large areas of chemical space and have advantages as starting points for drug development, their binding affinities are typically very low. Indeed, only three of the five compounds that were identified to bind to $AgrA_C$ by NMR analyses disrupted DNA binding in electrophoretic mobility shift assays. For these compounds, inhibition was detected only at millimolar concentrations. The low affinity of ligand binding provides a likely explanation for the lack of inhibition observed with the other two compounds.

The five compounds identified by fragment screening as $AgrA_C$ inhibitors all bind to a highly conserved region that forms a short helix at the C terminus of the AgrA protein. To date, there are 211 AgrA protein sequences across all Staphylococcal strains in the Uniprot knowledgebase database. Conservation of 100% sequence identity is observed for Ser231,

Val232, Arg233, Asn234, Lys236 and Lys237 that are identified as residues in the compound-binding site. Val235 is conserved in 97% of all Staphylococcal sequences, although isoleucine or leucine amino acids are observed at this position in *S. siminae*, *S. pseudintermedius*, *S. intermedius* and *S. carnosus* strains. The C-terminal isoleucine residue (Ile238) is conserved in 94% of the AgrA protein sequences but is replaced by a lysine in *S. aureus* Mu3 and Mu50 strains or by valine in *S. lugdunensis* or *S. saprophyticus* strains. Other modifications at the AgrA C terminus give rise to short extensions. However, these extensions result in a reduced or completely deficient *agr* phenotype.[40] Given the high degree of sequence conservation of the compound-binding site, it is reasonable to expect that any compound binding to this surface would be able to inhibit *agr* quorum sensing across all Staphylococcal strains.

The fragment compounds identified in this study bind with millimolar affinity to the $AgrA_C$ protein. The low affinity reflects the small size of the compounds that populate the library used in the screen. Despite the low affinity of initial hits, fragment screening has previously been shown to provide a useful starting point for the development of higher affinity inhibitors of protein:DNA interactions.[21,46] It is commonly observed that features of the original low affinity fragment compound, identified by the initial screen, are maintained in the compounds that ultimately enter clinical development[47–49]. Further research is required to expand the compounds that have been identified as AgrA inhibitors, to improve their affinity and specificity. The apo $AgrA_C$ crystal structure and NMR assignments that have been established in this research will provide valuable information to guide future compound development.

Acknowledgments

We thank G. V. T. Swapna for her assistance setting up the NMR experiments and G. T. Montelione for use of the NMR facility within the CABM Structural Bioinformatics Laboratory.

Funding Sources

This work was supported in part by National Institutes of Health grant R37 GM047958 to A.M.S.

ABBREVIATIONS

AIP	autoinducing peptide
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
NMR	nuclear magnetic resonance
NOE	Nuclear Overhauser Effect
PEG	polyethylene glycols
rms	root mean square

References

1. Talan DA, Krishnadasan A, Gorwitz RJ, Fosheim GE, Limbago B, Albrecht V, Moran GJ. Comparison of *Staphylococcus aureus* from skin and soft-tissue infections in US emergency department patients, 2004 and 2008. Clin Infect Dis. 2011; 53:144–149. [PubMed: 21690621]
2. Dohin B, Gillet Y, Kohler R, Lina G, Vandenesch F, Vanhems P, Floret D, Etienne J. Pediatric bone and joint infections caused by Panton-Valentine leukocidin-positive *Staphylococcus aureus*. Pediatr Infect Dis J. 2007; 26:1042–1048. [PubMed: 17984813]

3. Kollef MH, Micek ST. *Staphylococcus aureus* pneumonia: a “superbug” infection in community and hospital settings. Chest. 2005; 128:1093–1097. [PubMed: 16162690]

4. Petti CA, Fowler VG Jr. *Staphylococcus aureus* bacteremia and endocarditis. Infect Dis Clin North Am. 2002; 16:413–435. x–xi. [PubMed: 12092480]

5. Todd J, Fishaut M, Kapral F, Welch T. Toxic-shock syndrome associated with phage-group-I Staphylococci. Lancet. 1978; 2:1116–1118. [PubMed: 82681]

6. European Centre for Disease Prevention and Control. Annual Report of the European Antimicrobial Resistance Surveillance Network (EARS-Net). Stockholm: ECDC; 2011. Antimicrobial resistance surveillance in Europe 2010.

7. Queck SY, Jameson-Lee M, Villaruz AE, Bach TH, Khan BA, Sturdevant DE, Ricklefs SM, Li M, Otto M. RNAIII-independent target gene control by the *agr* quorum-sensing system: insight into the evolution of virulence regulation in *Staphylococcus aureus*. Mol Cell. 2008; 32:150–158. [PubMed: 18851841]

8. Koenig RL, Ray JL, Maleki SJ, Smeltzer MS, Hurlburt BK. *Staphylococcus aureus* AgrA binding to the *RNAIII-agr* regulatory region. J Bacteriol. 2004; 186:7549–7555. [PubMed: 15516566]

9. Janzon L, Lofdahl S, Arvidson S. Identification and nucleotide sequence of the delta-lysin gene, *hld*, adjacent to the accessory gene regulator (*agr*) of *Staphylococcus aureus*. Mol Gen Genet. 1989; 219:480–485. [PubMed: 2622452]

10. Bronner S, Stoessel P, Gravet A, Monteil H, Prevost G. Variable expressions of *Staphylococcus aureus* bicomponent leucotoxins semiquantified by competitive reverse transcription-PCR. Appl Environ Microbiol. 2000; 66:3931–3938. [PubMed: 10966411]

11. Dunman PM, Murphy E, Haney S, Palacios D, Tucker-Kellogg G, Wu S, Brown EL, Zagursky RJ, Shlaes D, Projan SJ. Transcription profiling-based identification of *Staphylococcus aureus* genes regulated by the *agr* and/or *sarA* loci. J Bacteriol. 2001; 183:7341–7353. [PubMed: 11717293]

12. Janzon L, Arvidson S. The role of the delta-lysin gene (*hld*) in the regulation of virulence genes by the accessory gene regulator (*agr*) in *Staphylococcus aureus*. EMBO J. 1990; 9:1391–1399. [PubMed: 2328718]

13. Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, Moghazeh S. Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. EMBO J. 1993; 12:3967–3975. [PubMed: 7691599]

14. Abdelnour A, Arvidson S, Bremell T, Ryden C, Tarkowski A. The accessory gene regulator (*agr*) controls *Staphylococcus aureus* virulence in a murine arthritis model. Infect Immun. 1993; 61:3879–3885. [PubMed: 8359909]

15. Booth MC, Atkuri RV, Nanda SK, Iandolo JJ, Gilmore MS. Accessory gene regulator controls *Staphylococcus aureus* virulence in endophthalmitis. Invest Ophthalmol Vis Sci. 1995; 36:1828–1836. [PubMed: 7635657]

16. Gillaspy AF, Hickmon SG, Skinner RA, Thomas JR, Nelson CL, Smeltzer MS. Role of the accessory gene regulator (*agr*) in pathogenesis of staphylococcal osteomyelitis. Infect Immun. 1995; 63:3373–3380. [PubMed: 7642265]

17. Mayville P, Ji G, Beavis R, Yang H, Goger M, Novick RP, Muir TW. Structure-activity analysis of synthetic autoinducing thiolactone peptides from *Staphylococcus aureus* responsible for virulence. Proc Natl Acad Sci USA. 1999; 96:1218–1223. [PubMed: 9990004]

18. Montgomery CP, Boyle-Vavra S, Daum RS. Importance of the global regulators Agr and SaeRS in the pathogenesis of CA-MRSA USA300 infection. PLoS One. 2010; 5:e15177. [PubMed: 21151999]

19. Wright JS 3rd, Jin R, Novick RP. Transient interference with staphylococcal quorum sensing blocks abscess formation. Proc Natl Acad Sci USA. 2005; 102:1691–1696. [PubMed: 15665088]

20. Galperin MY. Telling bacteria: do not LytTR. Structure. 2008; 16:657–659. [PubMed: 18462668]

21. Hajduk PJ, Huth JR, Fesik SW. Druggability indices for protein targets derived from NMR-based screening data. J Med Chem. 2005; 48:2518–2525. [PubMed: 15801841]

22. Sidote DJ, Barbieri CM, Wu T, Stock AM. Structure of the *Staphylococcus aureus* AgrA LytTR domain bound to DNA reveals a beta fold with an unusual mode of binding. Structure. 2008; 16:727–735. [PubMed: 18462677]

23. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 1997; 276:307–326.
24. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]
25. Terwilliger TC. SOLVE and RESOLVE: automated structure solution and density modification. Methods Enzymol. 2003; 374:22–37. [PubMed: 14696367]
26. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]
27. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213–221. [PubMed: 20124702]
28. Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD. ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. J Biomol NMR. 1995; 6:135–140. [PubMed: 8589602]
29. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR. 1995; 6:277–293. [PubMed: 8520220]
30. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins. 2005; 59:687–696. [PubMed: 15815974]
31. Kay LE, Keifer P, Saarinen T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. J Am Chem Soc. 1992; 114:10663–10665.
32. Grzesiek S, Dobeli H, Gentz R, Garotta G, Labhardt AM, Bax A. ^{1}H, ^{13}C, and ^{15}N NMR backbone assignments and secondary structure of human interferon-gamma. Biochemistry. 1992; 31:8180–8190. [PubMed: 1525157]
33. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. Biochemistry. 1994; 33:5984–6003. [PubMed: 7514039]
34. Furihata K, Shimotakahara S, Tashiro M. An efficient use of the WATERGATE W5 sequence for observing a ligand binding with a protein receptor. Magn Reson Chem. 2008; 46:799–802. [PubMed: 18537106]
35. Mulder FA, Schipper D, Bott R, Boelens R. Altered flexibility in the substrate-binding site of related native and engineered high-alkaline *Bacillus subtilisins*. J Mol Biol. 1999; 292:111–123. [PubMed: 10493861]
36. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010; 31:455–461. [PubMed: 19499576]
37. Eletsky A, Heinz T, Moreira O, Kienhofer A, Hilvert D, Pervushi K. Direct NMR observation and DFT calculations of a hydrogen bond at the active site of a 44 kDa enzyme. J Biomol NMR. 2002; 24:31–39. [PubMed: 12449416]
38. Fejzo J, Lepre CA, Peng JW, Bemis GW, Ajay Murcko MA, Moore JM. The SHAPES strategy: an NMR-based approach for lead generation in drug discovery. Chem Biol. 1999; 6:755–769. [PubMed: 10508679]
39. Lepre CA, Peng J, Fejzo J, Abdul-Manan N, Pocas J, Jacobs M, Xie X, Moore JM. Applications of SHAPES screening in drug discovery. Comb Chem High Throughput Screen. 2002; 5:583–590. [PubMed: 12470255]
40. Traber K, Novick R. A slipped-mispairing mutation in AgrA of laboratory strains and clinical isolates results in delayed activation of *agr* and failure to translate δ- and α-haemolysins. Mol Microbiol. 2006; 59:1519–1530. [PubMed: 16468992]

41. Fowler SA, Stacy DM, Blackwell HE. Design and synthesis of macrocyclic peptomers as mimics of a quorum sensing signal from *Staphylococcus aureus*. Org Lett. 2008; 10:2329–2332. [PubMed: 18476747]

42. Mansson M, Nielsen A, Kjaerulff L, Gotfredsen CH, Wietz M, Ingmer H, Gram L, Larsen TO. Inhibition of virulence gene expression in *Staphylococcus aureus* by novel depsipeptides from a marine photobacterium. Mar Drugs. 2011; 9:2537–2552. [PubMed: 22363239]

43. Otto M, Sussmuth R, Vuong C, Jung G, Gotz F. Inhibition of virulence factor expression in *Staphylococcus aureus* by the *Staphylococcus epidermidis agr* pheromone and derivatives. FEBS Lett. 1999; 450:257–262. [PubMed: 10359085]

44. Geisinger E, George EA, Muir TW, Novick RP. Identification of ligand specificity determinants in AgrC, the *Staphylococcus aureus* quorum-sensing receptor. J Biol Chem. 2008; 283:8930–8938. [PubMed: 18222919]

45. Jensen RO, Winzer K, Clarke SR, Chan WC, Williams P. Differential recognition of *Staphylococcus aureus* quorum-sensing signals depends on both extracellular loops 1 and 2 of the transmembrane sensor AgrC. J Mol Biol. 2008; 381:300–309. [PubMed: 18582472]

46. Hajduk PJ, Dinges J, Miknis GF, Merlock M, Middleton T, Kempf DJ, Egan DA, Walter KA, Robins TS, Shuker SB, Holzman TF, Fesik SW. NMR-based discovery of lead inhibitors that block DNA binding of the human papillomavirus E2 protein. J Med Chem. 1997; 40:3144–3150. [PubMed: 9379433]

47. Artis DR, Lin JJ, Zhang C, Wang W, Mehra U, Perreault M, Erbe D, Krupka HI, England BP, Arnold J, Plotnikov AN, Marimuthu A, Nguyen H, Will S, Signaevsky M, Kral J, Cantwell J, Settachatgull C, Yan DS, Fong D, Oh A, Shi S, Womack P, Powell B, Habets G, West BL, Zhang KY, Milburn MV, Vlasuk GP, Hirth KP, Nolop K, Bollag G, Ibrahim PN, Tobin JF. Scaffold-based discovery of indeglitazar, a PPAR pan-active anti-diabetic agent. Proc Natl Acad Sci USA. 2009; 106:262–267. [PubMed: 19116277]

48. Park CM, Bruncko M, Adickes J, Bauch J, Ding H, Kunzer A, Marsh KC, Nimmer P, Shoemaker AR, Song X, Tahir SK, Tse C, Wang X, Wendt MD, Yang X, Zhang H, Fesik SW, Rosenberg SH, Elmore SW. Discovery of an orally bioavailable small molecule inhibitor of prosurvival B-cell lymphoma 2 proteins. J Med Chem. 2008; 51:6902–6915. [PubMed: 18841882]

49. Wyatt PG, Woodhead AJ, Berdini V, Boulstridge JA, Carr MG, Cross DM, Davis DJ, Devine LA, Early TR, Feltell RE, Lewis EJ, McMenamin RL, Navarro EF, O'Brien MA, O'Reilly M, Reule M, Saxty G, Seavers LC, Smith DM, Squires MS, Trewartha G, Walker MT, Woolford AJ. Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (AT7519), a novel cyclin dependent kinase inhibitor using fragment-based X-ray crystallography and structure based drug design. J Med Chem. 2008; 51:4986–4999. [PubMed: 18656911]

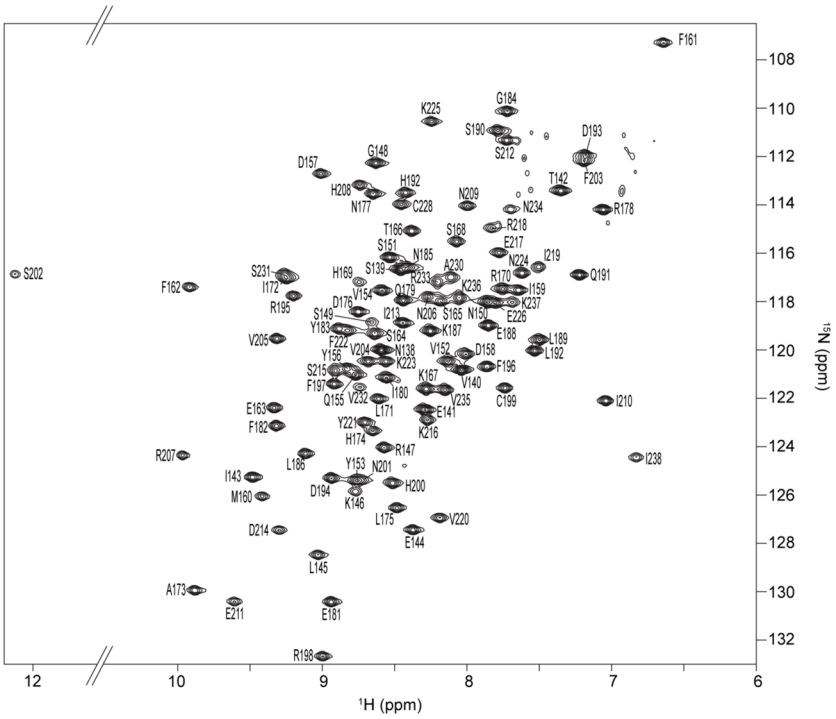

Figure 1.
2D ^{1}H-^{15}N HSQC spectrum of AgrA$_{C}$. Residue specific cross-peak assignments for the backbone amides are indicated.

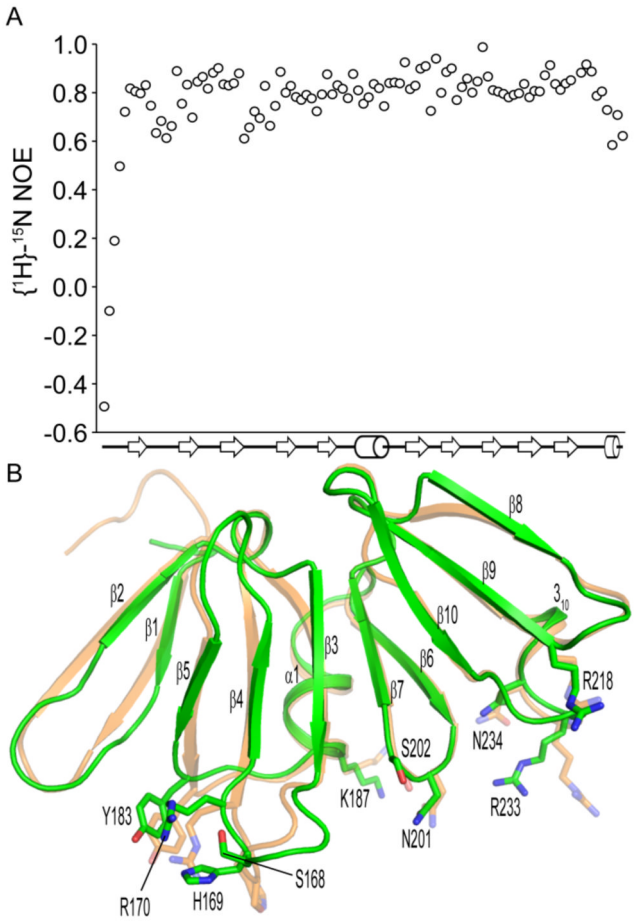

Figure 2.

Conformational differences between apo and DNA-bound $AgrA_C$. A, Steady-state heteronuclear $\{^1H\}^{15}N$-NOE values for each residue of $AgrA_C$ plotted versus the secondary structure topology of the protein. β strands and α helices are depicted as arrows and cylinders respectively. B, Superposition of the apo $AgrA_C$ crystal structure (green) with $AgrA_C$ in its DNA-bound state (orange).[22] The sidechains of amino acid residues within the DNA-binding surface that adopt different conformations between the apo and DNA-bound states are shown in stick format.

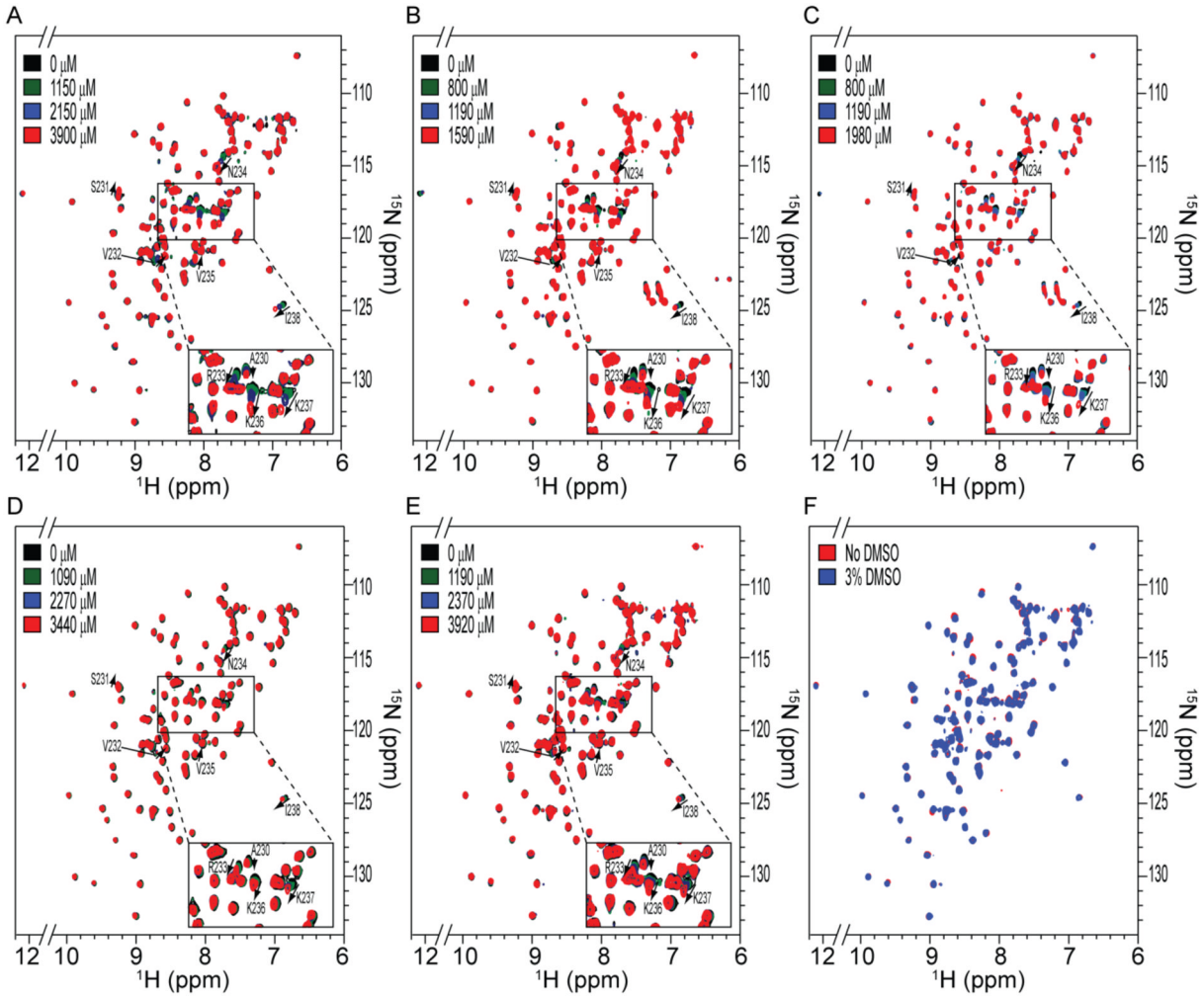

Figure 3.

Chemical shift perturbations observed upon binding of fragment compounds to $AgrA_C$. ^{1}H-^{15}N HSQC spectra of $AgrA_C$ recorded in the presence of the indicated concentrations of A, 4-phenoxyphenol, B, 9H-xanthene-9-carboxylic acid, C, 2-(4-methylphenyl)-1,3-thiazole-4-carboxylic acid, D, [5-(2-thienyl)-3-isoxazolyl]methanol and E, 4-hydroxy-2,6-dimethylbenzonitrile. The residue specific assignments for backbone amide cross-peaks that shift in a compound concentration dependent manner are highlighted. F, ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled $AgrA_C$ before and after the addition of 3% DMSO-d_6.

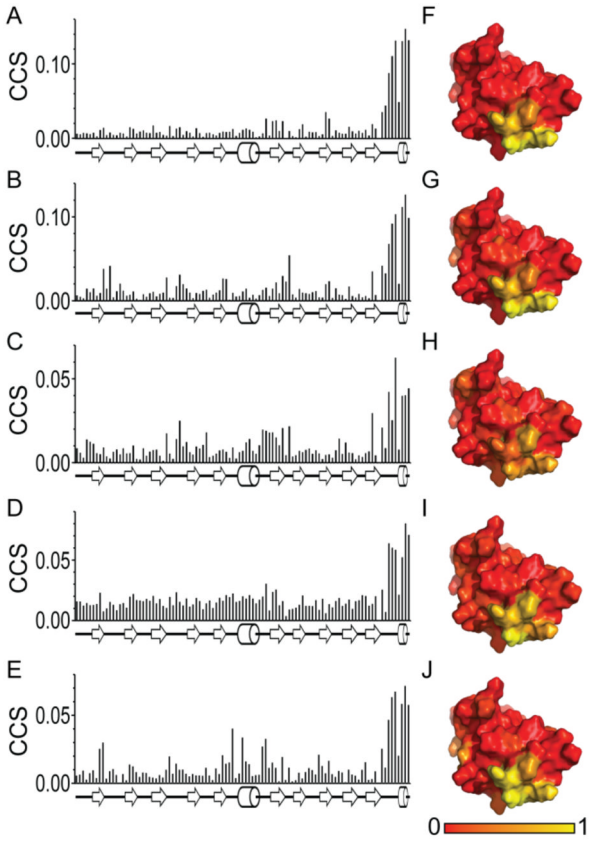

Figure 4.

Mapping the compound dependent chemical shift perturbations. Combined chemical shift perturbations of backbone amide resonances plotted versus the amino acid residue number for the addition of 4-phenoxyphenol *A*, 9H-xanthene-9-carboxylic acid, *B*, 2-(4-methylphenyl)-1,3-thiazole-4-carboxylic acid, *C*, [5-(2-thienyl)-3-isoxazolyl]methanol, *D*, and *E*, 4-hydroxy-2,6-dimethylbenzonitrile. The normalized combined chemical shift perturbations for *F*, 4-phenoxyphenol, *G*, 9H-xanthene-9-carboxylic acid, *H*, 2-(4-methylphenyl)-1,3-thiazole-4-carboxylic acid, *I*, [5-(2-thienyl)-3-isoxazolyl]methanol and *J*, 4-hydroxy-2,6-dimethylbenzonitrile are mapped onto the surface of the crystal structure of apo $AgrA_C$ with the region showing the greatest combined chemical shift perturbation colored in yellow.

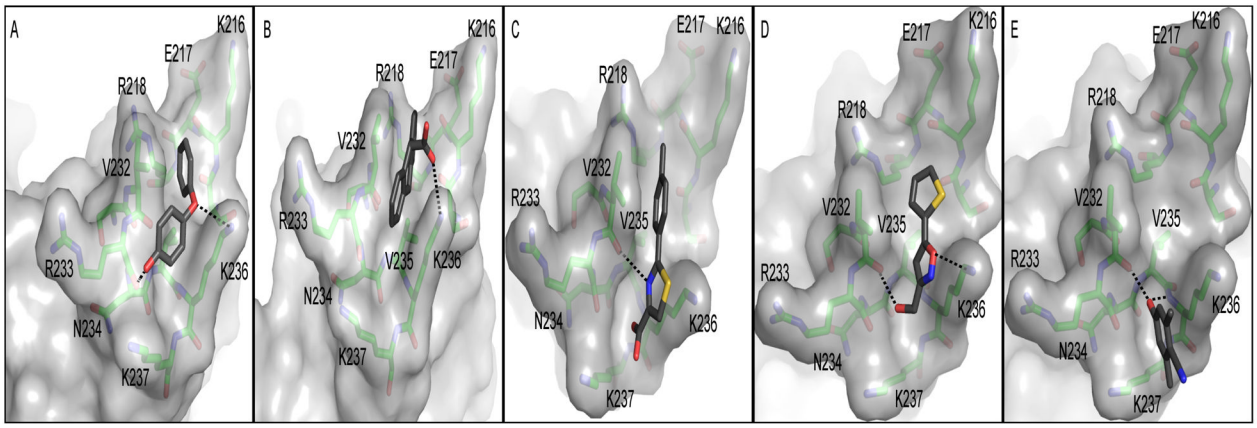

Figure 5.

In silico compound docking predictions. Autodock Vina predicted binding modes are shown for *A*, 4-phenoxyphenol, *B*, 9H-xanthene-9-carboxylic acid, *C*, 2-(4-methylphenyl)-1,3-thiazole-4-carboxylic acid, *D*, [5-(2-thienyl)-3-isoxazolyl]methanol and *E*, 4-hydroxy-2,6-dimethylbenzonitrile. Orientations of the protein are slightly different in each panel to optimize viewing of the predicted hydrogen bonds between the compound and protein, indicated by dashed lines.

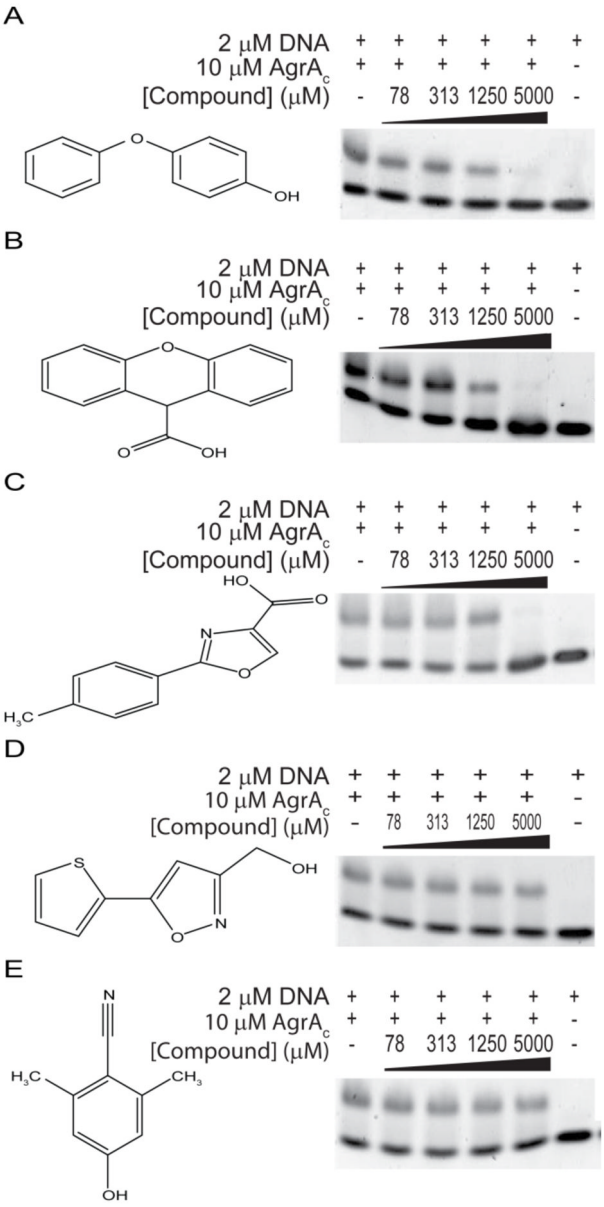

Figure 6.

Effects of compounds on AgrA DNA-binding activity. Electrophoretic mobility shift assays were performed with $AgrA_C$, its target DNA duplex and the indicated concentrations of *A*, 4-phenoxyphenol, *B*, 9H-xanthene-9-carboxylic acid, *C*, 2-(4-methylphenyl)-1,3-thiazole-4-carboxylic acid, *D*, [5-(2-thienyl)-3-isoxazolyl]methanol and *E*, 4-hydroxy-2,6-dimethylbenzonitrile. DNA incubated with or without $AgrA_C$ in the absence of compound are shown as positive and negative controls, respectively.

Table 1

X-ray data collection and refinement statistics (molecular replacement)

	$AgrA_C$[a]
Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	41.4, 45.9, 112.1
α, β, γ (°)	90.0, 90.0, 90.0
Resolution range (Å)	25.00-1.52 (1.55-1.52)[b]
R_{sym}[c]	0.045 (0.563)
$I/\sigma(I)$	32.9 (2.0)
Completeness (%)	97.2 (89.8)
Redundancy	3.3 (2.7)
Refinement	
Resolution (Å)	20.72-1.52
No. reflections	31964
Data cutoff	$\sigma(F) > 0$
R_{work}[d]/R_{free}[e]	0.180/0.209
No. atoms per asymmetric unit	
Protein	1702
Glycerol	18
Water	254
Average B-factors (A^2)	
Protein	20.5
Glycerol	32.3
Water	32.6
Rms deviations from ideality	
Bond lengths (Å)	0.006
Bond angles (°)	1.048

[a] A single crystal was used for the data collection.

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

[c] $R_{sym} = \Sigma_h|I_h - \langle I\rangle|/\Sigma_h I_h$, where I_h and $\langle I\rangle$ represent the diffraction intensity values of individual measurements and the corresponding mean values, respectively.

[d] $R_{work} = (\Sigma_h||F_O| - |F_C||)/\Sigma_h|F_O|$, where F_O and F_C are observed and calculated structure factor amplitudes, respectively.

[e] R_{free} was calculated for 5% of the randomly selected reflections of the data set that were not used in the refinement.