N-acetyl cysteine: A tool to perturb SARS-CoV-2 spike protein conformation

Utsab Debnath[1,2#], Amrita Mitra[3#], Varun Dewaker[4], Yenamandra S. Prabhakar[4], Raghu Tadala[5], Kiran Krishnan[5], Padmakar Wagh[5], Umashankar Velusamy[5], Chandru Subramani[6], Shubhra Agarwal[6], Sudhanshu Vrati[6], Aastha Baliyan[7], Anura V Kurpad[8], Parthasarathi Bhattacharyya[9], Amit Kumar Mandal[7]*

[1] School of Pharmaceutical Technology, Adamas University, Kolkata, West Bengal, India, Pin– 700126

[2] School of Health Science, University of Petroleum and Energy Studies, Dehradun, India, Pin - 248007

[3] Clinical Proteomics Unit, Division of Molecular Medicine, St. John's Research Institute, 100 ft Road, Koramangala, Bangalore – 560034

[4] Medicinal and Process Chemistry Division, CSIR-Central Drug Research Institute, Lucknow, India, Pin–226031

[5] Waters India Private Limited No. 36A, 2nd Phase, Peenya Industrial Area, Bengaluru, Karnataka,Pin- 560058

[6] Regional Centre for Biotechnology, Faridabad 121001, India

[7] Department of Biological Science, Indian Institute of Science Education and Research Kolkata, Nadia, West Bengal, India – 741246

[8] Department of Physiology, St. John's Medical College, Sarjapura Road, Bangalore - 5600034

[9] Institute of Pulmocare & Research, New Town, Kolkata, West Bengal, India, Pin – 7000156

[#] Both the authors contributed equally

*Corresponding Author:
Dr. Amit Kumar Mandal Associate Professor,
Department of Biological Sciences,
Indian Institute of Science Education and Research Kolkata Mohanpur, Nadia, West Bengal,
India-741246
E-mail: amitkm@iiserkol.ac.in

Abstract:

The infection caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) resulted in a pandemic with huge death toll and economic consequences. The virus attaches itself to the human epithelial cells through noncovalent bonding of its spike protein with the angiotensin-converting enzyme-2 (ACE2) receptor on the host cell. We hypothesized that perturbing the functionally active conformation of spike protein through reduction of its solvent accessible disulfide bond, thereby disintegrating its structural architecture, may be a feasible strategy to prevent infection. Proteomics data showed that N-acetyl cysteine (NAC), an antioxidant and mucolytic agent been widely in use in clinical medicine, forms covalent conjugates with solvent accessible cysteine residues of spike protein that were disulfide bonded in the native state. *In silico* analysis indicated that this covalent conjugation perturbed the stereo specific orientations of the interacting key residues of spike protein that resulted in threefold weakening in the binding affinity of spike protein with ACE2 receptor. Antiviral assay using VeroE6 cells showed that NAC caused 54.3% inhibition in SARS-CoV-2 replication. Interestingly, almost all SARS-Cov-2 variants conserved cystine residues in the spike protein. Our observed results open avenues for exploring *in vivo* pharmaco-preventive and therapeutic potential of NAC for Coronavirus Disease 2019 (COVID-19).

Key words: SARS-CoV-2, Spike protein, Disulfide bonds, N-Acetyl cysteine (NAC), Angiotensin-converting enzyme 2 (ACE2), Mass spectrometry, Proteomics, Molecular Docking, MD simulation, Antiviral assay, VeroE6

Introduction:

COVID-19 is caused by the highly pathogenic RNA virus, SARS- CoV-2.[1] Unlike SARS-CoV-1 and Middle East Respiratory Syndrome Coronavirus (MERS-CoV), the rapid rate of infection caused by SARS-CoV-2 with a mortality rate of 2-3% has resulted in a pandemic across the globe.[2] Phylogenetic analyses of the coronavirus genomes showed that SARS-CoV-2 belongs to the Betacoronavirus genus. The genome of SARS-CoV-2 is a single-stranded RNA consisting of about 30 kb nucleotides.[3] SARS-CoV-2 encodes for four major structural proteins, namely spike protein, membrane protein, envelope protein, and nucleocapsid protein.

The entry of the SARS-CoV-2 into human cells initiates through the transmembrane spike protein that forms trimers protruding from the virus cell surface.[3,4] The receptor-binding domain (RBD) of the spike protein binds strongly with the host ACE2 receptor to enter into host cells with a dissociation equilibrium constant (K_d) of ~15 nM.[1,4] ACE2 is a type I membrane protein expressed in multiple organs such as lungs, heart, kidneys, testis intestine, and also endothelial cells of arteries, with a primary physiological role in the maturation of angiotensin, a peptide hormone that controls vasoconstriction and blood pressure.[1]

An examination of the crystal structure of the spike protein and ACE2 complex showed that there are 13 hydrogen bonds and 2 salt bridges at the interface between RBD of the SARS-CoV-2 spike protein and the ACE2 receptor.[2] Stereo specific orientations of the interacting amino acid residues of the spike protein are provided by its structural architecture, where four intra-molecular disulfide bonds between cysteine residues contribute. Among these four disulfide linkages in the N-terminal domain (NTD) of the spike protein, three disulfide bonds (cys336–cys361, cys379–cys432 and cys391–cys525) stabilize the β sheet structure and the fourth disulfide bond (cys480–cys488) joins the loops at the distal end of the receptor binding motif that belongs to RBD of the spike protein.[2]

We hypothesized that the functionally active conformation of spike protein could be perturbed and rendered inactive through the reduction of its solvent accessible disulfide bond. In the present study, we have successfully reduced the most accessible disulfide bond (cys391-cys525) at the NTD of purified S1 subunit of SARS-CoV-2 spike protein through thiol/disulfide exchange mechanism, executed by N-acetyl cysteine (NAC), which is a commonly prescribed antioxidant and a mucolytic agent. Upon

covalent conjugation of NAC with cys525, a comparative *in silico* analysis showed significant perturbation in the stereo specific orientation of interacting residues of RBD of spike protein with ACE2 receptor. To explore the impact of this structural perturbation and the potential dysfunction of spike protein in live SARS-CoV-2, an antiviral assay was performed with NAC treated VeroE6 cells (kidney cell line).

Materials and Methods:

Materials

Purified N-terminal domain (NTD) of the S1 subunit of SARS-CoV-2 of spike protein was purchased from RayBiotech (Georgia, USA) where the region of expression of protein was from Val16 to Gln690 residues. N-ethyl maleimide (NEM), dithiothreitol (DTT), iodoacetamide (IAM) and N-acetyl cysteine (NAC) were purchased from Sigma (St. Louis, MO). Rapigest detergent was obtained from Waters (Milford, MA, USA). All solvents used were of LC−MS grade and all other chemicals used were of analytical grade.

Methods:

Solvent accessibility analysis of cysteine residues:

To analyze the solvent accessibility of cysteine residues connected via disulfide bonds, spike protein was extracted from the crystal structure of spike protein-ACE2 receptor (human) complex (PDB ID: 6LZG).[5] After selecting the protein, VEGAZZ based pocket analysis was carried out specifically for eight cysteine residues that are disulfide bonded in the spike protein using default parameters.[6]

Pocket identification and Molecular Docking:

To identify the probable binding pockets of ligand free spike protein, CASTp web server was used by considering solvent probe sphere 1.4 Å. The pocket identification analysis is based on recent theoretical and algorithmic results of Computational Geometry, which includes delaunay triangulation, alpha shape, and discrete flow.[7]

After selecting the pocket site, the docking study was performed in AUTODOCK 4.2.[8] Prior to the docking, NAC was generated in Sybyl-X 1.3.[9] Subsequently NAC was subjected to optimization for its geometry using the Powell energy minimization algorithm, Gasteiger-Huckel charges, and 0.001 kcal/(mol.Å) as convergence criteria.[10]

The protein coordinates of spike protein from the spike protein–ACE2 receptor (human) co-crystal (PDB code: 6LZG) was considered for investigating the binding mode of NAC. For docking study, the protein was prepared in SYBYL by making use of the same procedure as adopted in case of study molecule. Here, residue based docking method was carried out to dock the NAC molecules near to cys480 & cys488 residues of the ACE2 receptor binding domain of spike protein. The prepared spike protein and the corresponding study molecule were considered in AUTODOCK for the flexible docking experiments by previously reported methods.[11] Best docked conformers were collected from a population of 150 samples. Pymol and Discovery studio software were used to visualize the interactions of protein and docked molecules.[12]

Molecular Dynamics (MD) simulations:

In NAMD2.7 MD simulations were carried out on docking model to study the stability of these systems and dynamics of molecular interactions between spike protein and NAC.[13] To do this, the protonation state of histidine residues of the protein at physiological pH (7.4) were assigned using the H++ web server.[14] CHARMM22 force Field (with CMAP correction) was used to parameterize the protein.[15] This force field yielded improved dynamical and structural properties of proteins in MD simulations. Parameters of the ligands were generated via the ParamChem web server.[16] The MD simulations were carried out under periodic boundary conditions using the solvated (water box) protein. The prepared protein-ligand system was run for 100 ns using multi-core and CUDA supported NAMD package. The NPT technique (Pressure 1 atm; Temperature 310 K) with 12 Å as cut-off for non-bonded atom interactions, Particle Mesh Ewald algorithm for long-range electrostatic forces and Langevin's piston in 'on' condition were used to run the MD simulation.[17] The computation was executed with the standard parameter files (CHARMM22) of MD simulation, default arguments and an internal dielectric constant value 1.6.[18] The time step for

dynamics integration was set to 2 fs. The simulation trajectories were recorded at 1000 steps (2 ps) intervals. The system was adequately minimized for 70 ps followed by subjected to equilibration run for 100 ns. Visualization and evaluation of the trajectories were done in VMD 1.9.3.[19] Making use of standard tcl scripts the RMSD and Rg of the respective protein trajectories were analyzed.

Protein-protein docking:

Protein-protein docking study was carried via HDOCK web-server.[20] This exercise was performed for the crystal structure of ligand free spike protein with ACE2 receptor complex as well as MD simulated NAC docked spike protein with ACE2 receptor complex to understand the difference in their binding affinities towards ACE2 receptor. This study was based on a hybrid algorithm of template-based modeling and ab-initio free docking method. Prodigy server was used to determine ΔG and K_d values of the predictive protein-protein interaction contacts.[21] Pymol software was used to visualize the interactions between two proteins. Furthermore, NAC molecule was manually conjugated with free thiol group of cys525 residue by reducing the most solvent accessible disulfide bond (cys391-cys525) to form a new cys525-S-S-NAC linkage. This was executed on the basis of distance matrix analysis in between NAC and cysteine residues after MD simulation study. The experiment was carried out to check the difference in binding affinities with ACE2 receptor through additional protein-protein docking simulation.

Differential modification of the NTD of S1 subunit of spike protein with NEM and IAM:

64 μg of the NTD of S1 subunit of spike protein was dialyzed against 50 mM ammonium bicarbonate buffer, pH 7.4. The dialyzed protein was treated with NEM for 1 hr at 37 °C maintaining protein/NEM ratio 1/20 mol/mol. The NEM treated protein was subjected to a G-10 spin column to remove any unreacted NEM from the solution. Subsequently, the NEM treated protein solution was divided in two parts. The first part was directly subjected to overnight digestion with trypsin maintaining an enzyme/substrate ratio 1/10 mol/mol and served as control for the LC-MS analysis. The second part was treated with DTT maintained at 50 fold excess to the concentration of protein in the solution for 30 mins at 60 °C. The protein solution was then incubated with IAM for 1 hr in the dark at room temperature where IAM was maintained at 100 fold excess to the concentration

of protein in the solution. The differentially modified protein solution was subjected to overnight proteolytic digestion with trypsin maintaining enzyme/substrate ratio 1/10 mol/mol and used further for LC-MS analysis.

Modification of the NTD of S1 subunit of spike protein with NAC:

7 µM of the protein was incubated with NAC in 50 mM ammonium bicarbonate buffer pH 7.4 for 4 hrs at 37 °C where protein/NAC ratio was maintained at 1/10 mol/mol. After incubation, the NAC modified protein was subjected to G-10 spin column to remove any unreacted NAC from the solution. The protein in the solution was denatured using 0.2% Rapigest at 80 °C for 15 mins prior to trypsin digestion. Subsequently, trypsin was added to the protein solution at an enzyme/substrate ratio of 1/10 mol/mol and incubated overnight at 37 °C. Following this, the samples were acidified in 0.5% (v/v) formic acid (FA) and incubated at 37 °C for 90 mins to facilitate hydrolysis of Rapigest detergent. Samples were then centrifuged for 45 min at 4000 x g and the supernatant was isolated and subsequently used for LC-MS analysis.

LC-MS analysis:

LC-MS analysis was performed in Synapt G2-Si High Definition ESI Mass Spectrometer (HDMS, Waters) coupled with Acquity I class UPLC (Waters) as front end. The software, MassLynx™ 4.1 informatics solution was utilized for data acquisition and UNIFI 1.9.4 version for Data processing. Acquity UPLC peptide CSH C18 column (130Å, 1.7 µm, 2.1 mm X 150 mm, Waters) was used for the chromatographic separation step. The temperature of column was kept at 65°C and the autosampler was set at 10 °C. Mobile phase were comprised of 0.1% formic acid in waters (A) and 0.1% formic acid in acetonitrile (B).The flow rate was maintained at 0.2 mL/min. Injection volume of sample was 10 µL. The following gradient programme was used: 0–2 min, 95% A; 2–40 min, 57% A; 40–42 min, 20% A; 42-45 min, 20% A; 45-46min, 95% A, 46-60min, 95% A.

The ionization source of the mass spectrometer was ESI and operated in positive ionization mode with 40 V of sample cone voltage, and 3.0 kV of capillary voltage. Desolvation temperature was 350 °C, and ion source temperature was 120 °C. Desolvation gas flow and cone gas flow were 800 (L/Hr) and 50 (L/Hr), respectively. Data were acquired in MS^E continuum mode and mass was scanned in the range 50–1995 Da with trap ramp collision energy from 30 to 50 eV. For precise mass detection, leucine-enkephalin ($[M+H]^+ = 556.2771$) was selected as reference lock mass, and the lock-mass solution with the concentration

of 0.2 ng/mL.

Antiviral activity assay:

 The assay was done in a 96-well plate in triplicates for each sample. A total of $1x10^4$ Vero E6 cells were plated per well and incubated at 37 °C overnight for the monolayer formation. Cells were incubated with the different culture medium containing 1to 10 µM N-acetyl cysteine and the highest concentration (10 µM) did not show any cytotoxic. This was followed by the addition of SARS-CoV-2 (USA-WA1/2020 strain) at a 0.01 multiplicity of infection. Control cells are incubated in culture medium with remdesivir at 10 µM concentration. Plates were incubated at 37 °C and culture supernatant was harvested at 24 h and 48 h later. Viral RNA was isolated from 100 µl cell culture supernatant using PureLink Viral RNA/DNA Mini Kit (Invitrogen). The cDNA was prepared using ImProm-II Reverse Transcription System (Promega). Real-time PCR was performed in QuantStudio 6 Flex Real-Time PCR System using TB Green Premix Ex Taq II (TaKaRa). The following primers were used to quantify viral RNA levels: (a) Envelope (E) primers FP 5'-ACAGGTACGTTAATAGTTAATAGCGT-3' RP 5'-ATATTGCAGCAGTACGCACACA-3' (b) Nucleocapsid (N) primers FP 5'-GACCCCAAAATCAGCGAAAT-3' RP-5'-TCTGGTTACTGCCAGTTGAATCTG-3'. The Ct values for N and E gene sequence detection were determined and used for calculating the per cent virus inhibition with respect to the control.

Result and Discussion:

An attempt to impair the binding of the spike protein with human ACE2 receptor might be a feasible strategy to inhibit host infection with SARS-CoV-2. The spike protein interacts with ACE2 receptor through its RBD where the main interacting residues are following: Leu455, Tyr473, Tyr489, Gln493, Asn501, Tyr505 (Figure 1). In general, the tertiary/quaternary structure of a protein can be perturbed by reducing the disulfide bonds between cysteine residues that provide the structural architecture to the molecule. Crystal structure of spike protein (PDB ID: 6LZG) showed that out of nineteen cysteine residues, eight cysteines are covalently bonded by four disulfide bridges. Spike protein-ACE2 complex showed that the disulfide bond between cys480 and cys488 residues of spike protein is located closer to the interacting residues of ACE2 receptor in the complex, whereas other disulfide bonds are at a distant from the site of interaction (Supplementary figure S1). However, an analysis of the solvent

accessible surface area (SASA) of different cysteine residues in the crystal structure (PDB ID: 6LZG) indicated that among the four disulfide bonds of spike protein, the bond between cys391 and cys525 residues is most accessible to the solvent. (Supplementary table S1)

Reduction of surface exposed disulfide bonds in the protein can be achieved through thiol/disulfide exchange mechanism executed via free thiol group (-SH) of a reducing agent such as NAC, a drug that is commonly prescribed as an antioxidant and a mucolytic agent.[22] In the present study, spike protein was extracted from its complex with ACE2 receptor (PDB ID: 6LZG). Based on the SASA values of the disulfide bonds, pocket identification analysis indicated the following sites for the docking of NAC in the spike protein: one near the vicinity of RBD site, pocket 1 (cys480: 17.29 Å and cys488: 7.91 Å) and the other far apart from RBD site, pocket 3 (cys391: 57 Å and cys525: 44.74 Å). (Supplementary figure S2, Supplementary table S1). The solvent accessible area of pocket 1 near to the RBD site was smaller (74.09 $Å^2$) than that of pocket 3 (89.68 $Å^2$). However, during molecular docking, pocket 1 was considered for NAC binding as it was closer to the RBD site. The docking study showed that the H-bonding interaction occurred between NAC and following three residues of the spike protein, Pro479, cys480 and Asn481. In addition, hydrophobic interactions were observed between NAC and nearby residues including Thr478, Pro479, Phe486, cys488. These interactions resulted in a considerable binding affinity (ΔG = -7.32 Kcal.mol^{-1}). (Table 1) Subsequently, to evaluate the stability of the NAC-spike protein complex, Molecular Dynamics (MD) simulation was performed for 100 ns. MD simulation study demonstrated that the molecular trajectories of the spike protein and NAC were not close to each other, due to a shift of NAC from the docking site to the N-terminal site of the spike protein. The Radius of Gyration (RG) and Root Mean Square Division (RMSD) confirmed that these systems followed a harmonious swirl throughout the dynamics (Supplementary figures S3-S5, Supplementary video S1 and S2). After MD simulation, the coordinates of the most important ACE2 interacting amino acid residues changed significantly. Surprisingly, NAC became unglued from the docked site after a 3 ns run of the MD simulation and travelled across the solvent system. Finally, it became bound to the N-terminal site of spike protein. The distance between initial and final positions of NAC was 59.2 Å. The new location of NAC was close to the disulfide bond between the

cys391 and cys525 residues. The binding affinity of NAC with spike protein was improved by two strong H-bond interactions with the pocket residue Thr523. Docking study followed by MD simulation clearly showed that NAC can easily fit into the binding pocket 3 because of its better interaction with the larger pocket volume size in comparison with other binding pockets of the spike crystal protein (Supplementary figure S6). In addition, NAC demonstrated hydrophobic interactions with other pocket residues such as cys361, Val362, and Ala522. From the last trajectory of the MD simulation study, the distance between centroid of NAC with cys391 and cys525 were observed to be 4.4 Å and 3.2 Å respectively (Supplementary figure S7). Thus, it improved the probability of thiol/disulfide exchange between the free thiol group (–SH) of NAC and the cys391-cys525 disulfide bond.

To investigate the effect of NAC on the disulfide connectivity of spike protein *in vitro*, the purified NTD of spike protein was subjected to differential alkylation using N-ethyl maleimide (NEM) in the first step followed by the reduction of disulfide bonds of the NEM modified protein with dithiothreitol and finally alkylation of the reduced protein with iodoacetamide (IAM). Mass analysis of the trypsin digests of the differentially alkylated protein indicated that the cysteine residues within the proteins were in two different environments. When compared with the crystal structure of the spike protein in complexed with ACE2 receptor (PDB ID: 6LZG), the observed NEM versus IAM modification of cysteine residues in the mass spectra indicated that following cysteine residues were disulfide connected in the NTD of spike protein: cys391, cys525, cys336, and cys432. Figure 2, panel A shows two peptide fragments with 679.7 m/z and 790.7 m/z, containing two cysteine residues, cys525 and cys538 modified with IAM and NEM respectively. The observed results correlated with the crystal structure of the protein complex where cys538 was in the free form and cys525 was disulfide bonded with cys391. Panel B, figure 2 shows the proteolytic digest of the NAC treated NTD of spike protein, where the tryptic fragment consisting cys525 appeared with m/z 660.7 and the respective NAC conjugated peptide appeared with m/z 714.4. The tandem mass spectra of the molecular ion with 714.4 m/z in panel C, figure 2 indicates that the proteolytic fragment with m/z 714.4 in panel B originated from a stretch of amino acid residues 510-528 consisting of cys525 modified with NAC. All other modifications of disulfide bonded cysteine residues are shown in the supplementary figures S8-S10. In principle, NAC can form a covalent conjugate with a protein only through thiol/disulfide exchange process between its free sulfhydryl group and the disulfide bonded cysteine residues in the protein. The predicted dissociation equilibrium constant

(Kd) values of protein-protein interaction clearly revealed that upon NAC conjugation, the binding affinity of spike protein with ACE2 was decreased three-fold, with a concomitant decrease in the release of free energy (ΔG) by 5 Kcal.mol^{-1}(Table 1). The change in the number of polar-polar as well as polar-apolar/non polar interactions at the interface between two proteins provided the molecular insights into the decrease in binding affinity of NAC conjugate of spike protein with the ACE2 receptor (Table 1). In the *in silico* study, prior to MD simulation, the reduction of cys391-cys525 bond and subsequent conjugation between NAC with cys525 of the spike protein, the closest cys residue at 3.2 Å apart, was made manually by maintaining the protein coordinates. Upon reduction of the disulfide bond via NAC, the spike protein was unfolded to a significant extent. However, no similar type of unfolding event was observed for other three disulfide bonds of spike protein when those were separately broken. (Supplementary figure 11) This clearly indicated that the cys391-cys525 disulfide bond significantly contributed to the structural integrity of spike protein.

The foregoing information provides important insights into the cleavage of the disulfide bond between cys391 and cys525 residues of the spike protein by the action of NAC, which might impair the binding and consequently, inhibit the SARS-CoV-2 infection. After NAC conjugation with cys525, the energy minimization was carried out. The conformational change of the protein indicated that the unfolding of the spike protein led to an increase in the distance between C- and N-terminus of the NTD of spike protein from 9.9 Å, as per the crystal structure of the complex, to 130.2 Å (Figure 3). This resulted in a significant perturbation of the stereo specific orientation of the interacting residues in the RBD of spike protein with those of ACE2 receptor which eventually might be reflected in a decreased binding affinity of the spike protein with the ACE2 receptor. (Supplementary figure 12)

Antiviral assay in VeroE6 cells showed that NAC caused 54.3% inhibition of SARS-CoV-2 replication at 48 h post-infection with respect to E-gene, whereas the above inhibition was 42.9% with respect to N-gene. Remdesivir, a known SARS-CoV-2 inhibitor, showed 99.8% and 99.9% inhibition of the viral replication with respect to E-gene and N-gene at 48 h post-infection, respectively. The *ex vivo* data indicated that NAC might prevent the replication of SARS-CoV-2 modestly in human.

COVID-19 is primarily a respiratory disease. SARS-CoV-2 enters into human cells via binding with ACE2 receptor on the alveolar epithelium. Patients with the history of cardiovascular disease, diabetes and reduced immune functions are found to be at a higher risk. Severity of the disease has been found to be associated with the up-regulation of cytokines like interleukin (IL)-6, IL-10, interferon (IFN)-γ, tumour necrosis factor (TNF)-α, etc. This increase in the level of cytokines creates a cytokine storm and thereby causing significant lung damage, respiratory failure, cardiovascular alteration and multiple organ failure.[23,24] Exhaustion of T-cell population has been observed in COVID-19.[23] The redox imbalance that leads to pro-inflammatory states mediated by the expression of TNF-α, might result in T cell exhaustion. Human clinical trial demonstrated that the NAC, which is a precursor of reduced glutathione, improves the cellular redox status.[23] NAC has been reported to inhibit gene expression of TNF-α and IL-6.[23] NAC has also been used to prevent and treat microvascular thrombotic events in COVID-19.[25] The multimers of vonW-factor are disintegrated by NAC to act as 'medical revascularization' in patients of COVID-19 with intravascular thrombosis.[25] Based on its antioxidant and anti-inflammatory mechanism, Poe et. al., hypothesized that NAC might be a potential therapeutic molecule to treat COVID-19.[23] In the present study we proposed a different mechanism of action of NAC against SARS-CoV-2. The reduction of the solvent accessible disulfide bond followed by the conjugation of NAC leads to the perturbation of the functionally active structure of spike protein in SARS-CoV-2 and thereby reducing the infectivity of the virus. This effect might work in synergy with the other reported activities of NAC such as antioxidant and anti-inflammatory activity in combating COVID-19. Interestingly, almost all SARS-Cov-2 variants, discovered to date, have been found to conserve the cystine residues in their spike protein. Thus, all variants of SARS-CoV-2 might be expected to have identical pattern of the disulfide bonds. Therefore, the observed inhibition in viral replication indicates that it might be worth investigating the role of NAC in pharmaco-prevention and treatment of COVID-19 in a clinical setting.

Conclusion:

Both *in vitro* and *in silico* analysis indicated that the reduction of the disulfide bridge between cys391 and cys525 residues

followed by the covalent conjugation of NAC with cys525 of spike protein resulted in a significant perturbation that is critical in the interaction between spike protein and ACE2 receptor. The effect of this conformational change was reflected in the antiviral assay where incubation of the NAC treated VeroE6 cell line with SARS-CoV-2 resulted in the prevention in replication of the virus by 54.3% at 48 h post-infection. This action of NAC against COVID-19 might be fruitfully explored in a clinical setting.

Author contribution: UD designed and executed the experiments, analyzed data and wrote the paper. AM performed protein modifications and executed enzyme digestion, designed the mass spectrometry based experiments, analysed data and wrote the paper. VD has done MD simulation experiment. YSP has helped in designing MD simulation work and provided critical comments during manuscript preparation. RT, KK, PW and UV executed the mass spectrometry based experiments and analysed the data. CS, SA and SV designed the antiviral assay, performed the assay and wrote the relevant part of the paper. AVK and PSB and AB provided academic inputs, critical comments and wrote the paper. AKM conceived the idea, designed the experiment and wrote the paper.

Supplementary information: Supplementary information include (1) Distance between disulfide bonds (green color) of the viral spike protein with selected human ACE2 interacting residues (cyan color) present in the binding domain. (2) Predictive binding pockets of spike protein. Different colors indicate the position of binding pockets along with surface area for solvent accessibility. (3) Radius of Gyration (RG) of spike protein after MD simulation. (4) Radius of Gyration (RG) of spike-ligand (NAC) after MD simulation. (5) Root mean square fluctuation of ligand (NAC) during MD simulation (6) Figures of different binding sites of NAC before and after MD simulation. (**7**) Distance (in Å) of NAC with disulfide bond forming cysteine residues (cys391 and cys525) at N-terminal site of the spike protein after MD simulation. (8) Representation of NAC modification of the tryptic peptide containing cys336. Panel A shows the unmodified peptide that appeared with m/z 1020.5 $(M + 2H)^{2+}$. Panel B shows the NAC conjugated peptide that was observed with m/z 1101.01 $(M + 2H)^{2+}$. Panel C represents MS/MS spectra of the peptide with m/z 1101.01 with b and y ions labelled confirming that NAC was conjugated with cys336 in the peptide. (9) Representation of NAC modification of the tryptic peptide containing cys432. Panel A shows the unmodified peptide that appeared with m/z 737 $(M + 3H)^{3+}$. Panel B shows the NAC conjugated peptide that was observed with m/z 790.6 $(M + 3H)^{3+}$. Panel C represents the MS/MS spectra of the peptide with m/z 790.6 with b and y ions labelled confirming that NAC was conjugated with cys432 in the peptide. (10) Representation of NAC modification of the tryptic peptide containing cys391. Panel A shows the unmodified peptide that appeared with m/z 995.4 $(M + 2H)^{2+}$. Panel B shows the NAC conjugated peptide that was observed with m/z 1075.9 $(M + 2H)^{2+}$. Panel C represents the MS/MS spectra of the peptide with m/z 1075.9 confirming that NAC was conjugated with cys391 in the peptide. (11) Structural deformations after individual thiol formation at four different disulfide positions (across eight cysteine residues) of spike protein. (12) Changes in stereo specific orientation of the interacting residues in the RBD of spike protein after MD simulation. (Table S1) Cysteine residues based solvent accessibility analysis of spike protein using a crystal structure (PDB ID: 6LZG).

Supplementary Videos: Supplementary videos include (1) Removal of NAC from the docking site of the viral spike protein and (2) binding affinity of NAC with NTD of the viral spike protein during MD simulation up to 100 ns.

Notes: Authors declare no conflict of interest.

References:

(1) Renhong, Y.; Yuanyuan, Z.; Li, Y.; Lu, X.; Yingying G.; Qiang, Z. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*, **2020**, 367: 1444–1448.

(2) Jun, L.; Jiwan, G.; Jinfang, Y.; Sisi, S.; Huan, Z.; Shilong, Fan.; Qi, Z.; Xuanling, S.; Qisheng, W.; Linqi, Z.; Xinquan, W. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, **2020**, 581, 215–220.

(3) Junwen, L.; Yue, L.; Xiaolu, J.; Leiliang, Z. Spike protein recognition of mammalian ACE2 predicts the host range and an optimized ACE2 for SARS-CoV-2 infection. *Biochem Biophys Res Commun*. **2020**, 21, 526(1), 165–169.

(4) Alexandra C, W.; Young-Jun Park, M.; Alejandra, T.; Abigail, W.; Andrew, T.; McGuire, D. V. Structure, Function, and Antigenicity of the SARSCoV-2 Spike Glycoprotein. *Cell*, **2020**, 180, 1–12.

(5) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.Y.; Wang, Q.; Zhou, H.; Yan, J.; Qi, J. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell*, **2020**, 181, 894-904.e9.

(6) Pedretti A.; Villa L.; Vistoli G. VEGA - An open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming. *J Comput Aided Mol Des.*, **2004**, 18, 167-173.

(7) Binkowski, T. A.; Naghibzadeh, S.; Liang, J. CASTp: Computed Atlas of Surface Topography of Proteins. *Nucleic Acids Res*. **2003**, 31, 3352−3355.

(8) Morris, G.M.; Huey R, Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem*. **2009**, 30, 2785–

2789.

(9) Sybyl-X 1.3, St. Louis, 2010. http://www.tripos.com.

(10) Debnath, U.; Kumar, P.; Agarwal, A.; Kesharwani, A.; Gupta, S.K.; Katti, S.B. N-hydroxy- substituted 2-aryl acetamide analogs: A novel class of HIV-1 integrase inhibitors. *Chem Biol Drug Des*. **2017**, 90, 527–534.

(11) Seeliger, D.; de Groot, B.L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J Comput Aided Mol Des*. **2010**, 24, 417–422.

(12) MOE; The Molecular Operating Environment from Chemical Computing Group Inc., Montreal, Quebec, Canada; http://www.chemcomp.com.

(13) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, 26, 1781− 1802.

(14) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; & Onufriev, A. H++: A server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **2005**, 33(Web Server), W368–371.

(15) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; & Karplus,
M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, **1983**, 4(2), 187–217.

(16) https://www.paramchem.org/

(17) Dewaker, V.; Srivastava, P.N.; Verma, S.; Prabhakar, Y.S. Molecular dynamics study of HDAC8-largazole analogues co-crystals for designing potential anticancer compounds. *J Biomol Struct Dyn.* **2019**, 38, 1197-1213.

(18) Li, Y.; Cong, Y.; Feng, G.; Zhong, S.; Zhang, J. Z. H.; Sun, H.; Duan, L. The impact of interior dielectric constant and entropic change on HIV-1 complex binding free energy prediction. *Structural Dynamics* (Melville, N.Y.), **2018**, 5(6), 064101.

(19) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J Mol Graph Model.* **1996**, 14(1), 33–38.

(20) Yan, Y.; Zhang, D.; Zhou, P.; Li, B.; Huang S.Y. HDOCK: a web server for protein-protein and protein -DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res.* **2017**, 45, W365- 373.

(21) Xue, L. C.; Rodrigues, J. P.; Kastritis, P. L.; Bonvin, A. M.; Vangone, A. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics*, **2016**, 32(23), 3676-3678.

(22) Bürger, M; Chory1, J; Structural and chemical biology of deacetylases for carbohydrates, proteins, small molecules and histones. *Commun Biol.* **2018**, 217, 1-11.

(23) Poe, F.L.; Corn J. N-Acetylcysteine: A potential therapeutic agent for SARS-CoV-2. *Med Hypotheses*. 2020**,** 143, 109862-109865.

(24) Flora, S.D.; Balansky, R.; Maestra, S.L. Rationale for the use of N-acetylcysteine in both prevention and adjuvant therapy of COVID-19. *The FASEB journal*. **2020**, 34(10), 13185-13193.

(25) de Lizarrondo S.M., Gakuba C., Herbig B.A., Repessé Y., Ali C., Denis C.V., Lenting P.J., Touzé E., Diamond S.L., Vivien D., Gauberti M. Potent thrombolytic effect of N-acetylcysteine on arterial thrombi. *Circulation.* **2017,** 136(7), 646–660.
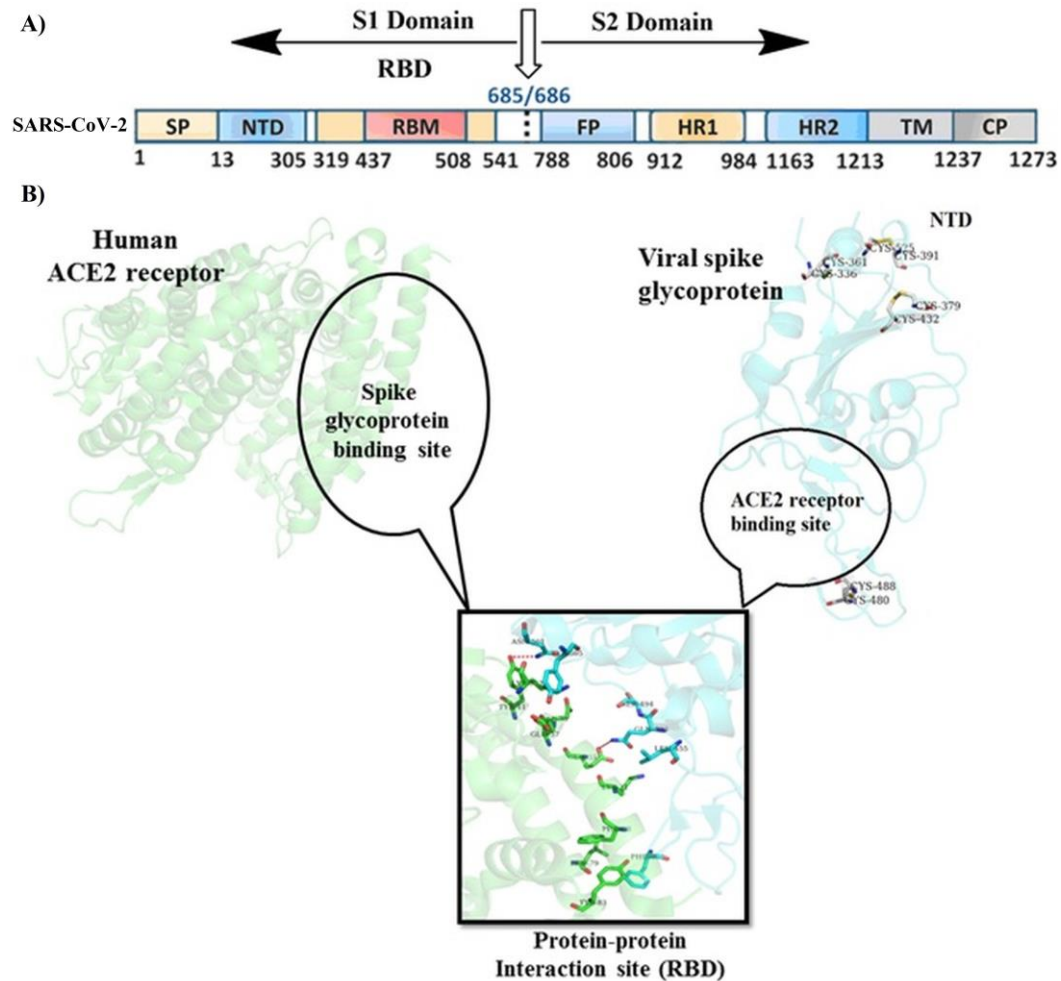
Figure 1: (**A**) Representing the position of RBD site at S1 domain of SARS-CoV-2 . (**B**) Selected important binding sites of protein-protein interaction in between ACE2 receptor (green color) and viral spike protein (cyan color). NTD indicates N-terminal domain where three disulfide bonds (-S-S-) are located within six cysteine residues (Cys336-Cys361; Cys379- Cys432; Cys391-Cys525). The other disulfide bond (-S-S-, Cys480-Cys488) is placed near to ACE2 receptor binding site.
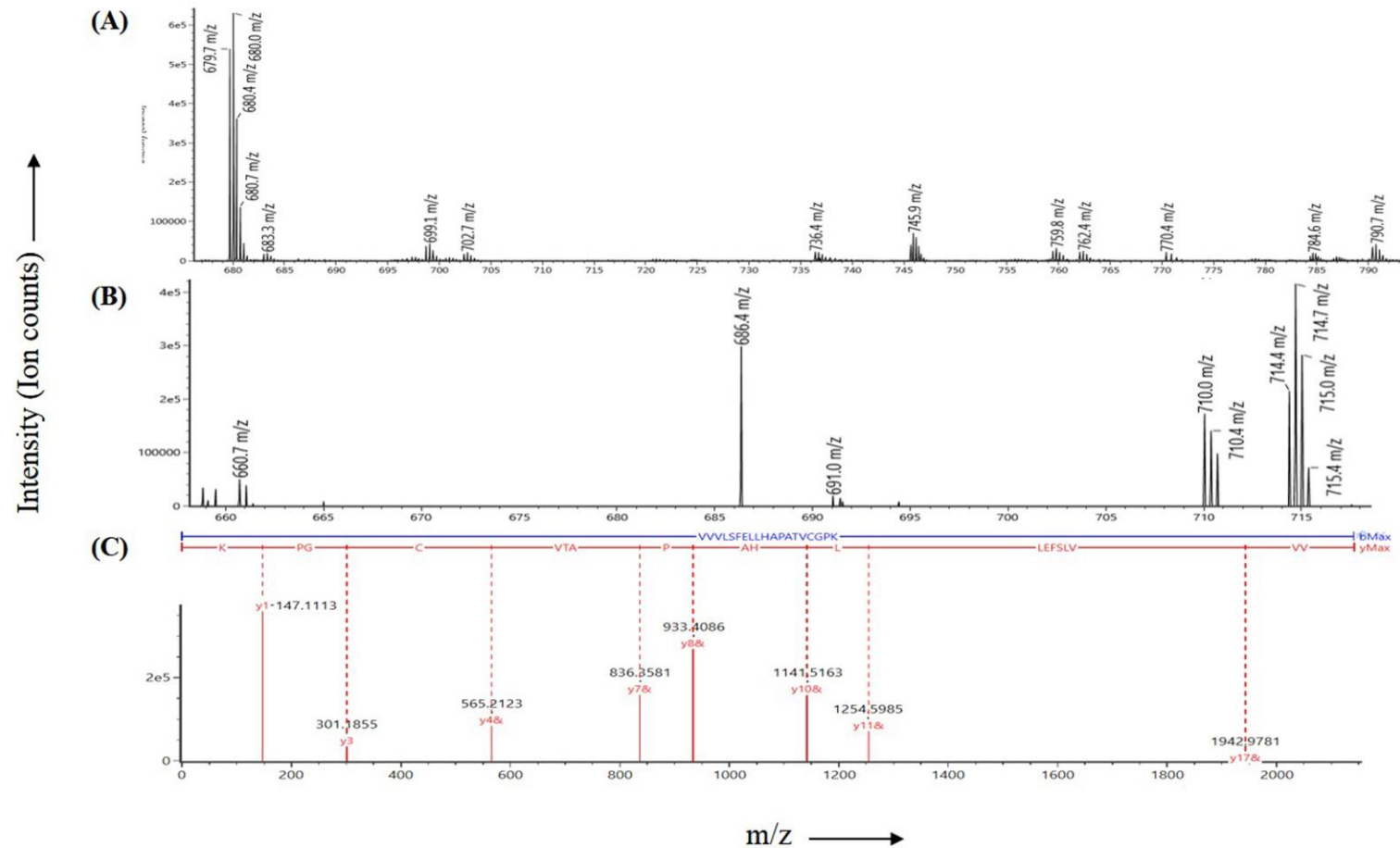
Figure 2: Representation of the differential chemical modifications of two peptides containing cys525 and cys538 residues respectively. Panel A shows the peptide with m/z 679.7 (M +3H)$^{3+}$ obtained upon conjugation of IAM to cys525, and peptide with m/z 790.7 (M +2H)$^{2+}$ obtained upon modification of cys538 with NEM. Panel B shows the tryptic peptide with m/z 714.4 (M +3H)$^{3+}$ obtained upon conjugation of NAC with cys525. Panel C represents the MS/MS spectra of the peptide with m/z 714.4 where the fragment ions, b and y, are labelled
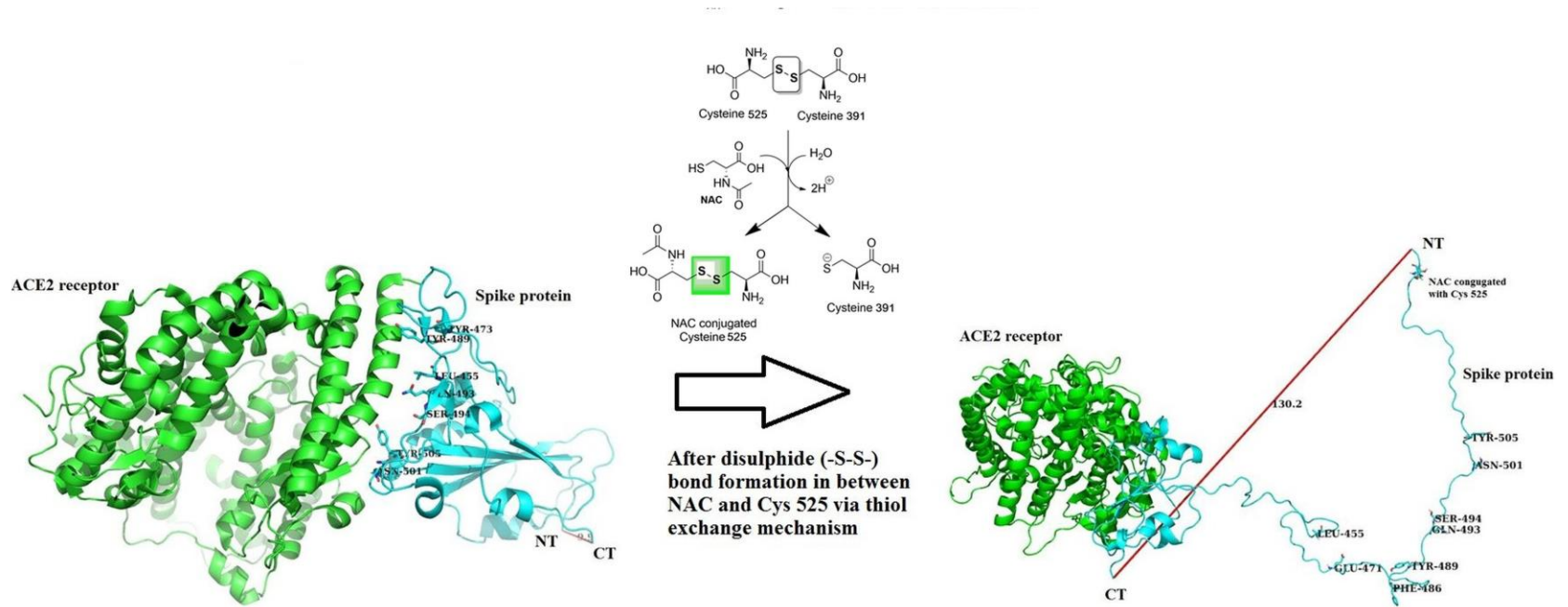
Figure 3: Difference in protein–protein interaction between ACE2 receptor and spike protein before and after NAC conjugation via thiol exchange mechanism: (A) shows interaction between the crytal forms of ACE2 receptor and NAC free spike protein whereas figure (B) shows *in silico* protein–protein interactions between ACE2 receptor and NAC conjugated spike protein. The red dotted lines (---) in (A) and (B) displays the distance between C terminal and N terminal of spike protein, before (9 Å) and after (130.2Å) NAC conjugation.

Table 1: NAC binding pocket analysis with the details of Details of protein-protein interaction contacts (ICs) after MD simulation study.

NAC binding pocket study after docking			NAC binding pocket study after MD simulation		
Pocket residues	H-bond interaction sites	Hydrophobic interaction sites	Pocket residues	H-bond interaction sites	Hydrophobic interaction sites
Cys488, Gly485, Thr478, Phe486, Pro479, Cys480, Asn481, Val483	Pro479, Cys480, Asn481	Phe486, Pro479	Asn334, Val362, Cys361, Cys391, Thr523, Ala522, Pro527, Cys525	Thr523	Val362, Ala522
Spike glycoprotein-ACE2 receptor(human) interaction after NAC conjugation (_in silico_ form)					
1	ΔG (kcal mol^{-1})				-7.48
2	K_d (M) at 25.0 °C				2.2E-09
3	Number of ICs (charged-charged)				3
4	Number of ICs (charged-polar)				7
5	Number of ICs (charged-apolar)				15
6	Number of ICs (polar-polar)				4
7	Number of ICs (polar-apolar)				14
8	Number of ICs (apolar-apolar)				6

Table of Content